S

Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools

Subhamoy Bhattacharya
University of Surrey, Guildford, UK

Synonyms

Case study of pile damage during the 1995 Kobe earthquake; Geotechnical analysis of pile foundation in liquefiable soils; Numerical Modeling; Pile foundations in liquefiable soils; Reliability of pile foundations; Seismic Analysis of Pile Foundations; Soil-Structure Interaction

Introduction

Piles are routinely used as foundations to support short-to-medium span bridges and buildings typically over four stories and other structures. Collapse and/or severe damage of pile-supported structures is still observed in liquefiable soils after most major earthquakes such as the 1995 Kobe earthquake (Japan), the 1999 Kocaeli earthquake (Turkey), and the 2001 Bhuj earthquake (India). The failures not only occurred in laterally spreading (sloping) ground but were also observed in level ground where no lateral spreading would be anticipated. A good discussion on the failure modes can be found in Bhattacharya and Madabhushi (2008).

The failures were often accompanied by settlement and tilting of the superstructure, rendering it either useless or very expensive to rehabilitate after the earthquake. Following the 1995 Kobe earthquake, investigations have been carried out to find the failure pattern of the piles. Piles were excavated or extracted from the subsoil, borehole cameras were used to take photographs, and pile integrity tests were carried out. These studies hinted at the location of the cracks and damage patterns for the piles. Of particular interest is the formation of plastic hinges in the piles. This indicates that the stresses in the pile during and after liquefaction exceeded the yield stress of the material of the pile, despite large factors of safety which were employed in the design. Hinges were found to have occurred at various depths along the pile: at the pile head, at the middle of the liquefiable layer, and toward the interface of liquefiable/non-liquefiable layer. In this context it must be mentioned that piles are currently designed with adequate factor of safety for geotechnical load-carrying capacity (maximum allowable load in the pile) and against bending failure. Bending moment can occur in a pile due to lateral loads arising from: (a) inertia load acting at the pile head and (b) kinematic loads from the ground. Further discussion on the dynamics of the problem can be found in Adhikari and Bhattacharya (2008).

Liquefaction of soil around the pile can affect a pile-supported structure in the following ways:

1. A pile will be laterally unsupported in the zone of liquefaction, i.e., no lateral restraint from the soil to the pile in the liquefied zone. As a result, a pile may be vulnerable to buckling instability if the axial load is high enough and the unsupported length of pile is sufficient.
2. The pile will lose its shaft resistance in the zone of liquefaction, i.e., no liquefied soil-pile friction, and as a result the axial load on the pile in the liquefiable zone will increase.
3. The time period of the structure will change due to liquefaction as the foundation becomes flexible. While calculating the period of a building under non-liquefied condition, it is assumed that foundation is rigid, and as a result only the dimensions of the buildings are adequate to obtain the period. However, as soon as a soil liquefies, the time period may increase.
4. The damping of a pile-supported structure also increases during and after liquefaction.

The above mechanisms were verified experimentally by carefully designed model tests and can be found in Bhattacharya (2003), Bhattacharya et al. (2004), Bhattacharya et al. (2005), and Lombardi and Bhattacharya (2014). Bhattacharya (2003) showed through dynamic centrifuge tests that axial load alone can cause a pile to fail if the surrounding soil liquefies in an earthquake and the mechanism being buckling instability.

Lombardi and Bhattacharya (2014) showed through high-quality shaking table tests that the time period of pile-supported structures will increase owing to liquefaction. In the experiments, the soil was liquefied progressively through a broadband white noise signal with increasing magnitude. Furthermore, they also showed that the overall damping ratio of the structures may increase in excess of 20 %. These have important design implications.

All current design methods, such as JRA (1996), NEHRP (2000), IS 1893 (2001), and Eurocode 8 (CEN 2004), focused on bending strength of the pile to avoid bending failure due to lateral loads (combination of inertia and lateral spreading). In contrast to these conventional design codes, which advocate bending mechanism as the main design consideration, recent research showed that an axially loaded pile can be laterally unsupported in liquefied soils and is susceptible to buckling failure.

Buckling instability under the interaction of axial and lateral loads can be a more critical design consideration because of its sudden nature and sensitivity to imperfection; see, for example, Dash et al. (2010). More recently, Bhattacharya et al. (2009) included the effects of dynamics on the combined axial and lateral loads on a pile foundation. Essentially, piles in liquefied soils may be better regarded as columns carrying lateral loads rather than laterally loaded beams.

In design, beam bending and column buckling are approached differently. Bending is a stable mechanism as long as the pile remains elastic and secondary failure (e.g., local buckling) is not a possibility. This failure mode depends on the bending strength (e.g., yield moment capacity and plastic moment capacity) of the member under consideration. In contrast, buckling is an unstable mechanism and it occurs suddenly and drastically when the elastic critical load is reached. It is the most destructive mode of failure and depends on the geometrical properties of the member, i.e., slenderness ratio, rather than the member strength. For example, steel pipe piles with identical length and diameter but having different yield strengths (e.g., 200 MPa, 500 MPa, and 1,000 MPa) will buckle at similar axial loads but can resist different amounts of bending. In other words, bending failure may be avoided by increasing the yield strength of the material, but it may not suffice to avoid buckling. To prevent buckling failure, there should be a minimum pile diameter depending on the depth of liquefiable soils. Therefore, designing against bending would not automatically satisfy buckling requirements. It is envisaged that there are plenty of existing pile-supported structures that may need retrofitting.

This entry therefore describes a probabilistic and a deterministic method to assess the reliability of pile foundation for a scenario earthquake. Before the methodology is described, a discussion is presented on the main loading on pile foundations during earthquakes in liquefiable soils.

Methodology

Different Stages of Loading on a Pile-Supported Structure During Earthquakes

Figure 1 shows the different stages of loading of a pile-supported structure during a seismic liquefaction-induced event. $P_{gravity}$ (Stage I) represents the axial load on the piles in normal condition. This can be estimated based on static equilibrium. This axial compressive load may increase/decrease further by $V_{inertial}$ due to inertial effect of the superstructure and kinematic effects of the soil flow past the foundation. $H_{inertial}$ is the inertial lateral loads due to the oscillation of superstructure (Stage II and III). Ground movement causes kinematic loads on the pile foundations. This load can be of two types: transient (during shaking, due to the dynamic effects of the soil mass) and residual (after the shaking ceased due to soil flow, often known as "lateral spreading") (Stage IV). The various forms of feasible failure mechanisms of pile foundations are shear failure, bending failure, buckling instability, and dynamic failure.

Formulation

Bhattacharya (2006) discusses the deterministic approach to determine the factor of safety of pile foundations against the buckling instability failure. Bhattacharya and Goda (2013) developed a probabilistic procedure for determining the occurrence of a buckling failure of existing piled foundations due to a scenario earthquake.

The methodology is based on assessing two length parameters:

1. Critical length of a pile denoted (H_C). Essentially this is the unsupported length of the pile that can sustain without collapse due to combined axial and lateral loading. This can be estimated based on manipulation of the Euler's buckling load equation considering the correct boundary condition of the pile. Specifically H_C depends on the type and dimension of superstructure (bridge or building), bending stiffness, axial load acting on the pile, dynamic characteristics of
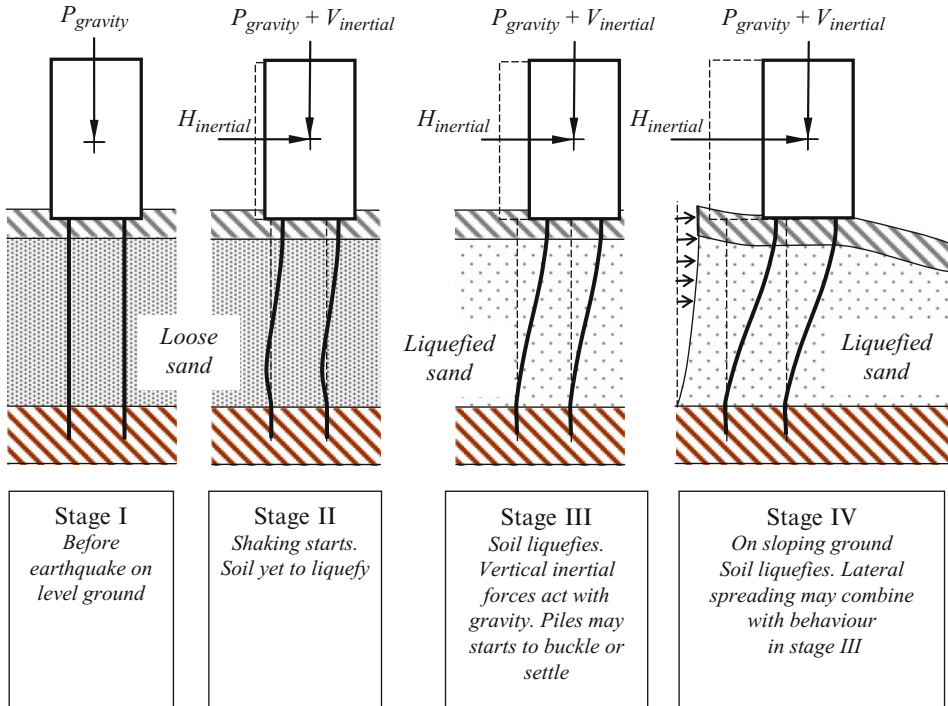
Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools, Fig. 1 Different stages of loading and failure mechanism of pile during earthquake

**Safety Assessment of
Piled Buildings in
Liquefiable Soils:
Mathematical Tools,
Fig. 2** Concept of critical
length (H_C) and
unsupported length (D_L)
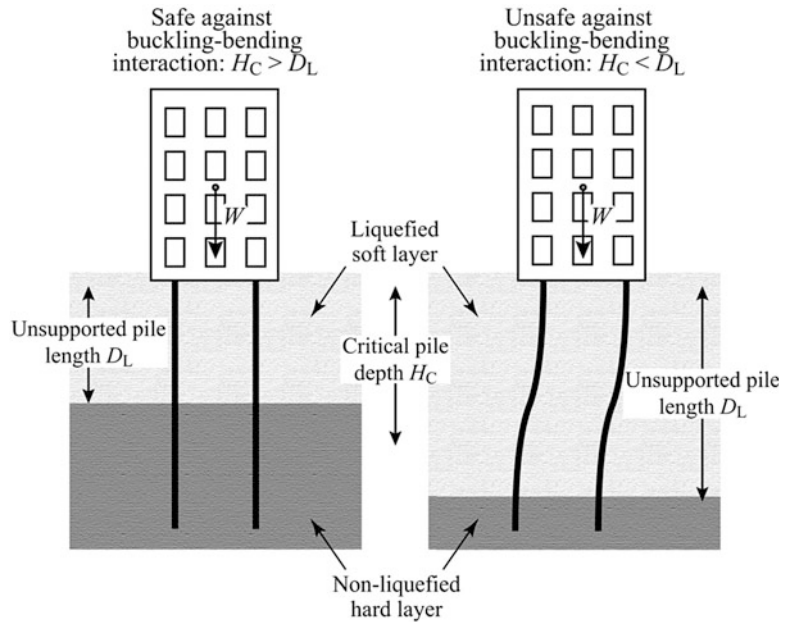
superstructure, and boundary conditions of the pile at the top and bottom of the liquefiable layer.

2. Length of unsupported pile due to liquefaction (D_L): this can be obtained based on the depth of liquefaction due to the design earthquake.

Figure 2 shows the concept of the two lengths. Comparing H_C with D_L, potential failure of the underground pile due to buckling is predicted when $H_C < D_L$ (see Figs. 1 and 2). In reliability terms, H_C is the capacity variable, while D_L is the demand term, and thus the failure criterion can be regarded as the limit state function:

$$g = (H_C - D_L) \quad (1)$$

Probabilistic and Deterministic Approach

In a deterministic approach, we obtain a single value of the two parameters and can obtain the factor of safety against failure. However, the advantages of assessing using probabilistic approach are (i) the outcome is expressed as a (estimated) likelihood of failure by taking into account various sources of uncertainty involved in the assessment; (ii) sensitivity analysis can be conducted to identify key factors that affect the outcome (the results are useful to improve the

methodology and tool); and (iii) the probabilistic framework provides a straightforward way to integrate geotechnical assessment tools into existing probabilistic seismic hazard methods as well as decision-support tools for implementing earthquake risk mitigation measures (Fig. 3).

Step-by-Step Methodology

The steps are:

(a) **Estimation of peak ground acceleration (PGA) at building/bridge site**: The assessment of D_L requires (i) estimation of ground motion parameters (typically, peak ground acceleration (PGA)) at a building site due to a specified scenario and (ii) assessment of liquefaction initiation potential at different depths along the pile length. Equation 2 is a typical equation to obtain PGA.

$$Log\,PGA_{bldg} = f\left(M_w, R_{bldg}, V_{S30,\,bldg}\right) + \varepsilon_{bldg} \quad (2)$$

where PGA_{bldg} is PGA value at the building site, M_w is the moment magnitude, R is the distance measured (typically closest distance to fault rupture plane), V_{S30} is the average shear-wave velocity in the uppermost 30 m in (m/s), and ε_{bldg} is the random error that is
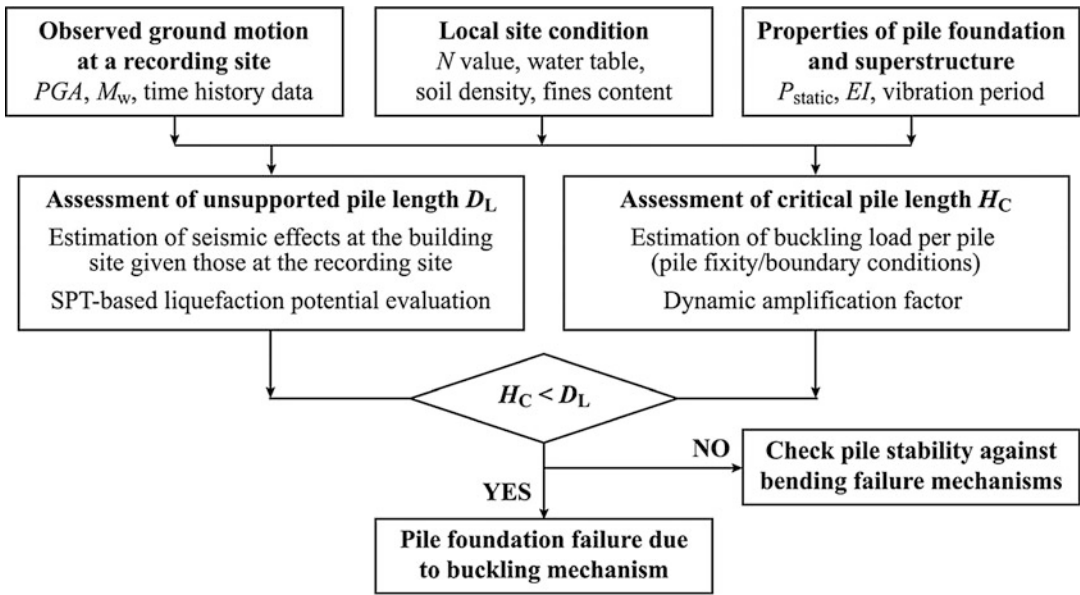
Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools, Fig. 3 Schematic diagram showing the methodology

modeled as a normal variable with zero mean and logarithmic standard deviation.

(b) **Assessment of the depth of liquefiable soil**: The next step is to conduct probabilistic liquefaction potential evaluation of a soil column (along a pile) at the building site to estimate D_L for a given seismic excitation level. Such assessment is often conducted by using simplified stress-based methods of Seed and Idriss (1971) based on standard penetration test (SPT) data, cone penetration test (CPT) data, and shear-wave velocity (V_S) data. Recently, probabilistic procedures for liquefaction initiation have been developed by considering different in situ measures for describing soil strength. Using a sophisticated Bayesian regression analysis and well-screened case

studies, Cetin et al. (2002, 2004) developed a statistical model for calculating the probability of liquefaction initiation based on SPT data, while Moss et al. (2006) developed a counterpart using CPT data. Using the first-order reliability method, Juang et al. (2005) developed a similar model based on V_S data. The significance of these developed models is that key uncertainties associated with the input data/parameters and the adopted models themselves are taken into account; the developed models can produce unbiased potential of liquefaction initiation and are useful to conduct probabilistic liquefaction hazard analysis (Goda et al. 2011). The probability of liquefaction initiation P_L at a depth of interest can be estimated as

$$P_L = \Phi\left(-\frac{N_{1,60}(1+0.004FC) - 13.32\ln(CSR_{eq}) - 29.53\ln(M_w) - 3.70\ln(\sigma_v'/P_a) + 0.05FC + 16.85}{2.7}\right)$$
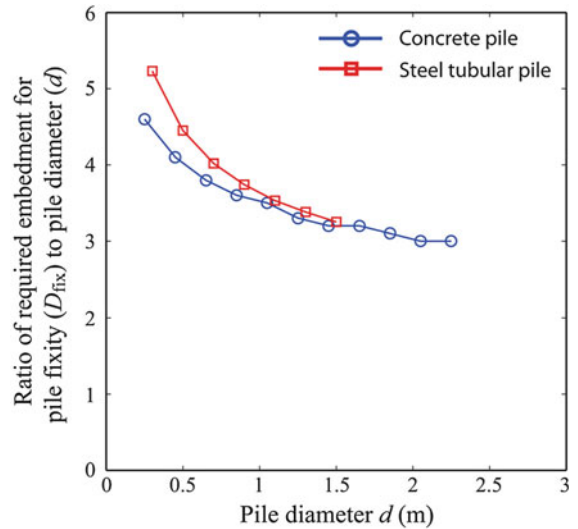
(3)

where Φ is the standard normal function, $N_{1,60}$ is the corrected SPT counts (but not

adjusted for fines content), FC is the fines content (in percentage), CSR_{eq} is the cyclic

Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools, Fig. 4 Depth required for fixity a pile

stress ratio (but not adjusted for the moment magnitude), and σ'_v is the vertical effective stress.

(c) **Unsupported length of the pile (D_L):** The next step is to determine the unsupported pile length (D_L) based on the liquefaction profile. D_L is equal to the thickness of liquefied soil layers plus some additional length necessary for fixity at the bottom of the liquefied soils. Typical calculations show that the fixity depth is about three to five times the diameter of the pile (see Fig. 4); see Bhattacharya and Goda (2013). If a relatively thin non-liquefied layer is overlain and underlain by thick liquefied layers, lateral restraint of the pile at the non-liquefied layer might not be expected, i.e., the pile is unsupported. In such a case, the thin non-liquefied layer can be ignored in determining D_L, and the unsupported length needs to be extended until a thick non-liquefied layer is encountered. However, if there is a complex soil profile with alternating liquefiable and non-liquefiable soil layers, more detailed analysis is required.

(d) **Assessment of the critical pile length H_C:** This section describes the mathematical background behind the rationale of obtaining the critical pile depth H_C. Before the onset of shaking, the static axial load P_{static} acts on

each pile beneath a building, assuming that each pile is equally loaded during static condition and neglecting any eccentricity of loading. During an earthquake, inertial action of the superstructure imposes the dynamic axial load on the piles, which will increase the axial load on some piles. These piles with increased axial loads may be vulnerable to buckling. An estimate of the maximum axial compressive load acting on a pile can be given by

$$P_{dynamic} = (1 + \alpha)P_{static} \qquad (4)$$

where α is termed as the dynamic axial load factor and is a function of type of superstructure, height of the center of mass of the superstructure, and characteristics of the earthquake shaking (e.g., frequency content and amplitude).

For buckling instability analysis, each pile needs to be evaluated with respect to its end conditions, i.e., fixed, pinned, or free. Each pile in a group of identical piles will have the same buckling resistance as a single pile. If a group of piles is fixed in a stiff pile cap and embedded sufficiently at the tip, as in Fig. 2, the pile group will buckle in side sway.

The elastic critical load of a single pile, P_{cr}, can be estimated as

Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools, Table 1   Values of K in Eq. 7

Boundary condition of the pile at the top and bottom of the liquefied layer		Effective length	K	Examples
Top	Bottom			
Fixed	Fixed [sufficient embedment at the dense layer]	$L_{\text{eff}} = 0.5H_c$	0.5	Pile groups with raked piles
Free to translate but restrained against rotation – sway frame	Pinned [insufficient embedment at the dense layer]	$L_{\text{eff}} = 2H_c$	2	See, for example, NFCH (Niigata Family Court House) building in Bhattacharya (2003)
Free to translate but restrained against rotation – sway frame	Fixed [sufficient embedment at the dense layer]	$L_{\text{eff}} = H_c$	1	Most cases fall under such category; see, for example, Fig. 2
Fixed in direction but free to rotate	Fixed [sufficient embedment at the dense layer]	$L_{\text{eff}} = 0.7H_c$	0.7	Pile groups with raked piles. Improper pile-pile cap connection
Fixed in direction but free to rotate	Pinned [less embedment at the dense layer]	$L_{\text{eff}} = H_c$	1	Pile groups with raked piles. Improper pile-pile cap connection
Free, i.e., unrestrained against rotation and displacement	Fixed [sufficient embedment at the dense layer]	$L_{\text{eff}} = 2H_c$	2	Piles in a row such as the Showa Bridge piles

$$P_{cr} = \frac{\pi^2 EI}{\left(L_{\text{eff}}\right)^2} \qquad (4)$$

where L_{eff} is the effective length, i.e., Euler's equivalent buckling length of a strut pinned at both ends, and EI is the bending stiffness of the pile. The effective length of the pile (L_{eff}) can be found in Table 1 or from any structural mechanics textbook or codes of practice. The unsupported length of the pile D_L is equal to the thickness of liquefiable soil plus some additional length necessary for fixity at the bottom of the liquefiable soil.

The applicability of the elastic critical load, as in Eq. 4, to pile buckling failure is an important factor. Experiments show that the actual failure load of a slender column is much lower than that predicted by Eq. 4. Rankine (1866) recognized that the actual failure involves an interaction between elastic and plastic modes of failure. Lateral loads and inevitable geometrical imperfection lead to creation of bending moments in addition to axial loads. Bending moments have to be accompanied by stress resultants that diminish the cross-sectional area available for carrying the axial load; thus the actual failure load is likely to be less than the elastic critical load,

i.e., $P_{\text{failure}} < P_{\text{cr}}$. Equally, the growth of plastic bending zones reduces the effective elastic modulus of the section, thereby resulting in the decreased critical load for buckling (i.e., capacity). Furthermore, these processes feed each other, and as elastic critical loads are approached, all bending effects are magnified. Stability analysis of elastic columns showed that if lateral loads in the absence of axial load would create a maximum lateral displacement δ_0 in the critical mode shape of buckling, then the displacement δ under the same lateral loads but with the concurrent axial load P is given by

$$\frac{\delta}{\delta_0} = \frac{1}{1 - P/P_{cr}} \qquad (5)$$

The term δ/δ_0 is the buckling amplification factor (i.e., amplification of lateral displacement due to the presence of the axial load). Figure 5 presents a graph of the buckling amplification factor plotted against the normalized axial load P/P_{cr}, where P denotes the applied axial load. It can be observed from Fig. 5 and Eq. 6 that if the applied load is 50 % of P_{cr}, the amplification of lateral deflection due to lateral loads is about two times. At these large deflections, secondary moments

**Safety Assessment of
Piled Buildings in
Liquefiable Soils:
Mathematical Tools,
Fig. 5** Buckling
amplification factor versus
normalized axial load
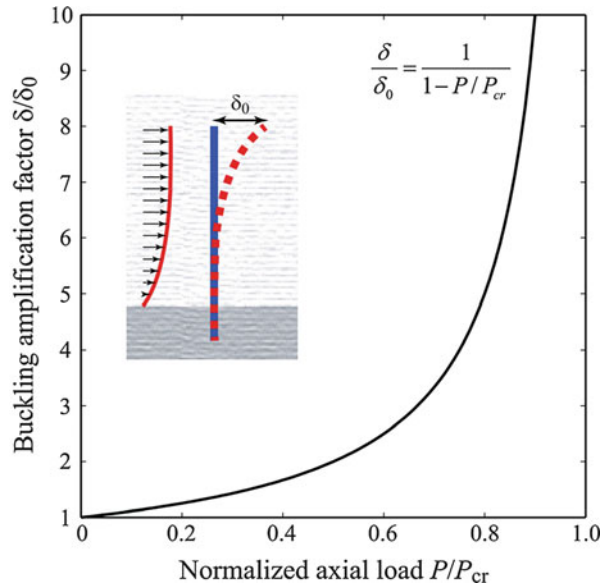
will be generated due to P-Δ moment, leading to more deflection. It is therefore important to remain in the linear regime and not in anyway near the asymptotic region, where the buckling amplification factor increases dramatically (e.g., $P/P_{cr} > 0.6$).

Moreover, it would also be unwise to use a factor of safety less than three against the Euler load of a pile, i.e., ($P/P_{cr} = 0.33$). Such consideration is consistent with general design practice where structural engineers use a factor of safety of at least three against linear elastic buckling to take into account the eccentricity of load, deterioration of elastic stiffness due to plastic yielding, and unavoidable imperfection. The actual failure load $P_{failure}$ is therefore some factor ϕ ($\phi < 1$) times the theoretical Euler's buckling load given by Eq. 5.

$$P_{failure} = \phi P_{cr} \qquad (6)$$

Based on the above discussion, it may be inferred that buckling instability is initiated at around 0.35, i.e., $\phi = 0.35$. It is noted that in reality, this factor depends on the axial load, imperfection of piles, and residual stress in the pile due to driving. The selection of ϕ is one of the significant sources of uncertainty in determining the critical pile depth H_C.

Determination of "Critical Depth" H_C

In the limit state condition of failure, $P_{dynamic} = P_{failure}$. For the type of structure shown in Fig. 2, $L_{eff} = H_C$ in Eq. 4. In order to generalize the boundary condition of the pile (i.e., pile head fixity condition with pile cap/superstructure and the fixity at the interface between liquefiable and non-liquefiable layers at deeper depths), one may write Eq. 7:

$$L_{eff} = K \times H_C \qquad (7)$$

where K is the column effective length factor (e.g., $K = 1$ for free lateral translation but restrained against rotation-sway motion). Values of K for other boundary conditions of the pile are given in Table 1.

With the abovementioned assumptions, Eq. 6 can be rewritten as

$$P_{dynamic} = \phi P_{cr} = \frac{\phi \pi^2 EI}{K^2 H_C^2} \qquad (8)$$

Rearranging Eq. 8 gives the estimate of the critical depth H_C for a pile as follows:

$$H_C = \sqrt{\frac{\phi \pi^2 EI}{K^2 P_{dynamic}}} = \sqrt{\frac{\phi \pi^2 EI}{K^2 (1 + \alpha) P_{static}}} \qquad (9)$$
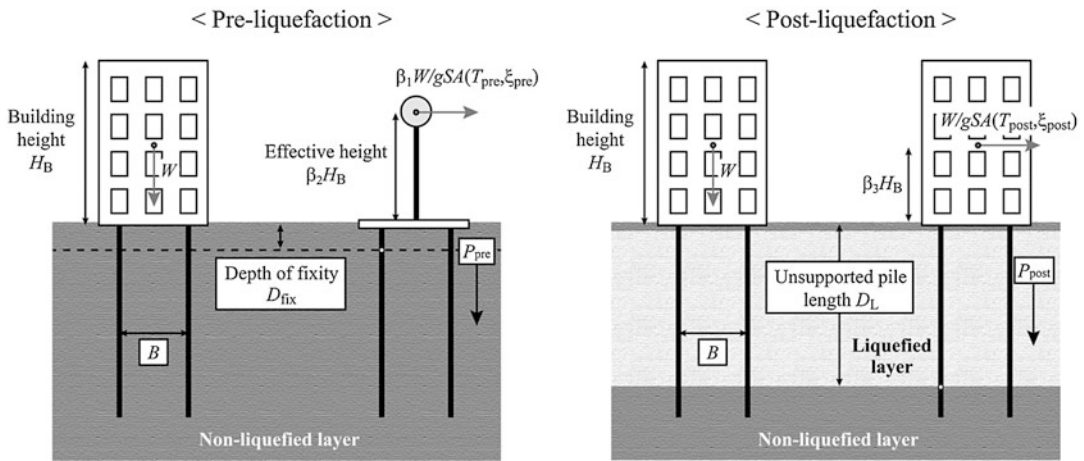
Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools, Fig. 6 Loading condition of pile foundation in pre-liquefaction and post-liquefaction stages

Therefore, the assessment of H_C is based on the calculation of the critical buckling load for a pile foundation surrounded by liquefied soils (from which the pile cannot receive sufficient lateral support). Key parameters in Eq. 9 are ϕ, K, and α. Reasonable values of ϕ and K can be selected based on engineering judgment.

Based on the procedure described above, both D_L and H_C can be assessed probabilistically. The occurrence of pile foundation failure due to buckling mechanism is indicated if $H_C < D_L$. By sampling D_L and H_C many times, probabilistic assessment of liquefaction-induced pile foundation failure due to buckling mechanism can be carried out.

Formulation for Determination of Dynamic Load Amplification Factor (α)

One of the critical factors is the assessment of α. This can be facilitated by examining axial forces acting on a pile in pre-liquefaction and post-liquefaction situations as shown in Fig. 6. In a pre-liquefaction stage, the plane of fixity of a building with pile foundation surrounded by non-liquefied soils can be estimated by following a similar procedure as shown in Fig. 4, i.e., depth from ground surface to the plane of fixity is denoted by D_{fix} in Fig. 6 and is typically a few meters. By assuming that the natural vibration period of the building before liquefaction is T_{pre},

the maximum axial force acting on a pile due to inertia can be calculated as

$$P_{pre} = \frac{\beta_1(W/g) \times SA(T_{pre}, \xi_{pre}) \times (D_{fix} + \beta_2 H_B)}{(N_p/2) \times B}$$

(10)

where W is the total weight of a building (note: $W = P_{static}N_p$), $SA(T_{pre}, \xi_{pre})$ (g) is the spectral acceleration at T_{pre} with damping ratio ξ_{pre} (typically, 2–5 %), β_1 is the coefficient to account for modal mass for the fundamental vibration mode (typically, 0.8–0.9), β_2 is the coefficient to account for the effective height where the inertia due to the modal mass acts (typically, 0.65–0.75), H_B is the height of building, N_p is the number of piles (assuming that an equal number of piles are positioned in two rows), and B (m) is the foundation width between the two rows of piles (note: the foundation width B is along the direction where axial force is induced by overturning moment due to lateral inertia). T_{pre} can be estimated by using an empirical equation, such as $T_{pre} = 0.09 H_B/B^{0.5}$; see Anderson et al. (1952) also adopted in IS 1893.

In a post-liquefaction situation, the building is supported by piles that have relatively long unsupported lengths of D_L (as evaluated based on the liquefaction initiation analysis). In this

case, the natural vibration period in a post-liquefaction stage, T_{post}, can be calculated as

$$T_{post} = 2\pi \sqrt{\frac{W/g}{N_p \times 12EI/D_L^3}} \qquad (11)$$

where $12EI/(D_L)^3$ is the lateral stiffness of each pile. Then, the maximum axial force acting on a pile is given by

$$P_{post} = \frac{(W/g) \times SA(T_{post}, \xi_{post}) \times (D_L + \beta_3 H_B)}{(N_p/2) \times B} \qquad (12)$$

where $SA(T_{post}, \xi_{post})$ (g) is the spectral acceleration at T_{post} with damping ratio ξ_{post} (typically, 10–30 % representing the damping of liquefied soil) and β_3 is the coefficient to account for the effective height where the inertia acts in a post-liquefaction condition (typically, 0.5). The underlying assumption of Eqs. 11 and 12 is that the building is considered as a rigid mass and the pile provides the primary lateral stiffness to the building. It must be mentioned that this approach is very simple, and many uncertainties, such as effects of vertical inertia, timing of max inertia and timing of loss of lateral support due to liquefaction, and potential effects due to lateral spreading, are not taken into consideration.

By defining $\alpha = \max(P_{pre}/P_{static}, P_{post}/P_{static})$, α is given by:

$$\alpha = \frac{2}{gB} \max \big[\beta_1 SA(T_{pre}, \xi_{pre})(D_{fix} + \beta_2 H_B), \\ SA(T_{post}, \xi_{post})(D_L + \beta_3 H_B) \big] \qquad (13)$$

The additional information needed for evaluating α is the estimated spectral acceleration values at vibration periods T_{pre} and T_{post} with damping ratios ξ_{pre} and ξ_{post} (note: T_{pre} and T_{post} are random variables; in particular, T_{post} is significantly affected by liquefaction initiation analysis).

A case study: **collapse of a 5-story RCC building in Kobe during the 1995 Kobe earthquake**: Several buildings with pile foundation had to be demolished due to severe liquefaction damage during the 1995 M_w6.9 Kobe earthquake. Because of the dramatic consequences, several detailed post-earthquake investigations were conducted to examine the cause and failure mechanism of these cases (Tokimatsu et al. 1997). For illustration, one of such case studies (Uzuoka et al. 2002; Bhattacharya 2006) is focused upon. A 5-story reinforced concrete frame building (total height H_B = 14.5 m) was located at 6 m from the quay wall on a reclaimed fill in the Higashinada area of the Kobe City; the distance from the rupture plane to the building site was about 5 km (Fig. 7). The Kobe earthquake caused a lateral displacement of 2 m to the quay wall toward the sea, and the building was tilted by 3° due to lateral spreading. The schematics of the post-earthquake investigation of the building and pile foundation are shown in Fig. 8. At the building site, significant lateral spreading was observed (about 1.0–1.5 deformation/movement of the ground toward the sea; Tokimatsu et al. 1997).

The building was supported on 38 hollow pre-stressed concrete piles (there were two pile rows separated by 7.5 m and 19 piles were aligned in each row); the pile length was 20 m with exterior and interior diameters of 0.4 and 0.24 m, respectively. Figure 9 shows variations of soil profiles and SPT N counts with depth at the building site. The site has fill/sand layers with relatively low N counts (e.g., 2–9 m and 12–16 m), which are susceptible to liquefaction (i.e., saturated sand layers with low strength); the average shear-wave velocities in the uppermost 12 and 30 m are estimated to be about 147 and 216 m/s, respectively. The water table level was about 2 m below ground surface. The post-earthquake investigation by Tokimatsu et al. (1997) indicated that soil layers shallower than 9 m were liquefied (based on a simplified stress method). Moreover, the borehole logging data shown in Fig. 8 suggest that the sandy silt
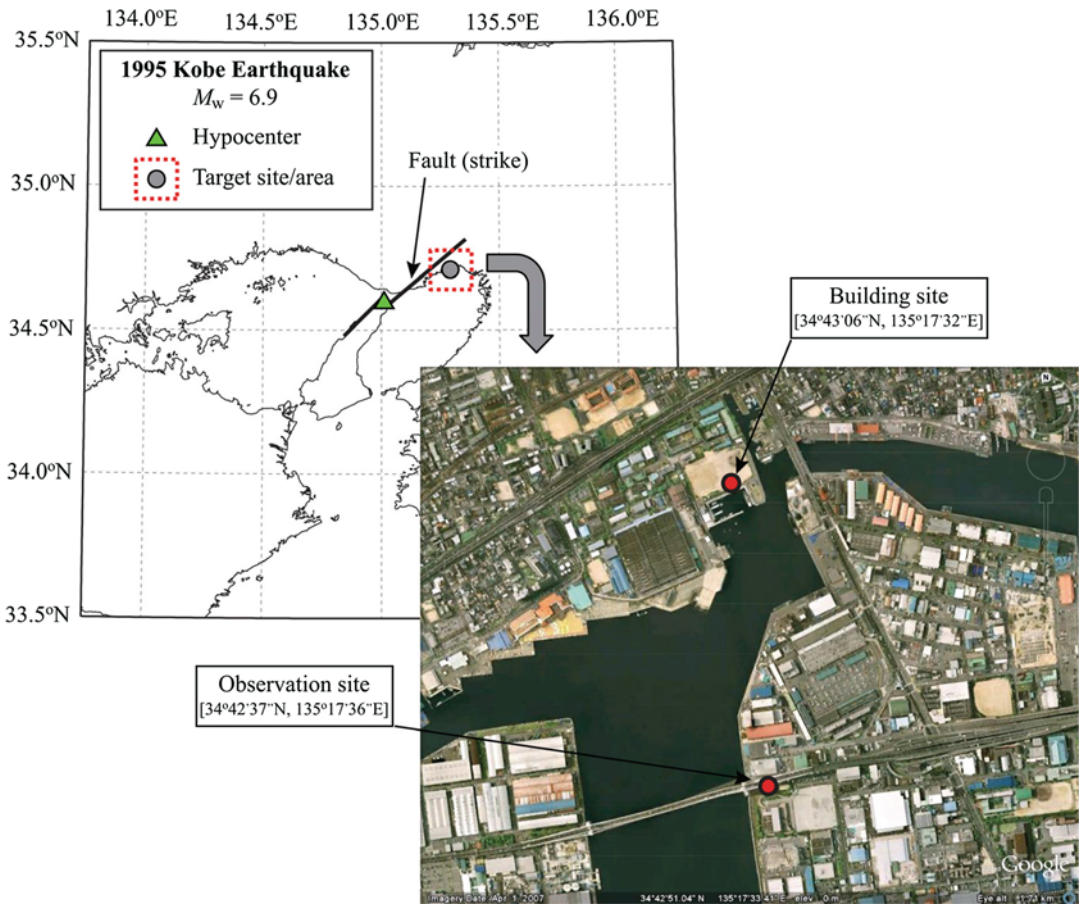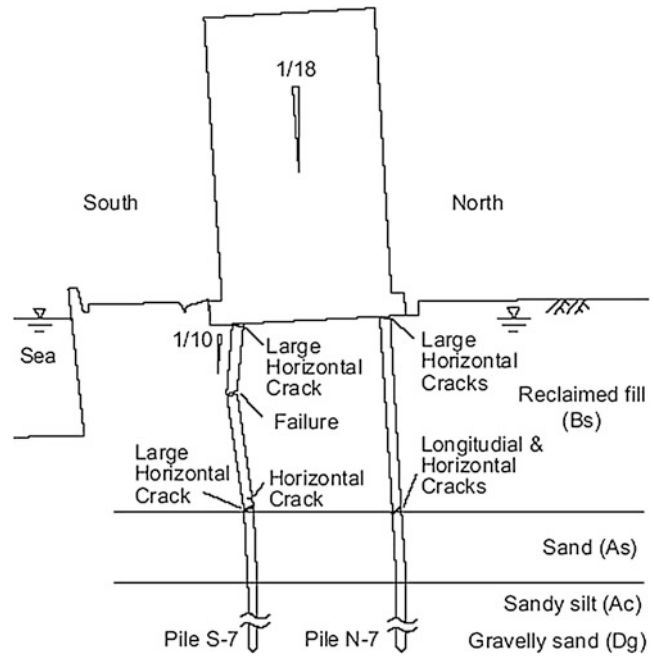
Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools, Fig. 7 Location of the building site and observation site

layer between 12 m and 16 m, having low strength, may be disturbed by strong ground motion; thus this layer may not offer much fixity to the pile.

The nearby ground motion recording was obtained at the Higashi Kobe Bridge, which is away about 0.9 km from the building site; distance to the rupture plane for the observation and building sites is about 5.2 and 4.4 km, respectively. The site condition at the observation site is similar to that at the building site (typically, NEHRP site class D or E). The recorded acceleration time histories at the observation site are shown in Fig. 9 (Public Works Research Institute 1995).

The 5 % damped response spectra of the two horizontal components and their geometric mean are presented in Fig. 9 and are compared with the median GMPE (Ground Motion Prediction Equations) by Zhao et al. (2006); this relation is used as a representative regional model throughout this study to estimate ground motion parameters at the building site. The comparison of the calculated response spectra with the Zhao et al. relation indicates that the observed response spectra have less spectral content at vibration periods less than 1.0 s, while they contain rich spectral content at vibration periods greater than 1.0 s. The responses at short vibration periods are likely

Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools, Fig. 8 Post-earthquake investigation of a case study (Tokimatsu et al. 1997)

Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools, Fig. 9 Boring log of the soil at the building site
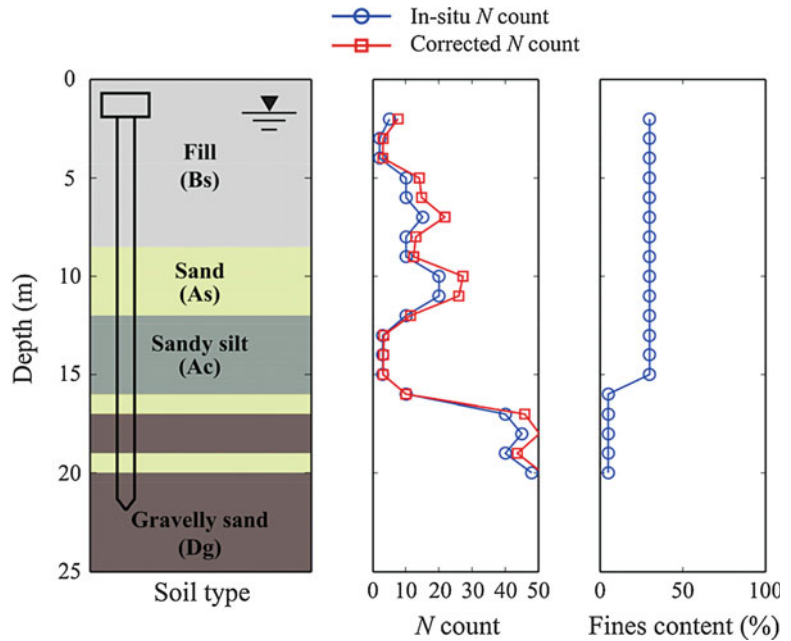
to be affected by nonlinear site amplification (i.e., de-amplification), and at such a site, liquefaction-induced ground failure may be expected. Further to note, geographical positions of the fault plane, hypocenter, and the observation site (Fig. 6) are of typical "near-fault motions" due to forward directivity (Mavroeidis and Papageorgiou 2003); this can be corroborated by large response spectra
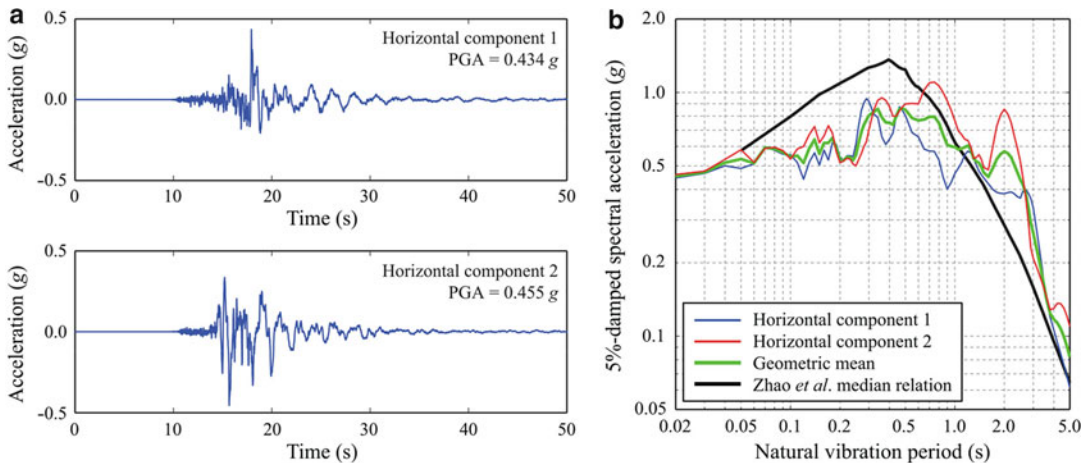
Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools, Table 2 Summary of probabilistic information of input parameters

Parameter	Mean	Coefficient of variation	[Lower and upper limits]	Distribution type
Moment magnitude M_w	6.9	0.1[a]	[6.6, 7.2]	Normal
Peak ground acceleration PGA (g)	Equation 2		–	Lognormal
Water table level (m)	2.0	–	[1.0, 3.0]	Uniform
FC (%)	Fig. 9	0.1	–	Lognormal
N count	Fig. 9	0.15	–	Lognormal
Vertical total stress σ_v (Pa)	–[b,c]	0.1	–	Lognormal
Vertical effective stress σ'_v (Pa)	–[b,c]	0.15	–	Lognormal
Pre-liquefaction period T_{pre} (s)	0.5	0.1	–	Lognormal
EI of a pile (MNm^2)	32.35	0.1	[24.26, 48.53]	Lognormal
Static axial force per pile P_{static} (kN)	412	0.1	[309, 618]	Lognormal

[a]This is the standard deviation

[b]It depends on the water table

[c]Dry and wet soil densities are set to 1.76 and 1.92 g/cm^3

Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools, Fig. 10 Ground motion time history (**a**) and 5 % damped response spectra (**b**) at the Higashi Kobe Bridge observation site

values at long periods (Fig. 9b) and by inspecting velocity time histories of the acceleration data where large velocity pulses are clearly visible. Table 2 provides a summary of the input parameters used in the analysis, and Fig. 11 shows the probability density function for D_L and H_C which shows the high likelihood that the building collapsed due to buckling of the piles (probability of buckling failure is 0.943). Further details of this methodology can be found in Bhattacharya and Goda (2013) (Fig. 10).

Safety Assessment of Piled Buildings in Liquefiable Soils: Mathematical Tools,
Fig. 11 Comparison of probability mass functions for D_L and H_C
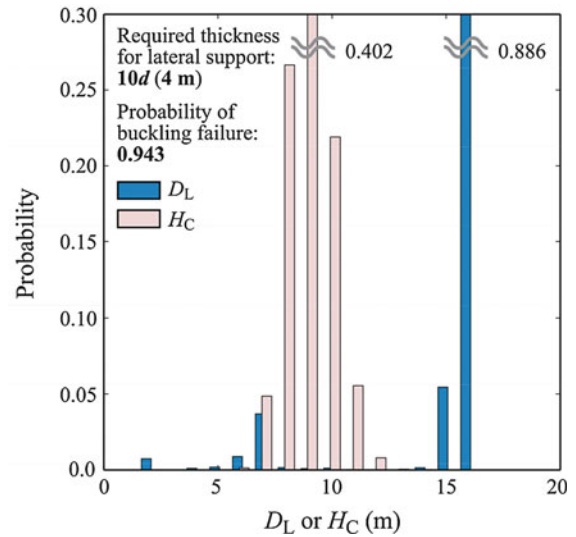
Summary

A probabilistic-based method to reassess such safety has been formulated in this entry. This method can easily be coded in a program, such as MATLAB or Fortran. This method checks the stability of the foundation against buckling instability at full liquefaction, i.e., when the soil surrounding the pile is at its lowest possible stiffness. Two parameters, namely, "critical depth of the pile H_C" and the "unsupported length of the pile due to liquefaction D_L," are estimated. Critical depth is a function of axial load acting on the pile (P), flexural stiffness of the pile (EI), and the boundary condition of the pile above and below the liquefiable soil. On the other hand, D_L mainly depends on the earthquake characteristics, soil profile, and ground conditions. A case study is considered to illustrate an application of the methodology.

Cross-References

▶ Damage to Buildings: Modeling
▶ Geotechnical Earthquake Engineering: Damage Mechanism Observed
▶ Seismic Collapse Assessment
▶ Seismic Reliability Assessment, Alternative methods for
▶ Seismic Vulnerability Assessment: Reinforced Concrete Structures

References

Adhikari S, Bhattacharya S (2008) Dynamic instability of pile-supported structures in liquefiable soils during earthquakes. Shock and Vibration 16(6):665–685

Anderson AW, Blume JA, Degenkolb HJ, Hammill HB, Knapik EM, Marchand HL, Powers HC, Rinne JE, Sedgwick GA, Sjoberg HO (1952) Lateral forces of earthquake and wind. Trans ASCE 117:716–80

Bhattacharya S (2003) Pile instability during earthquake liquefaction. Ph.D Thesis, University of Cambridge, U.K.

Bhattacharya S (2006) Safety assessment of existing piled foundations in liquefiable soils against buckling instability. ISET J Earthquake Technol 43:133–147

Bhattacharya S, Goda K (2013) Probabilistic buckling analysis of axially loaded piles in liquefiable soils. Soil Dyn Earthquake Eng 45:13–24

Bhattacharya S, Madabhushi SPG (2008) A critical review of methods for pile design in seismically liquefiable soils. Bull Earthquake Eng 6:407–446

Bhattacharya S, Madabhushi SPG, Bolton MD (2004) An alternative mechanism of pile failure in liquefiable deposits during earthquakes. Geotechnique 54:203–213

Bhattacharya S, Madabhushi SPG, Bolton MD (2005) Reply to the two discussions on the paper "An alternative mechanism of pile failure in liquefiable deposits during earthquakes". Geotechnique 55(3):259–263

Bhattacharya S, Adhikari S, Alexander NA (2009) A simplified method for unified buckling and free vibration analysis of pile-supported structures in seismically liquefiable soils. Soil Dyn Earthquake Eng 29:1220–1235

CEN (2004) Eurocode 8, design of structures for earthquake resistance – Part 1: general rules, seismic actions and rules for buildings, EN 1998-1:2004. Comite Europeen de Normalisation, Brussels, Belgium

Cetin KO, Der Kiureghian A, Seed RB (2002) Probabilistic model for the initiation of seismic soil liquefaction. Struct Safety 24:67–82

Cetin KO, Seed RB, Der Kiureghian A, Tokimatsu K, Harder LF Jr, Kayen RE, Moss RES (2004) Standard Penetration Test-based probabilistic and deterministic assessment of seismic soil liquefaction potential. J Geotech Geoenviron Eng 130:1314–1340

Dash SR, Bhattacharya S, Blakeborough A (2010) Bending-buckling interaction as a failure mechanism of piles in liquefiable soils. Soil Dyn Earthquake Eng 30:32–39

Goda K, Atkinson GM, Hunter JA, Crow C, Motazedian D (2011) Probabilistic liquefaction hazard analysis for four Canadian cities. Bull Seismol Soc Am 101:190–201

IS 1893 (2001) Indian Standard for Seismic Design

JRA (1996) Japanese Road Association, Specification for Highway Bridges, Part V, Seismic Design

Juang CH, Yang SH, Yuan H (2005) Model uncertainty of shear wave velocity-based method for liquefaction potential evaluation. J Geotech Geoenviron Eng 131:1274–1282

Lombardi D, Bhattacharya S (2014) Modal analysis of pile-supported structures during seismic liquefaction. Earthquake Eng Struct Dyn 43(3):119–138. doi:10.1002/eqe.2336

Mavroeidis GP, Papageorgiou AS (2003) A mathematical representation of near-fault ground motions. Bull Seismol Soc Am 93:1099–1131

Moss RES, Seed RB, Kayen RE, Stewart JP, Der Kiureghian A, Cetin KO (2006) CPT-based probabilistic and deterministic assessment of in situ seismic soil liquefaction potential. J Geotech Geoenviron Eng 132:1032–1051

NEHRP (2000) National Earthquake Hazards Reduction Program (NEHRP, 2000): commentary for Federal Emergency Management Agency (FEMA, USA 369) on seismic regulations for new buildings and other structures

Public Works Research Institute (1995) Strong-motion acceleration records from Public Works in Japan (No. 21). Ministry of Construction, Tsukuba, Japan

Rankine WJM (1866) Useful rules and tables. London

Seed HB, Idriss IM (1971) Simplified procedure for evaluating soil liquefaction potential. J Soil Mech Found Div 97:1249–1273

Tokimatsu K, Ohoka H, Shamoto Y, Asaka Y (1997) Failure and deformation modes of piles due to liquefaction-induced lateral spreading in 1995 Hyogoken-Nambu earthquake. J Struct Constr Eng, AIJ 495:95–100

Uzuoka R, Sento N, Yashima A, Zhang F (2002) 3-dimensional effective stress analysis of a damaged group-pile foundation adjacent to a quay wall. J Jpn Assoc Earthq Eng 2:1–14

Zhao JX, Zhang J, Asano A, Ohno Y, Oouchi T, Takahashi T, Ogawa H, Irikura K, Thio HK, Somerville PG, Fukushima Y, Fukushima Y (2006) Attenuation relations of strong ground motion in Japan using site classification based on predominant period. Bull Seismol Soc Am 96:898–913

SAR Images, Interpretation of

Antonio Iodice and Daniele Riccio
Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione, Università degli Studi di Napoli Federico II, Napoli, Italy

Introduction

A synthetic aperture radar (SAR) is a remote sensing microwave imaging system which consists of a radar sensor mounted on a moving platform, such as an airplane or a satellite. As the platform flies along an approximately straight line, the radar emits microwave pulses at fixed rate (pulse repetition frequency (PRF)) and receives corresponding returns (echoes) backscattered by the illuminated scene. A SAR is able to distinguish points at different distances from the line of flight based on different delays of their returns; in addition, points at different positions along a direction parallel to the line of flight are distinguished by forming a very long (and therefore very directive, with a very narrow beam) synthetic array. This is obtained by properly combining pulses received by the sensor at different positions along the line of flight, so that the synthetic array length is equal to the length of the portion of line of flight such that a given ground point remains within the real antenna beamwidth. In this way, a two-dimensional (2D) image is obtained, which is the projection

S

of the scene onto the plane containing the look direction (range direction) and the line of flight (azimuth direction). This is at variance with optical images, which are the projection of the scene onto the plane perpendicular to the look direction.

A SAR is an active sensor (i.e., it illuminates the scene), so that it can work during both the day and the night; in addition, it relies on microwaves, which can propagate through clouds, so that it can image the Earth's surface even in the presence of a cloud cover.

An important parameter characterizing an imaging sensor is its resolution, which is the minimum distance between two points such that they can be distinguished by the sensor. For a SAR sensor, the resolution along the range direction is $c/(2B)$, where c is the speed of light and B is the pulse bandwidth, whereas the resolution along the azimuth dimension is $\lambda r/(2X)$, where λ is the electromagnetic wavelength, r is the sensor-to-ground distance, and X is the synthetic array length. In its usual acquisition mode ("strip map" mode), the SAR radar antenna is constantly pointed along a direction perpendicular to the line of flight, forming a significant angle ("look angle") with the nadir direction (i.e., SAR is a "side looking" sensor); see Fig. 1. In this case, the synthetic antenna length X is $\lambda r/L$, where L is the real antenna effective azimuth length, so that the azimuth resolution turns out to be $L/2$. Note that the SAR resolution is independent of the

sensor-to-scene distance, at variance with the one of optical sensors. In order to obtain a higher (better) resolution, although with a smaller illuminated scene, a "spotlight" acquisition mode can be used, in which the SAR antenna beam is steered during the flight to constantly illuminate a given spot on the ground. In this way, a longer synthetic array can be obtained, this implying a better azimuth resolution. The resolution of modern spaceborne SAR systems spans from about 10 m to fractions of meter.

A SAR image provides information on the imaged scene which is in some sense complementary with respect to that provided by an optical image. In fact, while the intensity of a pixel in an optical image mainly depends on the chemical properties of the surface of the imaged objects, the intensity of a pixel in a SAR image depends on electromagnetic properties (permittivity and conductivity) of imaged objects and on their roughness at wavelength (i.e., centimetric) scale: smooth surfaces (calm water, concrete or asphalt surfaces, etc.) appear as dark areas on the image, whereas surfaces with increasing roughness appear as increasingly bright areas.

Finally, another important peculiarity of SAR sensors with respect to optical ones is their "coherent" nature: in fact, they are able to emit a coherent radiation and to measure not only the intensity of the received signal but also its phase. This allows using interferometric techniques (interferometric SAR (InSAR) and differential interferometric SAR (DInSAR)) to obtain terrain topography and to monitor small terrain movements and tomographic techniques (SAR Tomography) for the three-dimensional reconstruction of imaged objects. In addition, using two orthogonally polarized transmitting and/or receiving antennas, information can be extracted by observing how the polarization of the backscattered wave is modified with respect to the transmitted one (SAR Polarimetry). However, the coherent nature of SAR also causes the appearance of the "speckle" noise, which gives a "salt-and-pepper" look to SAR images: a macroscopically homogeneous area appears to be composed of pixels of randomly varying intensity. Speckle noise can be reduced, at the expense of
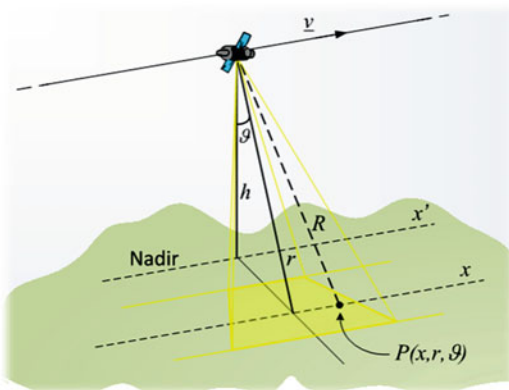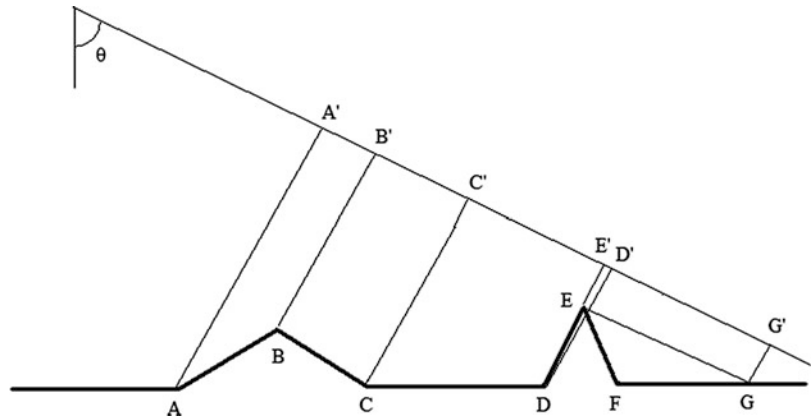

SAR Images, Interpretation of, Fig. 1 Geometry of SAR acquisition

SAR Images, Interpretation of, Fig. 2 Geometric distortions. Foreshortening-compression (A'-B'), foreshortening-dilation (B'-C'), layover (E'-D'), shadow (D'-G')

geometric resolution, by averaging over adjacent pixels (multi-look image), or by proper filtering.

By summarizing, SAR sensors have significant advantages with respect to optical ones (day-and-night, all-weather capabilities, coherent nature), but SAR images are more difficult to be visually interpreted than optical ones, due to both geometrical (i.e., image projection plane including the look direction) and radiometric (i.e., involved dependence of image intensity on terrain electromagnetic and roughness properties, speckle) issues. Fundamentals on SAR systems and applications can be found, e.g., in Ulaby et al. (1986), Elachi (1988), and Curlander and McDonough (1991). ▶ InSAR and ▶ SAR Tomography are the subjects of other entries of this encyclopedia, whereas SAR Polarimetry is analyzed in detail in Lee and Pottier (2009).

## Foreshortening, Layover, Shadow, and Geocoding

Due to the fact that a SAR image is the projection of the scene onto a plane including the look direction, geometric distortions on SAR images are very different from those experienced in optical images.

If terrain slope is smaller than the look angle θ, the resolution cell on the ground is compressed with respect to the horizontal terrain case if the surface is tilted toward the sensor; otherwise it is dilated (see Fig. 2.). This effect is termed "foreshortening."

If the surface is tilted toward the sensor and its slope is larger than the look angle, then there is an "inversion" of the SAR geometry: the positions along the range direction of peaks and bases of hills or mountains are exchanged, and the sides of hills or mountains are "folded" onto the valleys in front of them (so that a single pixel corresponds to both an area on the hill's side and an area on the valley, and a very bright area on the image appears) (see Figs. 2 and 3). This effect is called "layover."

Finally, if the surface is tilted away from the sensor and its slope is larger than 90° minus the look angle, then a portion of the surface is not illuminated (see Fig. 2), and no return is present in that portion of the SAR image: a "shadow" appears.

If SAR platform trajectory and terrain topography are known, abovementioned geometric distortions can be mitigated by a post-processing step, called "geocoding," that allows representing the SAR images in a standard cartographic map projection. In this way, the image can be, for instance, easily integrated in a GIS. Note however that, although foreshortening can be corrected by geocoding, layover and shadow effects imply a loss of information that cannot be recovered by this post-processing step. In addition, geocoding implies an interpolation process that may alter the image information content: therefore, in some applications it may be preferable to extract the physical parameter of interest directly from the image in SAR native geometry and then to geocode the obtained final

S
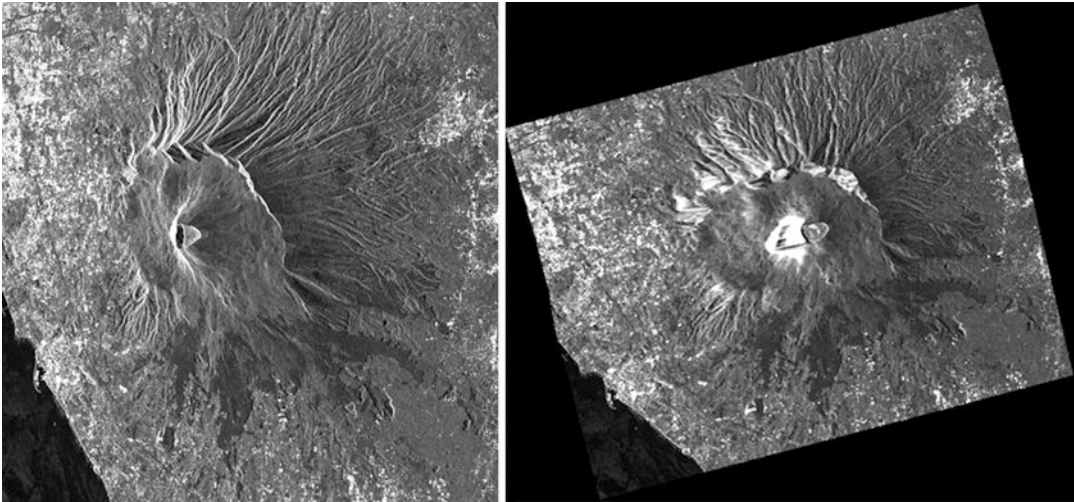
SAR Images, Interpretation of, Fig. 3 COSMO/SkyMed SAR image of the area of Mt. Vesuvius, Italy (*left*). Near range is on the left. Geocoded version of the same image (*right*). The very bright area near the crater corresponds to a layover area

SAR Images, Interpretation of, Fig. 4 Elements on a building facade forming dihedral and trihedral structures
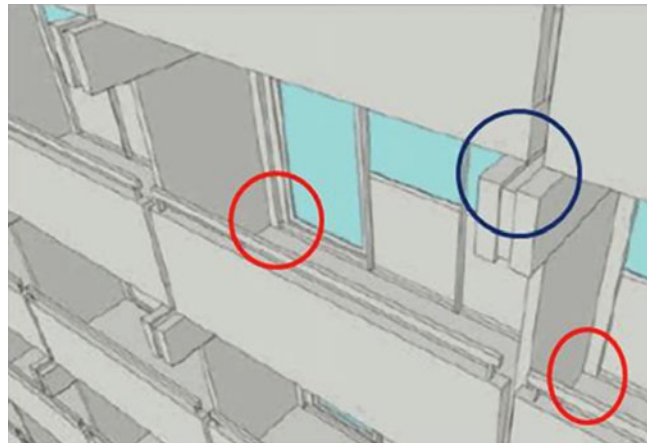

map (Guida et al. 2008; Di Martino et al. 2012). An example of SAR image before and after geocoding is shown in Fig. 3.

Interpreting SAR Images of Buildings

Geometric distortions described in the previous section are particularly severe in SAR images of urban areas, which are of prominent interest if, for instance, a fast post-event earthquake damage assessment is needed. With the launch of COSMO/SkyMed and TerraSAR-X missions, very-high-resolution (VHR) SAR images of urban areas have become routinely available. In particular, in the spotlight acquisition mode, COSMO/SkyMed SAR sensors are able to obtain a resolution even better than 1 m. Accordingly, in principle a lot of information on objects present in the urban scenario can be extracted from such images; however, due to the above-cited severe geometric distortions and due to the involved interaction between incident electromagnetic wave and imaged scene, direct interpretation of VHR SAR images is not straightforward. It is easy to realize (see Figs. 4 and 5) that such
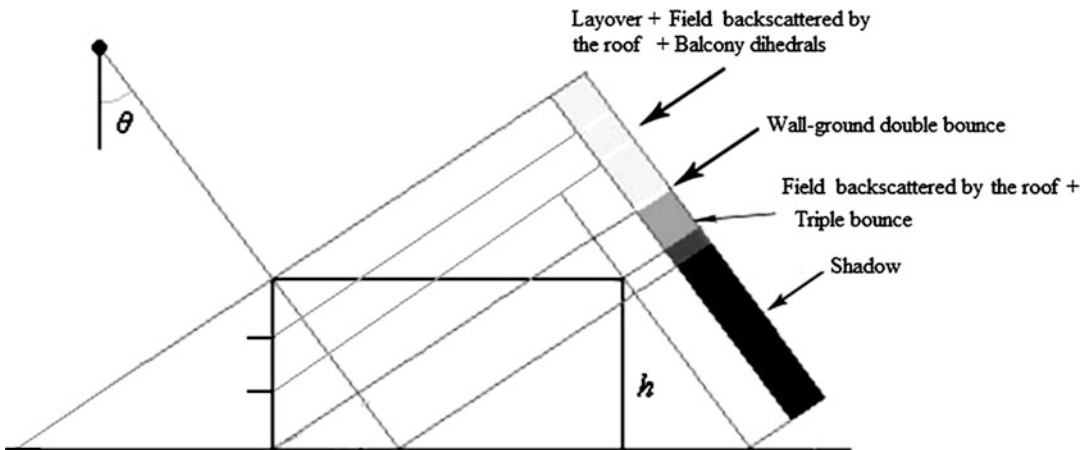
SAR Images, Interpretation of, Fig. 5 SAR image formation for a single building

images are dominated by the combination of single scattering from terrain and buildings and multiple scattering from dihedral and trihedral structures. For moderate- and low-resolution systems, they may be simultaneously present in a single resolution cell. Nonetheless, for very-high-resolution SAR systems, the resolution cell is so small that dihedral and trihedral returns dominate with respect to the single scattering background. Accordingly, a realistic description of VHR SAR amplitude images can be obtained by considering sparse brilliant points, or lines, over a dark background: the positions of these brilliant points and lines can be considered randomly distributed, unless they belong to a building facade, in which case an ordered, periodic spatial distribution is expected.

In fact, if we consider a single building, as illustrated in Fig. 5, by moving from near to far range, we first find a layover area in which each pixel is the superposition of two or three contributions: one from the ground, one from the vertical wall, and possibly one from the roof. Ground and roof are usually so smooth that corresponding contributions are weak, whereas often, on the vertical walls, dihedral or trihedral structures of balconies and windows are present (see Fig. 4) whose contributions to the backscattered signal are significant. In fact, with regard to dihedral structures, it is easy to realize that, if the structure is aligned with the SAR line of flight, all double-

bounce paths have the same length, corresponding to twice the path length from the sensor to the internal edge of the dihedral, and hence they reach the sensor simultaneously. Accordingly, a very bright line appears on the image. Similarly, if we consider a trihedral structure, for a wide range of structure orientations, all triple-bounce paths have the same length, corresponding to twice the path length from the sensor to the internal corner of the trihedral, and hence they reach the sensor simultaneously. Accordingly, a very bright point appears on the image. Since balconies and windows usually are spatially distributed in an ordered way, they form periodic patterns of brilliant lines and/or points within the layover area of the SAR image of the building.

By continuing to move from near to far range, the end of the layover area is marked by a bright line representing the double-bounce return from the large dihedral structure formed by the vertical wall and the ground. This line is particularly evident if the wall is aligned with the sensor line of flight. Then, a dark area is present, including the very weak triple-bounce return (wall-ground-wall and/or ground-wall-ground) and, possibly, the weak return from the roof. And, finally, we find a very dark shadow area.

An example of VHR SAR image of an urban area is reported in Fig. 6a. This is a 1-m resolution TerraSAR-X image of Naples, Italy. Near range

SAR Images, Interpretation of, Fig. 6   TerraSAR-X image of Naples, Italy (**a**). Near range is on the left. Excerpt of the previous image showing the area of Piazza del Plebiscito (**b**). An optical image of the same area (**c**)

is on the left. In Fig. 6b an excerpt of the previous image is reported, and in Fig. 6c an optical image of the same area is shown. The square at the center of this area is Piazza del Plebiscito, and the building at its right side is the Palazzo Reale (Royal Palace). In agreement with our previous discussion, in correspondence of the façade of

this building, by moving from left to right (i.e., from near to far range), we can distinguish three bright lines, followed by a very bright line and by a dark area. The three bright lines correspond to an architectural structure at the roof edge and to two lines of balconies, whereas the very bright line corresponds to the wall-ground

double-bounce return. Finally, the dark area corresponds to building roof and shadow.

Model-Based Interpretation of SAR Images

Above presented description allows a qualitative interpretation of SAR images of natural and urban areas. Quantitative information on the geometry (i.e., distances between objects, building size, floors' height, etc.) can be also obtained from such an analysis if look angle and pixel spacing are known. However, in order to obtain quantitative relations between image intensity and scene properties, a deeper analysis is needed. First of all, the scene must be described in terms of parameters of interest: for instance, for a natural scenario, soil moisture and composition, terrain roughness, vegetation biomass, etc. Then, a direct electromagnetic scattering model must be used to express the backscattered field in terms of such scene parameters. Inversion of this model allows retrieving the scene parameters of interest from SAR images of the scene.

This field is currently the subject of intense research activity, and the description of scattering models and retrieval algorithms goes beyond the scope of the present work. However, a review of scattering models can be found in some textbooks (e.g., Ulaby et al. 1986; Tsang et al. 2000; Franceschetti and Riccio 2007) and scientific papers (e.g., Fung et al. 1992; Franceschetti et al. 2002). Examples of retrieval algorithms can also be found in the scientific literature, both for natural scenes (Iodice et al. 2011; Di Martino et al. 2012) and urban areas (Guida et al. 2008, 2010).

RGB Compositions

Visual interpretation of SAR images can be made easier by combining different SAR images of the same area to create a false color image.

For instance, if a SAR polarimetric system is employed, a combination of different polarimetric channels can be used. Available channels are HH (i.e., transmit a horizontally polarized electromagnetic field and measure the horizontally polarized component of the received field), VV (i.e., transmit a vertically polarized electromagnetic field and measure the vertically polarized component of the received field), and HV or VH (i.e., transmit a horizontally polarized electromagnetic field and measure the vertically polarized component of the received field, or vice versa). For a wide range of scattering surfaces (reciprocal scatterers), HV and VH channel returns are equal, so that only one of the two can be actually used. A very useful combination consists of HH + VV, HH-VV, and HV channels, and it is called a "Pauli decomposition" (Lee and Pottier 2009). In fact, it turns out that the sum of HH and VV returns is dominated by single scattering from rough surfaces (soil surfaces, sea surfaces), the difference of HH and VV returns is dominated by double scattering (terrain-building walls, or ground-tree trunks), and HV return is dominated by volumetric scattering or extremely rough surface scattering (vegetation). Accordingly, an RGB color image can be obtained by loading the HH + VV signal onto the blue channel, HH-VV onto the red channel, and HV onto the green one. Accordingly, blue areas on the image will correspond to bare or little vegetated soils, or sea; red areas to built-up areas or trees with little foliage; and green areas to very vegetated soils or forests. Intermediate colors will correspond to pixels containing combinations of the previous targets. An example of false color SAR image obtained by using the Pauli decomposition is reported in Fig. 7.

Another possibility is to load, onto the three different color channels, SAR images of the same area acquired at different times. This allows to easily identify areas subjected to changes between two different acquisitions. For instance, in Fig. 8 a false color SAR image of an area in a semiarid region (Tougou basin, Burkina Faso) is shown. An image acquired during the dry season is loaded onto the blue channel, another image acquired in the wet season is loaded onto the green channel, and the interferometric coherence between the two acquisitions (see ▶ InSAR and

S

SAR Images, Interpretation of, Fig. 7 False color image obtained by performing a Pauli decomposition on the polarimetric SAR image of an agricultural area

SAR Images, Interpretation of, Fig. 8 False color SAR image of the Tougou basin area (Burkina Faso)

A-InSAR: Theory) is loaded onto the red channel. Accordingly, different colors can be interpreted as follows (Amitrano et al. 2015):

Black areas: permanent basin water (dark on both images, low coherence)

Blue areas: wet season basin water (dark on wet season image, intermediate on dry season image, low coherence)

Green areas: wet season vegetation (bright on wet season image, intermediate on dry season image, low coherence)

Red or white areas: man-made objects, village (intermediate or bright on both acquisitions, high coherence)

Blue-green intermediate color, high intensity (cyan): trees (bright on both acquisitions, low coherence)

Blue-green intermediate color, intermediate intensity (Prussian blue or dark green): bare soils (intermediate on both acquisitions, intermediate coherence)

Note that the one described above is just an example: different RGB compositions can be employed, according to the considered application and to the scene characteristics that the user is interested to highlight.

Summary

SAR sensors have significant advantages with respect to optical ones: day-and-night, all-weather capabilities, and the possibility of measuring terrain topography and monitoring small terrain movements, due to its coherent nature. In addition, SAR images of virtually any area of the Earth surface are today routinely available, due to the different SAR satellite missions currently in orbit. Therefore, their use has a huge potential impact on a number of applications, among which the fast post-event earthquake damage assessment.

However, visual interpretation of SAR images by a human operator requires that the latter is properly trained to get used to the peculiar characteristics of SAR images: geometric distortions, involved dependence of image intensity on terrain electromagnetic and roughness properties, and speckle noise. Visual interpretation can be made easier by properly combining different images to form a color image (RGB composition).

Finally, automatic quantitative interpretation of SAR images requires the availability or the development of electromagnetic scattering models and of corresponding retrieval algorithms. This field is currently the subject of an intensive research activity.

References

Amitrano D, Di Martino G, Iodice A, Riccio D, Ruello G (2015) A new framework for SAR data RGB multitemporal representation: rationale and products. IEEE Trans Geosci Remote Sens 53:117–133

Curlander JC, McDonough RN (1991) Synthetic aperture radar: systems and signal processing. Wiley, New York

Di Martino G, Riccio D, Zinno I (2012) SAR imaging of fractal surfaces. IEEE Trans Geosci Remote Sens 50:630–644

Elachi C (1988) Spaceborne radar remote sensing: applications and techniques. IEEE Press, New York

Franceschetti G, Riccio D (2007) Scattering, natural surfaces and fractals. Academic, Burlington

Franceschetti G, Iodice A, Riccio D (2002) A canonical problem in electromagnetic backscattering from buildings. IEEE Trans Geosci Remote Sens 40:1787–1801

Fung AK, Li Z, Chen KS (1992) Backscattering from a randomly rough dielectric surface. IEEE Trans Geosci Remote Sens 30:356–369

Guida R, Iodice A, Riccio D, Stilla U (2008) Model-based interpretation of high-resolution SAR images of buildings. IEEE J Sel Topics Appl Earth Observ 1:107–119

Guida R, Iodice A, Riccio D (2010) Height retrieval of isolated buildings from single high resolution SAR images. IEEE Trans Geosci Remote Sens 48:2967–2979

Iodice A, Natale A, Riccio D (2011) Retrieval of soil surface parameters via a polarimetric two-scale model. IEEE Trans Geosci Remote Sens 49:2531–2547

Lee JS, Pottier E (2009) Polarimetric radar imaging: from basics to applications. CRC Press, Boca Raton

Tsang L, Kong J, Ding K (2000) Scattering of electromagnetic waves – theory and applications. Wiley, New York

Ulaby FT, Moore RK, Fung AK (1986) Microwave remote sensing, active and passive. Artech House, Norwood

SAR Tomography for 3D Reconstruction and Monitoring

D. Reale and G. Fornaro
Institute for Electromagnetic Sensing of the Environment, National Research Council, Naples, Italy

Synonyms

Cosmo-Skymed; Differential interferometry; DInSAR; Displacement measurement; Infrastructures monitoring; Multidimensional SAR imaging; SAR tomography; Synthetic-aperture radar, SAR; TerraSAR-X

S

Introduction

Synthetic-aperture radar (SAR) is one of the most important Earth remote sensing sensors whose applications have grown dramatically in the recent years. It provides images at microwaves with resolution comparable to that of optical systems, but with the crucial advantage of all-time, day/night, and all-weather, imaging capability. Similarly to classical surveillance radars, SAR measures the distance (range) from sensor to target: resolutions of the order of meters are achieved through the pulse compression of large bandwidth (frequency-modulated) signals. Very high resolution in the along-track direction is achieved as well through the coherent combination of target echoes received over the illumination interval, thus implementing a virtual (synthetic) array of antennas (i.e., a very large antenna) by exploiting the movement of a very small antenna mounted on board airplanes or satellites. The latter feature turns SAR systems to be imaging radars.

SAR data are nowadays used in many areas of environmental risks monitoring situations such as flooding, glaciers, land cover, and forest monitoring (Curlander and McDonough 1991; Moreira et al. 2013). Among all, one of the primary applications of SAR is the 3D reconstruction and monitoring of the Earth surface displacements through the use of interferometric techniques.

SAR interferometry (InSAR) exploits the coherent properties of the sensor, i.e., the capability to accurately control not only the envelope but also the phase of the transmitted radiation. SAR images are in fact complex data characterized by an amplitude measurement (envelope), related to the backscattering properties of the scene, as well as a phase signal related to both phase of the backscattering coefficient and to the distance of the target from the sensor to an accuracy of the order of the wavelength (centimeters at microwaves). Similarly to the human vision system, acquiring images with a slight angular diversity allows SAR to be sensitive to the 3D scene properties, i.e., to estimate the topography of the observed scene. Topography is not accessible in a single SAR acquisition because the imaging process returns only a 2D projection of the 3D reality. The Shuttle Radar Topography Mission (SRTM) in the last decade has represented the first case of an extensive use of InSAR for the generation of a worldwide (except for the poles) digital elevation model (DEM), that is, a digital topography map of the Earth surface (Van Zyl 2001). SRTM DEM (90 m spacing DTED-1 standard) has been extensively used in many applications. Nowadays, the TerraSAR-X/TanDEM-X mission (launched in 2007 and 2011) is providing a "refreshing" with higher resolution (12 m spacing DTED 3 standard) of the Earth DEM on a global scale by exploiting the simultaneous acquisitions of two twin SAR sensors flying in a close formation (Krieger et al. 2007). On the other hand, differential interferometry (DInSAR) takes advantage of the very high precision of radar systems in measuring phase to estimate differential displacements of the area imaged at different time instants with an accuracy of the order of a fraction of the used wavelength. DInSAR is today routinely used to estimate displacements induced by large earthquakes as well as to monitor volcanic activities producing ground movements, subsidence caused by water and/or oil extraction, mining activities, and also slow moving landslides (Massonnet et al. 1993; Carnec et al. 1995; Fornaro and Franceschetti 1999; Crosetto et al. 2005).

The availability of long-term data archives of former C-Band ESA ERS1/2 and ENVISAT SAR satellites has pushed the development of multipass interferometric processing techniques which coherently process large dataset of tens of images. In this way intrinsic limitations of classical, single pair, DInSAR, as the presence of atmospheric phase contribution and unwrapping procedures needs, affecting the unambiguous estimation of the useful deformation signal, are overcome. These techniques are mainly categorized based on the assumption on the scattering on the ground. From one hand, the approaches referred to as Persistent Scatterer Interferometry (PSI) prioritize the spatial resolution and uses all available baselines for accurate monitoring of "strong" (i.e., persistent) scatterers, typically located on anthropic structures, exhibiting

a temporal response stability over the whole observation period (Ferretti et al. 2000, 2001). On the other hand, multipass DInSAR techniques as the Small BAseline Subset (SBAS) limits the processing to interferograms characterized by short temporal separation as well as reduced angular diversity and makes use of a spatial multilooking to enhance the signal quality at the expense of a spatial resolution loss. SBAS techniques, known also as DInSAR stacking, are devoted to the monitoring of wide areas including rural zones (Berardino et al. 2002; Ferretti et al. 2011).

PSI and SBAS approaches have been used to investigate several aspects of risks for monitoring of coseismic and post-seismic deformation corresponding to several major earthquakes, volcanoes deformation, landslides, etc., as well as for the monitoring of buildings and infrastructures (Cascini et al. 2007; Arangio et al. 2013).

A recent advance for the technological viewpoint in the application to building reconstruction and monitoring is provided by SAR tomography which extends the SAR imaging concept to the third dimension of height: SAR tomography is also known as multidimensional SAR imaging technique due to the capability of full 3D imaging plus time monitoring. As for PSI, SAR tomography exploits full resolution data with angular diversity but for the use of the whole complex measured data then introducing the virtual antenna array along the height: the larger height antenna extent allows reducing again, as for the azimuth, the antenna beamwidth and reconstructing with finer resolution the backscattering along height (3D imaging) (Reigber and Moreira 2000; Gini et al. 2002; Fornaro et al. 2005). This latter allows, from one side, estimating height parameters of scatterers with better accuracy but, above all, detecting the presence of possible multiple scattering mechanisms which may interfere within the same radar spatial resolution cell. This interference, known as the layover effect, is a direct consequence of the imaging principle of radar systems which discriminates scatterers in distance: in the typical SAR side-looking geometry, in presence of steep topography, as for buildings, walls, and in

general vertical surfaces, it happens that scatterers located at different heights may be sensed by the radar at the same distance and therefore their returns imaged in the same pixel. A pictorial explanation of the layover induced on a building, as well as its effect on a very high (1 m) resolution TerraSAR-X amplitude image, is provided in Fig. 1: backscattering returns from targets located on the left part of the roof are imaged at first, and then the contributions from the facade and finally the base of the building are imaged in far range. Consequently, the building is tilted toward the sensor in the resulting SAR image: notice that because of the very high resolution, returns from the facades are spread over a large number of pixels in which contribution from ground is also expected. The vertical synthetic aperture exploited by SAR tomography allows improving resolution and tightening responses of the different interfering scatterers and then giving a chance to detect and localize separately each scattering mechanism (Fornaro and Serafino 2006). This capability has a major importance in the processing of urban areas, where the presence of buildings causes very frequent occurrence of layover between building facades, surrounding structures and ground. It is worth to note that this feature is own of the tomographic approach: none of interferometric approaches, including the PSI, can counteract this interference because of the leading assumption of only one scattering mechanism per pixel and of the phase-only signal model which does not cope with an imaging viewpoint.

As a direct extension of DInSAR, differential tomography (4D imaging) has been also proposed: it extends the 3D imaging to measure also the deformation parameters of scatterer in the focused 3D space (Lombardini 2005; Fornaro et al. 2009a; Zhu and Bamler 2010a). Along the same lines of PSI, also time series explaining the temporal evolution of deformation can be extracted, even separately for each interfering scatterers possibly exhibiting different deformation behaviors (Fornaro et al. 2009).

SAR technology is also evolved to specifically accomplish requirements and improve monitoring performances of multipass techniques.
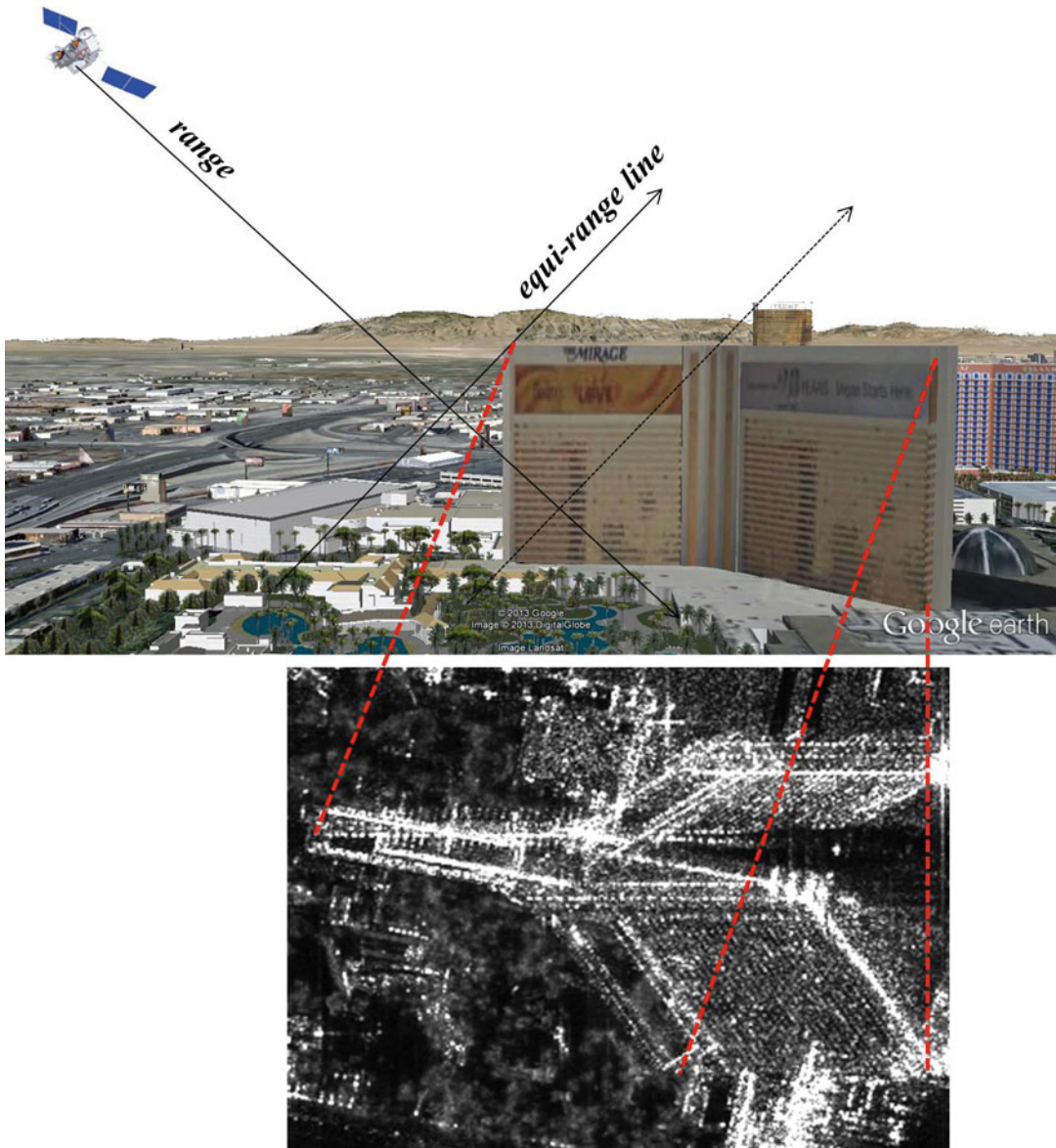
SAR Tomography for 3D Reconstruction and Monitoring, Fig. 1 Pictorial illustration of the layover distortion on SAR images induced by the side-looking imaging principle of radar and evidence on a 1 m spatial resolution TerraSAR-X image: closest targets located on the roof are imaged in near range respect to those on the base of the building. Optical image courtesy of Google

Recent X-Band sensors as the TerraSAR-X and the Cosmo-Skymed constellation are providing images with resolution never achieved in the past, in the meter/submeter regime with reduced revisiting times. As far as the latter aspect is concerned, the Cosmo-Skymed mission is worldwide the largest constellation for civilian application: it is composed by four small-size satellites that ensure an average revisiting time of 4 days at the maximum acquisition rate, which dramatically impacts in unexpected emergency situations when immediate imaging of damaged areas is required. High resolution allows capturing much more information from the scene,

SAR Tomography for 3D Reconstruction and Monitoring, Fig. 2 3D visualization of scatterers detected by the SAR tomography for the building imaged in the amplitude image of Fig. 1 (Reale et al. 2011a). *Colors* are set according to the estimated height

+640m +740m
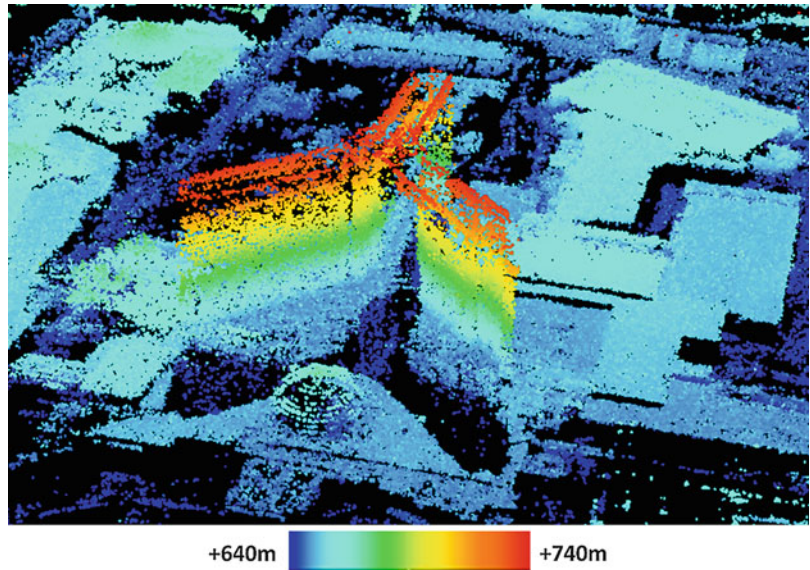
notably increasing the density of monitored scatterers (Gernhardt et al. 2010). At the same time, the layover becomes a major issue in the imaging of urban areas: the finer is the resolution, the larger will be the number of pixels affected by the layover induced by the buildings facades. Multidimensional imaging have demonstrated to be an effective tool in resolving the distributed layover on very high-resolution data, allowing fully separating contribution from ground and facade for the accurate 3D reconstruction of buildings. An example of the capability of multidimensional imaging for 3D building reconstruction and layover solution is provided in Fig. 2 which represents the 3D point cloud of the scatterers detected over the building imaged in Fig. 1 by the SAR tomography on a dataset of 25 TerraSAR-X very high-resolution spotlight images (Reale et al. 2011a). The precise reconstruction proves the effectiveness of this advanced processing which takes benefit of the detailed imagery provided by 1 m resolution acquisition capability.

This chapter is intended to introduce the principal concepts of the multidimensional imaging approach. A brief introduction on classical interferometric approach is firstly provided to introduce the concepts of angular and temporal diversity in SAR imagery and their relationship with physical parameters of interest as topography and surface movements. SAR tomography is introduced as an extension which turns the interferometric processing into an imaging problem, allowing both improving performances in estimation of parameters as well as allowing extracting further information from the data with respect to interferometry as the separation and detection of interfering scatterers in layover areas. Finally, examples of the application of both multipass DInSAR (small scale), and tomographic (full resolution) processing carried out on recent Cosmo-Skymed data are provided to point out the potentiality of the joint use of these technique for a complete monitoring of risk situations at different scales, from regional up to the level of the single infrastructure, spatial scales.

The SAR Interferometry Background

SAR imaging allows discrimination of targets along the azimuth and range directions. Assuming SAR sensors to fly locally rectilinear trajectories (airborne) or orbits (spaceborne), azimuth x (directed along the sensor velocity vector) and range r (distance orthogonal from the flight track)

represent two coordinates of the natural radar cylindrical reference system with the axis coincident with the flight track. By using large bandwidths, reaching the order of hundreds of megahertz, modern sensors distinguish targets in range with a resolution degree that reaches the meter/submeter scale. In addition, a high azimuth resolution capability of the final 2D images is achieved by synthesizing antennas in the order of kilometers through the exploitation of the intrinsic motion of the platform along its orbit (Curlander and McDonough 1991).

The height information, which is not accessible in a single SAR image, can be estimated through the interferometric concept. As in any electromagnetic coherent system, the phase information is related to the traveled path and therefore to the distance of the scene from the imaging radar. Classical SAR interferometry exploits the phase difference of, at least, two images acquired with an angular diversity induced by a slight orbital offset (spatial baseline), to retrieve the 3D localization of ground scatterers. Conversely, temporal separation (temporal baseline) is exploited by the differential SAR interferometry to measure possible displacements along the radar's line-of-sight (LOS) occurring at each acquisition epoch. The winning aspect of interferometry is the very high precision of radar in the estimation of the phase values, whose accuracy is of the order of a fraction of the wavelength: this allows estimating movements with sub-centimetric accuracy, using C- and recently X-Band radars from the space. Classical DInSAR has been extensively applied to measure large deformation caused mainly from earthquakes or volcanic activities. However, the presence of additional phase disturbing contributions does not allow pushing accuracy to a millimeter accuracy through the exploitation of just few images. As stated above, the interferometric phase difference $\Delta\varphi$ of a radar system working with a wavelength λ is related to the difference δr in the traveled path of the signal forming the two images. This phase difference is composed of multiple contributions associated with different source (Fornaro and Franceschetti 1999):

$$\Delta\varphi = \frac{4\pi}{\lambda}\delta r = \frac{4\pi}{\lambda}(\delta r_z + \delta r_d) + \Delta\varphi_a + \Delta\varphi_o + \Delta\varphi_n$$

(1)

The first term $\delta r_z = (b/r)s$, with s being the slant height (orthogonal to the azimuth/range imaging plane corresponding to the master image) ($s = z/\sin(\vartheta)$, ϑ is the look angle), accounts for the distance variation induced by the presence of topography on the ground which plays a role in the presence of an imaging parallax measured by spatial baseline b. Conversely, δr_d is the distance variation associated with possible deformation signal (LOS component) measured among the two acquisition epochs. Subsequent terms play a disturbance role. The term $\Delta\varphi_a$ is due to the propagation delay variation between the two acquisitions induced by changes in atmosphere. Slowing produces a time delay which is mapped into a range variation. The atmospheric propagation delay (APD) exhibits a spatial correlation over hundreds of meters: it is typically on the same level of deformations which are also often spatially correlated; $\Delta\varphi_o$ and $\Delta\varphi_n$ are associated with orbital inaccuracies and noise. In applications devoted to the estimation of the scene topography, simultaneous acquisitions are preferable because both deformation and atmospheric contributions are absent: this was the case of the Shuttle Radar Topography Mission (SRTM) which employed a dual-antenna system to reconstruct the digital elevation model (DEM) of the Earth and recently of the TerraSAR-X/TanDEM-X mission which employs a pair of twin satellites flying in close formation to produce a better resolved, 12 m spatial resolution, DEM. Differently, when deformation is of interest, as in differential interferometry, repeated passes over the exactly same orbit should be required to avoid the impact of topographic phase contribution. Indeed, this requirement is problematic to enforce, and then DInSAR interferograms are produced by subtracting from the original interferogram an estimation of δr_z evaluated from an external DEM (typically the SRTM DEM): such operation, referred to as zero baseline steering (ZBS), aims at eliminating or at least mitigating the fringes corresponding to

the topography. The fringe pattern retrieved by the external DEM is usually referred to as "synthetic interferogram." Estimation of deformation is however affected by the presence of atmospheric contribution, and then classical two-pass DInSAR configuration is usually applied to estimate predominant deformation caused, f.i., by large earthquakes.

The assumption leading to the differential phase in Eq. 1 is that the scene backscattering involved in the complex conjugate interferometric products are the same in both the images. Changes lead to the presence of noise contribution in the interferometric phase $\Delta\varphi$ which is referred to as decorrelation (Bamler and Hartl 1998; Fornaro and Franceschetti 1999). The coherence measures the degree of decorrelation: it measures the modulus of the cross-correlation index between the two images, which is a measure of the linear predictability of the two random variables corresponding to the master and slave SAR image values in the given pixel. The coherence is a product of several decorrelation effects, the most important being associated with the change of the imaging geometry due to the spatial orbit offset, the temporal separation between the acquisitions and the thermal noise. Spatial decorrelation is caused by the change of the imaging angle which determines a change of the backscattering coefficient, because of the presence in each resolution cell of a large number of scattering sources. The temporal decorrelation is due to the change of the scene backscattering properties over the time. This decorrelation source is of main importance in repeat pass interferometry, especially with systems operating at higher frequencies (f.i., X-Band). It is critical over the sea and in vegetated areas where the growth and in general the change of vegetation lead to strong variation of the backscattering coefficient. The last term, the thermal decorrelation, is due to the presence of thermal noise in the receiving apparatus. It is particularly evident in areas characterized by very low scene backscattering. Other decorrelation sources are associated with variations of the imaging aspect angle (Doppler centroid decorrelation) and to processing artifacts.

It is worth pointing out that what is really measured in interferometry is only a restricted version, corresponding to the modulus 2π of $\Delta\varphi$: one of the most critical steps of the interferometric processing is therefore represented by the unwrapping procedure which is required to retrieve the exact, unrestricted, absolute differential phase from whom quantity of interest can be estimated. Decorrelation effects, as well as large phase discontinuities, f.i., associated to steep height variations, affect the reliability of phase unwrapping algorithms (Ghiglia and Pritt 1998).

Multipass SAR Interferometry

Satellites regularly repeat orbits over the time. As a consequence, stacks of multipass acquisitions, characterized by angular and temporal diversity, are available in remote sensing archives for most of the Earth surface. In order to achieve a higher accuracy in the estimation of the deformation, modern advanced DInSAR processing algorithms jointly process all the images in stacks of multitemporal acquisitions: this processing allows in fact to provide a discrimination between the atmospheric and deformation as well as to cancel possible residual topography components.

Among all, interferometric techniques can be distinguished in two main categories which are characterized by complementary assumptions about the ground scattering: the multipass DInSAR stacking techniques and the Persistent Scatterers Interferometry (PSI). The class of the DInSAR stacking methods is a direct extension of the classical two-pass DInSAR technique. It assumes the scattering to be spatially distributed over the resolution cell and is based on the exploitation of both only small (temporal and spatial) baseline interferograms (hard baseline thresholding) and of a spatial multilook in order to limit the effects of decorrelation and reduce the phase noise in the interferograms, as for the Small BAseline Subset technique (Berardino et al. 2002). It is tailored to the monitoring of wide areas and scattering mechanisms that exhibit decorrelation including rural areas with
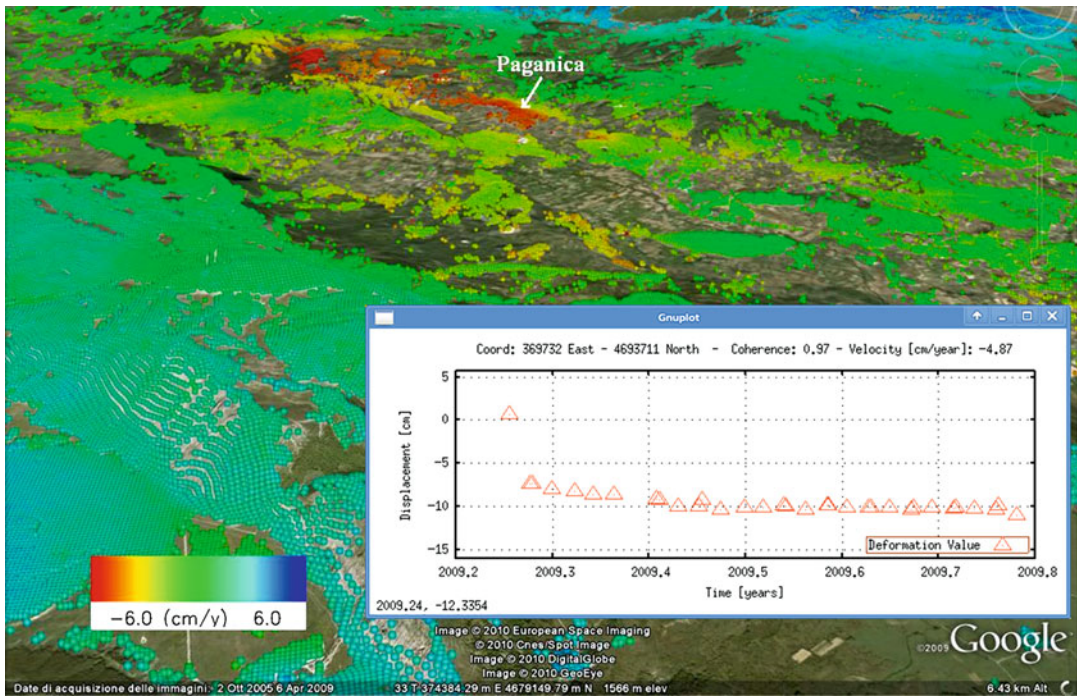
SAR Tomography for 3D Reconstruction and Monitoring, Fig. 3 Post-seismic deformation velocity map and co/post-seismic time series of a point in the Paganica area close to L'Aquila obtained by processing a dataset of 33 Cosmo-Skymed images acquired between April 4 and October 13, 2009 (Reale et al 2011b)

slow temporal correlation losses. This technique relies on the inversion of the linear system relating interferometric (differential) phase values after phase unwrapping to the phase values on each acquisition. The separation of displacement and atmospheric contribution is carried out by exploiting their statistical characterization in terms of spatial and temporal correlation. DInSAR stacking methods analyze interferograms typically generated by pairing acquisitions characterized by small spatial baselines and temporal separation, thus limiting the decorrelation. They play a favorable role in the design of a two-scale processing: besides effectively counteracting decorrelation and phase noise, the spatial multilooking allows also to under-sample the interferograms in the image spatial (azimuth and range) coordinates, thus reducing the amount of data and consequently the computational effort for the analysis of large areas. Therefore, this class of algorithms is particularly suitable for the analysis of large areas at lower resolution

(small scale). Additionally, through the analysis of multiple interferograms, the atmospheric phase contribution can be estimated and compensated from the data. This latter, coupled with the removal of the background deformation signal occurring on a small scale, are used to phase calibrate the full resolution data for processing at large scale, i.e., at the level of single building and infrastructure.

A typical product of the processing through DInSAR stacking techniques is reported in Fig. 3: it represents the deformation mean velocity map, superimposed to a Google Earth image, corresponding to the slope of the deformation time series for each pixel selected by looking at a quality index measuring the temporal consistency of measurements after phase unwrapping. The processing has been carried out through the application of the Enhanced Spatial Differences (ESD) technique which extends the classical SBAS approach by exploiting a model for the phase differences between adjacent pixels to

counteract the effect of atmospheric contribution, thus supporting the phase unwrapping step and increasing the coverage and accuracy of the retrieved deformation measurements (Fornaro et al. 2009b). The result is relevant to the dataset of Cosmo-Skymed acquisitions over the area of L'Aquila, in the central part of Italy, struck by the 2009 earthquake. During the emergency acquisition plan, the Cosmo-Skymed constellation, at that time composed by three of the four satellites nowadays operatively, intensified the acquisition to its highest possible rate (almost 1 acquisition every 5 days on average) to acquire in only 6 months datasets on ascending and descending orbits on different beams including a sufficient number of images for multipass differential interferometric processing. The result presented in Fig. 3 is relevant to the processing of 33 H-image ascending acquisitions (beam 09 corresponding to an incidence angle of about 40°) taken between April 4 and October 13 2009. Colors in the map move from blue, associated to movements toward the radar los (uplift), to red, associated to movements away from the sensor (subsidence). The velocity map (evaluated only on the post-seismic dates from April 12 to October 13) overlaid to a Google Earth image shows an area of subsidence with a rate exceeding the limit of 6 cm/year affecting the city of Paganica, in the eastern part of L'Aquila. In the same image, the plot of the time series of a point in the subsidence area shows the jump associated to the main shock (April 6) and the exponential decay of the subsidence with the aftershocks. Measurements were shown to be in full agreement with traditional leveling (Reale et al. 2011b) and GPS measurements (D'Agostino et al. 2012).

The second class of multipass techniques, the Persistent Scatterer Interferometry, as for SAR tomography, works at the highest spatial resolution to determine the deformation of single dominant scatterers typically associated with man-made structures (dihedral and trihedral of walls edges, poles, gratings, etc.) (Ferretti et al. 2000, 2001). In this case, to achieve also high accuracy in the estimation of the localization of scatterers, no limitations on the spatial baseline are introduced. Similarly to DInSAR stacking

and differently from SAR tomography, PSI uses only the phase information and assumes the presence of a single persistent scatterer per resolution cell retaining correlation over the time (persistent scatterers). The use of the model however presumes the compensation of phase contributions such as the atmospheric phase delay. Such a compensation can be carried out either by analyzing the phase on persistent scatterers (PS) candidates, which are strong scatterers where the phase is less affected by noise or by using the coarse resolution product of previous stacking techniques. In the latter case, a good practice is also to subtract the low-resolution, spatially correlated, deformation so to obtain also a zero deformation steering to carry out high-resolution analysis on residual phase signals: the following model is assumed for the vector $\boldsymbol{\varphi}$ collecting the compensated phase values in the N available acquisitions at full resolution

$$\boldsymbol{\varphi} = \frac{4\pi}{\lambda} \frac{\mathbf{b}}{r} s + \frac{4\pi}{\lambda} \boldsymbol{d}(s, \, \boldsymbol{t}) + \boldsymbol{\varphi}_n \qquad (2)$$

where $\mathbf{b}$ is the N-dimensional vector of spatial baselines and $\boldsymbol{d}(s, \, \boldsymbol{t})$ is the vector collecting the displacements measured at the acquisition instants collected in the vector $\boldsymbol{t}$. Following the compensation for low-resolution components, deformations are divided into the linear, described by the average (mean) velocity v of the pixel corresponding to the slope of the time series wrt to the epochs, and nonlinear $\boldsymbol{d}_{nl}$ addenda, i.e., $\boldsymbol{d}(s, \, \boldsymbol{t}) = v\boldsymbol{t} + \boldsymbol{d}_{nl}$.

PSI assumes nonlinear terms having a small amplitude and carries out, for each pixel, a measure of the correlation of the measured signal with the model in Eq. 2, through the maximization of the scalar product in the following, which returns also an estimation of the (s, v) parameters:

$$C = \max_{(s, v)} \frac{1}{N} \left| \left(e^{j\boldsymbol{\varphi}}\right)^H e^{j\frac{4\pi}{\lambda}\left(\frac{\mathbf{b}}{r}s + v\boldsymbol{t}\right)} \right| \qquad (3)$$

Only pixels for which C is above a fixed threshold are labeled as persistent scatterers and therefore the algorithm provides the temporal series as a product.

In the recent literature, the SqueeSAR approach has been proposed as an extension of the PSI technique to handle the impact of target decorrelation of distributed scatterers (Ferretti et al. 2011). SqueeSAR performs a phase filtering of the interferograms by exploiting the correlation matrix estimated from the data. SqueeSAR extends PSI to partially correlated (i.e., decorrelating) scatterers: the algorithm looks for a persistent scatterer (PS) mechanism equivalent to the real distributed scatterer, i.e., it assumes the response of the equivalent scatterer characterized only by phase variations with constant amplitude. Similarly to DInSAR stacking, SqueeSAR is tailored to the analysis of rural areas, however it does not perform a hard threshold on the baselines but rather it uses in a weighted way all the interferograms.

SAR Tomography System Model

SAR tomography is a step forward, respect to PSI: both are designed to monitor, at the full available spatial resolution, the deformation affecting ground scatterers. The key difference is however the different assumptions of the nature of the scattering which reflects into the interpretation of the received signal. As PSI processes only the phase information interferograms and therefore assumes the presence of a single scattering mechanism, SAR tomography removes this latter hypothesis and considers the complex value of each image pixel measured at the generic nth acquisition as the superposition of multiple elementary backscattering contributions distributed along the elevation s (Reigber and Moreira 2000; Gini et al. 2002; Fornaro et al. 2005). Differential tomography exploits the multitemporal characteristics to allow tomography to also monitoring deformation of scatterers. Particularly, a Fourier expansion of the deformation term $d(s, t)$ is introduced and, assuming the atmospheric phase delay caused by the propagation in atmosphere being compensated through a preliminary multipass DInSAR processing (f.i., SBAS), the measured signal at the generic antenna is therefore modeled as (Lombardini

2005; Fornaro et al. 2009a; Zhu and Bamler 2010a):

$$g_n = \iint_{I_s\,I_v} \gamma(s,v) e^{j\frac{4\pi b_n s}{\lambda\,r}} e^{j\frac{4\pi v t_n}{\lambda}} ds\,dv + w_n \quad (4)$$

which shows that, but for the noise term w_n, a 2D Fourier transform (FT) relationship stands between the data g_n and the backscattering distribution γ in the elevation/velocity (s, v) domain with respect to the Fourier conjugate variables $\xi_n = -2b_n/(\lambda r)$, associated with the elevation s, and $\eta_n = -2t_n/\lambda$, associated with the velocity v. In particular, as a consequence of the deformation term Fourier expansion, $\gamma(s, v)$ plays in v the role of the spectrum of the motion-related signal for elevation s. For linear deformation, the spectral velocity coincides with the deformation rate, i.e., $d(s, t_n) = v t_n$, and then v is usually referred to as deformation mean velocity, whereas for more complex motion, v identifies the (velocity) harmonic involved in the motion.

The tomographic problem consists of the estimation, in each image pixel, of the scene backscattering distribution $\gamma(s, v)$ starting from the N samples g_n and involves, in the most general case, the inversion of Eq. 4, that is, a spectral analysis of the data. This analysis moves the interferometric processing toward an imaging problem approach which extends the classical azimuth compression concepts widely known in the SAR 2D image focusing also for the third (elevation) dimension. Large antenna spans are coherently processed to achieve narrow responses and improve height resolution to the order of meters, allowing separating backscattering from source which are located at different heights. The scattering sources in the pixel can be spatially concentrated (compact scatterers), as for the layover in urban area where scatterers are typically located on the roofs and facades of building and interfere with those at lower heights, f.i., on the ground, or can be distributed along the elevation as for application in forest scenario where separation of ground level from canopy is of interest (Reigber and Moreira 2000; Cloude 2006; Tebaldini 2010).

The 4D model in Eq. 4 represents the most general imaging model which can be

particularized under specific conditions: in case of simultaneous acquisitions implying the absence of temporal diversity, as well as in case of absence of deformation, the signal model reduces to a 1D Fourier transform, and then the backscattering profile $\gamma(s)$ is of interest. This introduces to the 3D imaging framework. On the other hand, whereas uniform motion mostly applies for classical risk situations associated to slow, long-term deformation phenomena, more complex behavior can be also taken into account: the analysis of latest X-Band SAR data points out a higher sensitivity to small changes as those caused by the thermal dilation of materials (Reale et al. 2011a; Zhu and Bamler 2011). Moreover, since revisiting times are reduced with respect to the former generation of SAR sensors, the time requested to collect a sufficient number of images for reliable application of tomographic processing reduces to the order of 1 year, whereas typical C-Band ENVISAT and ERS acquisitions spanned a temporal interval of observation spanning several years. The reduced observation times may imply possible correlation with linear deformation behavior and impair the estimation of the deformation mean velocity (Reale et al. 2013). Even so, the sensitivity to thermal dilation can be exploited by SAR tomography as well. By extending the deformation model to account also for a second contribution linearly related to the average temperature T_n at the acquisition instants, i.e., $d(s, t_n) = vt_n + kT_n$, the order of the tomographic imaging can be extended (5D imaging) to estimate also a coefficient k which measures the expansion along the line-of-sight for each degree of temperature change (Zhu and Bamler 2011; Reale et al. 2013). In application to the monitoring of strategic infrastructures, this strategy allows estimating the stress induced by the temperature changes over the different segments of the structures (Fornaro et al. 2013).

SAR Tomography Imaging Algorithms

Multidimensional SAR imaging algorithms proposed in the literature typically work on a discretized version of the model in Eq. 4: letting $\boldsymbol{\gamma} = [\gamma(s_0, v_0), \ldots, \gamma(s_{M-1}, v_{M-1})]^T$ be the vector that collects the $M = M_s \times M_v$ samples of $\gamma(s, v)$ at the discrete points, hereafter called bins, (s_m, v_m), with $m = 0, \ldots, M - 1$, belonging to the $M_s \times M_v$ elevation/velocity discretization grid (T defines the transposition operator), and $\boldsymbol{g} = [g_0, \ldots, g_{N-1}]^T$ and $\boldsymbol{w} = [w_0, \ldots, w_{N-1}]^T$ the vectors collecting, for each pixel, the measured complex data and the noise contribution at each acquisition, respectively, the FT operator in Eq. 4 can be rewritten in the discrete case as

$$\boldsymbol{g} = \mathbf{A}\boldsymbol{\gamma} + \boldsymbol{w} \qquad (5)$$

where $\mathbf{A} = [\boldsymbol{a}_0, \ldots, \boldsymbol{a}_{M-1}]^T$ is the $N \times M$ system matrix collecting the steering vectors associated with each discretization bin synthetically defined as $\boldsymbol{a}_m = \boldsymbol{a}(s_m, v_m)$, whose generic element is $[\boldsymbol{a}_m]_n = \exp[-j2\pi(\xi_n s_m + \eta_n v_m)]/\sqrt{N}$.

Several techniques can be used to implement the imaging, that is, the inversion of Eq. 5 that leads to the estimation of the backscattering distribution in the elevation/velocity plane (s, v). Each is characterized by a different trade-off between simplicity, computational efficiency, sidelobes reductions, and super-resolution capability.

The beamforming (BF) technique represents the classical method to perform the inversion of Eq. 5: it makes use of the conjugate operator $\mathbf{A}^H$, with H being the Hermitian (conjugate transpose) operator, to profile the backscattering along the elevation bins (Fornaro et al. 2009a):

$$\hat{\boldsymbol{\gamma}} = \mathbf{A}^H \boldsymbol{g}. \qquad (6)$$

Once the backscattering in the (s, v) plane has been estimated, scatterers are selected by looking for strong peaks in $\hat{\boldsymbol{\gamma}}$. In this context, a tool for the effective selection of reliable scatterers is required: since in real data the useful information is corrupted by noise, a detection stage is required to control the false alarm rate, defined as the probability to declare the presence of a scatterer, whereas the scatterer is not really present on the ground. With reference to the case of a single

scatterer pixel, a test statistic based on the generalized likelihood ratio test has been proposed in the literature: it provides at the same time the maximum likelihood estimation of the (s, v) pair and declares the presence (hypothesis $\mathcal{H}_1$) or absence (hypothesis $\mathcal{H}_0$) of the scatterer for a given probability of false alarm by exploiting a test statistic which is strictly related to the BF reconstruction and is expressed as (De Maio et al. 2009):

$$\max_{(s,v)} \frac{|\boldsymbol{g}^H \boldsymbol{a}(s,v)|}{\|\boldsymbol{g}\| \|\boldsymbol{a}(s,v)\|} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} T \qquad (7)$$

where T is the detection threshold, belonging to the $[0, 1]$ interval and set according to the desired level of false alarm. It is worth noting that for single scatterers, the test statistic represents the highest peak of the normalized BF reconstruction and the argument of the maximization is just the

ML estimation of the (s, v) parameters. The GLRT in Eq. 7 allows also demonstrating the effectiveness of an imaging approach, as in SAR tomography, with respect to the classical interferometric processing. The PSI technique, in fact, exploits a similar test but for the use of only the phase information of each element of the vector $\boldsymbol{g}$ (Ferretti et al. 2000, 2001): the exploitation of the whole complex, amplitude plus phase, data as in Eq. 7 provides a significant increase, with respect to PSI, of the detection probability for a fixed false alarm rate.

Typical products achieved by the sequence of low-resolution DInSAR stacking techniques for the derivation of coarse-scale deformation and calibration of data for the subsequent full resolution processing is reported in Figs. 4 and 5. The processing involved a dataset of 29 Cosmo-Skymed images acquired over the city of Rome, Italy, from April 2011 to October 2012 in H-image mode which provides images with 3 m
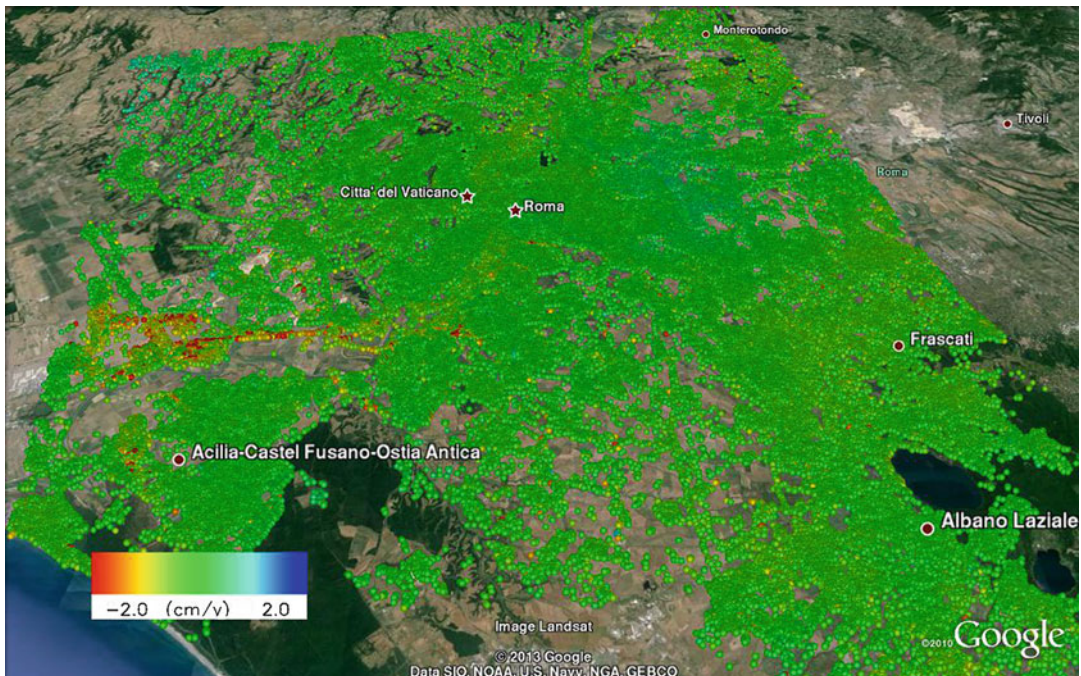
SAR Tomography for 3D Reconstruction and Monitoring, Fig. 4 Small-scale deformation mean velocity map estimated by the ESD technique on a dataset of 29 Cosmo-Skymed images over the city of Rome, Italy.

Colors are set on the estimated velocity and correspond to movements toward the sensor (uplift) in blue and away (subsidence) in *red*. Optical image courtesy of Google
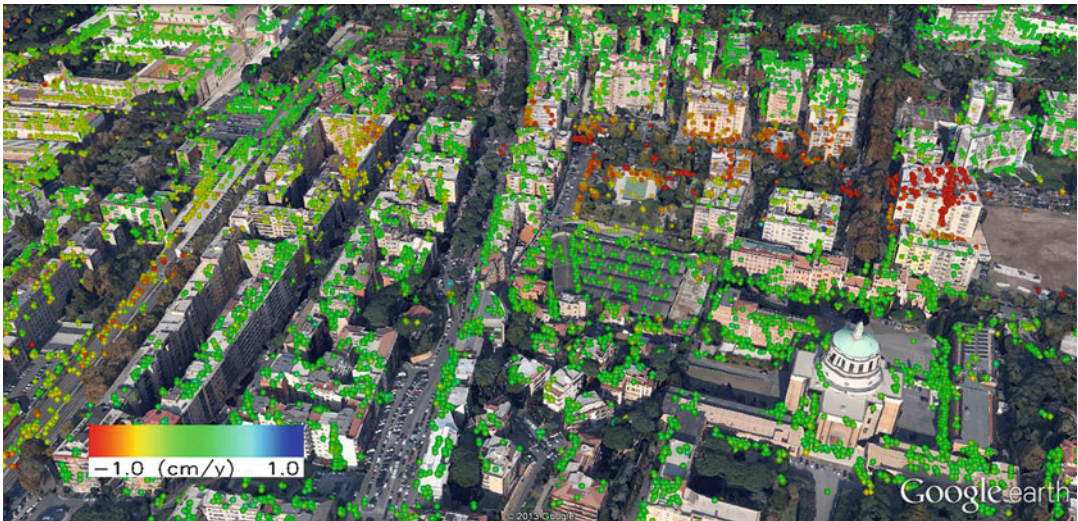
SAR Tomography for 3D Reconstruction and Monitoring, Fig. 5 3D visualization of the scatterers, represented as *dots*, detected by the GLRT after 4D beamforming reconstruction. *Colors* are set according to the estimated deformation mean velocity. The area is relevant to the Grotta Perfetta area in Rome, close to the Tevere River. Optical image courtesy of Google

of spatial resolution. According to the SBAS approach, a total number of 80 interferograms has been generated by imposing maximum spatial and temporal baseline spans of 1,000 m and 150 days, respectively. A multilook has also been carried out through the use of a 16×16 pixels moving average filter followed by a subsampling of the same factor. In Fig. 4 it is shown the deformation mean velocity map, superimposed to a Google Earth image, corresponding to the slope of the deformation time series for each pixel selected by looking at a quality index measuring the temporal consistency of measurements after phase unwrapping. As for the L'Aquila dataset, the low-resolution processing has been carried out through the ESD technique. Colors in the map move again from blue (uplift) to red (subsidence). The results highlight the presence of a distributed deformation pattern in the western part of the city which largely affects also the Rome airport. The large number of detected scatterers can be appreciated, although visualization constraints of the Google Earth environment impose a reduction of the total number of points, in this case of a factor 5. The dynamic range also hides some deformation signals occurring in the central part of the city, which is of interest for the application of full resolution tomographic analysis provided in a following section.

Figure 5 provides a close view, in a Google Earth framework, of the results of the application of 4D imaging and subsequent GLRT detection on the Cosmo-Skymed dataset over Rome introduced before. It allows attesting the capabilities of imaging approaches in the single building monitoring. The area, in the Rome city center, is affected by a severe subsidence induced by the consolidation of the alluvial sediments of the Tevere river. Each dot corresponds to a detected scatterer after geocoding from the original radar geometry to the natural geographic coordinates system: the correct estimation of heights allows effectively positioning scatterers, as they overlap the 3D models provided by Google Earth. Furthermore, the colors correspond to deformation velocity rate estimated for each scatterer: several buildings in the area suffer for large subsidence rates which cause extensive structural damages affecting their stability (Arangio et al. 2013).

S

The detection problem can be straightforwardly extended to the case of multiple scatterers per resolution cell, which may occur in layover areas. In this framework, a detector which tests the presence of possibly a maximum of two scatterers based on a sequential use of the GLRT discussed above has been also presented (Pauciullo et al. 2012). With reference to the separation of interfering scatterers within the same image pixel, in application to a real acquisition scenario, the acquisition geometry poses limitation on the imaging capabilities of BF: the baseline distribution defines the unambiguous elevation range that is the maximum extension allowed in elevation direction to avoid aliasing phenomena. Given Δb_n be the spatial separation between the successive antennas, the unambiguous elevation interval will be $\Delta_s = \lambda r/(2\Delta b)$, where Δb is the average value of Δb_n in case of nonuniform baseline distribution. Moreover, the final elevation resolution of BF cannot exceed the Rayleigh resolution given by $\Delta_s = \lambda r/(2B)$ with B being the total baseline span. In the same way, unambiguous velocity interval and velocity resolution can be also defined as $\Delta_v = \lambda/(2\Delta t)$ and $\Delta_v = \lambda/(2T)$, where Δt and T are the average temporal separation and the total temporal span, respectively. The elevation resolution plays a key role in the capability of distinguishing possible multiple scattering contributions within the same pixel: it implies the minimum separation of scatterers as they can be distinguished as separate scatterers. Latest satellite SAR missions as the TerraSAR-X are characterized by a narrow radius of the orbital tube, leading to small baseline excursions that result in final height resolution of some tens of meters. In application to urban areas, such a resolution can be very restrictive for the solution of the interference due to the layover. Finally, in all the practical applications, baselines are far from being uniformly distributed, thus BF gives poor reconstruction performances in terms of sidelobes and leakage in the (s, v) point spread function. Alternative strategies can be adopted to improve both the resolution performances and also the quality of the reconstruction of $\hat{\gamma}$ in the presence of highly uneven baseline distributions and at the same time achieve some super-resolution capabilities, i.e., the possibility to push the height resolution below the inherent Rayleigh limit.

The use of the singular value decomposition (SVD) of the operator $\mathbf{A}$ in Eq. 5 allows regularizing the inversion by restricting the solution space and benefiting of the inclusion of very limited a priori information on the expected scene elevation extent. The regularization, obtained through the so-called Truncated SVD, allows avoiding noise amplification and inversion instabilities and hence generally provides a better sidelobe reduction and as well as slight super-resolution than plain BF (Fornaro et al. 2009a).

Compressed sensing (CS) is a recent technique used in linear inversion problems for signal recovery that takes benefit of the hypothesis that the signal to be reconstructed have (in some basis) a sparse representation, i.e., a small number of nonzero entries. Under certain assumptions of the measurement matrix, the signal can be reconstructed from a small number of measurements. SAR tomography in urban areas is a favorable application scenario for CS due to the fact that for typical operative frequencies, the scattering occurs only on some scattering centers associated with ground, facades, and roofs of ground structures (Zhu and Bamler 2010b; Budillon et al. 2011).

Summary

Available high-resolution synthetic-aperture radar imaging sensors are capable of providing, in a systematic regular basis, images of single buildings and ground targets with very high spatial details. Persistent Scatterers Interferometry techniques have been already shown to largely benefit of the use of very high-resolution data to monitor buildings. Despite the spatial resolution increase, the steepness of the topography corresponding to vertical structured targets

generates critical distortion effect, the most critical ones are shadow and layover. These are major impairing sources in the analysis of SAR images corresponding to urban areas. SAR tomography represents a powerful technique which allows the implementation of a radar scanner from the space with large 2D antennas to scan the details of building and overcoming problems of layover to generate dense point-cloud measurements of buildings. Multidimensional SAR imaging, based on the concept of SAR tomography, is a tool that represents the most advanced method in the 3D reconstruction and monitoring of buildings. PSI and SAR tomography with very high-resolution sensors provide a unique too for application of spaceborne microwave radar imaging to urban areas, which is expected in the near future to play a key role in vulnerability and damage assessment of buildings and infrastructures.

Cross-References

▶ Building Monitoring Using a Ground-Based Radar
▶ InSAR and A-InSAR: Theory
▶ SAR Images, Interpretation of
▶ Urban Change Monitoring: Multi-temporal SAR Images

References

Arangio S, Calò F, Di Mauro M, Bonano M, Marsella M, Manunta M (2013) An application of the SBAS-DInSAR technique for the assessment of structural damage in the city of Rome. Struct Infrastruct Eng. doi: 10.1080/15732479.2013.833949

Bamler R, Hartl P (1998) Synthetic aperture radar interferometry. Inverse Probl 14:1–54

Berardino P, Fornaro G, Lanari R, Sansosti E (2002) A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. IEEE Trans Geosci Remote Sens 40:2375–2383

Budillon A, Evangelista A, Schirinzi G (2011) Three-dimensional SAR focusing from multipass signals using compressive sampling. IEEE Trans Geosci Remote Sens 49:488–499

Carnec C, King C, Massonnet D (1995) Measurement of land subsidence by means of differential SAR interferometry to sites of small extent. Geophys Res Lett 23:3579–3582

Cascini L, Ferlisi S, Peduto D, Fornaro G, Manunta M (2007) Analysis of DInSAR Data via geotechnical criteria. Rivista Italiana di Geotecnica 4:50–67

Cloude SR (2006) Polarization coherence tomography. Radio Sci 41:RS4017. doi:10.1029/2005RS003436

Crosetto M, Crippa B, Biescas E (2005) Early detection and in-depth analysis of deformation phenomena by radar interferometry. Eng Geol 79:81–91

Curlander JC, McDonough R (1991) Synthetic aperture radar – system and signal processing. Wiley, New York

D'Agostino N, Cheloni D, Fornaro G, Giuliani R, Reale D (2012) Space-time distribution of afterslip following the 2009 L'Aquila earthquake. J Geophys Res 117: B02402. doi:10.1029/2011JB008523

De Maio A, Fornaro G, Pauciullo A (2009) Detection of single scatterers in multi-dimensional SAR imaging. IEEE Trans Geosci Remote Sens 47:2284–2297

Ferretti A, Prati C, Rocca F (2000) Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry. IEEE Trans Geosci Remote Sens 38:2202–2212

Ferretti A, Prati C, Rocca F (2001) Permanent scatterers in SAR interferometry. IEEE Trans Geosci Remote Sens 39:8–20

Ferretti A, Fumagalli A, Novali F, Prati C, Rocca F, Rucci A (2011) A new algorithm for processing interferometric data-stacks: squeeSAR. IEEE Trans Geosci Remote Sens 49:3460–3470

Fornaro G, Franceschetti G (1999) SAR interferometry, Chapter IV. In: Franceschetti G, Lanari R (eds) Synthetic aperture radar processing. CRC Press, Boca Raton

Fornaro G, Serafino F (2006) Imaging of single and double scatterers in urban areas via SAR tomography. IEEE Trans Geosci Remote Sens 44:3497–3505

Fornaro G, Serafino F, Lombardini F (2005) 3D multipass SAR focusing: experiments with long-term spaceborne data. IEEE Trans Geosci Remote Sens 43:702–712

Fornaro G, Reale D, Serafino F (2009a) Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers. IEEE Trans Geosci Remote Sens 47:224–237

Fornaro G, Pauciullo A, Serafino F (2009b) Deformation monitoring over large areas with multipass differential SAR interferometry: a new approach based on the use of spatial differences. Int J Remote Sens 30:1455–1478

Fornaro G, Reale D, Verde S (2013) Bridge thermal dilation monitoring with millimeter sensitivity via multidimensional SAR imaging. IEEE Geosci Remote Sens Lett 10:677–681

S

Gernhardt S, Adam N, Eineder M, Bamler R (2010) Potential of very high resolution SAR for persistent scatterer interferometry in urban areas. Ann GIS 16:103–111

Ghiglia DC, Pritt MD (1998) Two-dimensional phase unwrapping: theory, algorithms, and software. Wiley-Interscience, New York

Gini F, Lombardini F, Montanari M (2002) Layover solution in multibaseline SAR interferometry. IEEE Trans Aerosp Electron Syst 38:1344–1356

Krieger G, Moreira A, Fiedler H, Hajnsek I, Werner M, Younis M, Zink M (2007) TanDEM-X: a satellite formation for high-resolution SAR interferometry. IEEE Trans Geosci Remote Sens 45:3317–3341

Lombardini F (2005) Differential tomography: a new framework for SAR interferometry. IEEE Trans Geosci Remote Sens 43:37–44

Massonnet D, Rossi M, Carmona C, Adragna F, Peltzer G, Fiegl K, Rabaute T (1993) The displacement field of the Landers earthquake mapped by radar interferometry. Nature 364:138–142

Moreira A, Prats-Iraola P, Younis M, Krieger G, Hajnsek I, Papathanassiou KP (2013) A tutorial on synthetic aperture radar. IEEE Geosci Remote Sens Mag 1:6–43

Pauciullo A, Reale D, De Maio A, Fornaro G (2012) Detection of double scatterers in SAR tomography. IEEE Trans Geosci Remote Sens 50:3567–3586

Reale D, Fornaro G, Pauciullo A, Zhu X, Bamler R (2011a) Tomographic imaging and monitoring of buildings with very high resolution SAR data. IEEE Geosci Remote Sens Lett 8:661–665

Reale D, Nitti DO, Peduto D, Nutricato R, Bovenga F, Fornaro G (2011b) Post-seismic deformation monitoring with the COSMO/SKYMED constellation. IEEE Geosci Remote Sens Lett 8(4):696–700

Reale D, Fornaro G, Pauciullo A (2013) Extension of 4-D SAR imaging to the monitoring of thermally dilating scatterers. IEEE Trans Geosci Remote Sens. 51:5296–5306

Reigber A, Moreira A (2000) First demonstration of airborne SAR tomography using multibaseline L-band data. IEEE Trans Geosci Remote Sens 38:2142–2152

Tebaldini S (2010) Single and multipolarimetric SAR tomography of forested areas: a parametric approach. IEEE Trans Geosci Remote Sens 48(5):2375–2387

Van Zyl J (2001) The Shuttle Radar Topography Mission (SRTM): a breakthrough in remote sensing of topography. Acta Astronaut 48:559–565

Zhu X, Bamler R (2010a) Very high resolution spaceborne SAR tomography in urban environment. IEEE Trans Geosci Remote Sens 48:4296–4308

Zhu X, Bamler R (2010b) Tomographic SAR inversion by L1 norm regularization – the compressive sensing approach. IEEE Trans Geosci Remote Sens 48:3839–3846

Zhu X, Bamler R (2011) Let's do the time warp: multicomponent nonlinear motion estimation in differential SAR tomography. IEEE Geosci Remote Sens Lett 8:735–739

School Seismic Safety and Risk Mitigation

Marla Petal[1], Ben Wisner[2,3], Ilan Kelman[4,17], David Alexander[2], Omar-Dario Cardona[5], Djillali Benouar[6], Sanjaya Bhatia[7], Jitendra Kumar Bothara[8,9], Amod Mani Dixit[8], Rebekah Green[10], Ram Chandra Kandel[11], Tracy Monk[12], Bishnu Pandey[13], Janise Rodgers[14], Zeynep Türkmen Sanduvaç[15] and Rajib Shaw[16]

[1]Risk RED (Risk Reduction Education for Disasters), Los Angeles, CA, USA
[2]University College London, London, England
[3]Oberlin College, Oberlin, OH, USA
[4]Institute for Risk and Disaster Reduction and Institute for Global Health, University College London, London, England
[5]Universidad Nacional de Colombia, Manizales, Colombia
[6]University of Bab Ezzouar, Algiers, Algeria
[7]UNISDR Recovery Platform, Kobe, Japan
[8]NSET, Kathmandu, Nepal
[9]Miyamoto Impact, Christchurch, New Zealand
[10]University of Western Washington, Bellingham, WA, USA
[11]Toronto, ON, Canada
[12]Families for School Seismic Safety, Vancouver, Canada
[13]University of British Columbia, Vancouver, Canada
[14]GeoHazards International, Menlo Park, USA
[15]Risk RED, Istanbul, Turkey
[16]University of Kyoto, Kyoto, Japan
[17]Norwegian Institute of International Affairs, Oslo, Norway

Synonyms

Comprehensive school safety; Critical infrastructure; Safe school construction; Safe school facilities; School construction

Introduction

Access to education is a basic human right. It is enshrined in Convention on the Rights of the

Child (1990), the World Declaration on Education for All (in 1990), and the World Education Forum (WEF 2000). It is one of the Millennium Development Goals for the decade starting in 2005 and continues to be part of the "post-2015" development agenda. Education is strongly associated with poverty reduction, and there are strong global and national drives to implement it. The Global Partnership for Education has 29 national partners, supporting the implementation of universal, free, quality basic education in 57 partner developing countries. In GPE's 2012–2015 Strategic Plan, the first of its four strategic goals is "All children have access to a safe, adequately equipped space to receive an education with a skilled teacher." However, none of its monitoring indicators mention safety. Since 2004, the GPE has contributed to build, rehabilitate, and equip close to 53,000 classrooms (GPE 2014). However, up until at least 2013, there was no systematic due diligence with respect to disaster-resilient construction. In the rush to fulfill the right to education, are children being put at risk?

This entry assesses seismic threats to schools and reviews incidents of children and teachers killed by structural failure of school buildings as well as structural damage to schools and near misses. It reviews progress, good practices, and lessons learned based on these threats. The entry goes on to overview school vulnerability, global experiences in seismic-resistant school construction, and retrofit. A summary of progress in school seismic safety along with a recommended systematic all-hazards approach to comprehensive school safety set the stage to make the case for continued advocacy for school seismic safety.

The Threat

According to estimates made by the Center for International Earth Science Information Network at Columbia University in 2006, there are more than 100 million school-aged children exposed to significant seismic risk throughout the world (New York Times 2008). In 2004, 10 of the 16 contributors to this article initiated recording of case studies on progress and struggles for school seismic safety which are updated in the entry "▶ School Seismic Safety: Case Studies." They set out the magnitude of concern and main arguments for advocacy in an unpublished article. At the time, the authors posited the gruesome estimate of "4,800 school children killed by earthquake-related school collapse or severe damage over the next decade... It might be reasonable and prudent to plan to avoid a loss of student life in earthquakes of somewhere between 2,000 and 5,000 in a 10 year period." At the time it was written, this estimate seemed to the authors to be somewhat alarmist.

The following year, shortly after the unprecedented destruction caused by the Indian Ocean earthquake and tsunami, 168 countries agreed to the 2005–2015 Hyogo Framework for action. Over the course of this 10 year period, this dire predication has been exceeded fourfold as the result of only two major earthquakes during school hours: the 8 October 2005 Kashmir earthquake which killed more than 18,000 students, in addition to staff, in schools, and the 12 May 2008 Sichuan earthquake which killed more than 5,300 students, in addition to staff, in their schools (UNISDR 2008).

In the powerful earthquake and massive tsunami on 11 March 2011 in northern Japan, schools themselves were by and large structurally sound and resisted earthquake damage, but tsunami-retaining walls were breached as the tsunami was larger than expected and land subsidence had not been factored in. Disaster drills and practice of "tendenko" (automatic tsunami evacuation), by many school children, saved many lives. Some schools provided vertical evacuation, and many survived at evacuation and shelter centers. But instances of confusion occurred and many school pupils and teachers also died. Today, the students now displaced by the resulting nuclear disaster recognize this neglected threat as the most catastrophic of all. The international community is virtually silent on this threat.

In common with other infrastructure, school buildings are subject to damage and collapse in earthquakes. Many of these have resulted in children killed while being educated (Table 1).

S

Pictorial evidence of historic earthquake damage to schools is available in the NISEE, Earthquake Engineering Online Archive.

There have also been many cases when an earthquake destroyed school buildings when they were not in session, and thus deaths and injuries were narrowly avoided (Table 2). However the severe impact on continuity of education and the potential magnitude of loss of life in these events further highlight the importance of ensuring the seismic safety of schools.

Making the Case for School Seismic Safety

Many public buildings and different sorts of critical infrastructure are threatened by earthquakes. The case can be made for giving priority to

School Seismic Safety and Risk Mitigation, Table 1 Children killed by structural failure of school buildings

Date/local time (Source)	Location/magnitude[a]	Consequences/schools	Consequences/children
12 Jan 2010 16:53 (CNN 2014)	Port-au-Prince, Haiti M 7.0	MoE estimates 4,992 schools affected (23 % of the nation's schools)	Deaths and injuries unknown. Many children with disabling injuries. Some schools were holding their third shifts. Est. 1.3 m children and youth affected
12 May 2008 14:28 (COGGS 2008)	Wenchuan, China M.7.9	175 schools (7,000 classrooms) in Sichuan and Shaanxi provinces were destroyed	>5,300 school children died in dozens of schools In the Beichuan Middle school, 1,300 of 2,999 students and teachers died
6 Mar 2007 11:00 (COGGS 2008)	Western Sumatra M 6.4	The wall of a primary school collapsed. Fire followed. Up to 329 schools affected by several earthquakes (2005–2010)	4 primary school children died
8 Oct 2005 St. 08:50 (UNISDR 2008)	Kashmir, Pakistan, and India M 7.6	More than 10,000 schools collapsed 80 % of Mahesehra's 2,749 66 % of Batagram's 678, and 37 % of Abbottabad's 1,829 public schools were destroyed or seriously damaged	>18,000 school children died >50,000 school children were seriously injured
1 May 2003 03:20 (Rodgers 2012)	Bingöl, Turkey M 6.4	4 school buildings collapsed. Only the dormitory was occupied	84 students killed and 114 survived in the dormitory
24 February 2003 10:03 (COGGS 2008)	Bachu, Xinjiang, China M 6.4	900 classrooms collapsed	Students were outside in physical education at the time of the earthquake. At least 20 students killed in one middle school collapse
31 October 2002 11:40 (COGGS 2008)	San Giuliano di Puglia, Molise, Italy M 5.9	San Giuliano infant school collapsed	26 children and 3 adults killed. 35 children rescued alive from the building but some reports suggest that one child died later
26 January 2001 Friday 08:16 Republic day holiday (COGGS 2008)	Gujarat, India M 7.6	1,884 school buildings collapsed. 5,950 classrooms destroyed. 36,584 unfit for instruction	971 school children and 31 teachers were killed in school activities. 1,051 students and 95 teachers seriously injured 32 children died at Swaminarayan School

(continued)

School Seismic Safety and Risk Mitigation, Table 1 (continued)

Date/local time (Source)	Location/magnitude[a]	Consequences/schools	Consequences/children
13 February 08:22 and 13 January 2001 (COGGS 2008)	El Salvador M 6.6	85 schools damaged beyond repair. In aftershock 22 preschoolers and their teacher were killed	50 % of fatalities were children
9 July 1997 15:24 (COGGS 2008)	Cariaco, Venezuela M 7.0	Two out of five school buildings collapsed. Four reinforced concrete buildings had serious structural defects	46 students killed
10 May 1997 12:57 (COGGS 2008)	Ardekul, Iran M 7.3	Elementary school collapsed	110 young girls were killed
1992 (COGGS 2008)	Erzincan, Turkey M 6.9	6-story medical school collapsed	62 students were killed.
7 December 1988 11:41 (COGGS 2008)	Spitak, Armenia M 6.8	380 children and youth institutions destroyed. 105 of 131 in Spitak and Leninakin destroyed	Likely thousands of school children killed. At least 400 children died in the collapse of a Dzhrashen elementary school
27 July 1976 03:42 (COGGS 2008)	Tangshan, China M 7.8	Most school buildings destroyed	2,000 students killed in the dormitory of the College Mining Institute
13 April 1949 11:58 (COGGS 2008)	Olympia, Washington, USA M 7.1	10 schools destroyed, 30 damage. Spring break	2 children in school were killed
31 October 1935 (COGGS 2008)	Helena, Montana, USA M 6.2	Newly built secondary school wing collapsed	2 students killed. Classes not in session, could have been much worse
10 March 1933 Long Beach (COGGS 2008)	Long Beach, California, USA M 6.4	70 schools destroyed. 120 with major damage. Classes held in tents for 2 years. First legislation for safe school construction	2 children died in gymnasium collapse. Spring break, classes not in session, could have been much worse

[a]Magnitudes from USGS or Wikipedia

schools from three perspectives: Duty bearers have moral and legal obligations to fulfill children's rights to both safety and survival and educational continuity. In more affluent countries, the cost benefits of investments in public safety, the importance of safeguarding development investments, and preventing educational disruption are undisputed. And, the uses of school buildings as multipurpose community centers and disaster shelters, even when children are not harmed, have cascading social and economic consequences beyond the replacement cost of school buildings themselves. In most cases public discussion and debate on these issues tend to mix these ethical and pragmatic arguments.

Human Rights Argument

The human rights argument suggests that no society should tolerate a choice between the safety of children's lives and their education. The right to

S

School Seismic Safety and Risk Mitigation, Table 2 School structural damage from earthquakes

Date (Source)	Location/magnitude[a]	Structural and educational impacts
2 July 2013 14.37 (Pandey 2013)	Aceh, Indonesia M 6.2	966 schools affected
4 April 2010 15:40 (Rodgers 2012)	California, USA, and Baja, Mexico M 7.2	Significant structural damage to several schools in Mexico. Significant nonstructural damage to several schools in the USA. Cost of repair almost 20 % of annual budget for one school district School was on spring break. Nonstructural damage would have caused injuries and blocked egress. In California hazardous asbestos from collapsed walkways and mercury from light fixtures closed schools for extended periods
30 September 2009 17:16 (Rodgers 2012)	Padang, West Sumatra, Indonesia M 7.6	2,164 severely damaged, 1,447 moderately damaged, 1,137 lightly damaged School was recently dismissed for the day. Temporary school buildings of timber frame and corrugated steel
2 September 2009 14:55 (Pandey 2013)	West Java, Indonesia M 7.0	716 schools affected
21 September 2009 14:53 (Rodgers 2012)	Mongar, Bhutan M.6.1	91 schools affected: 6 destroyed, 17 required major repair, 44 required partial repair, 24 required minor repair (cost 12.9 m USD) plus damage to boarding schools, water, and sanitation School was dismissed early for holiday. Temporary learning facilities inadequate for weather
6 April 2009 03:32 (Rodgers 2012)	L'Aquila, Italy M 6.3	78 schools had extended closures and 12 partial closures
12 September 2007 18:10 (Pandey 2013)	Bengkulu, Indonesia M 8.5	240 schools affected (2005–2010)
15 August 2007 23:40 (Rodgers 2012)	Pisco, Peru M 8.0	116 schools were severely damaged. 478 classrooms were needed to restore school activities
27 May 2006 05:53 (Rodgers 2012)	Yogyakarta, Indonesia M 6.3	Yogyakarta: 2155 educational facilities damaged or destroyed; Central Java: 752 damaged or destroyed. Damage and losses estimated at 1.7 trillion Indonesian Rupiah
26 December 2004 early am (UNISDR 2008)	Indonesia, Sri Lanka, Maldives, Thailand M 9.1–9.3	School earthquake and tsunami damage combined: Indonesia – 750 destroyed 2,135 damaged. Sri Lanka – 51 destroyed, 100 damaged. Maldives – 44 destroyed or damaged. Thailand – 30 destroyed
26 December 2003 5:26 (COGGS 2008)	Bam, Iran M 6.6	67 of 131 schools collapsed. The remaining 64 were heavily damaged and unusable 33,000 students were affected
22 September 2003 12:45 (UNISDR 2008)	Puerto Plata, Dominican Republic M 6.4	50 public schools damaged, 140 classrooms impacted 18,000 students were without classrooms
21 May 2003 19:48 (COGGS 2008; OECD 2004)	Boumerdes, Algeria M 6.8	130 schools damaged beyond repair. 753 schools extensively damaged or destroyed The earthquake occurred out of normal school hours, so children were not at school. Cost of school rehabilitation $79 million+
24 February 2003 (COGGS 2008)	Xinjiang, China	Dozens of schools collapsed The earthquake struck 27 minutes before thousands of children would have been in classrooms
25 April 2002 22:41 (Rodgers 2012)	Tbilisi, Georgia M 4.5	Approximately $8 million US in school damage. No collapses; 1 school with very heavy damage; 35 with substantial damage; 68 with moderate damage; 98 with negligible or slight damage

School Seismic Safety and Risk Mitigation, Table 2 (continued)

Date (Source)	Location/magnitude[a]	Structural and educational impacts
21 September 1999 1:47 (COGGS 2008)	Chi-Chi, Taiwan M 7.7	51 schools collapsed. 786 schools nationwide were damaged. 22 % of schools and 71 % of post-secondary institutes damaged The earthquake happened in the middle of the night, so no one was in the building. Cost of repair and reconstruction $1.3 billion
June 23 2001 15:33 (COGGS 2008)	Arequipa, Peru	98 school buildings seriously damaged School not in session on Saturday
17 August 1999 3:02 (COGGS 2008)	Kocaeli, Turkey M 7.6	43 schools were damaged beyond repair. 381 minor to moderate damage In Istanbul 60 km away 35 schools were unsafe and demolished School was not in session but was suspended for 4 months. In Istanbul 131 schools were closed temporarily, for inspection
25 January 1999 13:19 (OECD 2004)	Pereira and Armenia, Colombia M 6.2	74 % of schools in Pereira and Armenia were damaged School was not in session
20 August 1998 (UNISDR 2008)	Udayapur, Eastern Nepal M 6.6	1,200 schools heavily damaged or destroyed. 6,000 affected
17 July 1998 00:19	Papua, New Guinea M 7.0	Schools destroyed
9 July 1998 5:19 (COGSS 2008)	Faial, Azores, Portugal M 6.2	Schools damaged School was not in session
20 May 1998	Afghanistan/Tajikistan M 6.6	Unknown
12 November 1996 15:33 (COGSS 2008)	Nazca, Peru	93 school buildings seriously damaged
1996 (OECD 2004)	Temouchent, Algeria M 5.6	6 schools destroyed, 17 moderate damage, 36 light damage
17 January 1995 5:46 (COGSS 2008)	Hanshin-Awaji, Japan M 6.9	54 buildings school damaged beyond repair. Extensive earthquake and fire damage to 4,500 educational buildings. ¥94 billion. School was not in session
1994 (OECD 2004)	Beni Chourgrane, Aleeria M 5.6	4 schools destroyed, 16 moderate damage, 30 light damage
17 January 1994 04:31 (FEMA 2011)	Northridge, California, USA M 6.7	24 of 127 affected schools suffered significant structural damage. Suspended lighting and ceiling systems were damaged in 1,500 buildings 2, 204 schools were used as shelters. Had this occurred during the school day, significant injuries and lack of safe egress for thousands would have resulted. The Los Angeles Unified School District, amongst others, embarked on projects for nonstructural mitigation, now the responsibility of school maintenance personnel
25 March 1993 (COGSS 2008)	Scott Mills, Oregon, USA	Part of masonry school building collapsed Spring break, school was not in session
17 October 1989 17:04 (EERI 1990)	Loma Prieta, California, USA M 6.9	7 schools in three districts and one headquarters sustained severe damage. 1,544 were schools surveyed. Total value of damage $81 million USD
10 October 1989 (OECD 2004)	El Asnam, Algeria M 7.3	70–85 schools suffered extensive damage or collapsed The earthquake occurred out of normal school hours, so children were not at school
20 August 1988 4:39 (COGSS 2008)	Bihar, India, and Udaypur Nepal M 6.6	950 school buildings were damaged in Bihar 6,000 schools destroyed in Nepal School was not in session

School Seismic Safety and Risk Mitigation, Table 2　(continued)

Date (Source)	Location/magnitude[a]	Structural and educational impacts
8 November 1988 (UNISDR 2008)	Yunan, China	1,300 schools destroyed in earthquake
19 September 1985 :17 (COGSS 2008)	Mexico City, Mexico M 8.0	137 school buildings collapsed, 1,687 school buildings were damaged Schools were not yet open
2 May 1983 23:42 (COGSS 2008)	Coalinga, California, USA	Extensive nonstructural damage noted
10 October 1980 13:25 (OECD 2004)	El Asnam, Algeria M 7.3	70 schools totally destroyed, 25 moderately damaged School was not in session
9 February 1971 06:01 (State of California 2009)	Sylmar, California, USA M 6.6	Only 4 of 1,544 buildings surveyed suffered severe damage. Nearly all damage was nonstructural School was not in session
31 May 1970 4:23 (COGSS 2008)	Chimbote, Peru M 7.9	6,730 classrooms collapsed and hundreds seriously damaged
27 March 1964 (COGSS 2008)	Alaska, USA	Primary school destroyed by an earthquake-induced landslide. Half of Anchorage's schools were significantly damaged The earthquake struck on a holiday, Good Friday, so schools were closed
1963 (COGSS 2008)	Skopje, Macedonia	44 schools (57 % of urban stock) were damaged 50,000 students affected. Sunday, school not in session
21 July 1952 4:52 (COGSS 2008)	Kern County, California, USA M 7.3	20 schools damaged or destroyed (most built before 1933). Significant nonstructural damage also noted
4 March 1952 (USGS 2003)	Sapporo, Japan	400 schools collapsed in Sapporo
10 March 1933 (COGSS 2008)	Long Beach, California, USA	70 schools collapsed The earthquake hit early in the evening after children had left for the day which saved their lives. Five students were killed in a gymnasium
3 February 1931 (Dowrick and Rhoades 2004)	North Island, New Zealand	Several schools were severely damaged The earthquake happened at mid-morning during school playtime when the children were outdoors enjoying the summer weather. Some students were killed, but the death toll could have been several hundreds
17 June 1929 10:17	Murchison, New Zealand	College tower and dormitory roofs collapsed School was not in session
18 April 1906 05:12	San Francisco, California USA	28 schools burned in fire. 41 schools damaged or destroyed Classes were not in session

[a]Magnitudes from USGS

Bibliographic references on many structural impacts on schools are available on the internet (Rodgers 2012)

life and the right to education are both recognized human rights, and both should be met. This argument takes on additional salience in view of the current international effort to increase school enrollment and attendance by girls, disabled children, and children of the very poor and marginalized groups in society.

Around the world, at least 100 million children of school age do not attend school representing about 14 % of the world's children (UNESCO 2004). Providing facilities to educate them requires construction of schools and rapid expansion of building programs. The Education for All campaign originally

hoped to enroll 24 million of these children in a decade. Millennium Development Goals (MDGs) specifically aim to "[e]nsure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling."

In 2004 it was estimated that more than 7,500 new schools were needed within the next 3 years solely in Afghanistan, a country with a significant seismic hazard. It would be ironic and tragic if in the course of achieving one MDG, another is undermined.

Another MDG target is to reduce the under-five mortality rate by two thirds, between 1990 and 2015. On the one hand, the international community is seeking to save the lives of under-fives, only to put them at risk a few years later when they go to school.

Educational authorities changed with the construction and maintenance of schools are also the ones tasked with many other functions: They develop curricula, hire teachers, and choose educational resources such as textbooks and computers. School safety issues have to find a place in the capital, maintenance, and operation budgets of school buildings and school operation.

Retrofitting schools for seismic safety can be perceived to compete for funds with the rest of the educational process. The question facing decision makers can actually appear to be: "What is more important: an up-to-date textbook and good laboratory facilities now or a building that can withstand an extreme event which might or might not occur with the next few decades?"

Under most circumstances, young people do not lobby for their own rights to health and safety. Children cannot refuse to go to school because a building is unsafe. By law, they must attend school, though teachers, parents, and others may advocate on their behalf. Faculty and support staff in schools should also be concerned for their occupational safety and theoretically be natural advocates of safe school facilities. Yet there are no examples mentioned to date of teachers unions becoming involved in the issue of school disaster vulnerability.

Cost-Effectiveness Arguments

There are two forms that cost-effectiveness arguments may take. One asserts that the authority responsible for education incurs greater cost in the long run to repair and replace schools damaged by earthquakes than the cost of enforcing building codes and making sure that every new school is a safe school (or even of retrofitting older or poorly built schools). In some cases, replacement of unsafe schools is more cost-effective than repair (e.g., see entry "▶ School Seismic Safety: Case Studies," for examples from Algeria, Colombia, and Turkey). Notable studies of the benefits and costs of retrofitting schools in the USA, Italy, Mexico, and Peru have been published in the decade between 2004 and 2014.

A more ambitious and difficult case to make concerns the relative cost-effectiveness of investments in school seismic safety when compared to investing that money in other kinds of public health, safety, and welfare. In cases where child and infant mortality is high, longevity is shorter, basic vaccinations are not universal, or safe domestic water and sanitation facilities are inadequate, then the relative ranking of school safety as a cost-effective public health intervention may be low. Competition for public health funds could occur in trying to decide between clean water and vaccinations for everyone versus school seismic safety. In more affluent countries, the cost-effectiveness of saving lives in a future disaster usually has a high place among prioritized goals.

Of course, the physical safety of children, both in schools and out in the world at large, goes well beyond school seismic safety. HIV/AIDS, malnutrition, sexual violence, malaria, labor practices, and forced military service are day-to-day threats to the physical safety of many of the world's children. The small potential for an earthquake over the next century might appear to pale beside other concerns which daily kill many more children.

However, in places where school seismic safety is a prominent issue – such as Tehran, Vancouver, Kathmandu, Bogotá, and Wellington – a significant earthquake has a high

S

probability of happening during the lifetime of schools currently standing and, therefore, for the gradually changing cohort of children during the life of the building. If earthquakes happen with equal probability around the clock, then approximately a 6–23 % chance exists of schoolchildren being in the school during a damaging earthquake. Cost-benefit studies of seismic construction estimate that it would add about 5 % to the cost of building a school in the USA, and in other countries the highest estimates are about 15 %, making the expectation that "every new school be a safe school" a realistic expectation.

When a population at risk is predominantly children, depending on the country, each death represents 40–70 years of lost life and productivity, and each injury represents 40–70 years of potentially expensive medical care, such as for brain or spinal injuries. Fix schools and several generations of children are protected. Health economics and medical ethics agree that the greatest social benefit comes from investment in the health and capacities of children. Aside from saving lives, the cost of education interrupted, and the serious potential for drop out adds another cost factor that seismic safety could help avoid.

Argument from the School's Multiple Functions

The symbolic, cultural, economic, and political significance of schools as a community hub gives them an importance beyond merely being the site for educating children. Schools often play roles as central places for meetings and group activities, including literacy classes, religious services, political activities, and marriage ceremonies, particularly in rural areas where the school might be the only location big enough to hold such an event. Schools may also provide essential nutrition programs and serve as makeshift hospitals or vaccination centers even in normal times.

Where schools are the safest buildings in a community, they often serve as temporary shelter from storms and floods. They may be staging areas for first aid or rescue operation or other disaster response functions and even provide temporary housing, while still fulfilling their role as an education facility.

Thus schools have a value in the social fabric of a community, providing adult education, promoting public health, building and maintaining sustainable livelihoods, and protecting people. The monetary value of those social gains defies estimation but clearly adds value and further justifies investment in safe school construction and maintenance.

We know from many disasters the important role that schools play in anchoring and speeding community recovery. Rapid school re-opening has tangible benefits in terms of children who are safe, supervised, and progressing towards their educational goals. Intangible benefits of schools functioning normally following a disaster include the psychosocial support in the face of loss and change. The importance of operational continuity of schools is linked to community recovery.

To take another example, retrofitting can spread a message far beyond the school. When children see their school being seismically retrofitted, they may have and may be designed to have ripple effects on safer residential construction. However, this is by no means automatic, and just how to maximize school construction or retrofit experience into a wider learning opportunity is a promising line of pursuit. Schools certainly serve as community hubs for propagating the seismic safety messages. School seismic safety can not only protect a community's children but also educate communities to protect themselves.

Progress, Good Practices, and Lessons Learned

Assessing School Safety from Disasters, A Global Baseline Report (UNISDR 2012) found several consistent threats to safe school facilities:

- **Failure to assure every new school is a safe school:** Neither donors, governments, nor NGO associations have unequivocally committed to providing evidence or assurances or submitted to monitoring to assure that every new school is a safe school. Many small-scale

donors are particularly unaccountable and are not reached by the same accountability mechanisms and efforts of UN agencies and major international non-governmental agencies.

- **Multi-hazard awareness is often lacking:** In the construction of school facilities, there are many examples of fulfilling resilience to one hazard, while failing to mitigate against others – sometimes resulting in schools being dangerous in spite of good intentions or lying unused.
- **Impact of construction on education and family life not well understood:** School remodeling, retrofit, and replacement all have an impact on existing school programs and families. Planning these projects to minimize adverse impacts continues to be a concern.
- **Opportunity for construction and retrofit as an educational experience is untapped:** School construction and retrofit provide ideal opportunities for students and communities to learn the many principles of disaster-resilient construction to be applied throughout their communities. This opportunity is typically wasted as school sites are hidden from view and the experience is not used as a learning opportunity.
- **Lifeline infrastructure failures threaten school attendance:** Vulnerabilities in roads, bridges, and transportation systems must be prioritized when school attendance is threatened.
- **Failure to prioritize school re-opening jeopardizes community recovery:** Schools play a critical role in disaster recovery and community resilience where adults cannot return to work (UNISDR 2012).

The same study found consensus around the following core commitments required for safe school facilities: (1) Every new school must be a safe school. (2) Legacy schools should be prioritized for replacement and retrofit. (3) Lifeline infrastructure and nonstructural safety should be assessed locally and measures taken to mitigate [dangers]. (4) School furnishings and equipment should be designed and installed to minimize potential harm they might cause to school occupants.

The expert review process that was part of the *Guidelines for Safer School Construction* (INEE 2010) yielded identification of a rich set of enabling factors associated with successful and sustained programs for school structural safety that all school safety advocates need to consider awareness, community ownership, partnership and dialogue, quality assurance, appropriate technology, integrated education, cultivating innovation, encouraging leadership, and continuous assessment and evaluation.

Overview of School Building Vulnerability

Rodgers (2012) reviewed earthquake damage assessment reports through 2009, for 32 earthquakes globally and aggregated findings from 31 school building vulnerability assessments. Table 3 shows the most commonly cited sources of vulnerability from both sources.

The general lack of agreement between vulnerability assessment and damage data likely reflects fragmentary and typically inadequate efforts to collect school damage data following past earthquakes, as well as a tendency for vulnerability assessments to identify common characteristics (such as plan irregularities) that rarely lead to the severe damage noted in post-earthquake damage surveys and reconnaissance reports. More complete earthquake damage data would provide the best indicator of the vulnerability-creating characteristics more likely to cause severe damage, because many vulnerability assessments do not differentiate the severity of damage expected from observed deficiencies.

The sources of and characteristics of structural vulnerability can be summarized in terms of: configuration (large windows with partial height walls below create captive columns or narrow piers, large windows on one side, weak or soft stories, large rooms, buildings one bay wide often with irregular plans), building type (vulnerable forms of vernacular and engineered construction,

S

School Seismic Safety and Risk Mitigation, Table 3 Characteristics found in damage and vulnerability assessments

Characteristics	Cited in 25 % or more		Cited in 15–24 %	
	Damage assessments	Vulnerability assessments	Damage assessments	Vulnerability assessments
Captive columns due to partial height masonry infill walls under windows	✓	✓		
Non-ductile reinforced concrete frame construction	✓	✓		
Generally poor construction quality	✓			✓
Poor-quality engineered materials	✓			
Soft or weak story		✓		
General plan irregularity		✓		
Exterior falling hazards		✓		
Maintenance deferred or lacking		✓		
Inadequate doors, windows, halls/corridors, or stairs		✓		
Vulnerable masonry construction		✓	✓	
Lack of seismic design understanding by engineers			✓	
Interior architectural and contents hazards			✓	✓
Windows reducing solid wall area in masonry construction				✓
Torsion				✓
General vertical irregularities				✓

Rodgers (2012), pp. 4–5

safer traditional construction forms and practices abandoned, standard building plans with seismic deficiencies, heavy roofs), location (sites susceptible to ground failure, sites that amplify ground motions), construction practices (poor quality, unskilled or low-skilled local labor, reducing quality to save money or time), materials (poor-quality engineered materials, weak local materials), lack of construction inspection, lack of maintenance, subsequent modifications, falling hazards, and inadequate exit pathways (Rodgers 2012).

Underlying drivers create an environment conducive to the vulnerability-creating characteristics cited above. Published literature identifies the following: unregulated community-based construction, scarcity of resources, inadequate building codes or zoning, lack of code enforcement, corruption of enforcement mechanisms, unskilled or unaware building professionals, lack of accountability, lack of awareness, failure to prioritize school safety, and urgent need for large numbers of new schools (Rodgers 2012).

Overview of Global Experiences in Seismic-Resistant School Construction

Some of the major policy and programmatic endeavors to assure seismic resilient construction of schools, worldwide, as of 2013, have involved important steps such as providing risk maps for safe school site selection, construction guidelines, standards, and oversight and commitments to safe school construction in the context of both post-disaster reconstruction and new school construction to meet Millennium Development Goals.

The **provision of risk maps** for safe school site selection requires both national and subnational coordination and often several different agencies reporting on the full spectrum of geophysical and hydrometeorological risks and taking into account nuclear, biological, and chemical hazards. In Peru, a pool of trained consultants based in universities around the country are now available to advise Regional Education Offices on safe school site selection. They draw

from existing risk maps for 115 towns (UNISDR 2008).

In the area of **construction guidance and standards**, California's Field Act in 1933 stands as the starting point of the movement. The Act requited 15 % higher performance standards for new school construction and introduced stringent supervision. Legacy school construction was raised as a policy issue as early as 1938 (Garrison Act) but was not enforced until 1968. The oversight system involves structural plans prepared by engineers and approved by the Division of the State Architect. There is recurring on-site inspection and a final verification process.

The more common approach is the development of technical guidance for planning, design, construction, and local ongoing maintenance. There are numerous variations on this theme. For example, in the Philippines, in 2007, the Department of Education adopted the Principal-Led School Building Program approach where principals or school heads take charge of the implementation of management of the repair and/or construction. Assessment, design, and inspection functions are provided by Department of Education engineers who assist the principal during the procurement process. The Parent, Teacher, and Community Association and other community stakeholders are responsible for auditing procurements (INEE 2010). Interestingly, in Panama, it was the development and implementation of the maintenance guidance tool that paved the way for new school construction standards (UNISDR 2012).

There have been several examples of post-disaster commitment to "building back better," emerging from a general consensus following the 2005 Indian Ocean earthquake and tsunami, on the need use humanitarian assistance and reconstruction financing more responsibly. However, in the area of school seismic safety, these good intentions have only translated vaguely to measurable improvements in safe school construction. In Pakistan, 4 years after a devastating earthquake there, the National Education Policy 2009 section 5.5 addressed Education in Emergencies with several policy actions including requirements for school construction according to international standards (UNISDR 2012). Following the devastating 2010 earthquake in Haiti, many donors stated that the schools that they are supporting seismic, hurricane, and flood-resilient school reconstruction, though there is no program that monitors progress in this regard.

In Indonesia in 2009, the Center for Disaster Mitigation, Institute of Technology Bandung (CDM – ITB), and Save the Children International published a handbook of typical school design and a manual on retrofitting of existing vulnerable school buildings for the Aceh and West Sumatra Earthquake Response programs. The guidelines take into account lessons learned in safe school construction, weaknesses in oversight of local government construction, and the need to incorporate design of dual-purpose multi-hazard shelters. In 2014 they were considered ready for an update.

In the Philippines, following devastating typhoons in 2006, 99 disaster-resilient schools and 26 day-care centers were constructed with the support of the Department of Education engineers, school principals, and community members. The new buildings, with water and sanitation facilities, can also serve as evacuation centers with flexibility to accommodate large numbers of people for emergency shelter (Global Education Cluster 2011).

Following the 1999 Kocaeli earthquake, 820 of 1,651 schools that were 60 km away in Istanbul, were found to have sustained some damage. Thirty-five schools were replaced, 59 schools were strengthened, and 59 were repaired (COGSS 2008).

Clear warrants and commitments from donors IGOs or INGOs when it comes to safe school construction are still clearly much needed.

There have also been too few and/or too quiet **commitments to safe school construction** in the context of the Millennium Development Goals, in spite of the fact that the Global Partnership for Education states as its first strategic goal the provision of a quality basic education in a *safe* environment. The most important and notable has been in Uttar Pradesh, India, where 23.5 million children attend school in this moderate to severe seismic risk zone; 21,000 new school buildings

S

(30/day) were to be built in a 2-year period. In 2006–2007 the Elementary Education Department proposed to integrate earthquake-resilient design into all new school buildings. One primary school, two upper primary, and three additional classroom designs were prepared with detailed construction manuals. Disaster-resilient measures added 8 % to the construction costs. To cope with massive scale of the project, a cascading approach prepared 4 master trainers for each of 70 districts. These individuals trained 1,100 Junior Engineers and Education Officers. Ten thousand masons were also trained. In Uttar Pradesh every new school is now a safe school (UNISDR 2008).

Overview of Global Experiences in School Seismic Retrofit

In Sichuan, China, Prior to the 2008 Sichuan earthquake, school principal Ye Zhiping pestered local authorities until they consented to retrofit the buildings of Sangzao Middle School to improve their safety. He also initiated regular evacuation drills. The result of his efforts was that during the devastating earthquake, this school provided life safety for all of its students and staff.

The United Nations Center for Regional Development in Kobe began promoting school earthquake safety initiatives in 1999, in the process of resilience-building following the Hanshin-Awaji earthquake. A multi-country school seismic retrofit initiative (2005–2008) sought to make schools safer through self-help, cooperation, and education. The project engaged local communities, governments, and resource institutions in demonstration vulnerability assessments and school retrofit projects in four to six schools each, in Fiji, India, Indonesia, and Uzbekistan. In 2006, the state of Uttar Pradesh, in India, undertook large-scale disaster-resilient construction of new schools (Bhatia 2008). GeoHazards International also conducted small-scale screening and retrofit demonstration projects in vulnerable schools in Delhi, India (Rodgers 2012), and helped Bhutan's Ministry

of Education develop the process and tools for a nationwide school vulnerability assessment program, which is currently underway.

More than a dozen countries have developed approaches, conducted significant vulnerability assessments, and/or made commitments made to school retrofit since 2000. Several of these were inspired by unacceptable levels of damage experienced in recent large earthquakes. Many are instructive or inspiring, in terms of their scope, methods, and limitations. Looking at these regionally allows an overview of both limited scope and adequacy.

Middle East and North Africa: In Algeria, vulnerability assessment was done on 526 buildings in 190 schools across 9 municipalities in Algiers, using simple survey forms (Rodgers 2012). In Syria, UNDP is supporting earthquake school safety program incorporated into 5-year plan and institutions for disaster risk reduction are being consolidated (UNISDR 2012). The Arab League is currently considering a regional approach to disaster risk reduction, which will hopefully include a comprehensive approach to school safety.

North America: In British Columbia, Canada, Vancouver school buildings were surveyed in 1990 (Rodgers 2012). Responding to advocacy efforts of the local "Families for School Seismic Safety," in 2004, the provincial government committed $1.5 billion Canadian to ensure that BC Schools meet acceptable seismic life safety standards by 2019.

In the USA, there has been detailed assessment of 26 school buildings in Kodiak, Alaska, with recommendation for retrofit of four. In California, a desk assessment of 9.659 pre-1978 school buildings found 7,537 potentially vulnerable buildings. Twenty thousand uncertified projects have been mapped (California Watch 2011). The state of Oregon conducted collapse risk assessment of 2,185 K-12 school buildings using FEMA 154 rapid visual screening (RVS) and produced structural engineering reports for more than 300 buildings. South Carolina has completed a prioritization exercise on all public schools; six have been retrofitted. In Tennessee 49 buildings in 202 schools have been screened using ATC-21

plus local methods, and in Utah, RVS was used on a sample of 128 of 1,085 schools in the state (Rodgers 2012).

Latin America and the Caribbean: The Organization of American States began a commitment to school safety in 1992. A coordinated regional action plan was developed to benefit Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, and Panama. Development assistance donors and local organizations contributed to strategies and capacity to carry out retrofitting of educational facilities. School infrastructure experts from each country received training.

In Bogotá, Colombia, in 1997, seismic microzonation studies paved the way for seismic-resistant building codes in 1998. In 2000 the Directorate of Prevention and Attention of Emergencies in Bogotá found 434 of 710 schools vulnerable to earthquake damage, 3 in flood areas, and 20 in landslide-prone areas. Two hundred and one were prioritized for retrofit or replacement. Between 2004 and 2008, an investment of $460 m USD in school replacement, retrofit, and risk management promotion provided structural reinforcement of 172 schools, "nonstructural" risk reduction in 326 schools, and the construction of 50 new mega-schools, compliant with earthquake-resistance requirements. Three hundred thousand children are safer as a result (see entry "▶ School Seismic Safety: Case Studies" for case study of Colombia). In Ecuador, initial screening of 340 high-occupancy school buildings, and modified RVS of 60 most vulnerable, detailed analysis for 20, and retrofit designs for 15 have taken place (Rodgers 2012). In Lima, Peru, 28 schools in Barranco and 80 schools in Chorrillos were evaluated using ATC 21 RVS and EMS_98 estimation of damage potential (Rodgers 2012). A retrofit solution was developed to mitigate potentially devastating structural defect of "short columns." And in Venezuela, 50-year-old schools were identified as needing retrofitting in moderate and above seismic zones, whereas 20–30 year-old "box" schools only require retrofit in higher-risk zones. Practical retrofitting techniques were developed. As of 2007, 28,000 schools were being surveyed in a national program. Twelve schools were selected for pilot retrofits (Rodgers 2012).

Europe and Central Asia: In Europe, discussion has been robust in Italy and Portugal, innovations have been led by UNICEF and partners in Central Asia, and World Bank financing has supported Turkey to make significant progress in seismic safety (see entry "▶ School Seismic Safety: Case Studies" for case study on Turkey).

In Yerevan, Armenia, full assessments have been conducted by teams of dozens of people, mobilized from as many as seven different government agencies, over several days. Every year 40 of Yerevan's 200 schools are slated for special maintenance, upgrading, and retrofitting. It has been noted that a 2-person expert team spending 2 h per conducting a rapid assessment would require 6 FTE years to assess Armenia's 1,500 schools. In Kyrgyzstan, a national school safety assessment of over 3,000 learning facilities with support from USAID found that more than 80 % were vulnerable to earthquake damage. Public access to this information is made possible through an online portal (UNICEF 2011).

In Uzbekistan 1,000 school buildings were assessed, revealing that 51 % require demolition and replacement, 26 % require capital repair and reinforcement, and 27 % are life safe and require no intervention (Khakimov et al. 2007). Eleven design institutes participated in building code revision for school building construction. Typical designs were created for new schools of different sizes. A database of typical construction and technical decisions seismic reinforcement were developed. UNCRD provided financial and technical support for demonstration projects on reinforced concrete frame, masonry, and frame panel buildings. The incremental cost of seismic reinforcement was shown to be between 3 % and 14 % depending on intensity zone, type of construction, number of floors, capacity, and ground conditions (Khakimov et al. 2007).

In Italy, a substantial contribution comes in the form of an overall risk management framework developed for retrofit prioritization (Grant et al. 2007). Some schools have now been assessed in Emilia-Romagna (Rodgers 2012). Portugal has demonstrated an important

S

innovation by incorporating school vulnerability assessment and retrofit into its ongoing modernization program. At least 330 public school buildings have been assessed and retrofits designed (Rodgers 2012; UNISDR 2012).

The Istanbul Seismic Risk Mitigation and Emergency Preparedness Project (ISMEP) (with loans from World Bank and EIB) allowed for retrofitting of 250 schools and reconstruction of 36 schools in 2007–2008 with 600 more undergoing assessment and feasibility studies. In 2009 the remaining 450 schools were slated for retrofitting.

South Asia: Bhutan has begun a nationwide vulnerability assessment of school buildings. The first phase, covering 5 of Bhutan's 20 districts, began in 2013, with funding from UNICEF.

In India there are several examples of large-scale seismic vulnerability assessments: In Gujarat a modified RVS was conducted for 153 schools following the 2001 earthquake (Rodgers 2012). In Shimla, SEEDS India took a stepwise approach: Step one was low-cost mass scale RVS of school buildings. From these, a smaller number were selected for simplified vulnerability assessment using limited engineering analysis. The highest-risk buildings were identified for detailed vulnerability analysis (SEEDS 2006). Retrofitting designs were drawn up for 20 schools and implementation carried out in ten schools. Guidelines were developed for retrofit and training of local masons and engineers and delivery of skill training. "Nonstructural mitigation plans" were carried out in 20 schools. An awareness campaign reached out to all 750 schools, including nearly 100,000 students, 7,500 teachers and local builders, engineers, and officials (SEEDS 2006). The Government of India's National School Safety Program plans to seismically retrofit more than 40 schools throughout the country as demonstration projects. The National Center for Peoples' Action in Disaster Preparedness (NCPDP), GeoHazards International, and others also carried out school assessment and retrofit programs.

Nepal has also made some strides in both vulnerability assessment and retrofit planning. There are an estimated six million children and 140,000 teachers at risk of death and injury in schools. In the Kathmandu Valley, 643 schools (1,100 buildings) have been inventoried and 378 (695 buildings) surveyed for vulnerability. Seventy-five percent are expected to be damaged beyond repair, in a scenario earthquake. A school day earthquake would kill 29,000 children and teachers and injure 43,000 (Dixit et al. 2013). The MoE has planned to retrofit 900 schools in the Kathmandu Valley over 5 years (Dixit et al. 2012). In Lamjung and Nawalparasi, vulnerability screening has covered 745 and 636 buildings, respectively, some with detailed assessments (Rodgers 2012).

In Pakistan, in 2008, the Aga Khan Planning and Building Services, Habitat Risk Management Program in Northern Pakistan, used retrofitting of four schools to demonstrate structural and nonstructural seismic retrofitting, to train builders and to train female village youth in mapping, land-use planning, and disaster management (INEE 2010).

Southeast Asia: There has been relatively sparse activity when it comes to seismic safety of schools in Southeast Asia. It may be that the frequency of cyclones and flooding and even the threat of tsunami precede thoughts of earthquake risks. It may also be that the rapid pace of development and the increasing numbers of new children being brought into school have led to natural prioritization of safe new construction rather than retrofit. In the Philippines, local authorities are responsible for school construction. However, assessment, design, and inspection functions are provided by Department of Education engineers who assist the principal during the procurement process. The Parent, Teacher, and Community Association and other community stakeholders are responsible for auditing procurements. Earthquake, typhoon, flood, and even volcanic ashfall resilience must often be factored in (INEE 2010).

East Asia: School seismic safety has been on the agenda in Japan for many years, but it is only since 2005 that 125,000 public school buildings nationwide have been assessed by the Ministry of Education (MEXT) (Rodgers 2012). Sixty-two percent of these were constructed before 1981, when the current anti-seismic code enforcement began. About 25 % of schools are considered safe, but 48,000 older school buildings were found needing assessment or retrofitting. 10,000 of these were found to be at high risk of collapse in expected earthquakes. The Ministry of Education raised subsidies for vulnerable school buildings from 50 % to 67 % in 2008 when 229 billion JPY was allocated to meet the new goal of retrofitting of all highest-risk school buildings within 4 years.

Oceania: School seismic safety is also on the agenda in New Zealand, where a walk through survey of 21,000 buildings at 2,361 public schools in 1998 triggered a follow-up investigation in 2000 (Rodgers 2012). A World Bank GFDRR project demonstrated school retrofit in six schools in two districts (2008–2009).

Summary

In the course of the past decade, an approach to all hazards and all aspects of school safety has emerged in both the literature and practice of global advocacy. The Global Alliance for Disaster Risk Reduction and Resilience in the Education Sector (led by UNESCO, UNICEF, UNISDR, IFRC, INEE, Save the Children, Plan International, World Vision) use the shared Comprehensive School Safety framework. The framework takes a multi-hazard approach and addresses the many different factors needed to address safe school facilities, school disaster management, and disaster reduction education. While seismic vulnerability (and related secondary hazards) to school buildings are naturally of concern to earthquake engineers and many others, it is important to fit this into an all-hazard and comprehensive approach so that the solutions to seismic safety do not ignore coexisting vulnerabilities to cyclones, floods, and volcanic eruption nor conflated with the broader approach that also addresses disaster management and education (Global Alliance for DRRR in the Education Sector 2014).

Overall, the threat of earthquake damage to school buildings has not been sufficiently well appreciated. School safety issues have not featured in the major global campaign for increased school attendance ("Education for All" and the Millennium Development Goals). The full extent of the risk to school buildings and to students remains to be fully defined.

A global effort at mapping schools (by density of occupancy and quality of construction) in relation to seismic and other hazards has been proposed by the World Bank Global Facility for Disaster Risk Reduction and Recovery, to begin in 2014. The full impacts of earthquakes on the education sector cannot end with calculating the value of structural and nonstructural damage. The impacts on *children's education* are almost entirely unmeasured. Research is needed to understand how educational outcomes such as enrollment, attendance, and achievement are impacted by earthquakes.

There are strong arguments that support giving school seismic safety increased priority and a higher profile. An initial step in raising the visibility of this issue was the adoption of school safety as one of the focal points for advocacy in preparation for the Hyogo Framework for Action 2005–2015 adopted at the World Conference on Disaster Reduction held in Kobe, Japan, in January 2005. The development of the Comprehensive School Safety framework in 2013 has begun to articulate how school facilities safety can be understood within the wider context that includes school disaster management as well as risk reduction and resilience education. As a post-2015 agenda for both development and disaster risk reduction are currently under consideration, it

S

continues to be extremely important to raise the profile of school safety. In preparation for this, child-centered organizations have formed a Global Alliance for Disaster Risk Reduction and Resilience in the Education Sector.

Based in part on the case studies (see entry "▶ School Seismic Safety: Case Studies"), it seems evident that low-cost, accessible technology and design exists with which to build new schools and to retrofit existing ones. A community-based approach holds great promise involving many stakeholders, including local buildings, masons, contractors, etc. Promising demonstration and large-scale projects in Nepal, India, Turkey (see entry "▶ School Seismic Safety: Case Studies"), Central Asia, and the Caribbean islands all provide strong experience to build upon for case studies.

Case studies also make clear that child rights advocates, parents, and seismic safety experts together, lobbying for school seismic safety, can be extremely effective in achieving policy change, as case studies of British Columbia and Bogotá (see entry "▶ School Seismic Safety: Case Studies") show.

School seismic safety has been the subject of both research and policy since the 1933 Long Beach earthquake that spurred California's landmark *Field Act*, requiring that school construction meet seismic safety standards. As both seismic risk assessment and building codes have progressed, so too have expectations for selection of performance standards. However, globally, the application of these standards and codes falls short in several major respects: community-built schools are frequently constructed using high-tech materials intended for engineered construction, without the corresponding understanding, training, or supervision; where building codes exist they are not known, understood, or consistently applied; and safe site selection is frequently skipped, and site-specific hazards are not factored in. Privately built schools are often exempt from the same standards of construction as public schools. The need for programs and people that bridge the available engineering knowledge with scalable on-the-ground national programs is significant.

In 2009, *Guidance notes on safer school construction* (INEE 2010) was published to synthesize and kick-start systematic guidance. An important global resource for documents to guide safe school construction was initiated by UNESCO IPRED, immediately after the Haiti earthquake (UNESCO IPRED 2010). This resource database endeavors to compile both building codes and the now numerous documents produced by NGOs or at the national level with standard designs for safe school construction, and in some cases with construction guidance.

The past decade has seen several relevant scientific papers suggesting methods for vulnerability screening (e.g., in Italy, Grant et al. 2007), and detailing approaches to seismic retrofit. The challenge is whether or not the guidance and the science are put into practice. The written record does not suggest that these approaches are yet systematic, are supported with training, are monitored, or are applied to both public and private schools. Community-built and un-engineered construction has been addressed in far fewer publications and has not specifically addressed school construction.

There have been a small number of significant programmatic efforts to support seismic safety. UNICEF's regional office for Central and Eastern Europe and the Commonwealth of Independent States, with support from the World Bank and DIPECHO, has partnered with national governments in Central Asia and the South Caucasus to address school safety. Part of that work has included developing a broad regional framework for assessing and ranking school facilities based upon exposure and vulnerability to earthquakes and other natural hazards. Drawing upon INEE's *Guidance notes on safer school construction*, UNICEF elaborated a list of 17 simple indicators that local experts could use as part of a rapid visual assessment of school facilities in order to identify schools at risk of heavy damage in seismic events.

In 2012, engineers in Kyrgyzstan localized this framework and carried out a national school safety assessment of over 3,000 learning facilities with USAID funding. They reported to the national government that over 80 % of learning

facilities were vulnerable to damage in seismic events and provided public access to the assessment through an online portal. Similar national assessment strategies are being piloted in Kazakhstan, Tajikistan, Armenia, and Azerbaijan.

Similarly, UNCRD (UN Centre for Regional Development) showcased community-based comprehensive school earthquake safety in selected countries of Asia Pacific. Under the program "Reducing Vulnerability of School Children to Earthquakes," school communities carried out seismic retrofitting of their school buildings with expert guidance from Bandung Institute of Technology (ITB) in Indonesia. The retrofitting works in public schools were used for community awareness on earthquake safety through community visits in the school premises during construction time. Pilot school assessment and retrofitting in Fiji led to the National Disaster Management Office (NDMO) adopting school safety program under regular government that also developed seismic retrofit guidelines and mason's training manual. Tashkent city government (Hokimiyat) in Uzbekistan appraised neighborhood associations on schools retrofitting programs and used school constructions for training of engineers on seismic safety.

The United Nations International Strategy for Disaster Reduction (UNISDR) launched the 2006–2007 biennial awareness campaign "Disaster Reduction Begins in Schools." This was followed up in 2010 with the *Resilient Cities Global Campaign* for *One Million Safe Schools and Hospitals Campaign*. The 10-point checklist that 1,643 Mayors have signed on for, includes assessing and upgrading the safety of schools. These successes deserve praise but should not induce complacency. There is a long way to go with respect to school seismic safety.

Initial programs and guidance for safe school facilities have been provided by OECD (2004), UNCRD (2008), INEE/World Bank GFDRR/UNISDR (2010), and several other programs, with modest support of donors and lenders. These approaches experiences are now ripe for implementation at scale. These include regional hazard mapping and revision (where necessary), the potential for crowd-sourced mapping of local hazards; enforcement of seismic building codes by national, provincial, and local governments; training of engineers and significant capacity-building efforts to train local masons and other builders; and invention of more innovative models for funding reinforcement of schools.

It is important, however, not to fetishize the safety of school *buildings* and to take care not to separate the safety of the community of users and educational continuity planning, which is not limited to the buildings themselves. Neither should the focus be solely on fatality prevention. There is much similar work to be done to prevent disability and injury especially by securing the contents of the buildings and to assure educational continuity. All-school, participatory school disaster management planning, local risk assessment and risk reduction, mastery of emergency response skills, and regular drills to practice and improve readiness are important. A culture of safety is necessarily multifaceted, and activism in one area encourages changes in consciousness, expectations, and demands.

The enthusiasm for making education accessible to all does not absolve duty-bearers from assuring that school is safe from infrequent but high-impact hazards such as earthquake and various secondary hazards. It would be an ironic and tragic result if the achievement of one Millennium Goal (increased school attendance) is marred by increased death and injury of young people, thus setting back the achievement of another Millennium Development Goal (reduction of child mortality).

Cross-References

▶ "Build Back Better" Principles for Reconstruction
▶ Building Codes and Standards
▶ Earthquake Protection of Essential Facilities
▶ Earthquake Risk Mitigation of Lifelines and Critical Facilities
▶ School Seismic Safety: Case Studies

References

Bhatia S (2008) Earthquake-resistant school buildings in India. Urb Plan Des 161(4):147–149

California Watch (2011) K-12/public safety/spotlight/on shaky ground. 22 Dec 2011. Accessed 3 Feb 2012

CNN (2014) Haiti earthquake fast facts. http://edition. cnn.com/2013/12/12/world/haiti-earthquake-fast-facts/index.html. Accessed 6 Jun 2014

Coalition for Global School Safety (2008) School disaster reduction: think globally, act locally. 'Mother' slide presentation

Dixit A et al (2012) Institutionalization of school earthquake safety program in Nepal. In: Proceedings of the 15th WCEE. http://www.iitk.ac.in/nicee/wcee/article/WCEE2012_5260.pdf. Accessed 16 Feb 2014

Dixit et. al. (2013) Public School Earthquake Safety Program in Nepal, Geomatics, Natural Hazards and Risk, DOI: 10.1080/19475705.2013.806363

Dowrick D, Rhoades DA (2004) Risk of casualties in New Zealand earthquakes. Bull N Z Soc Earthq Eng http://www.nzsee.org.nz/db/Bulletin/Archive/38(2)0053.pdf. Accessed 5 Mar 2015

EERI (1990) Earthquake Engineering Research Institute, Oakland. Loma Prieta Earthquake 17 Oct 1989. Preliminary reconnaissance report. https://www.eeri.org/wp-content/uploads/USA_CA_LomaPrieta_PrelimReport.pdf. Accessed 16 Feb 2014

FEMA (2011) Reducing the risks of non-structural safety – a practical guide, 4th edn. Washington, DC. http://www.fema.gov/media-library/assets/documents/21405?id=4626. Accessed 19 Feb 2014

Global Alliance for Disaster Risk Reduction and Resilience in the Education Sector (2014) Comprehensive school safety. http://preventionweb.net/go/31059. Accessed 15 Jan 2014

Global Education Cluster (2011) Disaster risk reduction in education in emergencies: a guidance note for education clusters and sector coordination groups. http://preventionweb.net/go/20366. Accessed 16 Feb 2014

Global Partnership for Education (2014) Results and data. http://www.globalpartnership.org/results/. Accessed 16 Feb 2014

Grant DN, Bommer JJ, Pinho R, Calvi GM, Goretti A, Meroni F (2007) A prioritization scheme for seismic intervention in school buildings in Italy. Earthq Spectra 23(2):291–314

INEE/World Bank GFDRR/UNISDR (2010) Guidance notes on safer school construction. INEE (Inter-agency network for education in emergencies), New York. http://www.preventionweb.net/go/10478

Khakimov S, Nurtaev B, Tursunov K (2007) School earthquake safety program in Uzbekistan, slide presentation. Tashkent

New York Times (2008) School-age children in earthquake zones, 27 May 2008. http://www.nytimes.com/imagepages/2008/05/27/science/27school.graphic1.ready.html. Accessed 16 Feb 2014

OECD (2004) Keeping schools safe in earthquakes. Organisation for Economic Co-operation and Development, Paris. http://www.oecd.org/document/36/0,2340,en_2649_201185_33630308_1_1_1_1,00.html. Accessed 19 Feb 2014

Pandey B (2013) Political economy study on safe schools, World Bank, Global Facility for Disaster Reduction and Recovery

Rodgers JE (2012) Why schools are vulnerable to earthquakes. In: Proceedings, 15th world conference on earthquake engineering, Lisbon, 24–28 Sept 2012

SEEDS (2006) Earthquake Safety in Mountain Regions. School Earthquake Safety Initiative. Shimla, Delhi

State of California, Seismic Safety Commission (2009) The field act and its relative effectiveness in reducing earthquake damage in California's public schools. http://www.seismic.ca.gov/pub/CSSC%2009-02%20Field%20Act%20Report.pdf. Accessed 16 Feb 2014

UNCRD (2008) Reducing school vulnerability to earthquakes. United Nations Center for Regional Development, Hyogo, Japan. http://www.hyogo.uncrd.or.jp/publication/pdf/Report/2008SESI/UNCRD%20SESI%20publication%202008.pdf. Accessed 16 Feb 2014

UNESCO (2004) 'Education for all week 19–25 April 2004: the extent of the problem', information from UNESCO's (United Nations Educational, Scientific and Cultural Organization) http://portal.unesco.org/education/en/ev.php-URL_ID=28702&URL_DO=DO_TOPIC&URL_SECTION=201.html. Accessed 25 Jan 2014

UNESCO IPRED (2010) Resource/database, disaster preparedness and mitigation. http://www.unesco.org/new/en/natural-sciences/special-themes/disaster-preparedness-and-mitigation/natural-hazards/earthquakes/ipred/unesco-ipred-database/. Accessed 25 Jan 2014

UNICEF (2011) Toward safer schools – methodology for nationwide benchmarking of school safety http://www.unicef.org/ceecis/towards_safer_school_FINAL.pdf. Accessed 5 Mar 2015

UNISDR (2008) Disaster prevention for schools – guidance for education sector decision-makers. Geneva. http://www.preventionweb.net/go/7344. Accessed 25 Jan 2014

UNISDR (2012) Assessing school safety from disasters, a global baseline report. http://preventionweb.net/go/35274. Accessed 15 Jan 2014

USGS (2003) 'Magnitude 8.3 – Hokkaido, Japan region 25 Sep 2003 19:50:06 UTC'. http://earthquake.usgs.gov/recenteqsww/Quakes/uszdap.htm. Accessed 25 Jan 2014

USGS (2015) Historic World Earthquakes. http://earthquake.usgs.gov/earthquakes/world/historical.php. Accessed 5 Mar 2015

WEF (2000) UNESCO World Education Forum. http://www.unesco.org/education/efa/wef_2000/. Accessed 5 Mar 2015

School Seismic Safety: Case Studies

David Alexander[1], Sanjaya Bhatia[2], Djillali Benouar[3], Jitendra Kumar Bothara[4,5], Omar-Dario Cardona[6], Amod Mani Dixit[4], Rebekah Green[7], Manu Gupta[8], Ram Chandra Kandel[9], Ilan Kelman[10,17], Tracy Monk[11], Bishnu Pandey[12], Marla Petal[13], Zeynep Türkmen Sanduvaç[14], Rajib Shaw[15] and Ben Wisner[1,16]

[1]University College London, London, England
[2]UNISDR Recovery Platform, Kobe, Japan
[3]University of Bab Ezzouar, Algiers, Algeria
[4]NSET, Kathmandu, Nepal
[5]Miyamoto Impact, Christchurch, New Zealand
[6]Universidad Nacional de Colombia, Manizales, Colombia
[7]University of Western Washington, Bellingham, WA, USA
[8]SEEDS of India, Delhi, India
[9]Toronto, ON, Canada
[10]Institute for Risk and Disaster Reduction and Institute for Global Health, University College London, London, England
[11]Families for School Seismic Safety, Vancouver, Canada
[12]University of British Columbia, Vancouver, Canada
[13]Risk RED (Risk Reduction Education for Disasters), Los Angeles, CA, USA
[14]Risk RED, Istanbul, Turkey
[15]University of Kyoto, Kyoto, Japan
[16]Oberlin College, Oberlin, OH, USA
[17]Norwegian Institute of International Affairs, Oslo, Norway

Synonyms

Education sector; Examples; Policies; Policy; Progress; Schools; Seismic safety

Introduction

The case studies in this entry accompany the previous entry on School Seismic Safety and Mitigation. They tell the stories of a variety of national efforts to improve school seismic safety. The contributors to the case studies are engineers and parents, social workers, and international development specialists. They examine policy, advocacy, vulnerability, and solutions. They contain observations about stepwise progress, motivation, political will, technical approaches, innovations, moderate successes, and long roads ahead. There are more stories to be added. The intention is to provide school seismic safety advocates with both elements of inspiration and way points on a road map with many options to consider.

The case studies and their contributors are:

Algeria – Djilali Benouar
Canada – Tracy Monk
China – Sanjaya Bhatia
Colombia – Omar Dario Cardona
India (Delhi, Shimla) – Manu Gupta
India (Uttar Pradesh) – Sanjaya Bhatia
Italy – David Alexander
Japan – Rajib Shaw
Nepal – Amod Mani Dixit, Jitendra Kumar Botendra, Ram Chandra Kandel, and Bishnu Pandey
Turkey – Marla Petal and Zeynep Türkmen Sanduvaç
USA (California) – Marla Petal and Rebekah Green

It is important to note that school *seismic* safety should not be isolated from school safety from disasters and climate change impacts in general. It cannot be considered responsible to approach the rights of children from our narrow professional silos. It is incumbent upon all advocates for children to see the matter from their point of view, that is, from the perspective of all of the physical threats to their right to survival and safety and from all of the threats to their right to education and development.

Similarly, key stakeholders in the seismic safety of schools (engineers, architects, builders) must make the effort to think not primarily about the *structures*, but rather about the people who

use the structures. The users must also have safe access and egress. For the users, infill walls that fall out of plane and unsecured nonstructural building elements or building contents are a far greater threat than their "minor damage" designation suggests. If the building is going to be one of the strongest structures in the community, then it may also need to be planned to be as a cyclone shelter or to double as a shelter for people displaced after an earthquake. In that case, water and supply storage and extra sanitation facilities have to be considered. And the job is not complete when the key is handed over. Buildings that do not come with a user's manual and a school maintenance calendar stand little chance of being safely maintained. Thus, the concern for school seismic safety does not begin and end with the structures themselves, but must take a holistic approach and, of course, include the user community.

Algeria

Ninety percent of Algeria's population of 30 million is concentrated in a band about 60 km wide and 1,200 km long along the Mediterranean coast on the African and Eurasian tectonic plate boundary. This region has repeatedly experienced moderate-to-strong earthquakes. During the twentieth century, earthquakes claimed at least 10,000 lives, injured about 27,000, and made about half a million homeless. In addition to building collapse due to earthquakes, considerable damage from liquefaction and landslides was observed.

School buildings have also suffered considerable damage in earthquakes, varying according to the period during which they were built: (1) those degraded through aging and lack of maintenance, built during the colonization era (1830–1962), account for about 30 % of the school building stock; (2) those built after independence, during the 1970s, with rapidly growing population and democratization of educational opportunity (when primary school became free and compulsory), when school construction accounted for the largest single item in Algeria's budget, were designed and built without taking into account

seismic risk; and (3) those built with technical supervision after 1983 and the introduction of Algeria's seismic building code in 1981.

Schools in Algeria are all state owned and were built by the government. The government adopted one typical structure for all schools that could be duplicated easily across the country. The standard architectural design of schools involves two fundamental elements: the classroom and the circulatory corridors. Classrooms are 7×4 m and circulatory corridors are 2.5 m wide. These are far from those of an ideal seismically resistant structure as recommended by Algeria's own seismic codes. Other standard design elements also unfortunately reduce the resilience of these school structures.

Numerous reports show the deficiencies in design, construction techniques, and materials (poor quality of concrete) with respect to particular earthquakes at El Asnam (1980), Chenoua-Tipaza (1989), Beni Chougrane-Mascara (1994), Ain Temouchent (1999), and Boumerdes-Algiers (2003) (Bendimerad 2004). The following typical damage to school buildings was recorded in recent earthquakes:

- Rupture of staircases
- Destruction of joints
- Destruction of short columns
- Damage in Masonry
- "Pancake" collapse due to weak columns, overly strong beams, and heavy roofs composed of reinforced concrete slabs

Such damage causes enormous financial loss to the government. For instance, after the Boumerdes-Algiers earthquake disaster of 2003, 100 primary schools had to be rebuilt completely for the sum of US$4.28 million and 253 rehabilitated for $10.65 million. In addition, 12 junior high schools were completely rebuilt for the sum of $10.28 million and 111 rehabilitated for $20.85. Also, 10 high schools were rebuilt for the sum of $21.42 million, and 58 were rehabilitated for $12 million.

So far these destructive earthquakes have occurred after school hours or on weekends, and thus, no loss of life or injuries have been recorded

at schools in Algeria. This good luck may have made the government and the civil society alike less aware of the high vulnerability of the schools and partially explains why there has so far been no implementation of a Ministerial instruction dating from 1989 that required application of "technical expertise and the eventual reinforcement of all public buildings and in particular schools and universities." Instead, the introduction of new materials such as reinforced concrete in the absence of proper seismic-resistant design, building codes, and enforceable regulations has increased the risk to structures and their occupants. Relatively minor reinforcements could reduce the potential for damage to these structures.

Canada

British Columbia's (BC) west coast is Canada's region of highest seismic hazard. Two-thirds of the province's 3.9 million people live within the zone of highest risk. The region has experienced ten moderate-to-large earthquakes since 1870. In recent millennia, an earthquake on the order of the largest magnitude experienced globally in the last 100 years has struck approximately every 500 years. In addition to potential building collapse induced by ground shaking, significant damage from liquefaction, tsunami, and landslides are expected.

Older BC schools were built from some of the most seismically vulnerable materials – in the early 1900s, unreinforced masonry and then, in the mid-1900s, non-ductile concrete frame. Thus, in many communities, the school might be one of the buildings at highest risk for earthquake damage. A formal risk assessment of all BC school buildings was initiated in June 2004 with the full report due in October 2004. Initial estimates from the provincial government suggest that 800 of BC's schools might need some form of seismic upgrading.

In Vancouver, BC's largest city with a core population of about 560,000 and a metropolitan population of about two million, a 1989 rapid seismic risk assessment found that 30 % of the city's school buildings were at high risk of experiencing structural damage in an earthquake, and 15 % were at moderate risk (Taylor 1989). Between 1990 and 2004, 11 schools have been upgraded, so if the current pace continues, this work would be completed in 2064.

While the seismic hazard facing schools in greater Vancouver is similar to that in Seattle, Washington, school safety has not been a priority on the Canadian side of the border. Indeed, differences in seismic hazard mapping techniques used by Canadian and US geoscientists may actually underestimate the risk to Vancouver schools.

The current National Building Code of Canada ranks buildings according to their priority as critical infrastructure. The higher the number assigned, the higher the priority. Average houses are assigned an Importance Factor of 1.0, schools are designated 1.3, and hospitals, police stations, and prisons are assigned 1.5. Schools – unless they are designated as post-disaster shelters – get a lower priority than hospitals, police stations, and prisons. Vancouver City Council is funding the seismic upgrading of community centers so that they could be used as post-disaster receiving centers. Due to the differences in funding sources, some community centers are being upgraded, while nearby schools, which might be less seismically resistant, are sometimes not.

In general in BC there is high awareness of earthquake risk. For example, the City of Vancouver has seismically upgraded its water supply system and bridges, and the electric utility, BC Hydro, is systematically upgrading its buildings and infrastructure, including dams. Even some provincially run liquor outlets have been seismically upgraded. With seismic safety so clearly on the policy agenda in many sectors in British Columbia, why have public schools received so little attention?

The answer is that in BC funding for public school seismic upgrades has been part of the capital budget of the provincial Ministry of Education. Local school boards oversee this work and must proactively request provincial funding for projects that they deem to be high priority. Advocacy work by one of the authors on behalf of Families for School Seismic Safety British

Columbia (FSSS) identified and tackled concerns with this funding system.

First, there was no district-to-district standardization of approach. Each local school district was individually approaching the issue of seismic safety for only their schools. FSSS pressed the government to unify the approach taken by centralizing expertise. The earthquake engineering community, through its professional association, the Association of Professional Engineers and Geoscientists of BC, is now advising the government on standardized, peer-reviewed methods for assessing and addressing seismic risk to schools.

Second, local authorities were concerned that informing parents of the seismic risks to local schools could cause panic and could be politically damaging. FSSS's and others' work has ensured that Vancouver parents are now well informed about the issue and are actively involved in trying to solve it. This process did not cause panic. Instead, bringing parents into the consortium has yielded an active and effective lobbying group.

Finally, there is the problem of setting priorities. The primary concern of school boards – quite rightly – is the day-to-day education of children. Seismic safety of a school building does not lead to improved education, so school boards have sometimes had difficulty making the issue a high priority. FSSS is trying to help public officials see school seismic safety as an infrastructure, public health, and human rights issue and to obtain new funding from outside of the Ministry of Education, that is, from provincial and national, authorize with mandates in those areas. The aim is that this work be seen as an infrastructure project for children and not perceived as competing for funding with their day-to-day educational needs.

Ultimately, the two basic human rights of children, to an education and to physical safety, should not be competing for the same funds. The expert community is now driving the initiative and the government appears to be listening. Many positive steps have been taken in BC, but there is much work left to be done.

China

Following the 2008 Sichuan earthquake in which at least 15,000 children lost their lives in schools, the Ministry of Education and the Ministry of Construction and National Development and Reform Commission (NDRC) jointly released the Design Instructions for School Planning and Construction after the Sichuan earthquake. These standards require that school sites are assessed before the schools are built in accordance with national regulation, performance objectives are determined by the country-level government, schools are built or retrofitted to meet performance objectives, and schools' furnishings and equipment are designed and installed to minimize potential harm they might cause to school occupants. The quality-monitoring bureau leads monitoring on the safety of equipment installation.

In Sichuan in the spirit of "building back better," the investigation, design, construction, supervision, inspection, and acceptance of school construction are conducted in line with relevant national construction standards. In Sichuan, steps were also taken to make sure that there are mechanisms to ensure that schools' maintenance is financed and executed. From 2000 to 2005 the first and second session of school renovation and maintenance was conducted. After 2006 a long-term mechanism for school building maintenance was put into place. MOE and UNICEF collaborated and prepared and revised the National Guidelines for Safe School Construction and Management and also collaborated to develop construction standards for kindergartens and preschools.

In 2009 the Ministry of Education (MoE) initiated a 3-year national "School Construction Safety Programme" to upgrade the safety of primary and middle school buildings all over the country with the aim of making schools the safest places in China. The program has several key elements:

- To screen and assess the quality of all school buildings across the country, to understand the qualities of buildings resistant to local disaster

risks, and to input the data and information of the assessment into a database

- To understand disaster risks in the regions where the schools are located, such as determining whether local disaster risks come from floods, landslides, earthquakes, or rainstorms threatening the safety of the school buildings
- To determine whether to repair, strengthen, or reconstruct school buildings that have not reached official standards based on the intensity of the identified disaster risks in the region where the schools are located
- To allocate funds and to start the construction work to upgrade the primary and middle school buildings that are at risk (International Recovery Platform 2010)

Colombia

The capital city of Colombia, Bogotá, is the most important political, administrative, economic, and cultural center of the country and has one million school children. Bogotá's population was estimated to be around 7.6 million in 2013. As a result of social investment over the past decade, more than 12 % live below the poverty line.

Among the most common hazard events affecting Bogotá are earthquakes and landslides. Although there has not been a severe earthquake in Bogotá since 1917, there is certainly the potential for one. Also, elsewhere in Colombia, 74 % of the schools in the cities of Pereira and Armenia suffered damage in the 1999 earthquake. Fortunately this occurred during the lunch hour, when no children were in the school buildings.

Several risk identification methods were put in place in the city prior to 2004. These include compilation of records of hazard events, generation of hazard maps, studies of physical and social vulnerability, and studies of environmental degradation. One of the means of reducing risk from earthquakes and landslides in Bogotá is the assessment of seismic risk of bridges, hospitals, and schools. This has become a core part of the city's economic and social development plan. Of these assessment programs, the best-known

is the Department of Education's effort to identify school seismic risk and to reinforce schools.

Much of the educational infrastructure in Bogota is more than 50 years old and does not meet minimal standards of safety. The Department of Education commissioned a systematic review of schools that ran from 1997 to 2003 (Secretaría de Educación del Distrito Capital de Santafé de Bogota 2000). This study covered approximately 2,800 buildings at 706 schools serving roughly 54 % of the student population in Bogotá. The other 46 % of the student population attends private schools and was not covered in this review. The review found that 434 of the schools presented high risk to students. Some 772 buildings at these schools fell into this category (16 %). The study also found that 60 schools had buildings in immediate and urgent need of reinforcement.

During the next city administration, from 2004 to 2008, after a detailed technical explanation on the need of schools' retrofitting, the city mayor decided to implement a retrofitting program of the 200 most vulnerable schools. A special risk analysis was performed on each building that was identified and prioritized. From this analysis, the structural reinforcement requirements were defined according to the seismic building code updated in 1998, with new special provisions for schools. The comprehensive improvement program had to be adjusted. Taking into account other technical, urban, economic, and environmental issues, many schools were not retrofitted: 67 schools were demolished and full restitution was made for a total of 107 new schools. This additional program was called the "50 Macroschools plan" whose goal was to provide an educational infrastructure of maximum specifications and supplement the retrofitting and integrated improvement program. At the end, due to the costs involved in reducing vulnerability for 434 vulnerable schools, 201 schools were considered in critical condition according to risk studies that were conducted to prioritize and rank the schools. The reinforcement of these buildings had a cost about US$ 200 million and the total program including the new schools was about

US$ 430 million. Additionally, this program was based on the implementation of a teaching strategy to incorporate risk management into the culture. Both structural and nonstructural objectives were implemented to obtain a comfortable and safe school environment and a high-quality education service.

Assessing and reducing the risk to schools in Bogotá took place in a more general planning and management context. For example, zones of high risk of landslide, where no mitigation works are possible, are declared to be protected land. Human occupation is restricted in these areas as well as those considered at high risk to floods. In 2003 it was estimated that some 185,000 people lived in informal settlements in a total of 34,230 informal housing units. In Bogotá there are 173 illegal settlements that account for 14 % of the total land area. The city administration has developed a massive legalization program since 1995, thus reducing the number of informal settlements from 1,451 to its current number, an eightfold reduction in less than 10 years.

Nevertheless, as much as 60 % of the population of the city lives in informally constructed dwellings. While most of these are located in legal settlements, they still represent a challenge to seismic safety. The year 2000 land-use master plan for Bogotá contains hazard and risk maps that determined land use, details of special treatment for high-risk areas, and arrangements for issuance of building permits, as well as protection plans for utilities and services. The city also relies on community-based networks to control illegal land occupation and has developed a large-scale relocation program for families living in high-risk conditions.

The city of Bogotá has disaster risk reduction at the center of its planning process, and in this context school seismic safety ranks very high. Having diagnosed the scale and urgency of the problem in Bogotá, steps were being taken to reinforce the most hazardous school buildings. The challenges the city still faces include completing the program, making the retrofitting of the second priority group of vulnerable schools, extending its school safety program to private schools that cover nearly half the school-aged population, and accelerating the rate of school reinforcement to improve the safety of children and teachers. Bogota has had different governance problems during the last two administration periods. Unfortunately, two mayors have been removed or suspended by the national general attorney due to corruption and inefficiency. Due to these situations and perhaps changing priorities in risk management, the school safety retrofitting program was not continued.

India: Delhi

NGO partners SEEDS and GeoHazards International (GHI), working with the Government of Delhi, demonstrated earthquake nonstructural risk reduction in a public school. The school welfare committee comprised of faculty, staff, and local community members learned to identify the nonstructural building elements and building contents that could fall, slide, or collide during a likely Delhi earthquake, as well as fire and evacuation hazards. They were exposed to simple low-cost techniques for reducing these risks (moving some items, fastening others) and came up with innovative solutions of their own. The logic of regular fire and earthquake drills became readily apparent to these new stakeholders. A handbook for schools on nonstructural risk reduction developed by the NGO partners and published by the Government of Delhi provides a new resource for generalizing these lessons.

India: Shimla

A small-scale demonstration project for school retrofit was carried out by SEEDS of India and GeoHazards International. Structural assessment of school buildings was carried out using a filtering method: The first step was low-cost mass scale Rapid Visual Assessment Survey of school buildings for potential seismic hazards. Based on these surveys a smaller number were selected for Simplified Vulnerability Assessment using limited engineering analysis. The highest-risk

buildings were identified for detailed vulnerability analysis. Retrofitting designs were drawn up for 20 schools and implementation of retrofit carried out in 8 schools. Guidelines were developed for retrofit and training of local masons and engineers and delivery of skill training. "Nonstructural mitigation plans" were carried out in 20 schools. An awareness campaign was designed to reach all 750 schools in the region including nearly 100,000 students, 7,500 teachers, and local builders, engineers, and officials (SEEDS 2006).

India: Uttar Pradesh

There are 23.5 million children attending school in this moderate-to-severe seismic risk zone. As part of the Education for All campaign, the State Government of Uttar Pradesh constructed 82,000 additional elementary school classrooms and 7,000 buildings in 2006–2007. To ensure seismic resilience of the buildings, UNDP provided earthquake engineers who examined the blueprints for the schools and modified the design to integrate seismic resilience. The marginal cost increase of 8 % to assure seismic safety was funded by the government.

For effective implementation of the new modified designs, training and orientation programs were initiated by the government, supported by the local UNDP office, building the capacity of 40 architects, over 200 engineers, and over 10,000 masons. To ensure transparent monitoring, the designs were widely circulated to the local communities where the schools were constructed, so they could monitor the quality, along with support of departmental engineers. To complete the safety of the schools, school-level safety committees were established, school emergency plans developed, and mock drills became a part of the school program.

In 2006–2007 the Elementary Education Department proposed to integrate earthquake-resilient design into all new school buildings. To prepare for this, one design of primary school buildings and two upper primary and three additional classroom designs were prepared with detailed construction manuals. The disaster-resilient measures added 8 % to the construction costs. To cope with the massive scale of the project, a cascading approach prepared 4 master trainers for each of the 70 districts. These individuals in turn conducted trainings for 1,100 fellow Junior Engineers and Education Officers. Ten thousand masons were also trained. This program ensures that *every new school will be a safe school*. The problem of preexisting stock of 125,000 unsafe school buildings in need of retrofit remains to be tackled (Bhatia 2008).

Italy

Of the 8,102 municipalities in Italy, all are regarded as "seismically active." Up until 2003, 2,965 of them (representing 40 % of the land surface and 45 % of the population) were considered "highly seismic" category. As a result, new construction in this category must observe stringent anti-seismic building codes. Subsequently, a more sophisticated classification was introduced, based on a 50-year recurrence interval and local estimates of peak ground acceleration (PGA). This had the effect of increasing the areas classified as highly seismic. Whereas in previous classifications, some municipalities were effectively regarded as aseismic, that is no longer the case and all 8,102 are now considered to be in a seismic zone to a greater or lesser extent. Italy bases its seismic classification on historical records and calculated return periods. Where these are a poor reflection of seismic hazard, it can underestimate the earthquake threat. Hence, severe damage occurred in northern Emilia and southern Lombardy in the earthquakes of May 2012, an area that had not had a major seismic disaster since 1574.

Several hundred municipalities are faced with the highest seismic risk in the Strait of Messina (including eastern Sicily) and southern Calabria (the Aspromonte) areas. The Apennine Mountains, which form the "backbone" of the Italian peninsula, are, in a tectonic sense, divided into blocks, which means that seismicity varies significantly from one locality to another. However, the vulnerability of buildings, including schools,

S

is almost universally high: modest local taxation revenues inhibiting retrofitting and maintenance. The largest seismic event of the past century remains that in Avezzano, in 1915.

An event that for Italians most encapsulates the seismic risk to schools was that of 31 October 2002 at 11:40 a.m. where in the Apennine town of San Giuliano di Puglia (population 1,195), the infants' school collapsed onto a class of small children (Augenti et al. 2004). Twenty-six small children and three teachers were crushed to death, and 35 children were rescued and lived. The building had been constructed with regional development funds in the early 1960s and had had its roof renewed a year before the earthquake. Evidently, the roof, of reinforced concrete with a ring beam, was too rigid and too heavy for the underlying structure, a concrete frame building with hollow-brick infill. There were signs that the quality of the cement was poor and the reinforcing steel was not used as it should have been. Moreover, despite mounting evidence that the Molisan Apennines are significantly affected by periodic earthquake activity, the revision of local building codes to take account of the new data on seismic risk cannot be applied as easily to existing structures as it can to ones that are about to be built (Augenti et al. 2004, p. S258).

The school at San Giuliano di Puglia succumbed because it contravened simple, well-known laws of dynamic response in structures affected by seismic acceleration. Inertial forces applied to a heavy roof sitting upon a weak frame structure amount to a recipe for tragedy.

Consider the schools of the Lunigiana, a surprisingly remote mountainous area of northwestern Tuscany. The Lunigiana has a sparse and dispersed population. Children attend elementary schools in the villages and secondary schools in major population centers, which they reach by bus or car. Many of the school buildings were constructed in the 1950s and 1960s to cater for the postwar population boom, and, in a rural area of relative economic stagnation, they have neither been built to be fully anti-seismic nor retrofitted. Indeed, in the minor population centers, they are decidedly dilapidated, nor do their staffs seem to have much interest in repeatedly practicing evacuation drills. Yet the area awaits a magnitude 6 earthquake, which it is predicted may kill up to 120 people and injure more than a thousand. How many of them will be school children? This situation is typical of the seismic risk that affects highland Italy.

Elsewhere in Italy, much more progress has been made. The civil protection departments of several regional governments have introduced comprehensive *Scuola sicura* (safe schools) programs, notably in the northern regions of Lombardy, Piedmont, and Emilia-Romagna and in the autonomous Region of Sicily. The programs involve a combination of structural measures and school disaster management efforts, such as evacuation drills and lessons in civil protection. In many of the major cities, fire brigades and volunteer civil protection services are heavily involved in the programs, with public-private partnerships supporting attractive safety literature for school children.

Despite these developments, as in other seismic countries so in Italy, the building stock of schools continues to age and the civil protection educators must fight against the indifference of teachers, principals, and administrators. In many respects, mass mortality in Italian schools during recent seismic events has been avoided mostly by the lack of major earthquakes during school hours, a situation that will not prevail forever. Given an overwhelming need to upgrade the seismic performance of schools in Italy, the response of the national government has been to rank the buildings in terms of the deficit between design requirements (a function of the rules that prevailed at the time they were built) and the latest assessment of peak ground acceleration (PGA). Priority funding is given to those schools that have the greatest "PGA deficit" (Grant et al. 2007). However, the problem of unsafe school stock is simply too expensive to solve in the short to medium term.

Japan

In general, Japan is understood to be a leader in evaluating seismic risk and in implementing building codes for seismic-resilient construction.

Two publications available in English are MEXT's school seismic retrofit handbook (MEXT 2008a) and school nonstructural reference book (MEXT 2008b). In addition to high seismic performance standards for schools, following the 1995 Kobe earthquake, Japan also began providing guidance for mitigation of hazards due to building nonstructural elements and contents in schools. Nonetheless, the East Japan earthquake and tsunami [of 11 March 2011, with magnitude 9.0 earthquake off coast of East Japan], destroyed 6,284 in the affected region with different damage levels. Most of these were affected by tsunami waves, rather than earthquake. This was due to the location of the school buildings [proximity to coastal areas], the layout and structure of the buildings, and the subsidence of local tsunami retention walls. At the immediate aftermath, some schools [with higher stories] were used as temporary evacuation sites, and later people were rescued by helicopters. In some schools, located in higher ground, people took shelter, which lasted till 6 months in some cases, which caused serious disruption of school education.

The disaster pointed out several dimensions of role of schools and disaster education: (1) Schools can play an important public infrastructure of the community; however, the structural safety of the building needs to be linked to operational [including supplies of emergency kit] and locational issues. (2) School-community linkage is an important element, and local communities played an important role along with school teachers for the management of the evacuation sites. (3) While it is unavoidable to use the school as shelter, the continuity of education in emergency is a crucial issue. (4) School-based community recovery emerged as an effective concept, where the reconstruction of school building was linked to enhanced community cohesion (Shaw and Takeuchi 2012).

Nepal

A risk estimate for Kathmandu Valley, the economic, political, and technological hub of Nepal, based on a scenario earthquake similar to the 1934 Bihar-Nepal earthquake, suggests that more than six million children and 140,000 teachers are at risk in schools (Bothara et al. 2002). A survey of 900 public schools in greater Kathmandu Valley estimated that more than 75 % of school buildings would suffer severe damage beyond repair (estimated at US$7 million), and other 25 % would suffer repairable damage. In the absence of intervention, an estimated 29,000 children could be killed in their schools. With intervention 24,000 of these could be saved and buildings protected (NSET 2000).

A more recent assessment of school buildings in other parts of the country shows that more than 9,000 school buildings, more than 10 % of the total in Nepal, would suffer partial to complete collapse, resulting in very high casualties.

Most Nepalese school buildings are community built, by local craftsmen who have no formal training and are often illiterate. Technically trained people are not part of this process, unless it is funded by the government. Construction is characterized by the high degree of informality. The local availability of the construction materials, such as fired or unfired bricks, stone in mud mortar, and timber, controls the construction process. The use of modern materials such as cement, concrete, and steel bars is limited by affordability and accessibility and is confined to urban areas and areas accessible by transport.

In Nepal, there is no mandatory policy to control school design and construction. While some schools are supported by international donor agencies and/or the government requires design/drawing details, many are directly constructed by the communities without standard design criteria or technical supervision. Likewise, site-specific hazards are also not considered during the design and construction. Some design details are available, but they may not be entirely suitable for specific sites/locations. At most local levels, people lack the capacity to understand and implement the earthquake-safe construction method.

Low budgets for most school construction and lack of awareness and knowledge on the part of

S

graduate engineers of traditional and informal construction methodology result in most school buildings lacking earthquake resilience.

The National Society for Earthquake Technology in Nepal (NSET), a national NGO, conducted a program to strengthen existing school buildings and promote structural as well as nonstructural components of the school buildings for seismic safety, leveraging the decentralized, traditional, and informal approach to construction (Bothara et al. 2004). This program involved craftsman training, technology development and transfer, and community awareness raising. Many local masons became master masons. On-site master masons worked in residence, supported by visiting engineers with far-reaching effects. Shake-table demonstrations of typical versus seismic-resilient construction impressed communities with the effectiveness and feasibility of seismic-resistant measures. By raising awareness in schools, the entire community is reached because lessons trickle down to parents, relatives, and friends.

The approach developed took into account sociocultural and economic issues, with outreach to all stakeholders – school staff, students, local community, local clubs, and the local and central government. They have all been involved in the process so that they become aware of the risk and support the solution. School building construction was taken as an opportunity to train masons and to transfer simple but effective technology to others in the community, including house owners.

Following this approach, NSET retrofitted more than 40 schools, mostly of unreinforced masonry buildings. The program was found successful in transferring technology to local craftsmen who were quite keen to learn about the complete process and to adopt the technology. These masons became the propagators of the safety message in the vicinity of these schools and the replication of earthquake-resilient construction. The long-term sustainability of these impacts has yet to be assessed, but NSET's experience shows that seismic retrofitting and earthquake-resistant new construction are both affordable and technically viable.

Turkey

Turkey has more than eight million children attending schools in 64 provinces in the first- and second-degree seismic risk zones. The 1999 Kocaeli (moment magnitude Mw 7.4) and Duzce (Mw 7.2) earthquakes with approximately 20,000 fatalities raised awareness of the school safety question, and the 2002 (Mw 6.0) Afyon-Sultandagi and 2003 Bingöl (Mw 6.4) earthquakes kept awareness high.

During the 1999 earthquake in Kocaeli, 43 schools were damaged beyond repair, and 381 sustained minor-to-moderate damage. School was suspended for 4 months causing major disruption to the lives of families and children. In Istanbul, 60 km away, there was damage at 820 of the 1,651 schools. Damage at 131 of these sites necessitated at least temporary school closure. Thirteen were immediately demolished, and another 22 were later slated for demolition when retrofitting proved too costly. Fifty-nine schools were strengthened and 59 repaired.

In the Bingöl earthquake of 2003, out of 29 schools in the affected area, 4 school buildings collapsed completely, 10 were heavily damaged, 12 slightly or moderately damaged, and 3 undamaged.

Public schools in the Kocaeli earthquake fared better than residential buildings and private schools. Had children been at school during the Kocaeli earthquake, far fewer would have lost their lives. The fatality rate in residential buildings in the Kocaeli earthquake was 1.5 % in heavily damaged buildings and 16.5 % in totally collapsed buildings (Petal 2009). Similar damage in higher occupancy buildings of the same type would cause higher fatality rates. In the single example of the school dormitory in the Bingöl earthquake where 84 children died, the fatality rate was 44 %. Average risks are theoretical and don't occur. Instead, the reality is that either the school is not occupied and no one dies or it is occupied and the fatality rates are high, and the tragedy is wholly unacceptable.

There is much that is right with school construction in Turkey. As a result of an assigned

importance factor of 1.5, public schools are designed to withstand a 50 % increase in earthquake design loads. Schools have regular symmetrical structural designs, and those that are only one or two stories have fared well, for the most part meeting standards for life safety, if not continuous occupancy. The lethality of school buildings is almost entirely attributable to shoddy construction and is particularly lethal in taller buildings that may also have design defects.

For decades all public construction was under the authority of the Ministry of Public Works and Settlement. Earthquake building codes on the books since the 1930s were updated most recently in 1976 and 1998, yet the existence of these laws has not guaranteed the safety of construction. The reasons are numerous.

Beyond an undergraduate or graduate degree, there have been no independent or nonacademic professional qualifications, proficiency standards, continuing education requirements, or licensure for architects or engineers nor any qualifications for building contractors. There are also no guidelines for reliable and systematic building inspection during construction. Penalties for noncompliance with building codes are beset with bureaucratic and social impediments and often are simply not applied. Legal liability in some future event with low-frequency occurrence can hardly be a deterrent with so many to share blame. Public construction has also suffered from a standard (though not legally required) preference for the lowest bid in public tenders. The civil service employment system also lacks proficiency standards and qualifications for professional staff; so at the local level there is a wide variety in the capacity for project supervision and control. Wage and salary levels are low, and there has been opportunity for both favoritism and corruption. There are no ombudsman or advocacy services to support consumer whistle-blowers.

Istanbul provides a dramatic example of three overlapping tasks:

- Immediate response to damages caused by the 1999 earthquake

- Implementation of a comprehensive retrofitting and replacement for seismic risk mitigation
- Follow-through on an ambitious program of school expansion and construction initiated to respond to the acute shortage of class space occasioned by three additional years of compulsory education enacted in 1998

After the 1999 earthquake, the responsibility for school construction was shifted to the Ministry of Education's Division of Investments and Facilities (DIF). In turn, DIF appointed consultants from the private sector to oversee the new facility design and construction. DIF also developed standard designs for the new facilities, and new school construction was financed by a combination of government funds and charitable contributions raised by not-for-profit foundations. New construction and procurement laws also went into effect; however, the cumulative impact of these changes and pressures is not yet known (Gülkan 2004).

In Istanbul of all projects the highest priority is given to regional boarding schools, then to schools in the 12 highest-risk districts, and to those in proximity to the Marmara seacoast. The overall mitigation and retrofit effort targets more than 1,800 buildings that constitutes the 80 % of stock predating the 1998 Building Code. This ambitious program is budgeted for US$320 million (Yüzügüllü et al. 2004).

An additional problem in Turkey is that awareness of nonstructural hazards remains low. Classroom doors often open inward and shelving and laboratory equipment remains unfastened. However, concern that children advised to "drop, cover, and hold" might be injured by flimsy wooden desks led to production and distribution of 80,000 steel desks to more than 500 schools in the most vulnerable areas.

Schools are a well-distributed means of public education, and children can play a leading role in the dissemination of public safety messages. Thus, Professor Isikara, former head of KOERI, a major earthquake research institute in Istanbul, toured the country visiting schools, becoming known as "Grandpa Quake," and produced the

first children's books and popular educational and rap music cartoons for earthquake awareness. Both the Istanbul Governor's Office and KOERI's newly established Istanbul Community Impact Project (ICIP) produced handouts distributed to all school children. Schools also received books and CDs. At the national level an introduction to natural hazards was integrated into the primary school curriculum in Environmental Studies in 2002. Annual school-wide earthquake drills and preparedness and remembrance activities were initiated on 11 November 2001, to coincide with the Duzce earthquake anniversary.

Between 2001 and 2003, a cascading model of training and instruction called "ABCD Basic Disaster Awareness" was implemented by KOERI's Istanbul Community Impact Project. A curriculum was developed to address specific assessment and planning activities, physical risk reduction, and response preparedness measures to be taken prior to a disaster. This was a significant reorientation from previous "awareness" programs that began with what to do "during the shaking." A single full day of instructor training for 3,600 teachers was provided in collaboration with the Ministry of Education Provincial Directorates and outside donors. These teachers in turn communicated with 121,000 school personnel and through them with 1.68 million school children as well as with 700,000 parents. The project established an Internet-based monitoring system to monitor dissemination. Based on the success of this project, almost a decade later an even more ambitious scale-up was attempted.

In 2010 the Ministry of National Education is committed to taking the program nationwide with technical support from Risk RED, in a follow-up project supported by the American Red Cross. A distance-learning self-study curriculum was developed consisting of 1 course (10 lessons) in household disaster preparedness and 1 course (9 lessons) in school disaster management, with the goal of reaching 25,000 school-based instructors (Petal and Türkmen 2012). During the first year of deployment of the courseware in 2012, more than 79,000 MoNE employees completed one or more lessons. More than 65,600 users completed the School Disaster and Emergency

Management Course and more than 50,000 passed the final test. Almost 50,000 users have completed the Individual and Household Disaster Preparedness Course and more than 40,000 users have passed the final test. There were a total of 114,700 course completions, 92,800 final tests passed, and more than one million lessons were successfully completed by users. Within the first year, active users completed an average of almost 15 lessons each (Petal and Türkmen 2012).

In 2005, a loan to implement the Istanbul Seismic Risk Mitigation and Emergency Preparedness (ISMEP) was funded by the World Bank and European Investment Bank, to help prepare Istanbul for a probable earthquake in the Marmara Region. The project had broad aims to enhance the institutional and technical capacity of emergency management-related institutions, raise public awareness, assess priority public buildings for retrofit or reconstruction, and support building code enforcement (www. ipkb.gov.tr). "Component B" addressed seismic risk mitigation for priority public buildings. This included a feasibility study for retrofit of 1,128 education buildings on 796 school sites. Of these, 506 were strengthened and 148 were reconstructed. At the time of completion of the project, further 34 were slated for retrofitting and 27 for reconstruction.

USA: California

In 2008, seismic safety advocates in California launched the first Great Southern California ShakeOut. The now annual event has grown to include five western states and several central US states and has been conducted in cities in five other countries. In its first year four million children and adults participated through 207 school districts plus almost 750 individual schools (Risk RED 2009). In 2013 there were more than 9.6 million participants in the Great California ShakeOut drill. Globally there were almost 25 million registered participants in similarly inspired Great ShakeOut Drills. Of these almost 75 % were school-based participants (Earthquake Country Alliance 2013). This regular public

awareness event has heightened interest and concern in safe school facilities. With 3.6 million children enrolled in 262 public school districts in seven counties in Southern California, a major earthquake in the region could cause an unprecedented catastrophe for schools, children, and teachers.

School seismic safety has been a policy and a community concern in California since the 1933 Long Beach earthquake, and school emergency planning has been required statewide since 1984. With 75 years of public policy leadership to support school safety, new school construction standards are higher than those for regular buildings and come close to assuring life safety. An advice regarding non-structural mitigation measures (fastening furnishings, etc.) has been in place for 20 years and requirements for such mitigation have been in place for 10 years.

The 1933 Field Act implemented immediately after the Long Beach earthquake that year required that schools be built to 15 % higher performance standards than normal construction. In 1938 the Garrison Act required examination and improvements to pre-1933 construction, but went unenforced until 1968. The Uniform Building Code enacted in 1976 is now the current standard for safe school construction. The Field Act has been hailed as a high point in school seismic safety and California schools are considered the safest in the United States. Some school facility managers feel that its requirements are too stringent and too costly, and many seismic safety advocates feel that it does not go far enough. The Act requires that structural plans be prepared by licensed structural engineers and approved by an independent state agency (the Division of the State Architect (DSA)). Schools have continuous on-site inspection (rather than periodic), by a DSA-approved project inspector. Project architect and engineers must perform construction observation and administration, and a final verified report must be filed by the project architect, engineers, inspectors, testing labs, and the contractor (State of California, DSA 2007, 2009).

In 2007 the California Seismic Safety Commission found that: (1) The cost of compliance with the Field Act is incremental and minimal. (2) Timeliness, consistency, accuracy, and communication are being improved by the Division of the State Architect. (3) The exemplary performance of school buildings is directly attributable to the stringent seismic design provisions, plan review, field inspection, and testing required by the act and which go beyond the standard building codes. (4) All public schools should be covered. A 2009 study reported that in the four major earthquakes since the Field Act, there have been no public school collapses. The construction to Uniform Building Code, the special enforcement and quality control provisions, an oversight by the Office of the State Architect, and the 2003 publication guiding mitigation of nonstructural hazards are all judged to be successful. Nonetheless, the work is not yet complete.

In 1999 Assembly Bill 300 required desk assessment of 9,659 pre-1978 school buildings. The final report based on woefully incomplete records was released in 2002 and found 7,537 potentially vulnerable buildings requiring detailed seismic evaluation. The cost of retrofit was estimated at $4.5 billion (State of California 2002). Due to fear of planning and financial implications, details were not released to the public. In 2005 an investigative reporting series by California Watch finally achieved this (Risk RED 2009). In 2011, 20,000 uncertified projects were released on an interactive map (California Watch 2011).

A review of school seismic safety in California identifies four remaining areas of concern:

- There are still some 7,537 school buildings in California constructed before 1978 that are of questionable safety.
- Portable classrooms, which may account for one-third of all classrooms in California, may be particularly hazardous if not properly supported and fastened.
- Private schools are not currently required to meet these same construction standards as public schools.
- Nonstructural mitigation measures continue to require consistent application to protect children and adults from both injury and death.

• Each school district and private school is strongly recommended to conduct its due diligence and report on these issues transparently to parents, staff, and students, so that collective action can be taken to address these serious vulnerabilities. Neither fear nor California's persistent financial crisis in the education sector should be acceptable excuses for inaction (Risk RED 2009).

Summary

Most of these case studies have focused on the primary importance of safe school facilities, through both standards for new school construction and strategies for school vulnerability assessment and planning for retrofit and replacement. While sound earthquake engineering expertise is fundamental to advocacy, communications, planning, and execution of these efforts, it is also important to retain the perspective of the primary beneficiaries: children, teachers, and school communities. This necessitates going beyond the obvious: site selection, design, and construction. As some of these case studies indicate, it is also important to consider nonstructural mitigation, ongoing maintenance, safe access to school, school function as temporary emergency shelters, and even structural awareness education and the use of construction as an educational opportunity for children and communities. By taking this wider (and multi-hazard) view, the focus on safe school facilities overlaps with both ongoing school disaster management and with risk reduction education.

Cross-References

▶ School Seismic Safety and Risk Mitigation

References

Augenti N, Cosenza E, Dolce M, Manfredi G, Masi A, Samela L (2004) Performance of school buildings during the 2002 Molise, Italy, earthquake. Earthq Spectra 20(S1):S257–S270

Bendimerad F (2004) Earthquake vulnerability of school buildings in Algeria. In: Group meeting on earthquake safety of schools. OECD, Paris, pp 35–44

Bhatia S (2008) Earthquake-resistant school buildings in India. Urban Plan Des 161(4):147–149

Bothara JK, Guragain R, Dixit A (2002) Protection of educational buildings against earthquake. National Society for Earthquake Technology, Nepal

Bothara JK, Pandey B, Guragain R (2004) Seismic retrofitting of low strength unreinforced masonry non-engineered school buildings. Bull N Z Soc Earthq Eng 37(1):13–22

California Watch (2011) K-12/public safety/spotlight/on shaky ground. 22 Dec 2011. Accessed 3 Feb 2012

Earthquake Country Alliance (2013) Great shake out earthquake drills. http://www.shakeout.org. Accessed 30 March 2015

Grant DN, Bommer JJ, Pinho R, Calvi GM, Goretti A, Meroni F (2007) A prioritization scheme for seismic intervention in school buildings in Italy. Earthq Spectra 23(2):291–314

Gülkan P (2004) Seismic safety of school buildings in Turkey: obstacles impeding the achievable? In: OECD (ed) Keeping schools safe in earthquakes. Organization for Economic Cooperation and Development, Paris, pp 64–87

International Recovery Platform (2010) Case reports – China. United Nations International Strategy for Disaster Reduction. http://www.recoveryplatform.org/outfile.php?id=548&href=http://www.recovery-platform.org/assets/recoverycases_reports/china/CHINA_300910.pdf. Accessed 16 Feb 2014

MEXT (2008a) School seismic retrofit handbook http://www.nier.go.jp/shisetsu/pdf/e-taishinjirei.pdf. Accessed 30 March 2015

MEXT (2008b) School non-structural reference book http://www.nier.go.jp/shisetsu/pdf/e-jirei.pdf. Accessed 30 March 2015

NSET (2000) Seismic vulnerability of the school buildings of Kathmandu valley and methods for reducing it. Kathmandu Valley Earthquake Risk Management Project implemented by National Society for Earthquake Technology-Nepal, Kathmandu

Petal M (2009) Evidence-based public education for disaster prevention: the causes of deaths and injuries in the 1999 Kocaeli earthquake, VDM. Doctoral dissertation, UCLA, 2004

Petal M, Türkmen Z (2012) Dreams for Turkey: a case study of scale and reach of distance learning self-study for individual and household preparedness and school disaster management. http://www.riskred.org/activities/dreamsforturkey.pdf. Accessed 30 March 2015

Risk RED (2009) School disaster readiness: lessons from the First Great Southern California ShakeOut. http://www.riskred.org/schools/shakeout2009.pdf. Accessed 31 March 2015

Secretaría de Educación del Distrito Capital de Santafé de Bogota (2000) Análisis de vulnerabilidad sísmica de las edificaciones de la Secretaría de Educación del

Distrito y diseños de rehabilitación de algunas de ellas. Proyectos y Diseños, P&D, Ltda, Bogotá

SEEDS of India (2006) Earthquake safety in mountain regions. School Earthquake Safety Initiative, Shimla/Delhi

Shaw R, Takeuchi Y (2012) East Japan earthquake and tsunami: evacuation, communication, education and voluntarism, Research Publisher

State of California Department of General Services (2002) Seismic safety inventory of public schools (AB300 final report)

State of California, Seismic Safety Commission (2007) The field act and public school construction: a 2007 perspective. Accessed 30 March 2015

State of California, Seismic Safety Commission (2009) The field act and its relative effectiveness in reducing earthquake damage in California's public schools. Available online at http://www.seismic.ca.gov/pub/CSSC%2009-02%20Field%20Act%20Report.pdf. Accessed 16 Feb 2014

Taylor G (1989) Seismic assessment of Vancouver school buildings. TBG (Transit Bridge Group) Engineering Firm, Vancouver

Yüzügüllü Ö, Barbarosoglu G, Erdik M (2004) Seismic risk mitigation practices of school buildings in Istanbul. In: OECD (ed) Keeping schools safe in earthquakes. Organization for Economic Cooperation and Development, Paris, pp 176–183

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis

Nicola Nisticò and Alessandro Proia
Dipartimento di Ingegneria Strutturale e Geotecnica, Università La Sapienza, Rome, Italy

Synonyms

Nonstructural elements; Secondary Nonstructural Elements; Secondary Structural Elements.

Introduction

Seismic evaluation of a construction has to include, with a given level of accuracy, the interaction phenomena between (1) soil, (2) foundation, (3) structural part of the construction to be intended as the part in elevation, and (4) nonstructural part to be intended as a part of the construction with minor or no task to the structural capacity.

From a conceptual point of view, it could be easy to assert the following principle:

> the seismic evaluation of a construction has to be performed based on 1) the definition of Structural Resisting System (SRS) 2) a proper model implementation of the SRS 3) a proper analysis of the SRS.

Given that the previous principle could resolve the problem, the SRS verification remains one of the goals of the analysis process: it can be pursued with a probabilistic approach (ATC 58 2012) (1) defining the required performance (2) based on the predictable loss (direct and indirect) consequent to a given seismic event.

A general approach, in which the strategy can be framed, is the performance-based approach that (1) defines a given number of performance levels (PLs), (2) chooses a seismic level for each PL, and (3) requires a given performance for each PL. Examples of PLs are operational, immediate occupancy, life safety, and collapse prevention.

The performance-based approach delegates the applicator of it the definition or selection of the most appropriate tools to be applied for the (1) identification of the resisting system (RS), (2) structural analysis (modeling included), (3) capacity definition, and (4) verification.

The first step (RS identification) is not an easy task: the RS includes the soil and the construction that on the other hand can be split in structural and nonstructural elements which are not supposed to have any role in the global resistance of the construction with regard to neither of the so-called vertical loads or of the horizontal loads such as those that schematize wind and seismic actions.

The nonstructural elements, having no role in the seismic capacity, are generally considered as additional weight to be included in the mass evaluation, neglecting the structural interaction between them and the structural resisting system.

S

For many of the nonstructural elements, the absence of interaction can be considered realistic so that they can be considered as attached elements (from which the name attachment derives) that having a proper structure (mass, stiffness, structural capacity) have to be verified with regard to the seismic action transferred to them (from the resisting system).

The attachments (as previously introduced) are objects with their morphology so that they can be schematized with either continuous or discrete models opportunely connected to the structure they are attached to. Examples of attachments are (1) furnitures, (2) technical systems, and (3) art objects of a museum.

If the structure-attachment interaction is negligible as well as the soil-structure interaction, a cascade procedure can be adopted evaluating the seismic action (in terms of time histories or response spectrum) (1) at the base of the structure in elevation and then (2) transferring it at the points to which the nonstructural element is attached.

The cascade procedure, not considering the primary (SRS) and secondary (attachments) systems as a whole entity (PS system), cannot be adopted when the two systems are tuned; that means their periods are similar and the attachment could be acting as a tuned mass damper (TMD) for the primary system.

The seismic analysis of the attachments can be performed by means of different strategies (Villaverde 1997; Chen and Soong 1988) among which linear and nonlinear analyses can be included: the PS system can be analyzed as a global system with an evident computational effort.

The need of efficient and accurate methods to analyze the PS systems inspired methodologies (Igusa and Der Kiureghian 1985a, b) based on (1) modal synthesis, (2) perturbation theory, and (3) random vibrations.

The decoupling of the secondary system (from the P system) allows the evaluation of the seismic action in terms of Floor Response Spectra whose approach is similar to the approach that governs the decoupling between soil and structures: a Response Spectra is defined and applied at the base of the structure, including in it the effect of the propagation of the seismic action in the soil. Similarly, a spectra (FRS) is defined at the base of the attachment, including in it, with a cascade procedure, the effect of the propagation of the action at the soil (at first) and, subsequently, at the elevation structure.

The definition of a Response Spectra at the base of the attachment solves the problem since the attachment can be analyzed with traditional methodologies that are, for example, Seismic Modal Analyses or Time Histories Analyses based on acceleration histories compatible with the given Floor Response Spectrum.

Usually the effect of the propagation of the seismic event (from the soil to the attachment) is performed considering a linear behavior of the primary structure: this is supposed a realistic assumption for new conceived structure when operational and immediate occupancy PLs are considered.

The linear structural behavior of the principal system could be considered as nonrealistic in some cases, where the system's nonlinearity could produce effects (on the attachment) more severe than those evaluable with a linear behavior assumption (Chaudhri and Villaverde 2008).

The attachments, as discussed so far, are secondary elements that do not give any contribution to seismic resistance of the primary system and, in these terms, can be classified as secondary nonstructural elements (NSEs) to be distinguished from the secondary structural elements (SEs) that have a negligible role in supporting the seismic action but can have a specific role in transferring the vertical load to the foundation system.

Secondary elements, either attachment (secondary NSEs) or structural elements (SEs), are both systems having their structure, opportunely linked to the primary structure: they have to be adequately modeled in order to be analyzed with the strategies common to the seismic branch, such as (1) static analyses (linear and nonlinear), (2) modal response spectrum analyses, and (3) time domain analyses either linear or nonlinear.

Morphological and Phenomenological Aspects Versus Modeling and Analysis

Depending on the structural resisting system typology and the construction usage, the NSE typology can be wide (see Fig. 1, for the building case) and their taxonomy can be found in Taghavi and Miranda (2003) where a comprehensive database of nonstructural components is presented covering different aspects such as, among others, cost information from which is deducible that the structural cost of a building could be not relevant with respect to the global cost: the office buildings structural costs, even if relevant, are only 18 % of the total cost of construction that can be split in cost of (1) structural elements, (2) secondary structural elements, and (3) nonstructural elements such as the contents are. The cost of the nonstructural elements can be estimated to reach the 70 % of total construction costs if the hotel buildings are concerned, while it is lower in office buildings (62 %) and hospitals (48 %) where contents (such as medical equipment) can be estimated to be 44 % of the total cost.

Economic loss due to seismic nonstructural damage can be relevant: during the 1994 Northridge earthquake, the nonstructural damage was (Kircher 2003) about 50 % of the global building damage which was estimated to be $18.5 billion.
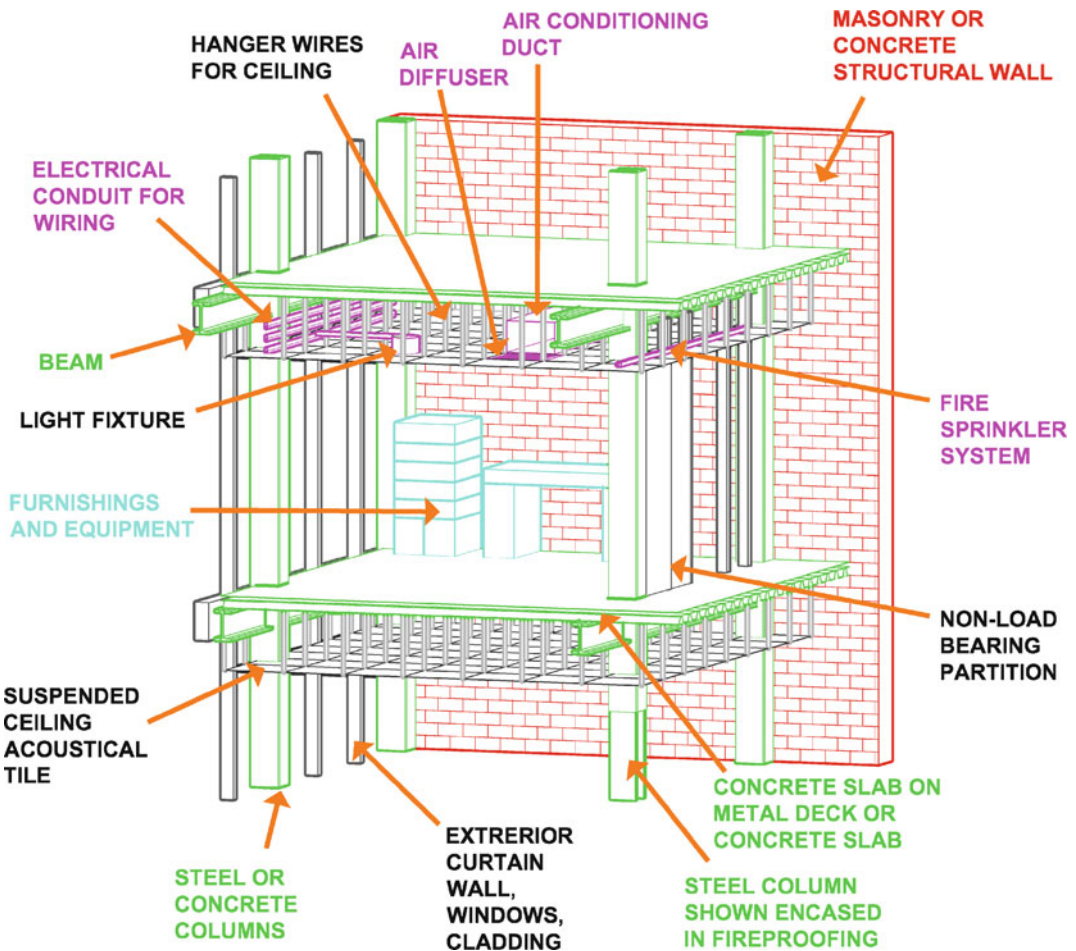
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 1 Typology of building nonstructural elements. Reproduced from FEMA 74 (2005)
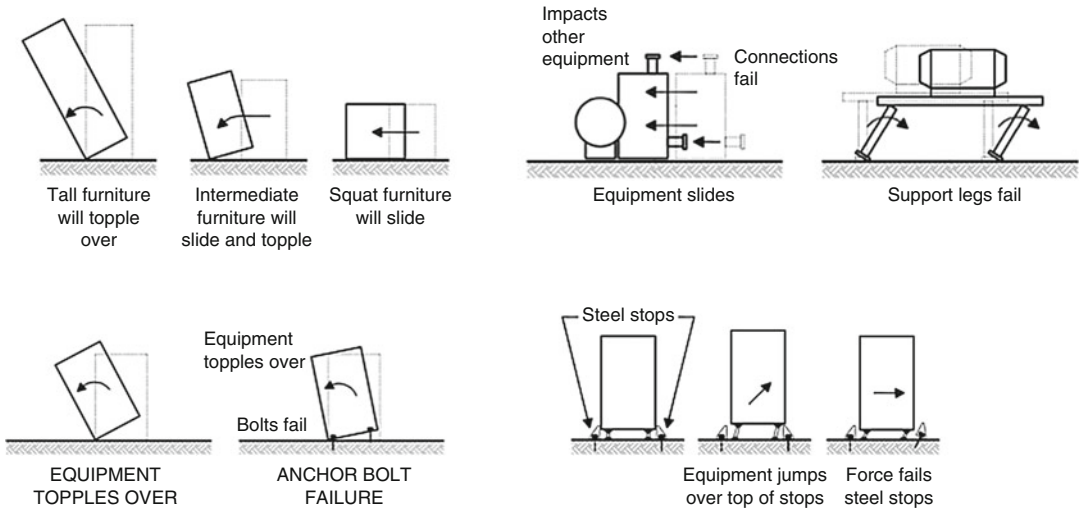
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 2 Principal rigid body mechanism

Most of the NSEs have limited seismic performance because they are not properly attached to the primary structures, so that, depending on their slenderness, they can (see Fig. 2) (a) topple over, (b) slide and topple, and (c) slide. The loss of capacity of the NSEs or their connections can cause damage to other equipment (see Fig. 2) and injury to people, so careful attention has to be paid to the design of the connection (see Fig. 2).

Unlike the old conceived NSEs, the new-generation elements can have a good seismic performance, thanks to the wide range of connections that can be adopted to link the NSE to the structure. Depending on the case at hand, the design can include (1) seismic joints opportunely designed to accommodate seismic displacements, (2) seismic isolators to reduce the acceleration level, and (3) dissipative device to reduce the level of acceleration, velocity, and displacement.

So the new-generation NSEs cannot be conceived without an adequate strategy for the connection design, an example of which is reported in Fig. 3 where the case of a machinery (for cement production) mounted on a steel structure *attached* on a reinforced concrete building has been reported: dissipative devices have been introduced at the base (of the steel structure) to

reduce the seismic action on both the machinery and building.

Referring (see Fig. 3) to the previously introduced example (where the equipment can be considered as an attachment of the primary structure), the following indications can be given for the seismic analysis:

1. If the NSE is rigidly connected to the structure and its mass (M_{NS}) is not negligible with respect to the building mass (M_S), a global analysis of the PS system is required. In case of modal spectra time history (TH) analysis, some approximation in the damping definition is needed due to the different damping of the attachment (steel structure) with respect to the reinforced concrete structure. More appropriate step-by-step TH analyses can consider the element damping, properly modeling it by means of dashpots when *nonclassical* damping is present.

2. If M_{NS} is negligible with respect to M_S, but the NSE period ($T_{NS)}$ is close to structural period (T_S), the so-called *tuning* happens and a global analysis of the PS system is required, not excluding positive effect.

3. If M_{NS} is negligible with respect to M_S, and T_{NS} is not close to T_S, the decoupling could be considered and the strategy for the analyses
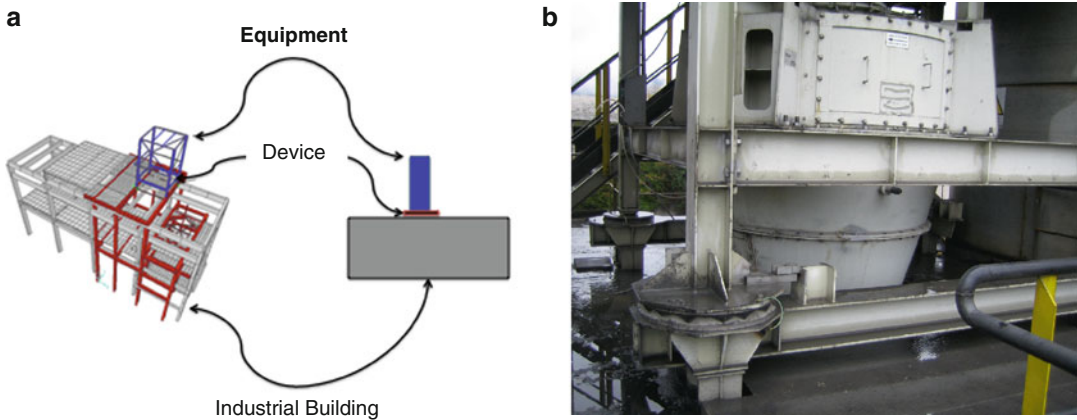
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 3 (**a**) Cement industrial building. Machinery mounted upon an *rc* structure. (**b**) Cement Industrial Building Applications: Italy

could be oriented to the definition of the seismic action at the level of the connection, considering the structural system as a stand-alone element subjected to a seismic action at the base of it.

The following strategy can be considered: (a) time history analyses applied either to the whole system or to the stand-alone attachment, considering the acceleration histories recorded at the level of the attachment connection and (b) spectrum-based analyses defining an acceleration spectrum consistent with the time histories recorded at the points where the attachment is connected to the structures: the generated spectrum is usually named Floor Response Spectrum (FRS) even if (as the case reported in Fig. 4) the FRS has been evaluated where the attachment is linked and not at the floor level.

The FRS definition is an important task for the NSE analysis, being the referred analysis tool to be adopted, due to its recognized simplicity in conjunction with good level of reliability (in those cases where the decoupling can be adopted). It is possible to affirm that the usually adopted acceleration spectrum *is* to structural analysis *as* the FRS *is to* the analysis of nonstructural elements, so that the FRS is a period-dependent function that can be evaluated for different value of the attachment damping given a specific soil and structure characterized by their own periods and dampings.

Differently from the secondary nonstructural elements, the secondary structural elements (SE) require different strategies of modeling and analysis. It is useful to introduce them as reported in CEN (2008): some structural elements (i.e., beams and columns) can be designed as seismic secondary elements, neglecting their contribution to the global seismic resistance so that their stiffness and strength can be neglected. As further specified in CEN (2008), the SE and the joint (that link them with structure) have to be designed considering (1) the vertical gravitational load, (2) displacement consequent to the seismic action, and (3) second-order effects that include the flexural moments evaluated considering the deformed element shape (P-Delta effect).

Clearly the previous definition of the SE supposes that they have a negligible influence in the global structural behavior.

Starting from the classification of the secondary elements in NSEs (attachments) and SEs, the following can be asserted: (1) given a construction, it includes a principal (P) and a secondary (S) structure; (2) if the S structure has a negligible influence on the P structure, the whole construction (PS) structural behavior can be decoupled; (3) the S structure can be classified in structural (SEs) and nonstructural (NSEs) elements; (4) the SEs have to be designed for vertical gravitational

S
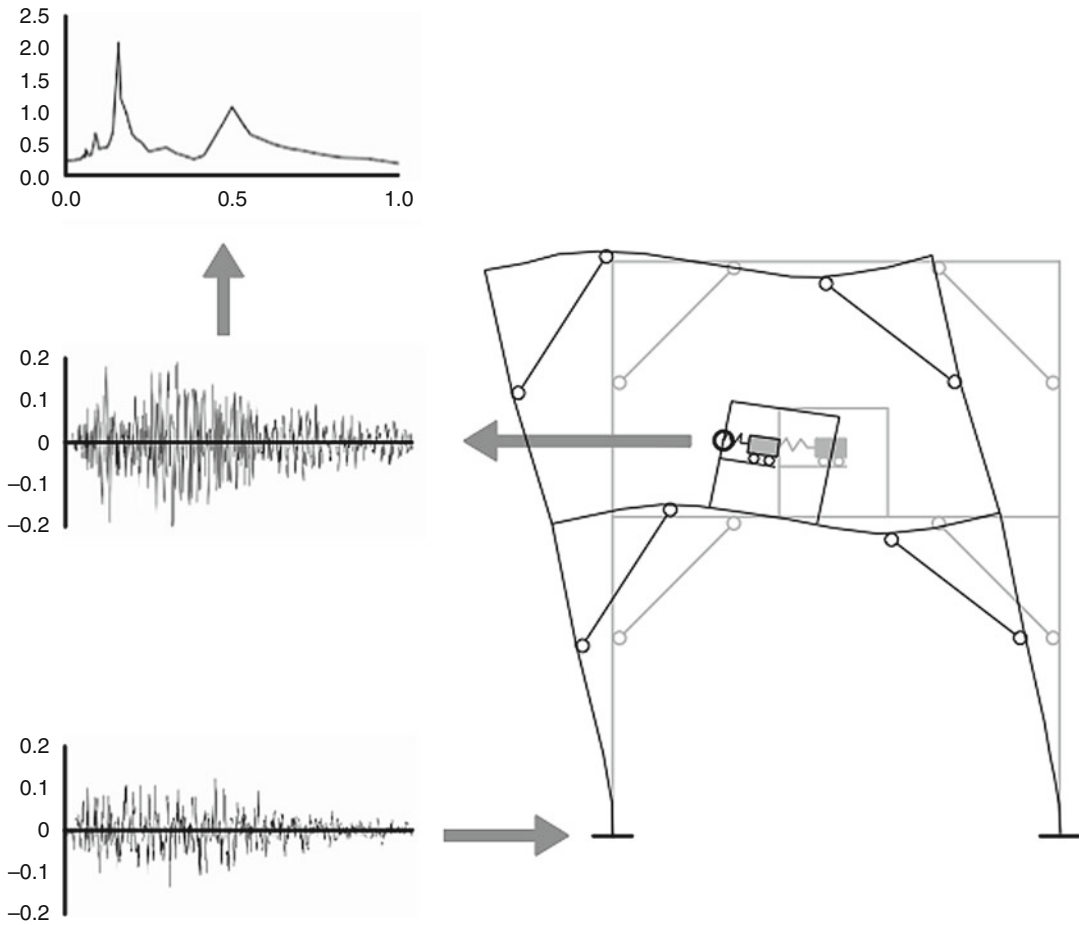
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 4 FRS generation

load (transferred from the PS to them, including the self-weight loads) considering the seismic-induced displacement (P-Delta effect included); (5) the NSEs have to be verified with regard to self-weight loads and seismic action transmitted by the P structure; (6) the secondary structure elements and their supports (links) have to be verified in order to avoid that their partial or total failure can induce injury to people or important objects; and (7) if the interaction between P and S systems is not negligible, a global PS analysis is required.

In the following, some of the principal characteristic regarding secondary structural and nonstructural elements will be described.

Secondary Structural Elements

Typical case of SE is internal and external building partitioning system (Glass Systems included): their seismic contribution is usually neglected in the seismic analysis (1) accepting (for severe earthquake) their damage and (2) imposing that out-of-plane collapse (see Fig. 5) has to be prevented.

The infilled partitioning system can have a role in transferring the vertical load even if they can have a minor contribution in seismic global capacity. If they have no role in the vertical load as well as in the seismic P structure capacity, their classification as NSE (attachments) is reasonable.

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 5 Molise (Italy): Seismic event occurred in 2002 (October 31 (M = 5.4) and November 1 (M = 5.3). Example of damaged infilled frame: in-plane and out-of-plane mechanisms

If no flexible joints are considered (in between the P and S systems they represent), the absence of collaboration with the P system is not judicable by means of qualitative considerations, but it can require a structural analysis of the PS system including them as structural elements. In this case, the designer can follow some suggestions such as those included in CEN (2008) that consider a structural system as SE if its global stiffness is lower than 15 % of the P system stiffness. The models concerning the infilled frames systems are well known.

The damage level in the partitioning system is usually controlled imposing a threshold to the interstory drift (see Table 1) as function of the performance level and construction usage (ASCE 2002).

Attachments

Typical cases of attachments are parapets, windows, partitioning systems, antennas, electrical power systems, and furnitures. Depending on their components, they can be sensitive to the seismic acceleration or deformation (see Table 2).

Modeling and Analysis

NSEs are elements characterized by their mass and stiffness, and independently of the seismic action they are subjected to, they can be modeled and analyzed based on FEM strategies considering either their linear or nonlinear behavior.

In general the NSE is a system composed by subsystems with a structural complexity (see Fig. 6) that can require 3D complex models to be calibrated by means of experimental tests (Fig. 6a) including identification strategies: dynamic tests can be carried out by means of shacking tables (Fig. 6c).

The experimental tests in support of modeling and analysis implementation are especially required either when the importance of NSE usage is considered strategic or when the cost of it justifies the experimental activity. In some cases, a qualification procedure can be required, generally ruled by international standard (Gilani et al. 1999; IEEE 2005).

Modeling

Modeling has to take into account all the components that give stiffness and strength

S

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Table 1 Drift control: usually adopted values as function of usage and performance level

Performance level	Damage description and downtime/loss	Drift control
Immediate occupancy: usually required for construction which usage is considered strategic	Negligible structural damage; essential systems operational; minor overall damage. **Downtime/loss: 24 h**	*0.3 %* (stiff joint), *0.6 %* (deformable joints)
Life safety	Probable structural and nonstructural damage; no collapse; minimal falling hazards; adequate emergency egress. **Downtime/loss: possible total loss**	*0.5 %* (stiff joint) *1.0 %* (deformable joints)
Collapse prevention	Several structural and nonstructural damage; incipient collapse; probable falling hazards; possible restricted access. **Downtime/loss: probable total loss**	*Not required*

contributions, including the connection elements that, if needed, have to be modeled as nonlinear elements.

In many cases, such as the bushing sketched in Fig. 6a, an accurate modeling requires informations about all the subcomponents (coil springs, valves) in terms of mass stiffness and strength. The needed informations are not usually known and the element investigation has to be supported by means of experimental tests devoted either to global information acquisition (frequencies, modal shapes) or to evaluation of the level of

performance given a defined seismic action. Experimental tests can include shaking table tests or static tests: this aspect is strictly linked to the qualification process (IEEE 2005).

Seismic Action Modeling and Structural Analyses

Seismic action can be simulated according to the usually adopted strategies that, for the case at hand, include (1) time histories (usually in terms of acceleration) and (2) response spectrum finalized either to modal analyses or to static linear or nonlinear pushover analyses.

Seismic level will depend on the referred performance level that (see Table 1) identifies the required performance associable to a seismic event with a given return period, to be defined based on cost-benefit analysis.

General rules valid for secondary elements are the following:

1. Mass and stiffness uncertainties have to be considered together with spatial distribution of seismic effect in case of extended SE systems.
2. Seismic effects on SE have to take into account, in general, both horizontal and vertical components to be evaluated based on a structural model of the principal system.
3. If the SE behavior can be decoupled from the principal system, the datum method for the evaluation of the peak acceleration at the SE is based on the Floor Response Spectrum (FRS) that given an SE element, with a defined structural period and damping, attached to a given part of a structure, having its mechanical properties, subjected to a given seismic event (*E*), allows to define the peak acceleration to which the element will be subjected when the seismic event (*E*) is transferred at the base of NS element.

Based on the knowledge of the FRS, one of the following methods can be adopted: (a) static equivalent forces (including nonlinear pushover analysis), (b) modal analysis, and (c) time history (linear or nonlinear) analyses based on accelerograms compatible with the FRS.

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Table 2 NSE classification (ATC/BSSC 1997) and element sensitivity with regard to acceleration and deformation

Component	Sensitivity		Component	Sensitivity	
	A	D		A	D
A. Architectural			B. Mechanical equipment		
1. Exterior skin			**1. Mechanical equipment**		
Adhered veneer	S	P	Boilers and furnaces	P	
Anchored veneer	S	P	General mfg. and process machinery	P	
Glass blocks	S	P	HVAC equipment, vibration isolated	P	
Prefabricated panels	S	P	HVAC equipment. Nonvibration isolated	P	
Glazing systems	S	P	HVAC equipment, mounted in-line with ductwork	P	
2. Partitions			**2. Storage vessels and water heaters**		
Heavy	S	P	Structurally supported vessels (category 1)	P	
Light	S	P	Flat bottom vessels (category 2)	P	
3. Interior veneers			**3. Pressure piping**	P	S
Stone, including marble	S	P	**4. Fire suppression piping**	P	S
Ceramic tile	S	P	**5. Fluid piping, not fire suppression**		
4. Ceilings			Hazardous materials	P	S
(a) Directly allied to structure	P		Nonhazardous materials	P	S
(b) Dropped, furred, gypsum board	P		**6. Ductwork**	P	S
(c) Suspended lath and plaster	S	P			
(d) Suspended integrated ceiling	S	P			
5. Parapets and appendages	P				
6. Canopies and marquees	P				
7. Chimneys and stacks	P				
8. Stairs	P	S			

A acceleration sensitive; *D* deformation sensitivity; *P* primary response; *S* secondary response

Floor Response Spectra-Based Evaluation

Floor Response Spectra are functions that define the response spectrum of a given response parameter (e.g., acceleration, velocity, displacement) as a function of period and damping of a given structure (attachment) localized at a given point of the construction.

The generally adopted technique for the FRS definition consists in (1) analyzing the P structure (to which the S structure is attached) in the time domain, considering n time histories (e.g., acceleration TH), (2) evaluating (for each TH), at a given point of the structure, the TH of the acceleration and the related response spectrum for a given damping value, and (3) defining one representative spectrum (based on the n available FRSs) having a given overcoming probability (usually a 50 % probability is considered): for

low values of n (e.g., minor than 7), an envelope spectrum has to be considered.

Usually the location of the attachment is not known in advance, so that the previous procedure can be applied considering p points obtaining p Response Spectra. For those points that are located at the same level (*floor*) of the P structure, a single spectrum can be evaluated (enveloping the Response Spectra), naming it Floor Response Spectrum.

If the goal is the evaluation of conservative FRSs, for each floor, a set of points has to be opportunely selected so that both translational and rotational effects are captured: they usually include the floor centroids and one or more corners for each floor.

It is worth mentioning that having defined the P system structural model, it is possible to

S

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 6 (**a**) Morphology of a bushing. (**b**) Bushing experimental test carried out at UC Berkeley (CA). (**c**) Cabinet experimental test carried out at UC Berkeley (CA)
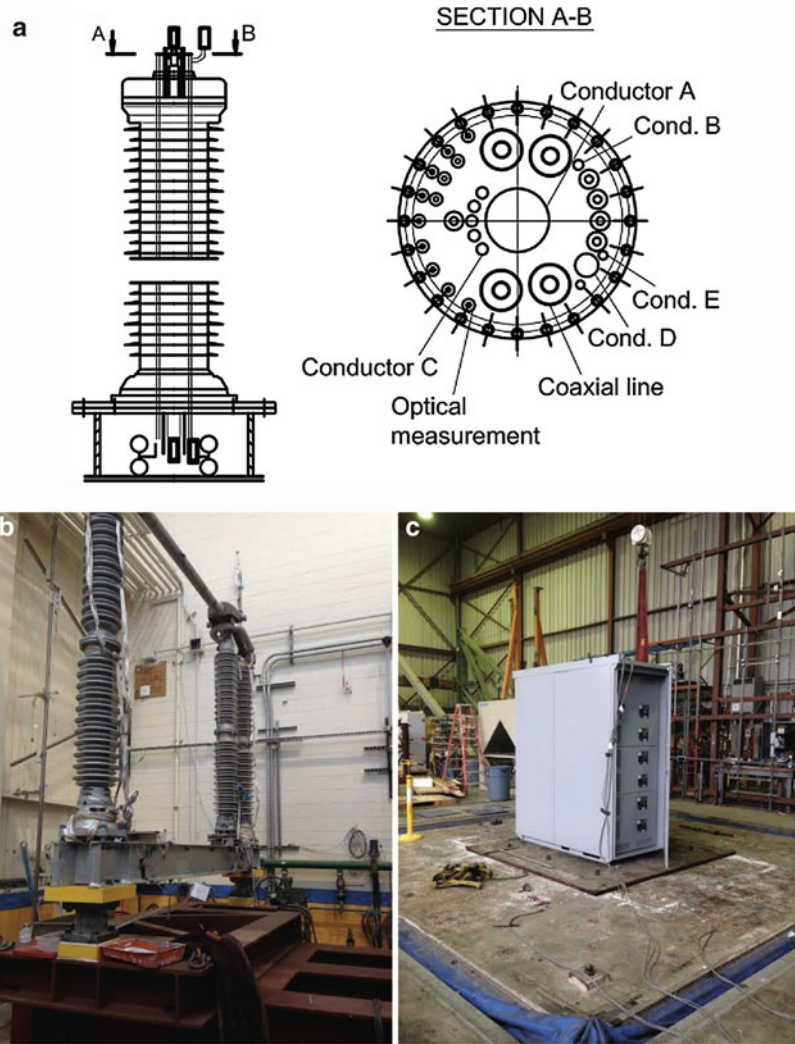
evaluate a transferring function H_p that (1) knowing the Fourier Transform (F_i, input FT) of a given accelerogram (2) allows the definition of the Fourier Transform of the TH acceleration at given point (F_o, output FT) so that (3) the inverse Fourier Transform of F_o gives the TH at the considered point that is the required information for the Response Spectrum evaluation.

Alternatively, if the Power Spectrum G_i of a given earthquake or of a family of earthquakes is known as well as the previous defined Input Transferring Function (F_i), the output Power Spectrum (G_o) is evaluable according to Eq. 1.

$$G_o(w) = G_i(w) \, |H_p(w)|^2 \qquad (1)$$

Knowing $G_o(W)$ and the transferring function (H_{SDOF}) of a single-degree-of-freedom system (SDOF), the SDOF Spectra Power Density is evaluable (see Eq. 2) and the related Response Spectrum is the required FRS.

$$G_{SDOF}(w) = G_o(w) \, |H_{SDOF}(w)|^2 \qquad (2)$$

The previously presented approaches are not usually adopted for conventional structures as they are time consuming, so that predictive expressions are proposed in literature or enforced in

international recommendations: given the peak ground acceleration, the floor peak acceleration is evaluated by multiplying the PGA by an analytical function, named S_a in the following.

The usually proposed functions (S_a) are based on (1) simplified expressions for the evaluation of the floor acceleration and (2) simplified shape functions representative of the required FRS.

It is worth mentioning that given a structure, knowing of it (1) the prevalent modal shape (Φ), (2) the prevalent period (T_s), and (3) the modal participation factor (Γ) if a seismic action is considered, the absolute structural acceleration ($\ddot{u}_i$) associated to the single modal coordinate (Φ_i) can be evaluated according to Eq. 3 where $R_S(T_S)$ is the value of the normalized acceleration spectrum, for a given value of the structural damping (ζ_S) at the prevalent period of the structure.

$$\ddot{u}_i = PGA \cdot \Gamma \cdot \{1 + [R_S(T_S) - 1] \cdot \Phi_i\} \quad (3)$$

Assuming a given analytical function (R_{FRS}), being it dependent on the period (T_{NS}) and the damping (ζ_{NS}) of the nonstructural element, the required FRS, associable to the Φ_i modal coordinate, is equal to

$$FRS_i = PGA \cdot \Gamma \cdot \{1 + [R_S(T_S) - 1] \cdot \Phi_i\} \\ \cdot R_{FRS} \cdot (T_{NS}) \quad (4)$$

Usually, Eq. 4 is simplified adopting (1) a constant value for $R_S(T_s)$ evaluated at the plateau of the acceleration spectrum (assumed in the range

of 2–3), (2) a simplified expression for the evaluation of modal displacement Φ_i of a given floor supposed to be equal to z/H where z is the level of the considered floor and H is the total construction height, and (3) a value of Γ between 1 and 1.5.

Based on the previous assumptions, Eq. 4 can be simplified as follows, having assumed $R_S = 3$, $\Gamma = 1$:

$$FRS_i = PGA \cdot \Gamma \cdot [1 + 2 \cdot (z/H)] \cdot R_{FRS} \cdot (T_{NS}) \quad (5)$$

In the following sections, some literature expressions (CEN 2008; FEMA 369 2001; AFPS 2007; KTA 2012) will be given, expressing them in terms of the normalized FRS (S_a) that corresponds to FRS_i evaluated for PGA = 1.

It has to be specified that in order to show the trend of the S_a functions, they will be plotted, contextualizing it to the simple structure described in the Case Study section.

CEN (2008)

The following Eq. 6, plotted in Fig. 7, is proposed, supposing $\zeta_S = \zeta_{NS} = 5\%$.

$$S_a = 3 \cdot \frac{1 + \dfrac{z}{H}}{1 + \left(1 - \dfrac{T_{NS}}{T_S}\right)^2} - 0.5 \geq 1 \quad (6)$$

Equation 6 can be evaluated assuming $T_{NS} = 0$, obtaining (1) the value of the expression

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 7 Amplification factor S_a for different values of z/H

••••• z/H = 0.24   ▬ ▬ z/H = 0.43   ▬ • ▬ • z/H = 0.62   ▬▬▬ z/H = 0.81   ▬▬ z/H = 1.00
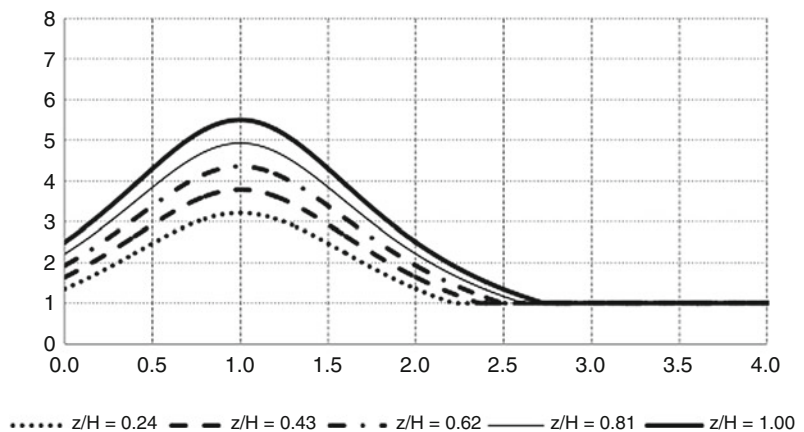
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 8 Amplification factor (S_a) shape versus T_{NS}/T_N

$\cdots\cdots$ z/H = 0.24 $-\ -\ -$ z/H = 0.43 $-\cdot\ -\cdot$ z/H = 0.62   ⎯⎯ z/H = 0.81   ▬▬ z/H = 1.00
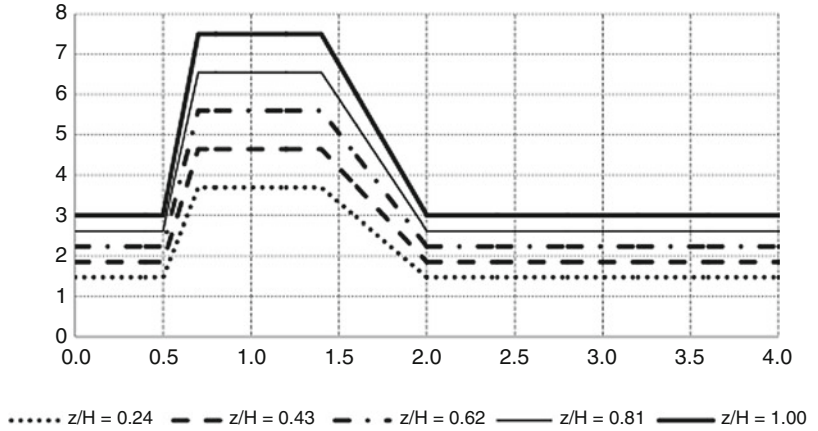
adopted for the evaluation of the normalized floor acceleration (Eq. 7a) and (2) the value of the adopted expression for the evaluation of R_{FRS} (Eq. 7b). So that Eq. 6 can be rearranged as reported in Eq. 7c.

It is possible to recognize that (1) a value of 2.5 has been assumed for the evaluation of R_s and (2) the expression adopted for the evaluation of R_{FRS} is supposed to be dependent on the normalized floor height.

$$\frac{\ddot{u}_i}{\mathrm{PGA}} = \left\{1 + [R_S(T_S) - 1] \cdot \frac{z}{H}\right\} = 1 + 1.5 \cdot \frac{z}{H} \tag{7a}$$

$$R_{FRS} = \frac{1 + \frac{z}{H}}{1 + 1.5 \cdot \frac{z}{H}} \cdot \frac{3}{1 + \left(1 - \frac{T_{NS}}{T_S}\right)^2} - \frac{0.5}{1 + 1.5 \cdot \frac{z}{H}} \tag{7b}$$

$$S_a = \left(1 + 1.5 \cdot \frac{z}{H}\right)$$
$$\cdot \left\{\frac{1 + \frac{z}{H}}{1 + 1.5 \cdot \frac{z}{H}} \cdot \frac{3}{1 + \left(1 - \frac{T_{NS}}{T_S}\right)^2} - \frac{0.5}{1 + 1.5 \cdot \frac{z}{H}}\right\} \geq 1 \tag{7c}$$

FEMA 369 (2001)

The following Eq. 8, plotted in Fig. 8, is proposed supposing $\zeta_S = \zeta_{NS} = 5\ \%$, where the values of R_{FRS} are reported in Table 3:

$$S_a = \left(1 + 2 \cdot \frac{z}{H}\right) \cdot R_{FRS} \tag{8}$$

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Table 3 Values of R_{FRS} as function of T_{NS}/T_N

T_{NS}/T_N	R_{FRS}
$T_{NS}/T_N < 0.5$ and $T_{NS}/T_N > 2.0$	1.0
$0.7 \leq T_{NS}/T_N < 1.4$	2.5
$0.5 \leq T_{NS}/T_N < 0.7$	($7.5 \times T_{NS}/T_N$) -2.75
$1.4 \leq T_{NS}/T_N < 2.0$	$6 - (2.5 \times T_{NS}/T_N)$

It is possible to recognize that (1) a value of 3.0 has been assumed for the evaluation of R_s and (2) the expression adopted for the evaluation of R_{FRS} is supposed to be independent on the normalized floor height.

AFPS (2007)

The following expression is proposed:

$$S_a = \sqrt{1 + \Gamma_S^2 \cdot R_S^2 \cdot \left(\frac{z}{H}\right) 2\alpha} \cdot R_{FRS} \tag{9}$$

where (1) α is a parameter to be calibrated in order to minimize the difference between the effective modal displacements (Φ_i) and the proposed simplified expression (z/H), (2) R_s is the value of the normalized structural acceleration evaluated at the fundamental structural period (T_S) for a considered value of the structural damping, (3) the participation factor (Γ_S), evaluable according to Eq. 10a, assumes a

maximum value of 1.6, if $\alpha = 1.5$ is imposed, and (4) R_{FRS} values are reported in Table 3 as a function of parameter A (see Eq. 10b) that takes into account the damping (ζ_{NS}) of the nonstructural element.

$$\Gamma_S = \frac{2\alpha + 1}{\alpha + 1} \quad (10a)$$

$$A = \frac{35}{2 + \xi_{NS}} \quad (10b)$$

It is worth mentioning that Eq. 9 derives from Eq. 4, with the following assumptions: (1) a unitary participation factor is considered for the ground acceleration, while the principal mode participation factor is considered according to Eq. 10a. (2) The spectral acceleration $R_s(T_s)$ is considered for the evaluation of the floor relative acceleration instead of the spectral relative normalized acceleration ($R_s(T_s) - 1$). (3) The ground acceleration and the relative structural acceleration are combined through the SRSS (Square Root of the Sum of the Squares) combination rule (Table 4 and Fig. 9).

KTA (2012)

The proposed expression does not give any information to evaluate the floor acceleration (a_g), but it only defines the amplification shape (R_{FRS}) reported in Fig. 10 (left), where the f-axis is the component frequency axis (in logarithmic scale) and f_1, f_n, and f_{limit}, respectively, are (1) lowest decisive eigenfrequency of the principal system

at the lower limit value in the variation range of the system parameters, however, not lower than the rightmost corner frequency of the highest plateau of the associated response spectrum; (2) highest decisive eigenfrequency of the principal system for the upper limit value in the variation range of the component parameters, however, not lower than the rightmost corner frequency of the highest plateau of the associated response spectrum; and (3) upper limit frequency of the associated response spectrum.

The maximum value of the amplification factor is reported in Fig. 10 (right), where D_1 and D_2 are respectively the damping ratios (in percentage of critical damping) of structural and nonstructural elements whose suggested values are reported in Table 5.

In order to compare the obtainable FRS with those previously discussed, the amplification factor (S_a) is plotted in Fig. 11, having assumed (1) a 5 % damping for both structural and nonstructural

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Table 4 Values of RFRS as function of T_{NS}/T_N

T_{NS}/T_N	R_{FRS}
$T_{NS}/T_N < 0.5$ and $T_{NS}/T_N > 2.0$	1.0
$2/3 \leq T_N/T_N < 3/2$	A
$0.5 \leq T_{NS}/T_N < 1.5$	$A - (A-1) \times \frac{\left(\log\frac{3}{2}\frac{T_{NS}}{T_S}\right)}{\left(\log\frac{3}{4}\right)}$
$2 < T_{NS}/T_N$	$K_T = A - (A-1) \times \frac{\left(\log\frac{2}{3}\frac{T_{NS}}{T_S}\right)}{\left(\log\frac{4}{3}\right)}$

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 9 Amplification factor (S_a) shape versus T_{NS}/T_N ($\alpha = 1.5$, $\Gamma_S = 1.6$, $R_S = 2.5$)

•••••• z/H = 0.24 ▬ ▬ z/H = 0.43 ▬•▬ z/H = 0.62 ——— z/H = 0.81 ▬▬▬ z/H = 1.00
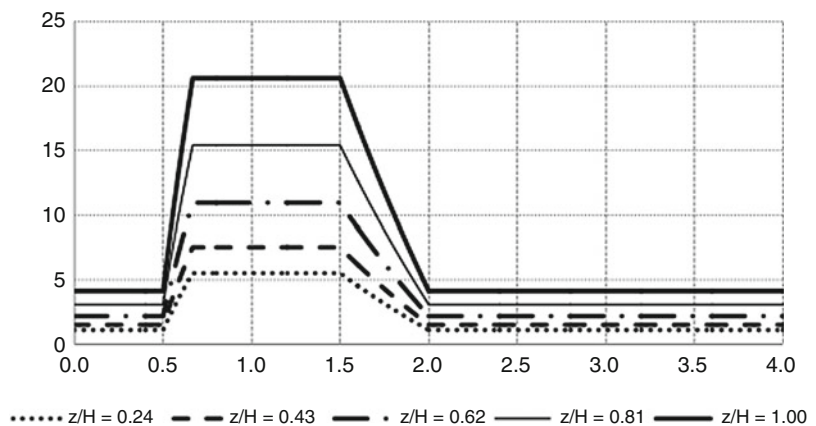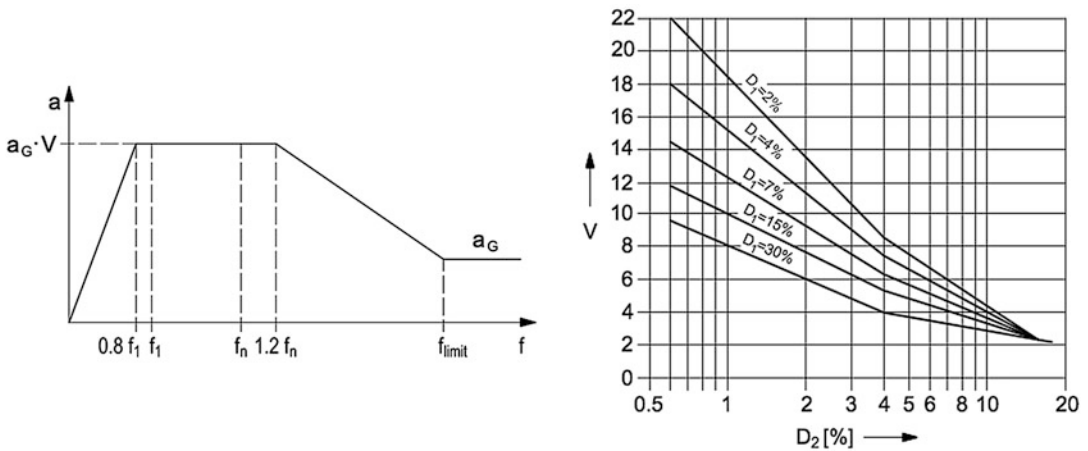
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 10 Amplification factor shape (*left*) and maximum amplification factor (*right*)

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Table 5 Suggested damping values (in percent of critical damping). Column A: to be adopted for verifying the load-carrying capacity and integrity and for determining the spectra. Column B: in the case of mechanically active components for which the functional capability is verified by a deformation analysis

Components	Damping ratios	
	A	B
Pipes	4	2
Steel with welded connections and welded components (e.g., vessels, valves, pumps, motors, ventilators)[a]	4	2
Steel with SL or SLP bolt connections (*SL – structural bolt connection with a borehole tolerance ≤ 2 mm; SLP – fitted bolt connection with a borehole tolerance ≤ 0.3 mm*)	7	4
Steel with SLV(P) or GV(P) bolt connections (*SLV(P) – preloaded fitted bolt connection; GV(P) – fitted friction-grip bolt connection*)	4	2
Cable support structures	10[b]	7
Fluid media	0.5	0.5

[a]If, on account of the design, deformations are possible only in small regions of the structure (low structural damping), the values as listed shall be halved (special cases)

[b]In well-substantiated cases, the damping ratio may be increased up to 15 %

elements, (2) the FEMA expression (see Eq. 11) for the evaluation of the normalized floor acceleration, (3) $f_1 = 6.66$ Hz that is the rightmost corner frequency of the acceleration plateau of CEN (2008) type 1 Spectrum (A soil), and (4) $f_n = 11.1$ Hz that is the highest decisive eigenfrequency of the principal system described in the Case Study section.

$$a_g = 1 + 2 \cdot \frac{z}{H} \qquad (11)$$

Verification

As stated in KTA (2012), the verification process has to regard (1) the load-carrying capacity in terms of strength, stability, and secure positioning (e.g., their protection against falling over, dropping down, impermissible slipping); (2) the integrity, that is, the capability of a component above and beyond its load-carrying capacity to meet the respective requirements regarding leak tightness and deformation restrictions; and (3) the functional capability, that is, the capacity of a system or component above and beyond its load carrying capacity to fulfill the designated tasks by way of its respective mechanical or electrical function.

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 11 Normalized floor acceleration based on Eq. 11
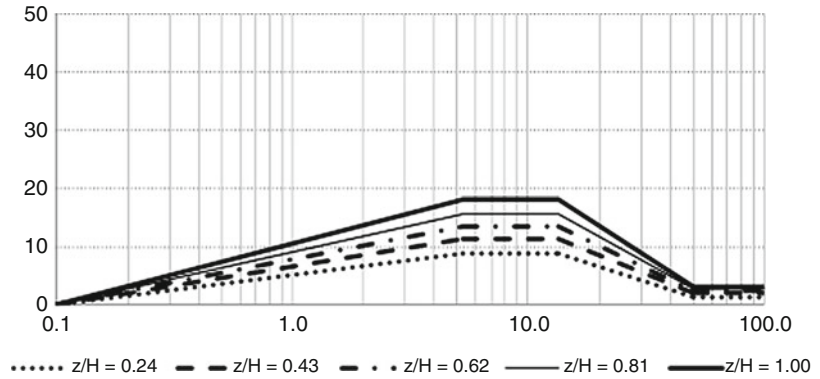
Depending on the importance of the element to verify and the material (conventional or nonconventional material), the verification process could include experimental tests either for the validation of the numerical model or for the qualification of the element itself. The verification procedure can include a) analysis, b) physical experiments, and c) analogies and plausibility considerations.

Based on the introduced classification that distinguishes secondary element in structural and nonstructural, the following criteria can be defined:

1. Secondary structural elements have to be verified with regard to the vertical loads transmitted from the P structure, opportunely combined with the other action considered to be contemporary to the seismic action. The connections have to be verified with regard to the seismic-induced action, including second-order effects such as those induced by the axial load in the deformed configuration (P-Δ effects).
2. Secondary nonstructural elements have to be verified with regard to the self-weight loads opportunely combined with the other actions considered to be contemporary to the seismic action.
3. For both types of elements (S and NS elements), the action supposed to act contemporary with the seismic action can be consequent to different events such as those pertaining to collisions, explosions, and fires.

The verifications have to consider potential damage induced to other elements, which loss of capacity could induce either human or economic loss.

The verification is performed checking that the element capacity will be greater than the demand, defined in terms of different mechanical properties (stresses, forces, displacement) depending on the adopted materials.

In order to define the design forces, the considered floor response acceleration spectra can be reduced to take into account the nonstructural element ductility. If the FRSs have been numerically evaluated, they have to be modified (see Fig. 12) to take into account the structural stiffness uncertainties: (1) an adequate plateau has to be imposed in correspondence of the structural period, (2) the linear envelope has to be properly introduced, and (3) the ductility of the nonstructural element can be considered, properly reducing the FRS (see Fig. 12b, c).

Case Study

The previously described procedures to determine FRSs will be applied to a steel frame system hosting a set of equipments, whose characteristics and localization are reported in Table 6 and Fig. 13 (left), reproduced from KTA (2012).

The maximum acceleration of each equipment can be evaluated by means of (1) time histories considering the interaction between the principal structure and the equipment or (2) FRSs

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 12 (**a**) Design spectra definition. Shape modification: alternative solution. (**b**) Design spectra definition. Design spectra based on ABCDEF spectrum. (**c**) Design spectra definition. Design spectra based on ABB'CDEF spectrum
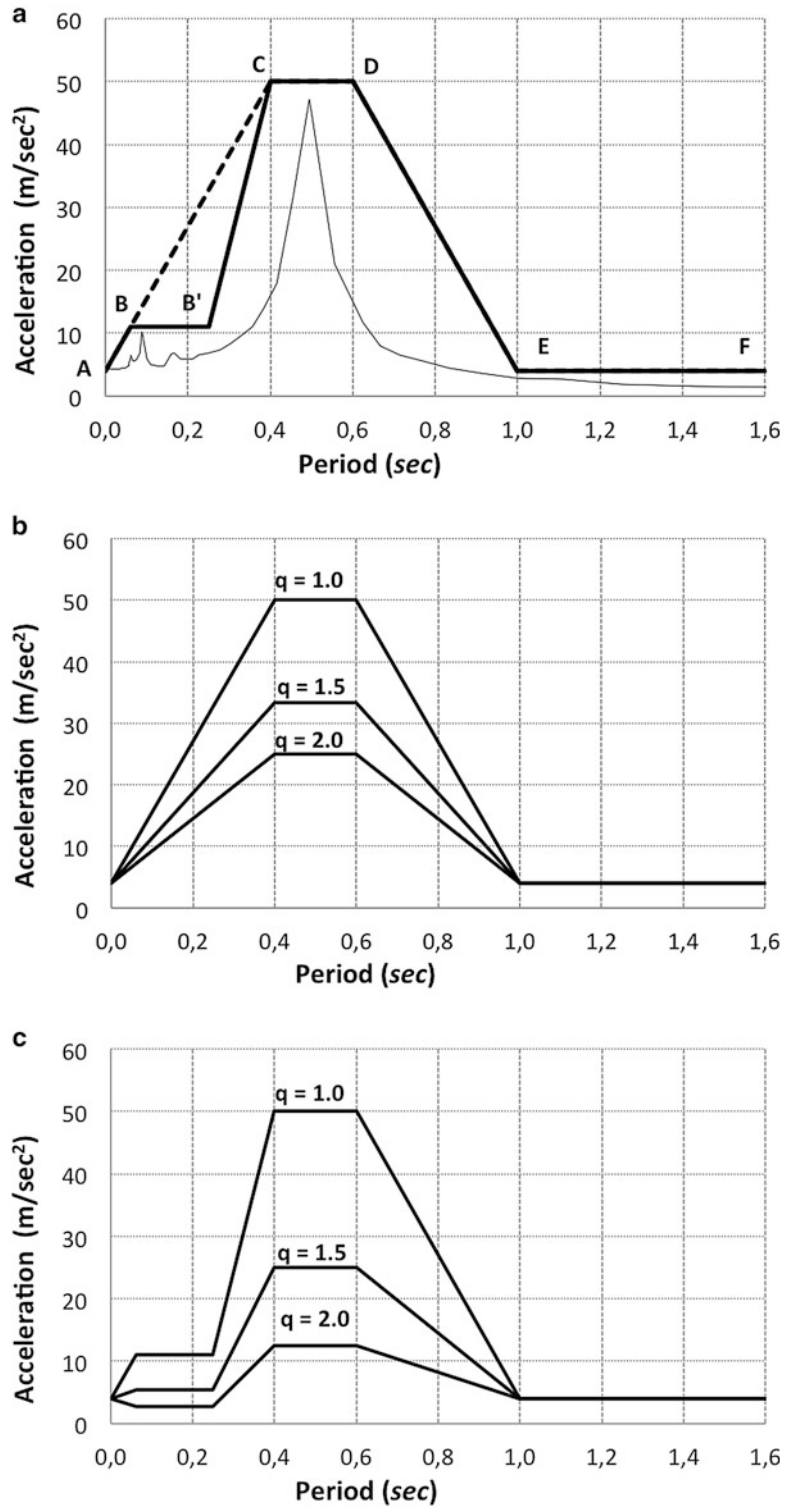
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Table 6 Equipment mass *(Ton)* and period *(sec)*

Equipment		Mass	Floor	Period
1		20	1	0.051
2		10	1	0.093
3,4,5		10	2	0.070
6		10	3	0.034
7		20	3	0.060
8		30	4	0.033
9,10		15	5	0.036

evaluated based on the previously described *cascade* procedure or on the predictive expressions already presented.

Time History-Based Evaluation of Equipment Accelerations

A detailed model of the PS system could include the secondary system modeled as reported in Fig. 13 (right): PS principal modal shapes are those reported in Fig. 14a. Alternatively equipments can be modeled by means of mass lumped at pertinent position of the floor as reported in Fig. 13 (center): PS principal modal shapes are those reported in Fig. 14b.

Both models have been analyzed by means of time histories, carried out (1) generating 7 accelerograms compatible with the acceleration spectra (PGA = 0.15 g) suggested in CEN (2008) for B soil and low-magnitude events (M < 5.5) (see Fig. 15), (2) considering a constant damping value of 2 % for the structural model and for the

equipments, (3) performing a dynamic modal TH analysis evaluating, for each accelerogram, the maximum absolute value of a given quantity (acceleration), and (4) averaging the maximum values obtained (at step 3) for each analysis.

If the detailed model with interaction (WI) is considered, the evaluated quantities are the mass accelerations of the single equipment.

Regarding to the lumped mass system, (1) for each considered accelerogram, an FRS has been generated (Fig. 16a): the FRS is relative to the acceleration of the top left floor corner (no sensible variations in FRSs have been observed if other floor points are considered); (2) having generated (for each floor) seven FRSs, the averaged FRS has been evaluated (Fig. 16b); and (3) for each equipment, depending on the floor it is attached to and its period (T_{NS}), the acceleration has been evaluated through the resulting FRS.

The results of the performed evaluation are reported in Table 7 where the acceleration of each floor and the acceleration of each equipment are reported, calculated with or without interaction.

Comparing the maximum acceleration evaluated considering the PS system with those evaluated through the cascade procedure, a significant difference (if the Equipment 2 is concerned) between the two approaches can be noticed: the difference is aspectable since the equipment period (0.093 *s*) is close to the period of the third modal shape (0.090 s) so that (see Fig. 16c) in a very small period range (in between 0.09 and 0.1 s), the acceleration ranges between 7.0 and 4.5 m/s – the already-mentioned *tuning* effect causes the equipment acceleration reduction if the complete PS system is analyzed in order to include the P-S interaction.

Analytical FRS-Based Evaluation of Equipment Accelerations

It has been already outlined that CEN proposal (CEN 2008) and FEM proposal (FEMA 369 2001) are based on a fixed value (5 %) of structural and equipment damping. So that only the proposal reported in AFPS (2007) and KTA (2012) will be considered in the following.
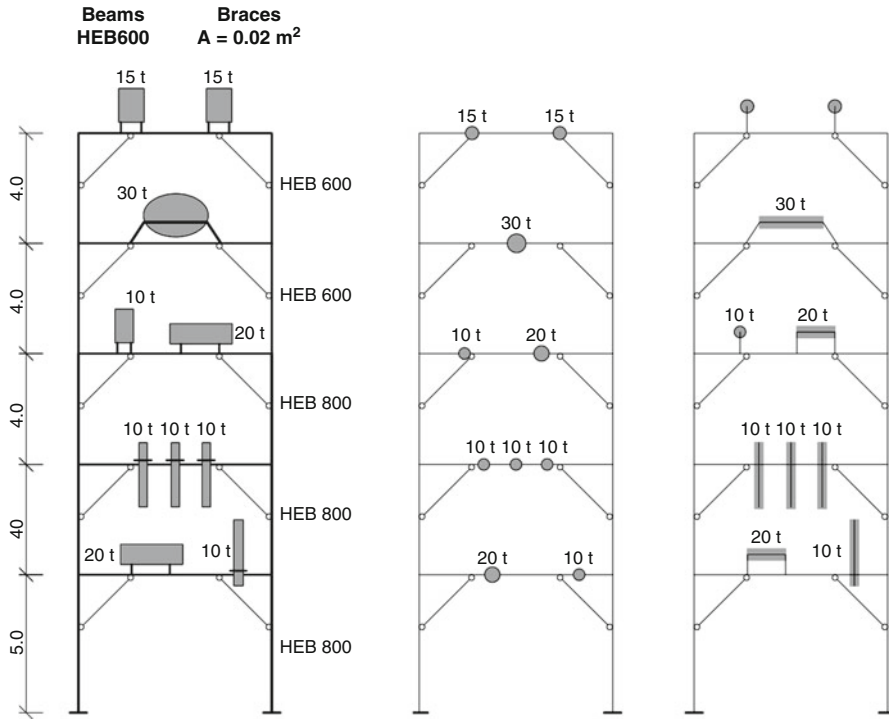
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 13 Case study: geometry (*left*) and models (*center*, *right*)
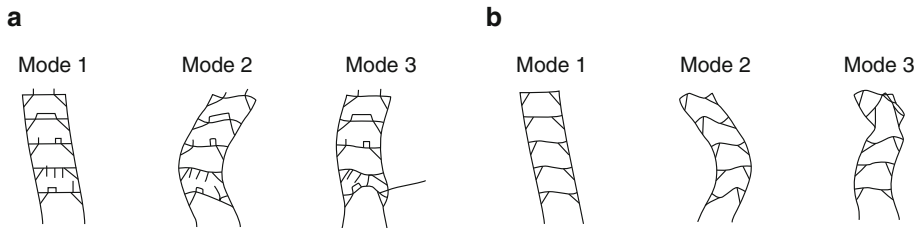
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 14 (**a**) Modal shape: direct modeling of the equipment. Periods (sec):

0.48, 0.158, 0.121. (**b**) Modal shape: equipment modeled as lumped masses. Periods (sec): 0.49, 0.159, 0.09

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 15 Target acceleration spectrum (PGA = 0.15 g) and spectra of the generated accelerograms
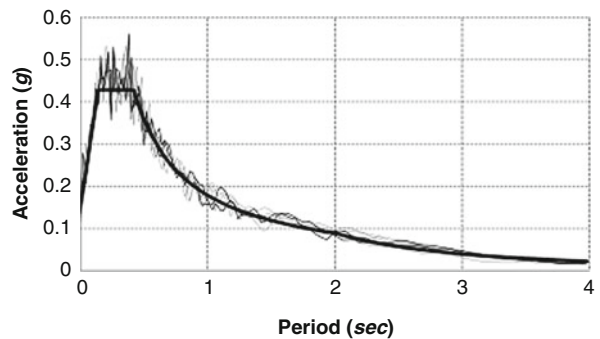
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 16 (**a**) Floor Response Spectra generations: scheme. (**b**) Generated Floor Response Spectra. (**c**) Generated Floor Response Spectra: 1° Floor; Period range 0.07–0.1 s
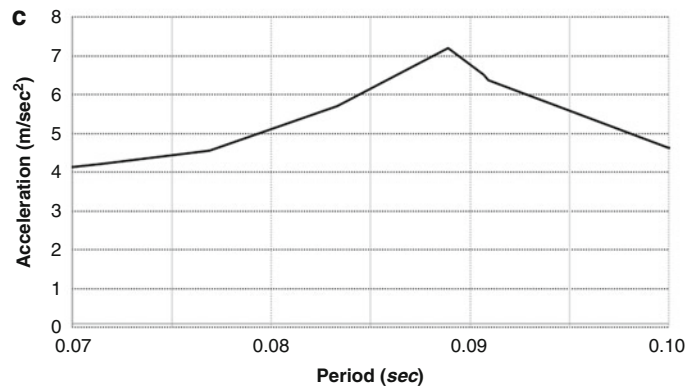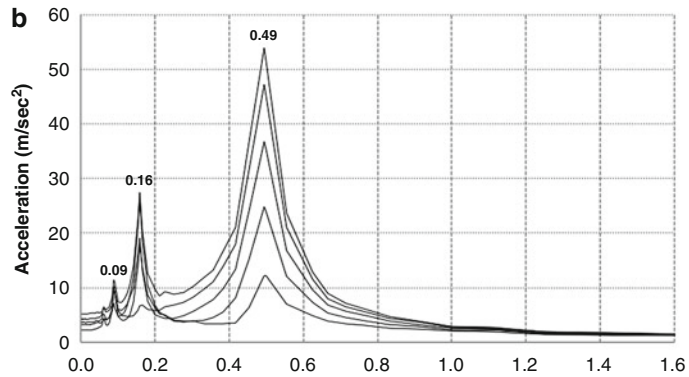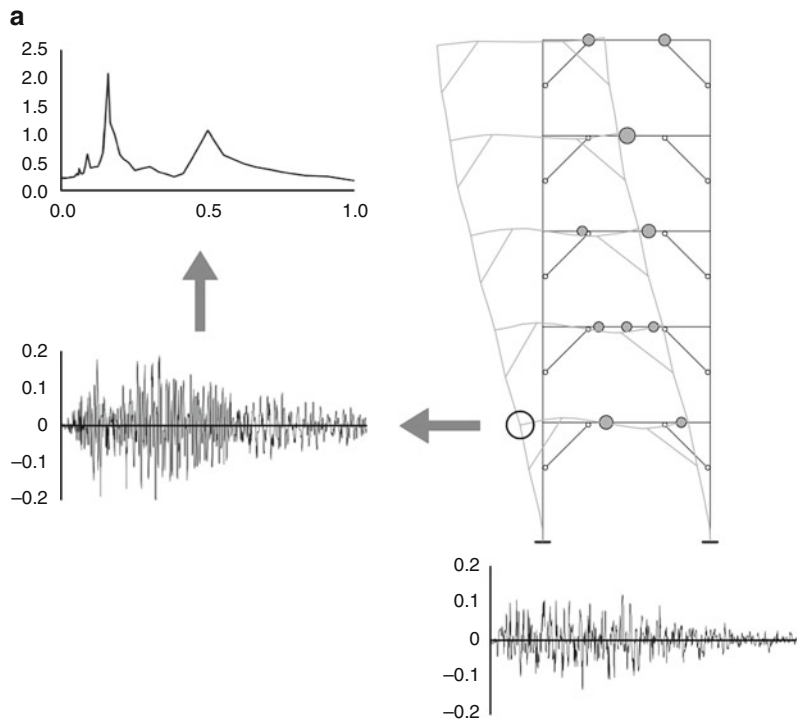
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Table 7 Floor acceleration (FA, m/sec^2) and equipment acceleration evaluated without (F_w/o) and with (F_w) interaction

N_E	N_F	Floor acceleration	FRS W/0	FRS W
1	1°	2.21	2.7	2.1
2	1°	2.21	6.5	3.2
3	2°	3.26	4.20	4.0
4	2°	3.26	4.20	4.0
5	2°	3.26	4.20	4.0
6	3°	3.65	3.70	3.8
7	3°	3.65	5.10	3.8
8	4°	4.28	4.30	4.3
9	5°	5.25	5.3	5.3
10	5°	5.25	5.3	3.3

It has to be specified that (1) concerning the KTA (2012) proposal, the floor acceleration evaluated by means of TH analyses (see Table 7) has been considered and (2) concerning the AFPS (2007) proposal, the FRSs have been evaluated by means of Eq. 9, evaluating the spectral acceleration corresponding to the first structural period (0.49 s) and a participation factor (Γ_S) equal to 1.6 that correspond to $\alpha = 1.5$ (see Eq. 10a).

It is clear (see Fig. 17a, b) that considering equipment periods close to the lowest structural periods, the KTA proposed expressions are more conservative while the AFPS expressions are less conservative (Tables 8 and 9).

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Fig. 17 (**a**) Generated FRSs. Acceleration versus period: (1) numerical simulation and (2) AFPS predictive equations evaluated based on Eq. 9 ($\alpha = 1.5$, $\Gamma_S = 1.6$, $R_S = 2.4$). (**b**) Generated FRSs. Acceleration versus frequency: numerical simulation and KTA predictive equations
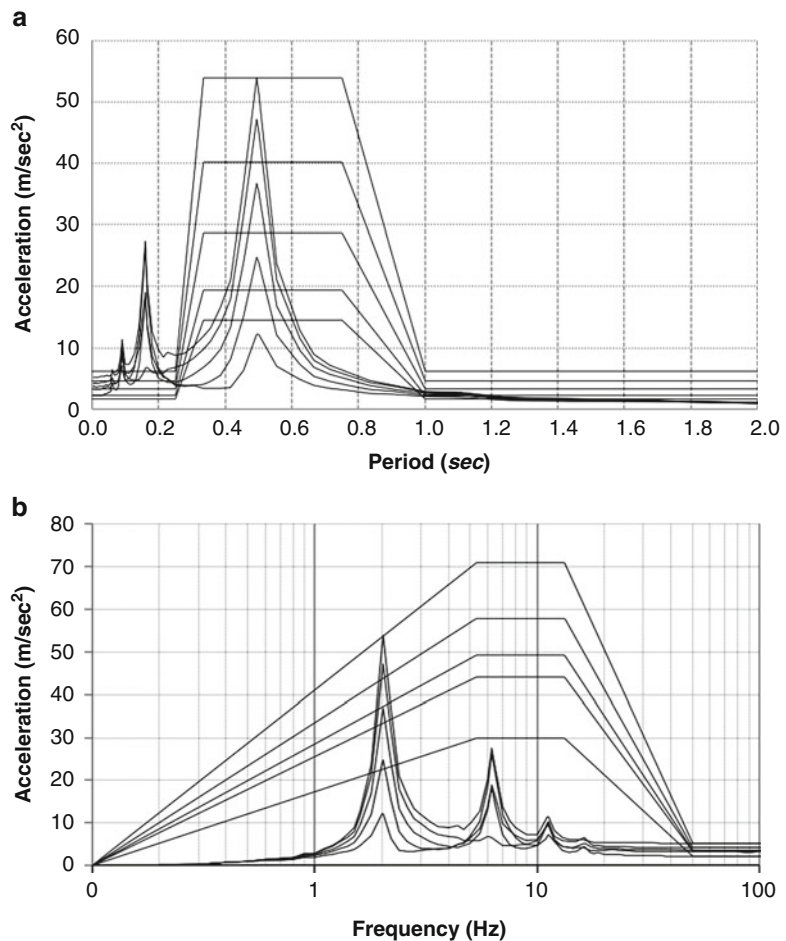
Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Table 8 Normalized floor displacement

Floor	MD	z/H
1°	0.21	0.24
2°	0.45	0.42
3°	0.68	0.62
4°	0.87	0.81
5°	1.00	1.00

Secondary Structures and Attachments Under Seismic Actions: Modeling and Analysis, Table 9 Equipment (Eq) acceleration

N_E	N_F	AFPS	KTA
1	1°	2.2	22.0
2	1°	2.2	34.0
3,4,5	2°	3.3	42.0
6	3°	3.7	22.0
7	3°	3.7	42.0
8	4°	4.3	25.0
9,10	5°	5.3	35.0

Summary

The chapter deals with the methodologies focused on the seismic analyses of the so-called *secondary* (sometimes *attachments*) elements that are part of a construction whose seismic resistance is delegated to a *primary* resistant structure.

Although secondary elements can be decontextualized from the primary resistant structures, they will be subjected to seismic action as well and, having their own structures, need to be modeled and analyzed by means of methods included in the general methodologies proper of seismic branch.

Among the methodologies usually adopted for secondary element analyses, the Floor Response Spectra (FRS)-based analyses become popular due to their recognized simplicity.

FRSs provide acceleration (consequently velocity and displacement) to which the secondary element (with a given period and damping) will be subjected to when *attached* (from which the alternative name *attachments* derives) to a given part of the structures such as a building floor (from which the name Floor Response Spectra derives).

Given that FRS generation could require onerous numerical analyses, simplified expressions are proposed in literature and discussed in the following together with the general methodologies tailored to secondary element modeling and seismic analyses.

Cross-References

▶ Behavior Factor and Ductility
▶ Building Codes and Standards
▶ Classically and Nonclassically Damped Multi-degree of Freedom (MDOF) Structural Systems, Dynamic Response Characterization of
▶ Code-Based Design: Seismic Isolation of Buildings
▶ Earthquake Risk Mitigation of Lifelines and Critical Facilities
▶ Equivalent Static Analysis of Structures Subjected to Seismic Actions
▶ European Structural Design Codes: Seismic Actions
▶ Modal Analysis
▶ Nonlinear Dynamic Seismic Analysis
▶ Performance-Based Design Procedure for Structures with Magneto-Rheological Dampers
▶ Response Spectrum Analysis of Structures Subjected to Seismic Actions
▶ Seismic Fragility Analysis
▶ Soil-Structure Interaction
▶ Spatial Variability of Ground Motion: Seismic Analysis
▶ Tuned Mass Dampers for Passive Control of Structures Under Earthquake Excitations

S

References

AFPS (2007) Guide methodologique pour la conception, l'installation et le diagnostic des equipment en zone sismique. Association Francaise de Genie Parasismique, Paris

ASCE 7-02 (2002) Minimum design loads for buildings and other structures. American Society of Civil Engineers, Reston

ATC 58 (2012) Seismic performance assessment of buildings. Applied Technology Council, Redwood City

ATC/BSSC (1997) NEHRP commentary on the guidelines for the seismic rehabilitation of buildings. Report no FEMA 274, prepared by the Applied Technology Council for the Building Seismic Safety Council, published by the Federal Emergency Management Agency, Washington, DC

CEN (2008) Eurocode 8: design of structures for earthquake resistance, part 1: general rules, seismic actions and rules for buildings. Comité Européen de Normalisation, Brussels

Chaudhri SR, Villaverde R (2008) Effect of building nonlinearity on seismic response of nonstructural components: a parametric study. J Struct Eng 134(4):661–670

Chen Y, Soong TT (1988) State-of-the-art review seismic response of secondary systems. Eng Struct 10:218–228

FEMA 369 (2001) NEHRP recommended provisions for seismic regulations for new buildings and other structures: part 2-commentary. Building Seismic Safety Council, National Institute of Building Sciences, Washington, DC

FEMA 74 (2005) Earthquake hazard mitigation for nonstructural elements. Field Manual, Emergency Management Agency, Washington, DC

Gilani AS, Whittaker AS, Fenves GL, Fujisaki E (1999) Seismic evaluation of 55° kV porcelain transformer. Pacific Earthquake Engineering Research Center, Berkeley

(IEEE) (2005) Recommended practice for seismic design of substations, Institute of Electrical and Electronics Engineers, New York

Igusa T, Der Kiureghian A (1985a) Dynamic characterization of two degree-of-freedom equipment-structure systems. J Eng Mech ASCE 111(1):1–19

Igusa T, Der Kiureghian A (1985b) Dynamic response of multiply supported secondary systems. J Eng Mech ASCE 111(1):20–41

Kircher CA (2003) It makes dollars and sense to improve nonstructural system performance. In: Proceedings of ATC 29–2 seminar on seismic d, performance, and retrofit of nonstructural components in critical facilities, Newport Beach, 23–24 Oct 2003

KTA (2012) Safety standards of the nuclear safety standards commission: design of nuclear power plants against seismic events; Part 4: components. KTA-Geschaeftsstelle c/o BfS, Willy-Brandt-Str. 5, 38226 Salzgitter

Taghavi S, Miranda E (2003) Response assessment of nonstructural building elements. Pacific Earthquake Engineering Research Center, PEER 2003/05 Sept

Villaverde R (1997) Seismic design of secondary structures: state of the art. J Struct Eng 123(8):1011–1019

Seismic Accelerometers

Gerardo Alguacil[1] and Jens Havskov[2]
[1]Instituto Andaluz de Geofísica, University of Granada, Granada, Spain
[2]Department of Earth Science, University of Bergen, Bergen, Norway

Synonyms

Broadband seismometers; Force-balance accelerometers; MEMS accelerometers; Passive seismometers; Seismic recorders; Servo-accelerometers; Strong ground motion

Introduction

Earthquake ground motion ranges in amplitude from a few nanometers (e.g., for a distant earthquake or a local microearthquake) to several meters close to the fault causing a big quake. It is very difficult to record that wide range of signals with a single type of instrument. Seismometers are designed to be very sensitive for detecting weak signals of ground motion and have a response proportional to ground velocity in a frequency band typically from 0.01 to 100 Hz. On the other hand, strong ground motion instruments are in general less sensitive and may manage on scale large ground motion amplitudes. This strong ground motion may reach peak accelerations above 2 g (g is gravity acceleration) and peak velocities higher than 3 m/s (Anderson 2010). The motion experienced at specific points in a building or structure may be even higher. Seismic accelerometers were developed for recording on scale vibrations up to such range of amplitude either on ground or in structures.

Due to instrumental and practical reasons, the preferred motion measure for strong ground motion is acceleration, since the inertial force acting on a structure due to a seismic action is proportional to ground acceleration. Nevertheless, some strong-motion velocimeters are commercially available, but their use is much less widespread.

Weak motion from local small earthquakes or moderate to strong earthquakes at regional or large distances are studied fundamentally on seismograms, proportional to ground velocity in a wide frequency band. However, even for strong motion, ground velocity and displacement are also of interest: Ground velocity is directly related to important parameters such as shake energy or soil liquefaction potential and is linked to damage on intermediate period structures and buried pipelines (e.g., Akkar and Bommer 2007). Ground displacement is significant for large structures, where differential displacements may cause damage, and from a seismological view, it is proportional to seismic moment, a source parameter defining the earthquake size. Therefore, it is desirable that velocity and displacement may be estimated from acceleration records.

Historical Overview

This section partly follows the review article by Trifunac (2009). See also "▶ Historical Seismometer."

Earlier attempts to record strong ground motion were made with the so-called seismoscopes (Fig. 1), most of which were limited to record the horizontal ground motion on a (curved) surface without a time reference. These instruments consisted of a pendulum, either normal or inverted, whose mass motion was drawn on a curved surface (usually a smoked glass) by a stylus.

As a result of the program of strong-motion earthquake investigation of the US Coast and Geodetic Survey and cooperative institutions (McComb and Ruge 1937), the first instrument designed to record the strong ground motion versus time was developed in the decade of 1930.
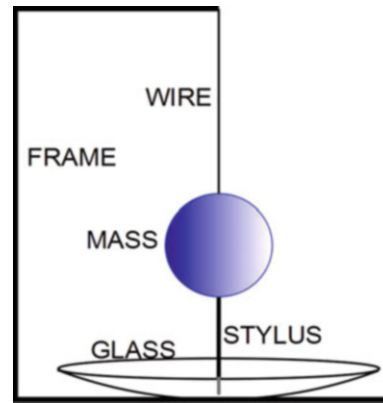
Seismic Accelerometers, Fig. 1 Simplified schematics of a seismoscope. A mass is suspended from the frame by a wire, and a stylus writes the relative motion on a spherical cap surface

This instrument was built with a mass suspended on torsion wires, based on the Wood-Anderson seismometer principle (known because it was used by Richter to define his magnitude scale). Damping was achieved – like for the Wood-Anderson seismometer – using a conductive copper mass in which the field of a permanent magnet induced parasitic currents. A more robust pivoted suspension and a spring were later incorporated. Recording was made on photographic paper on a drum, with a "starter" system built with an independent undamped pendulum closing an electric contact when a strong motion occurred. A clock mechanism interrupted the light beam to produce time marks on the record. The first strong ground motion accelerograms were obtained on three stations in 1933 from the Long Beach earthquake (McComb and Ruge 1937).

With analog records and the computational means available at that epoch, it was very difficult to estimate displacement from acceleration. Therefore, some seismometers with unity amplification and photographic recording were used to obtain displacement records (see "▶ Passive Seismometers").

A remarkable effort was made to test these first instruments with a shaking table to assess the independent motion of the components, zero shifts (instability of the mass rest position),

S

parasitic vibrations (instrument vibration modes different from fundamental), the ability to calculate velocity and displacement from acceleration records, and other features (McComb and Ruge 1937, Ruge and McComb 1943).

It was not until early 1960 that accelerographs become commercially available. Around that time the digital computation also made it possible to calculate the velocity and displacement histories from the corresponding accelerograms. In this decade, some seismic accelerometers were still of torsion wire or inverted pendulum types, but the first servo-accelerometers using the concept of force balance appeared (Reasemberg 1969; Eisenberg and McEvilly 1971). This concept (see later) is to apply a feedback force on the mass to keep it at rest relative to the ground. So the applied force is the mass times the ground acceleration. The main advantages are improved linearity and dynamic range (i.e., the relation between the maximal non-clipped signal amplitude and the minimal resolvable amplitude; see later). Recording was analog on photographic paper, but some prototypes used digital recording (Reasemberg 1969). The first digitizers used, however, had a too low resolution to match the dynamic range of servo-accelerometers.

Following the period of photographic recording, some models adopted analog recording on magnetic tape using frequency modulation, but digital recording was soon the norm due to its better characteristics. Presently all high-performance seismic accelerometers use force-balance sensors and high-resolution digital recording.

Basic Principles of Accelerometers

Seismometers and accelerometers are both based on the inertia principle.

In Fig. 2, a schematic view of the well-known passive damped oscillator is drawn. A mass m is fixed to one end of a spring of elastic constant k. The other end of the spring is fixed to a frame anchored to the ground. The motion is mechanically limited to the horizontal axis shown. When the ground moves a displacement
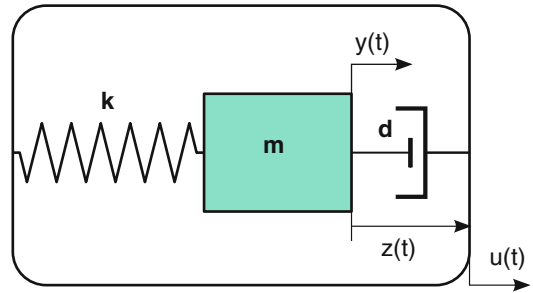
Seismic Accelerometers, Fig. 2 A schematic single degree of freedom oscillator. A spring makes a force proportional and opposed to the relative displacement of the mass $z(t)$, and a dashpot damps this motion with a force proportional to the relative velocity. The ground (and the frame) displacement is $u(t)$

$u(t)$ with respect to some inertial frame, the mass displacement is $y(t)$ in this inertial frame, and $z(t) = y(t) - u(t)$ is the mass displacement relative to the instrument casing or the ground. A dashpot represents a damping device (usually magnetic) that acts on the mass with a force proportional to its relative velocity. This damping avoids the oscillation of the system with its free period when it is excited. The two real forces acting on the mass are the spring force and the damping force, so the dynamics equation is (Havskov and Alguacil 2010)

$$-k \cdot z - d \cdot \dot{z} = m \cdot y(t) = m[\ddot{u}(t) + \ddot{z}(t)] \quad (1)$$

It is useful to write the equation as a function of the two coefficients experimentally measurable. The free or natural angular frequency ω_0 can be defined as

$$\omega_0 = \sqrt{\frac{k}{m}} = 2\pi f_0 = \frac{2\pi}{T_0} \quad (2)$$

where T_0 is the suspension-free period and its inverse f_0 is the natural frequency. The damping coefficient or damping fraction h is

$$h = \frac{d}{2\omega_0 m} \quad (3)$$

Rearranging Eq. 1 then gives

$$\ddot{z}(t) + 2h\omega_0 \cdot \dot{z}(t) + \omega_0^2 \cdot z = -\ddot{u}(t) \quad (4)$$

Both ω_0 and h can be measured, e.g., from the system transient response or by exciting the system with steady-state harmonics of frequencies in a band including f_0. For low frequencies, the mass velocity and acceleration terms in Eq. 4 are small, so mass displacement z is then proportional to the ground acceleration $\ddot{u}$.

The system response may be characterized in the Laplace transform domain. Let $Z(s)$ and $U(s)$ be the Laplace transforms of $z(t)$ and $u(t)$. By assuming null velocities and displacements at initial time, the Laplace transform of Eq. 4 becomes

$$s^2 Z(s) + 2h\omega_0 s Z(s) + \omega_0^2 Z(s) = -s^2 U(s) \quad (5)$$

The transfer function between the mass relative motion Z and the ground acceleration $s^2 U$ is then

$$T_a(s) \equiv \frac{Z(s)}{s^2 U(s)} = \frac{-1}{s^2 + 2h\omega_0 s + \omega_0^2} \quad (6)$$

The explicit frequency response function is obtained by substitution of s by $i\omega$ in Eq. 6, that is,

$$H_a(\omega) = \frac{-1}{-\omega^2 + 2h\omega_0 \cdot i\omega + \omega_0^2} \quad (7)$$

For a sinusoidal motion of arbitrary frequency $f = \omega/2\pi$, the modulus of this complex function represents the amplitude relation between the mass relative motion $z(t)$ and the ground acceleration $\ddot{u}(t)$, and its phase is the relative phase between them. Both are plotted in Fig. 3 as functions of frequency. The amplitude response is flat for acceleration up to the corner frequency f_0 and decays at higher frequencies as f^{-2}. The frequencies of interest are usually in the flat zone, so the instrumental correction for amplitude is simply a constant factor. Observe, nevertheless, that the phase response deviates from flatness even for frequencies well under the natural frequency. This response function is formally like the response of a second-order low-pass filter with cutoff frequency $f_0 = \omega_0/2\pi$.

Seismic Accelerometers, Fig. 3 *Top*: amplitude response of a mechanical accelerometer for ground acceleration. *Bottom*: phase response. The free period is $T_0 = 0.02$ s ($f_0 = 50$ Hz) and $h = 0.70$
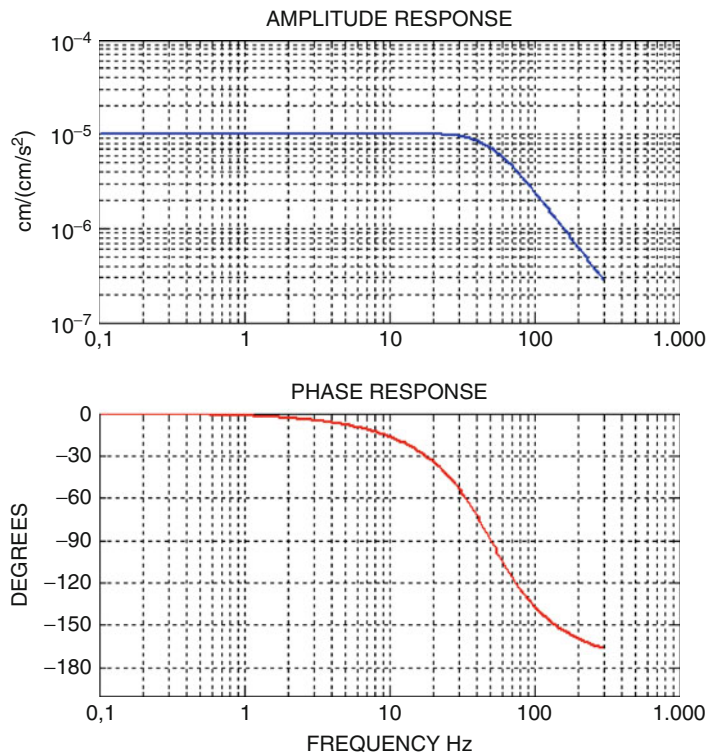
Displacement Transducer

The mass motion has to be measured and recorded by some device. In early accelerometers, a light beam was reflected in a mirror which rotated with the mass motion. Presently, almost all seismic accelerometers use a capacitive transducer, which gives a voltage output proportional to the mass displacement. This type of transducer is very sensitive – it can resolve displacements of the order of pm (10^{-9} mm).

Two types of capacitive transducers may be used: variable gap or variable area. The most used arrangement, Fig. 4, is a pair of capacitors with a common central moving plate and two fixed plates or vice versa (variable gap). When the central plate moves with the mass relative to the fixed plates, one of the capacities increases and the other decreases. The same happens with the variable area type. An identical sinusoidal or square signal with frequency of several kHz is fed to each capacitor with opposite sign. The signal amplitude at the common point of both capacitors is proportional to the capacitance

difference, and the phase depends on the sign of this difference. This may be demodulated with a phase-sensitive demodulator (PSD) circuit. While this is the most used technique, other approaches are possible.

The output of the detector is proportional to the capacitance difference. Let A be the overlapping area between plates and d the distance (gap) between them. If d is much smaller than the plate dimensions, the capacitance of a parallel plane capacitor is

$$C = \frac{\varepsilon \cdot A}{d} \qquad (8)$$

where ε is the dielectric permittivity (for air $\varepsilon = 8.85 \cdot 10^{-3}$ pF/mm). In the variable gap transducer, for a displacement x of the central plate, it may be easily shown that the capacitance difference is

$$C_1 - C_2 = \varepsilon \cdot A \cdot \frac{2x}{d^2 - x^2} \qquad (9)$$

which is not linear with x.
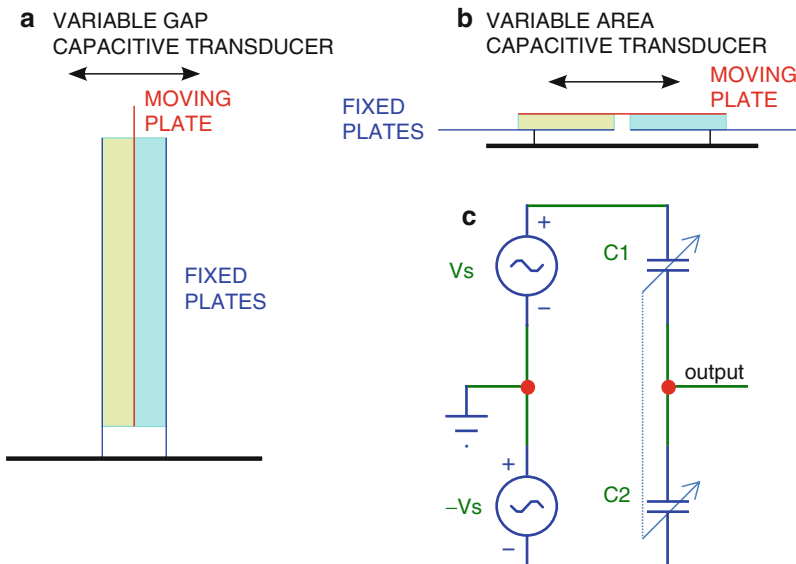
Seismic Accelerometers, Fig. 4 Schematics of variable capacitance displacement transducers. (**a**) Variable gap. As the central plate (*red*) moves right, the gap with the left fixed plate (*blue*) increases (capacitance of left capacitor, *yellow*, decreases), and the gap with the right plate (*blue*) decreases (capacitance of the right capacitor, *light blue*,

increases). (**b**) Variable area. The effective area of each capacitor (same colors as in **a**) varies in opposite sense when the central plate (*red*) moves horizontally. The gap is not at scale, since in practice it is very narrow in relation to plate dimensions. (**c**) Circuit schematics, see text

Seismic Accelerometers, Fig. 5 Normalized sensitivity of capacitive transducers of variable gap (*blue*) and variable area (*red*). A *straight line* is plotted for reference (*dotted*) with the slope of the variable gap at the origin
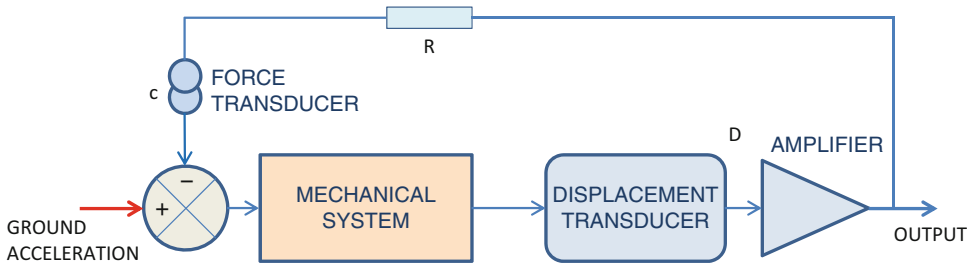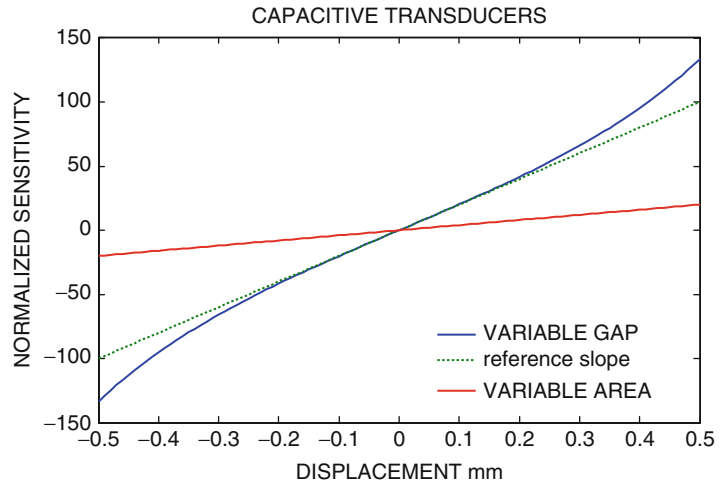
Seismic Accelerometers, Fig. 6 Block schematic of an FBA

In a variable area transducer with plate overlap dimensions $a \times b$, the capacitance difference is linear with the displacement x:

$$C_1 - C_2 = \frac{\varepsilon}{b} \cdot 2b \cdot x \qquad (10)$$

In Fig. 5, these functions – excluding ε – are plotted for an overlapping area of $A = 100$ mm^2 and gap distance of d $= 1$ mm and 0.5 mm for variable gap and variable area, respectively. Two features are clear: variable gap type is more sensitive but nonlinear, and variable area type is less sensitive but linear.

Variable gap transducers are often used in feedback systems, like force-balance accelerometers, since the mass displacement is held very small – within the linear zone – by the servo control.

Variable area transducers are used in some micro-electro-mechanical (MEM) accelerometers (see later) without feedback.

The Force-Balance Accelerometer (FBA)

Most present seismic accelerometers are of the force-balance type. The principle is to apply a feedback force on the inertial mass opposed to its motion, in such a way that this motion is reduced to a minimum. A simplified block schematic is shown in Fig. 6.

The ground acceleration produces an inertial force on the mass, like in the basic passive accelerometer of Fig. 1. The relative mass motion z is sensed with a displacement transducer, whose output is amplified, with a total factor D (volt/meter, V/m). The amplifier output v_o is applied
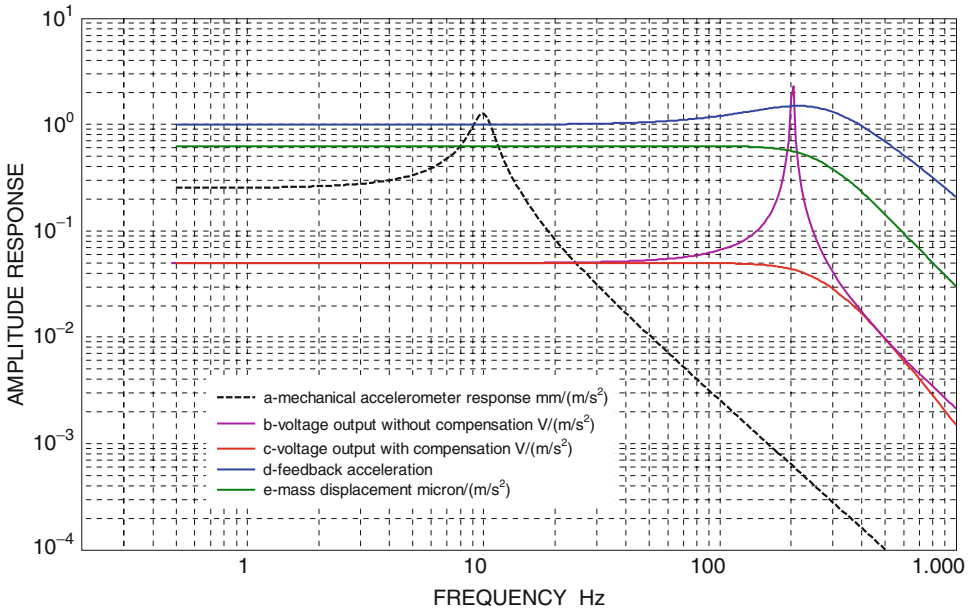
Seismic Accelerometers, Fig. 7 Several responses relative to ground acceleration. See text. Note the different units for each curve

through a resistor R to a force transducer (normally a coil-magnet system) with constant c (newton/ampere, N/A). The feedback force $-c \cdot v_o/R = -c \cdot A \cdot z /R$ is then applied to the mass. Equation 1 has now to be modified to

$$-k \cdot z - d \cdot \dot{z} - \frac{Dc}{R} \cdot z = m[\ddot{u}(t) + \ddot{z}(t)] \quad (11)$$

It is clear that the effect of feedback is to increase the effective stiffness of the system from k to $k + D \cdot c/R$. Arranging terms of Eq. 11 to get the form of Eq. 4, it is found that with feedback, a new effective "free" angular frequency is ω_f

$$\omega_f = \sqrt{\frac{k + D \cdot c/m}{m}} = \sqrt{\omega_0^2 + \frac{D \cdot c}{R \cdot m}} \quad (12)$$

which is higher than the "open-loop" free angular frequency ω_0.

On the other hand, the new damping coefficient with feedback h_f is decreased to

$$h_f = h\frac{\omega_0}{\omega_f} \quad (13)$$

This is not desirable, since a low damping will produce a ringing transient response or even an unstable system. The frequency response (Fig. 7) has the same form as Eq. 7, but in this case it will show a resonance peak.

Actually, the phase-sensitive demodulator (PSD) associated with the displacement transducer will often include a low-pass filter at frequencies above the seismic band, thus introducing an additional phase delay in this band.

Some additional circuitry is then required to improve the transient response and prevent a possible self-oscillation of the servo system. Several techniques for this are possible. A simple one consists of including a second feedback loop to add a small force proportional to the time derivative of the mass displacement, i.e., the mass velocity. This allows the feedback force to control the mass motion "in advance," and the servo loop is stabilized.

Figure 7 shows the responses relative to ground acceleration of several systems and output points of the systems. An under-damped mechanical accelerometer (without feedback

force) is drawn in curve *a*, in this example with a free oscillation frequency of 10Hz. Curve *b* corresponds to an FBA built with this system by adding a displacement transducer and a feedback force proportional to the mass displacement, which increases the resonance frequency. The damping is still too low and a strong resonance peak at 200 Hz appears. Curve *c* shows the voltage output response for the same FBA once stabilized by the addition of a small feedback force proportional to the mass velocity. Curve *d* shows the feedback acceleration response (i.e., the feedback force per unit mass), which is nearly equal to the ground acceleration (so the amplitude response is 1) up to the frequencies around the resonance, where the feedback force needs to be a little higher to stabilize the motion. Curve *e* plots the mass motion response measured in microns (μm) relative to ground acceleration in m/s^2. It is approximately 0.6 in the useful band, so a ground acceleration of 1 g will produce a mass motion of $9.8 \text{ m} \cdot \text{s}^{-2} \cdot 0.6 \, \mu\text{m}/(\text{ m} \cdot \text{s}^{-2}) \approx 6 \, \mu\text{m}$, a very small displacement, for which the capacitive transducer will be quite linear.

It is also illustrative to view the transient response of the FBA (Fig. 8). The input is a simulated ground acceleration pulse of 1 m/s^2 and with duration of 50 ms. It can be seen that the output voltage has only a small overshot, due to the compensation circuit in the feedback. So, there is an unavoidable small distortion affecting the higher frequencies, like any instrument with a limited bandwidth, but this will not affect the band of interest. The feedback acceleration has a little overshoot to counteract the resonance and keep the mass as stationary as possible. The mass displacement is also shown (with sign changed) and has an amplitude about 0.6 μm, as predicted by the frequency response in Fig. 7.

The FBA principle is the basis not only of servo-accelerometers but also of ▶ broadband seismometers (BB). These instruments have a flat frequency response for ground velocity in a band from ~0.01 to several tens of Hz. The feedback loop keeps the mass almost at rest relative to the ground, but the signal is integrated in the loop, and the output is taken from a point where it is proportional to ground velocity in a wide band. For more details, see the entry "▶ Principles of Broadband Seismometry."

Characteristic Parameters of Accelerometers

The performance of a seismic accelerometer may be characterized by several parameters.

Sensitivity: The relation between the voltage output and the input acceleration in V/m/s^2 or V/g (gravity normal acceleration = 9.807 m/s^2). Typical values are 0.5–5 V/g.

Input Full Scale (FS): The maximum ground acceleration before the instrument gets saturated. A large FS prevents saturation with strong ground

Seismic Accelerometers, Fig. 8 Transient response of the example FBA (Fig. 7) for a pulse in acceleration input of $1 \text{ m} \cdot \text{s}^{-2}$
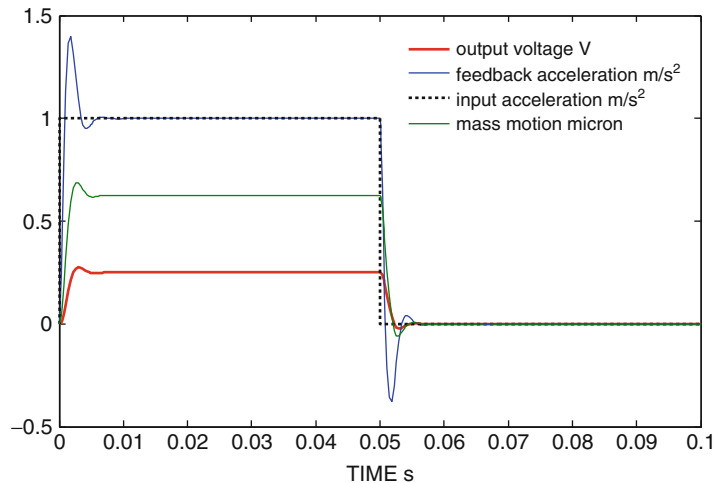
motion but at a cost of sensitivity. In general, for accelerometers to be installed in zones where large ground motion is expected, FS should be 2–4 g, but in low-moderate seismicity areas, 0.5–1 g would be suitable. Nevertheless, it should be kept in mind that some of the largest peak ground motions ever recorded (up to 4.36 g for Iwate-Miyagi 2008 earthquake; Yamada et al. 2009) have been produced by moderate-size earthquakes (Anderson 2010). Some accelerometers allow user setting its FS.

Resolution: It is the minimum acceleration amplitude that it may measure. In early accelerometer digital recorders, it was limited by the digitizer (A/D converter) resolution. In modern instruments, the digitizer is usually 24 bit, and the resolution is related to the self-noise level of both the accelerometer and the digitizer. So it is usually given as self-noise level, or it may be obtained from the dynamic range. Good-quality accelerometers have noise levels under 1 μm/s^2 *rms* (root mean square).

Dynamic Range: The relation (expressed in decibels, dB) between the maximum acceleration amplitude (FS) and the resolution. A ratio between two amplitudes a_1 and a_2 expressed in dB is $20 \cdot \log(a_1/a_2)$; for energy or power ratios, a factor 10 instead of 20 is used. The best accelerometers now have a dynamic range up to more than 150 dB, which means a ratio $10^{150/20} > 30 \cdot 10^6$, but most instruments achieve 120–140 dB.

MEMS Accelerometers

The same principles described so far are valid regardless the size of the accelerometer, but accelerometer sensors of micrometric size have specific characteristics, so a brief description follows.

Originally, the aim of reducing the size of accelerometer sensors was due to the manufacturing of integrated accelerometers for industry and navigation (airborne) applications. The space exploration requirements, among other applications, led to the adaptation of some of the techniques of manufacturing integrated circuits to the production of very lightweight and low-power-sensitive accelerometers.

In recent years, the industry has found a large number of consumer applications for this kind of accelerometers. This has led to the development of micro-sized, light and low-power sensors, suitable for mass production and thus inexpensive. These are the *micro-electro-mechanical system* (MEMS) devices. MEMS accelerometers are used for automotive industry (e.g., for air-bag triggering), inertial navigation, medical applications, game consoles, cellular phones, tablets and laptops, etc. Most of these devices have a poor performance for seismic recording, but a few commercial models are suitable for low-resolution strong ground motion recording and earthquake early warning (EEW) systems or for triggering shutoff procedures on critical facilities (trains, gas valves, lifts, critical machinery) in case of a strong earthquake (e.g., Yanada et al. 2002). Furthermore, some seismic instrumentation companies have developed MEMS accelerometers for seismic-grade recorders with comparable performance of conventional sensors. The idea is to build the mass-spring system within the multilayer structure of an integrated circuit (Fig. 9), using the same techniques as for electronic components and circuits.

The internal noise of a spring-mass system is due to the thermal-mechanical coupling by means of the damping dissipative forces. For instance, the air molecules interchange kinetic energy with the mass, which then "dances" randomly with a Brownian motion.

It may be shown (e.g., Aki and Richards 2002; Havskov and Alguacil 2010) that this noise is proportional (Eq. 14) to the mechanical damping coefficient h (not included the feedback effect) and inversely proportional to the suspended mass. The power spectral density (PSD) $\ddot{Z}_n^2$ of the mass acceleration noise (the PSD integrated between two frequencies gives the noise variance in this band) is

$$\ddot{Z}_n^2 = \frac{8 k_B T}{m} h \omega_0 \qquad (14)$$
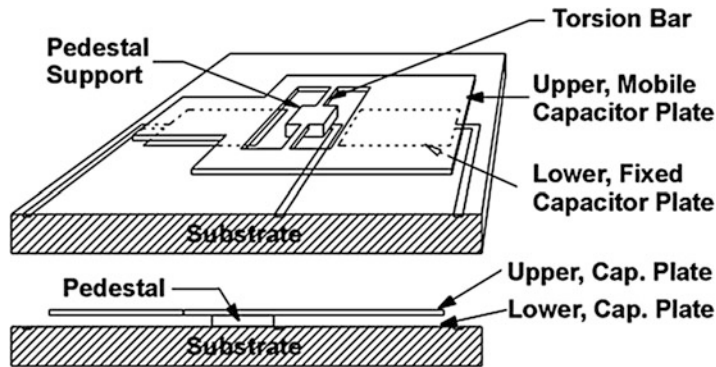
Seismic Accelerometers, Fig. 9 Principal elements of a MEMS (micro-electro-mechanical system) accelerometer with capacitive transducer. The mass is the *upper* mobile capacitor plate which can rotate around the torsion bars. The displacement, proportional to acceleration, is sensed with the variance in the capacitance. For high-sensitive applications, a feedback circuit is added which controls a restoring electrostatic force, thus making an FBA. The size of the sensor above is about 2 mm (Figure from www.silicondesigns.com/tech.html)

where k_B is the Boltzmann constant ($1.38 \cdot 10^{-23}$ J/K), T is Kelvin temperature, m the suspended mass, h the mechanical damping coefficient, and ω_0 the free angular frequency. This equation explains why a good part of the effort in improving MEMS accelerometer noise is focused (Walmsley et al. 2009) in increasing the proof mass and reducing damping.

Thus, for a micro mass, the open-loop damping should be kept as low as possible for an acceptable noise level at ambient temperature. High-purity silicon spring elements are almost perfectly elastic, and the spring is operating in vacuum to avoid air damping, so the mechanical damping can be held quite low.

Several strategies have been tested in prototypes for sensing the mass motion. Most commercial devices use some kind of variable capacitance. Other sensing designs are based on tunnel effect, optical diffraction and interferometry, piezo-resistivity, piezoelectricity, resonance, or electrochemical phenomena. Very small capacitive transducers are not practical in general, since the circuitry parasitic capacitances make the relative changes too small. Nevertheless, these parasitic capacitances may be reduced to a minimum in MEMS devices, since the associated circuit is integrated within the same chip (see, for instance, Li et al. 2001).

The two capacitive techniques described above are used: variable gap between plates and variable area. The first is quite nonlinear and usually requires electrostatic feedback to linearize the transducer response. Recently, electromagnetic feedback has also been proposed for MEMS accelerometers (Dwyer 2011). On the other hand, the variable area technique is more linear, and some devices using it operate in open-loop mode (e.g., Homeijer et al. 2011). Figure 10 shows a possible arrangement for a MEMS accelerometer.

MEMS sensors are now made with electrochemical sensors (Deng et al. 2013) with a noise PSD of 0.1 $(\mu m/s^2)/\sqrt{Hz}$ or 10 ng/$\sqrt{Hz}$ at 1 Hz, so these sensors could be used at medium noisy sites.

All of these devices have two distinct parts within the same package: the mechanical sensor and the associated electronics, typically implemented by an ASIC (application-specific integrated circuit). Some commercial models have a sigma-delta modulator (a device yielding a sequence of logical pulses whose average value is proportional to its analog voltage input) included (see "▶ Recording Seismic Signals"), and so its output is a pulse-density digital signal, able to be directly interfaced to a microcontroller or computer. The integrated MEMS has to be

Seismic Accelerometers, Fig. 10 A possible arrangement of the mass, spring, and capacitor plates in a MEMS accelerometer. All elements are in a layer a few microns thick
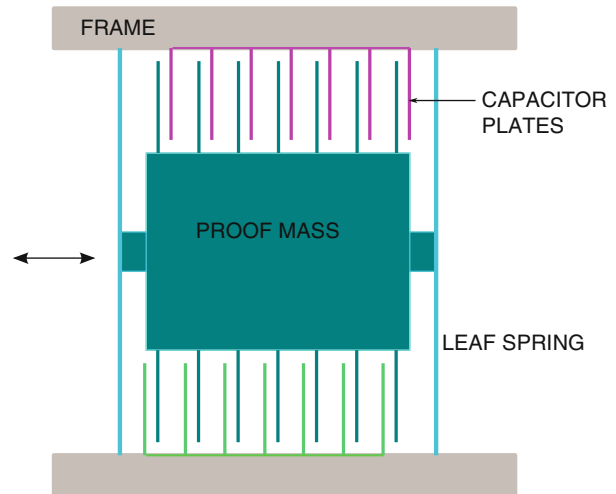
mounted in a printed circuit board (PCB) with some more electronics, at least a power supply unit, and optionally assembled in a casing with connectors, able to be fixed to another element. This assembly has to be rigid enough not to introduce parasitic resonances in the seismic band. Including electronics and housing, the total weight of these sensors might be around 0.1–0.5 kg. The sensor chip itself typically weights less than 1 g.

The (US) Working Group on Instrumentation, Siting, Installation, and Site Metadata of the Advanced National Seismic System (ANSS 2008) defined four classes of strong-motion stations – A, B, C, and D – in terms of performance. Class A has the highest performance, with resolution better than 7 μg and broadband dynamic range > 111 dB (≥20bit). Class B has a resolution between 7 and 107 μg and a dynamic range 87–111db (or 16–20bit). Class C resolution is 107–1,709 μg and dynamic range 63–87 dB (12–16bit). Class D has poorer performance than C. Currently (2014) there are no commercial MEMS accelerometers that fulfill the class A specifications, which are only met by classical macroscopic FBA devices. One of the reasons is the difficulty of achieving low noise levels with a so small suspended mass.

Several manufacturers use MEMSs for accelerograph recorders with class B performance, e.g., GeoSIG GMS-18 (www.geosig. com/productfile.html?productid=10319). And some MEMS manufacturers offer class B accelerometers, e.g., Silicon Designs 1221 (www. silicondesigns.com/pdfs/1221.pdf) has typical noise PSD of 5 μg/√Hz – an rms noise of 50 μg over 100 Hz bandwidth. Recent designs (Homeijer et al. 2011) report noise levels under 100 ng/√Hz or even 10 ng/√Hz above 1 Hz (Milligan et al. 2011) with new capacitance transducers operating without feedback. These performances make it even suitable for use as exploration sensors, substituting geophones.

Class C devices are cheaper and mostly used in consumer products. Those models with stable response and enough bandwidth may be suitable to be applied in seismic strong ground motion monitoring, for instance, instrumental intensity estimation (shake maps), structural response, earthquake early warning, and shut off of critical facilities.

As an example of this class of sensors, Fig. 11 shows a MEMS triaxial accelerometer LIS 344 ALH (www.st.com/web/en/catalog/ sense_power/FM89/SC444) from STMicroelectronics mounted on a small printed circuit board, which is sold as an evaluation board, since hand soldering the chip for prototyping is quite a difficult task. This chip accelerometer has a user-selectable full-scale ±2 g or ±6 g and comes with factory-trimmed sensitivity and offset. Its bandwidth may be selected by an external capacitor.

Seismic Accelerometers,
Fig. 11 A MEMS triaxial
accelerometer LIS
344 ALH from
STMicroelectronics
(*center*) mounted in a small
printed circuit with the
minimal external
components
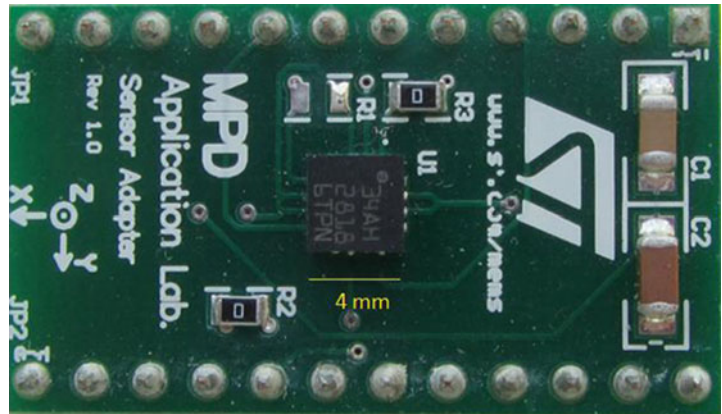
Seismic Accelerometers,
Fig. 12 *Upper*,
accelerogram obtained with
a class B accelerograph
from an earthquake of
moment magnitude
$m_w = 3.8$ at an epicentral
distance of 16 km. *Lower*,
the simulated record with
the MEMS sensor of Fig. 11
by adding the recorded self-
noise
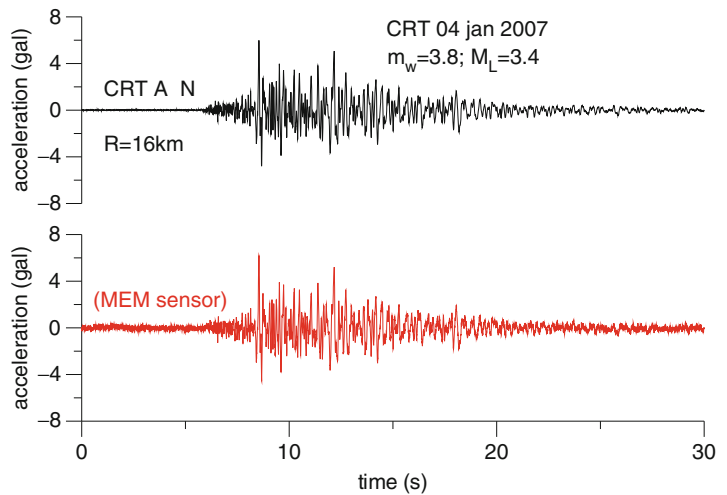
Figure 12 plots a small local earthquake recorded by a standard accelerometer and the simulated record with this sensor, using its real noise. The MEMS sensor shows a higher noise, but a useful signal is still available.

A comparative test of these sensors (Evans et al. 2014) shows that some of them could be very useful for low-cost seismic networks. Some models have performances suitable for strong ground motion recording with acceptable SNR to be used as class C seismic accelerometers.

Accelerometer Examples

Many seismic accelerometers commercially available are sold assembled with the corresponding recorder in a case, which is called an accelerograph. Some of them may be acquired separately, with an analog voltage output for each axis. For earthquake recording, the standard is a triaxial instrument. Practically all the recorders have GPS synchronization of internal clock. Some models have the possibility of Internet timing, for sites where GPS signal cannot be received.

Most accelerograph manufacturers offer several models with different performance and price. The models shown below are just examples and may not be representative of the best instrument offered by each company.

Figure 13 shows one of the accelerometers with highest dynamic range, the EpiSensor from Kinemetrics (www.kinemetrics.com/p-87-EpiSensor-ES-T.aspx), a triaxial FBA with

Seismic Accelerometers, Fig. 13 The EpiSensor accelerometer from Kinemetrics. *Left*, surface sensor package. *Right*, the case is removed to show the components
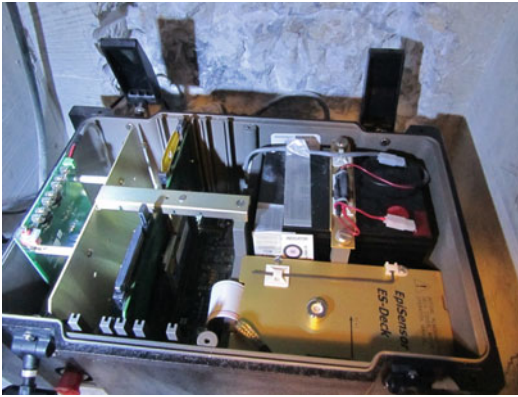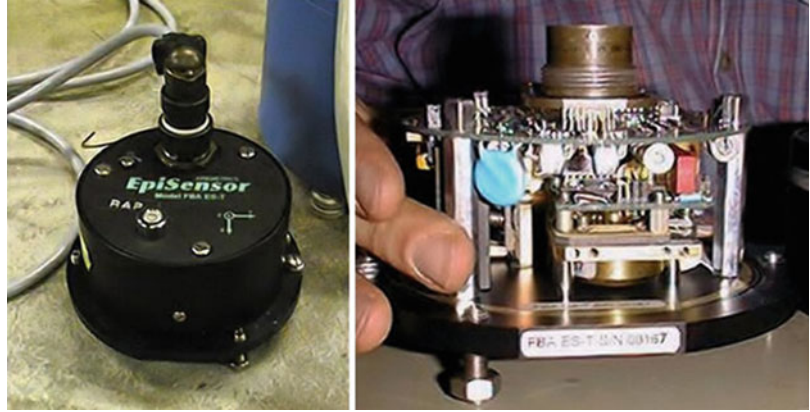
Seismic Accelerometers, Fig. 14 The Etna accelerograph from Kinemetrics with its lid removed. It has an internal triaxial accelerometer (EpiSensor). The hermetic case is made of fiberglass

Seismic Accelerometers, Fig. 15 CMG 5TD from Güralp Systems. This model can include a recorder and an Ethernet interface. The photo shows a unit installed in an underground gallery with the mounting base anchored to the concrete ground

full-scale selectable up to ±4 g. Also available is a uniaxial and with different packages, including borehole. It may be mounted inside a compact accelerograph package like Etna model (Fig. 14). This robust accelerograph model has been in the market for many years, so, in spite of the good accelerometer inside, its performance as recorder (dynamic range, communication interface, data storage) does not fully match the sensor specifications, and it may be considered a class B accelerograph. The same manufacturer offers new recorder models with higher performance.

Güralp Systems (www.guralp.com/products/instruments/cmg-5) CMG5 (Fig. 15) is an FBA

with high dynamic range that can be supplied with built-in digitizer, recorder, and communication facilities. It is a class A accelerograph. A borehole version is also available.

Figure 16 shows Nanometrics Titan (www.nanometrics.ca/products/titansma). It is one of the newest class A accelerographs, with internal FBA and high dynamic range in a very compact

Seismic Accelerometers, Fig. 16 Nanometrics Titan accelerograph. Its size is approximately 18 x 12 x 10 cm

Seismic Accelerometers, Fig. 17 FBA with MEMS technology AC-43 from GeoSIG. It weights 2 kg and its size is 19 x 11 x 9 cm (Figure from GeoSIG (www.geosig. com/AC-4x-id10357.html))

package. It has a web interface and can store data in a removable SD memory card.

Finally, two accelerometers with MEMS technology will be shown. In Fig. 17, the GeoSIG AC-43 (www.geosig.com/AC-4x-id10357.html) is a triaxial FBA with class B performance when mounted inside or connected to a recorder, such as the GSR-18, from the same company. The model 148–01 "QuakeRock" from REFTEK (Fig. 18)

Seismic Accelerometers, Fig. 18 The REFTEK 148–1 "QuakeRock" accelerograph, with MEMS sensor (Photo from REFTEK (www.reftek.com/products/motion-recorders-148-01.htm))

is a class C accelerograph with only event recording (not continuous) and limited dynamic range but may operate unattended for 2 years with two "D" size batteries. The internal clock is free running (no GPS).

Accelerograph Installation

Two kinds of accelerometer or accelerograph installations are possible: (a) free-field installation and (b) structural-monitoring installation. The purpose of the first is to record the ground motion unaffected by man-made structures. The acceleration records obtained from this kind of stations may be considered the base-level excitation of any building or structure in the zone, if the building itself does not interact with the ground at the station. The second type is an accelerograph in a building or engineering structure at different levels and positions to study and monitor the structure vibrations in response to ground seismic motions.

The main consideration is that the installation setup should not affect the accelerograph records.

S

For free-field stations, it is usual to build a small concrete pier on which the accelerometer (or accelerograph if the sensor is inside it) must be firmly anchored so a strong motion cannot move the instrument relative to ground! Most commercial instruments include a suitable base with anchoring holes or a similar system. Other materials for pier, like a table made with steel bars, are not suitable since it may resonate with very low damping in the seismic band of interest.

Free-field accelerographs should not be close to tall buildings that may modify the ground motion in their proximity. Actually some accelerograph stations are installed at the basement of buildings and are considered as free-field but may not truly fulfill this condition, due to the soil-structure interaction.

If the free-field installation is to be done on soil, the pier should have a suitable foundation to assure that it is well coupled to the ground, but not being so heavy as to modify the local soil dynamic response substantially. Free-field accelerographs must be protected with a cover or small hut from the weather. Provisions have to be made for power supply, GPS antenna, and communication, usually Ethernet via cable, satellite link, or cellular modem. And a fence around the installation would protect it from animals, human-made noises, and eventually vandalism. Accelerographs are not very sensitive to weak motions, but modern high-resolution instruments are capable of detecting human activity such as traffic or machinery working at short distance, so this kind of noise should be avoided as far as possible.

Structural-monitoring installations are usually done under cover, and additional weather protection is not required, but the instrument may have to be protected from human activity or other disturbances. Normally a pier is not required, since the sensor may be anchored directly to a structural element. GPS reception for time synchronism may be a problem if the station is far from the open sky (e.g., in a dam gallery), but there exist technical solutions: e.g., a GPS receiver may be outside, and the signal is "repeated"; the accelerograph may be synchronized via Ethernet or the accelerometer is installed on the site, and the recorder is separately installed near the open sky so the GPS antenna can be placed outside.

Summary

Seismic accelerometers sense the ground or structure seismic vibrations and, together with a suitable recorder, are called accelerographs. Most modern seismic accelerometers are of force-balance type (FBA), a servo system in which a feedback force is applied to the suspended inertial mass to keep its motion as small as possible. This improves the instrument linearity and dynamic range. Usually the mass motion is measured by a sensitive capacitive transducer.

MEMS accelerometers are integrated micromachined electromechanical devices widely used in industry that presently do not match the classical FBA performance, but are useful for some seismic applications.

Examples of commercial seismic accelerometers and accelerographs and some brief guidelines for the installation of free-field and structure-monitoring accelerographs are given.

Cross-References

▶ Passive Seismometers
▶ Principles of Broadband Seismometry
▶ Recording Seismic Signals
▶ Seismic Network and Data Quality

References

Aki K, Richards PG (2002) Quantitative seismology. University Science Books, Sausalito CA, 700 pp

Akkar S, Bommer JJ (2007) Empirical prediction equations for peak ground velocity derived from strong-motion records from Europe and the Middle East. Bull Seismol Soc Am 97:511–530. doi:10.1785/0120060141

Anderson JG (2010) Source and site characteristics of earthquakes that have caused exceptional ground accelerations and velocities. Bull Seismol Soc Am 100:1–36

ANSS Working Group on Instrumentation, Siting, Installation, and Site Metadata of the Advanced National Seismic System Technical Integration Committee (2008) Instrumentation guidelines for the advanced national seismic system, U.S. Geological Survey. Open-File Report 2008–1262, 41 pp

Deng T, Chen D, Wang J, Chen J, He W (2013) A MEMS based electrochemical vibration sensor for seismic motion monitoring. J Microelectromech Syst 99:920–923

Dwyer PD (2011) MEMS accelerometer. US Patent 8–065.915 B2, 29 Nov 2011

Eisenberg A, McEvilly TV (1971) Comparison of some widely-used strong-motion earthquake accelerometers. Bull Seismol Soc Am 61:383–397

Evans JR, Allen RM, Chung AI, Cochran ES, Guye R, Hellweg M, Lawrence JF (2014) Performance of several low-cost accelerometers. Seismological Research Letters 85:147–158

Havskov J, Alguacil G (2010) Instrumentation in earthquake seismology. Springer, Dordrecht, 358 pp

Homeijer B, Lazaroff D, Milligan D, Alley R, Wu J, Szepesi M, Bicknell B (2011) Hewlett Packard's seismic grade MEMS accelerometer. In: Proceedings of the MEMS, Jan 2011, Cancún, Mexico, pp 585–588

Li B, Lu D, Wang W (2001) Micromachined accelerometer with area-changed capacitance. Mechatronics 11:811–819

McComb HE, Ruge AC (1937) Tests of earthquake accelerometers on a shaking table. Bull Seismol Soc Am 27:325–329

Milligan DJ, Homeijer BD, Walmsley RG (2011) An ultra-low noise MEMS accelerometer for seismic imaging. In: Sensors 2011 IEEE, October 2011, Limerick, DOI 10.1109/ICSENS.2011.6127185, pp 1281–1284

Reasemberg P (1969) Non-linearity test of an accelerometer system designed for the seismic near-field measurement. Bull Seism Soc Am 59:1213–1225

Ruge AC, McComb HE (1943) Tests of earthquake accelerometers on a shaking table. Bull Seismol Soc Am 33:2–12

Trifunac MD (2009) 75th anniversary of strong motion observation-a historical review. Soil Dyn Earthq Eng 29:591–606

Walmsley RG, Kiyama LK, Milligan DM, Alley RL, Erickson DL, Hartwell, PG (2009) Micro-G silicon accelerometer using surface electrodes. In: IEEE sensors 2009 conference, October 2009, Christchurch, DOI 10.1109/ICSENS.2009.5398201, pp 971–974

Yamada M, Mori J, Heaton T (2009) The slapdown phase in high-acceleration records of large earthquakes. Seismol Res Lett 80:559–564. doi:10.1785/gssrl.80.4.559

Yanada T, Furukawa H, Ichida S, Takubo K, Shimizu Y, Koganemaru K, Suzuki T (2002) Earthquake sensor. Sens Update 11:159–216. DOI:10.1002/seup.200211101

Seismic Actions Due to Near-Fault Ground Motion

George P. Mavroeidis
Department of Civil and Environmental Engineering and Earth Sciences, University of Notre Dame, Notre Dame, IN, USA

Synonyms

Damping coefficient; Forward directivity; Near field; Near source; Permanent translation (fling); Response spectrum; Seismic ground excitation; Strength reduction factor; Time history

Introduction

Near-fault seismic ground motions are frequently characterized by intense velocity and displacement pulses of relatively long duration that clearly distinguish them from typical far-field ground motion records. This observation, along with its engineering significance, was first made with respect to the C02 record (Fig. 1a) generated by the 1966 Parkfield earthquake at a distance of only 80 m from the fault break (Housner and Trifunac 1967). The damage that the Olive View Hospital sustained during the 1971 San Fernando earthquake was also attributed to the effect of near-fault ground motions on flexible structures (Bertero et al. 1978). That was perhaps the first time that earthquake engineers linked the structural damage caused by an earthquake to the impulsive character of near-fault ground motions (Fig. 1b). However, it was not until the 1994 Northridge and the 1995 Kobe earthquakes that the majority of engineers recognized the destructive potential of near-fault ground motions and started considering methods to incorporate near-source effects into engineering design. Code provisions have historically been developed based on recorded ground motions not sufficiently close to the causative fault. Thus, the effect of near-fault pulse-like ground motions on the dynamic
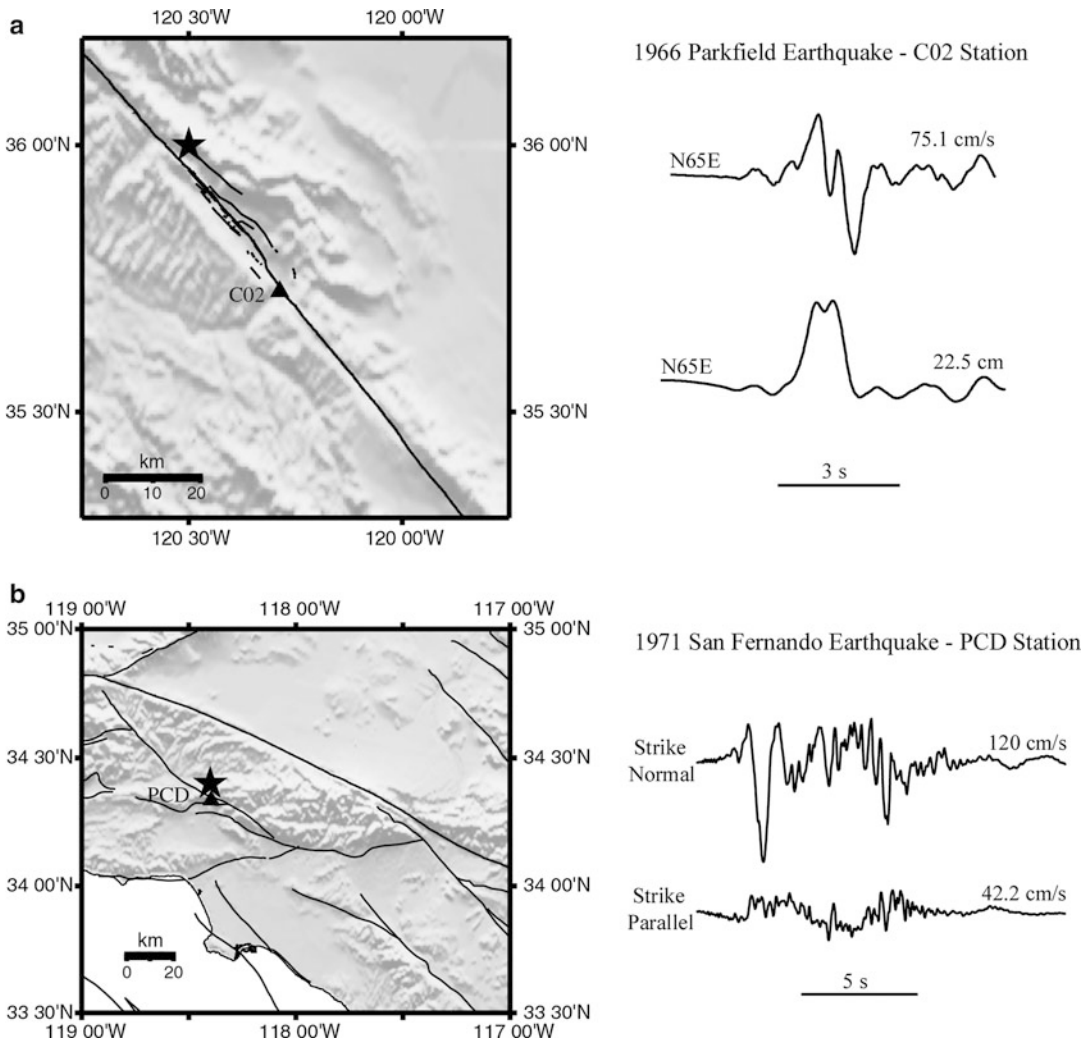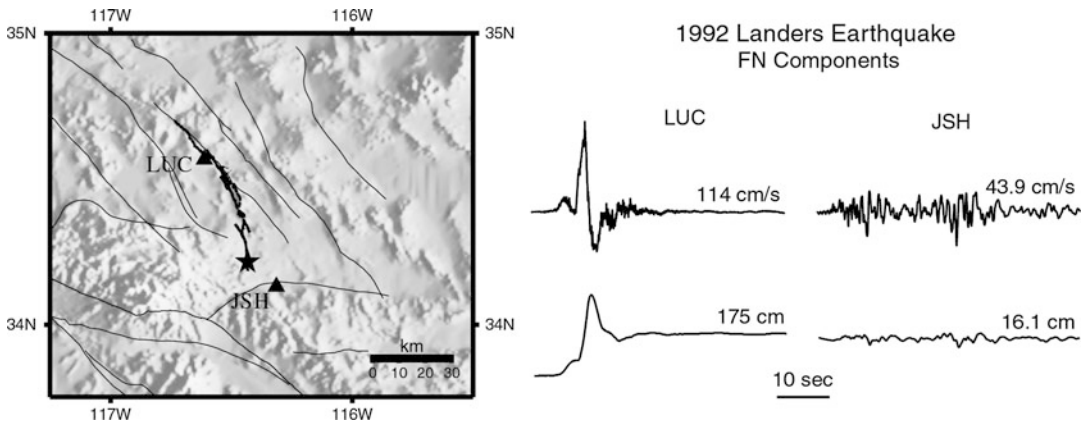
Seismic Actions Due to Near-Fault Ground Motion, Fig. 1 Characteristic examples of near-fault pulse-like ground motion records: (**a**) Station No. 2 (C02) record from the 1966 Parkfield, California, earthquake and (**b**) Pacoima Dam (PCD) record from the 1971 San Fernando, California, earthquake (Reprinted from Mavroeidis and Papageorgiou (2003). Copyright © 2003 Seismological Society of America)

response of engineering structures has received much attention over the past two decades.

This entry focuses on the description of seismic actions due to near-fault ground motions with particular emphasis on the following topics: (1) characteristics of near-fault ground motions, (2) effect of fault rupture parameters on near-fault seismic excitations, (3) synthesis of near-fault ground motion time histories for earthquake engineering applications, and (4) derivation of response spectra, strength reduction factors and damping coefficients for engineering analysis and design in the near-fault region. The material presented in this entry is primarily based on previous articles published by the author and Professor Apostolos S. Papageorgiou and is presented in a manner that provides established knowledge in the disciplines of engineering seismology and earthquake engineering to technically inclined and informed readers. It should be emphasized that this entry does not intend to be a specialized research article advancing the current state of

**Seismic Actions Due to Near-Fault Ground Motion,
Fig. 2** Characteristic examples of forward and backward
directivity from the 1992 Landers, California, earthquake

(Reprinted from Mavroeidis and Papageorgiou (2002).
Copyright © 2002 Earthquake Engineering Research
Institute)

knowledge or a review article summarizing the
vast amount of archived research literature on the
subject.

Characteristics of Near-Fault Ground Motions

Not all ground motion time histories recorded at
stations in the vicinity of a fault exhibit intense
velocity pulses. The existence of pulse-like
ground motions in near-fault records primarily
depends on the relative position of the station
that recorded the motion with respect to the direc-
tion of propagation of rupture on the causative
fault plane and on the magnitude and direction of
slip on that segment of the fault that is located in
the vicinity of the station. Whenever these ground
motion pulses do occur, they are typically caused
by the forward directivity and/or permanent
translation (fling) effects.

Forward directivity occurs when the fault rup-
ture propagates toward a site with a rupture
velocity that is approximately equal to the shear
wave velocity. In this case, most of the energy
arrives coherently in a single, intense, relatively
long-period pulse at the beginning of the record
representing the cumulative effect of almost all
the seismic radiation from the fault. Forward
directivity pulses are polarized in the fault-
normal direction for both strike-slip and dip-slip

faults. Figure 2 illustrates a characteristic exam-
ple of forward directivity from the 1992 Landers
earthquake. The fault rupture propagated to the
north along the indicated strike-slip fault. The
fault-normal velocity and displacement time his-
tories recorded at Lucerne Valley (LUC) station
(which is located in the forward direction with
respect to the propagation of rupture) are charac-
terized by intense pulse-like motions. In contrast,
the ground motion recorded at Joshua Tree (JSH)
station (which is located in the backward direc-
tion with respect to the propagation of rupture) is
relatively weak.

Permanent translation (fling) is a consequence
of permanent fault displacement due to an earth-
quake; it appears in the form of step displacement
and one-sided velocity pulse in the strike-parallel
direction for strike-slip faults or in the strike-
normal direction for dip-slip faults. In the latter
case, directivity and permanent translation
effects "build up" in the same direction. Figure 3
illustrates characteristic examples of permanent
translation (fling) from the 1999 Izmit earth-
quake. The fault-parallel velocity and displace-
ment time histories recorded at Yarimca (YPT)
and Sakarya (SKR) stations are affected by the
permanent displacement along the right-lateral
strike-slip North Anatolian Fault.

Even though emphasis has traditionally been
given to the investigation of forward directivity
and permanent translation (fling) effects, other

S
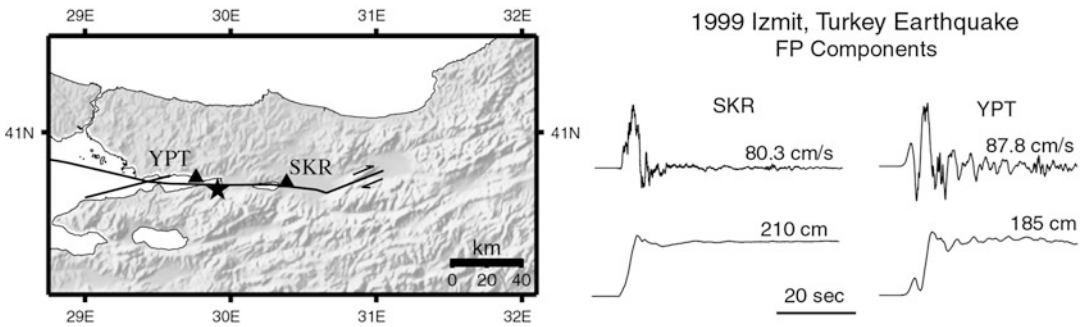
Seismic Actions Due to Near-Fault Ground Motion, Fig. 3 Characteristic examples of permanent translation (fling) from the 1999 Izmit, Turkey, earthquake

(Reprinted from Mavroeidis and Papageorgiou (2002). Copyright © 2002 Earthquake Engineering Research Institute)

conditions may also give rise to near-fault pulse-like motions. A comprehensive review of the factors that influence near-fault ground motions, along with a detailed list of references on the subject, has been presented by Mavroeidis and Papageorgiou (2002, 2003). Figure 4 illustrates a large number of actual near-fault ground motion records with "distinct" velocity pulses. These records are part of the near-fault ground motion database compiled by Mavroeidis and Papageorgiou (2003). It is evident that the pulse duration (or period), the pulse amplitude, as well as the number and phase of half cycles are the key parameters that define the waveform characteristics of the near-fault velocity pulses.

Effect of Fault Rupture Parameters on Near-Fault Ground Motions

The effect of fault rupture characteristics on near-fault ground excitations has been investigated by Mavroeidis and Papageorgiou (2010) using a kinematic modeling approach. In order to associate fault rupture characteristics (such as slip, rupture velocity, and state of stress) with near-fault ground motions, four well-documented seismic events (1979 Imperial Valley, 1985 Michoacan, 1989 Loma Prieta, and 1999 Izmit) were considered along with the concept of isochrones. An isochrone is the locus of all those points on the fault plane, the radiation of which arrives at a certain observer at a specified time.

Isochrones are frequently used in seismology to provide intuitive insight into factors that strongly influence the generation of strong ground motions. By plotting the S-wave isochrones on the fault plane of the investigated seismic events, the long-period velocity pulses of the near-fault ground motions can be directly associated with specific regions and characteristics of the fault rupture.

The results indicated that the seismic energy radiated from the high-isochrone-velocity region of the fault arrives at the receiver within a time interval that coincides with the time window of the long-period ground motion pulse recorded at the site. Furthermore, the near-fault ground motion pulses are strongly correlated with large slip on the fault plane locally driven by high stress drop. In addition, the local rupture velocity seems to be inversely correlated to the spatial distribution of the strength excess over the fault plane confirming findings of previous studies (e.g., Bouchon 1997). As an example, Fig. 5 illustrates time histories of near-fault ground motions and S-wave isochrones for selected stations of the 1979 Imperial Valley earthquake. These stations are located close to the ruptured fault where the effect of forward directivity was pronounced. The spatial distribution of the static slip offset and rupture time inferred by Archuleta (1984) and the spatial distribution of the static stress drop and strength excess calculated using the methodology proposed by Bouchon (1997) are also illustrated.
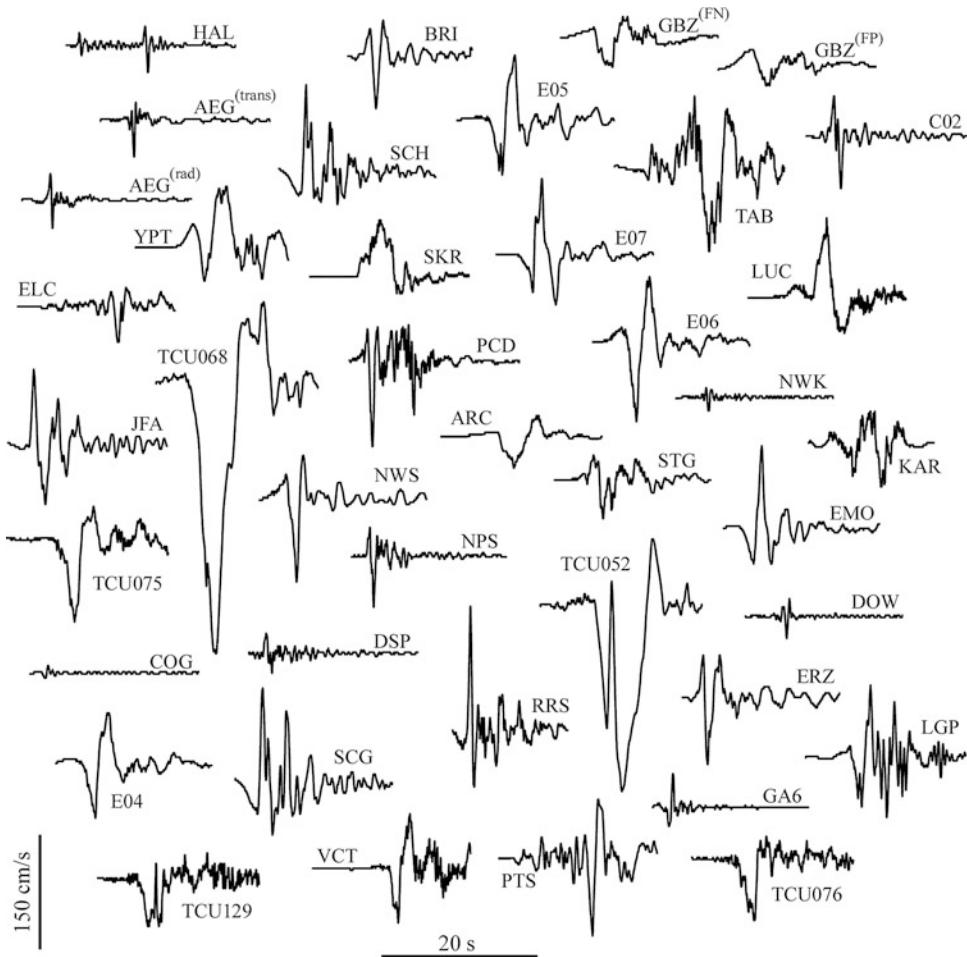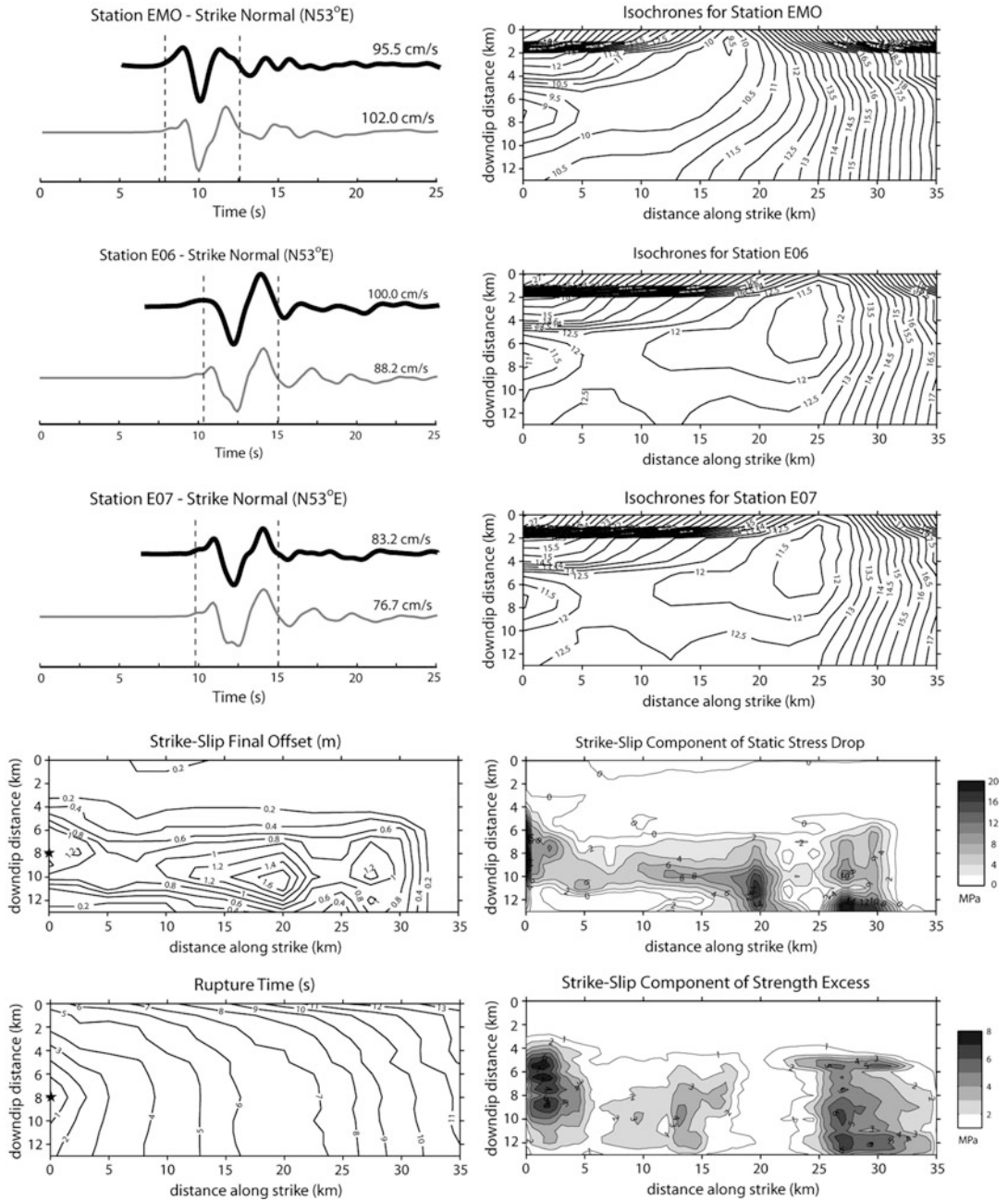
Seismic Actions Due to Near-Fault Ground Motion, Fig. 4 Strong motion records with "distinct" velocity pulses (Reprinted from Mavroeidis and Papageorgiou (2003). Copyright © 2003 Seismological Society of America)

Finally, it was found that for various events, the area of the fault that contributes to the formation of the near-fault pulse encompasses more than one patch of significant moment release (subevent) (e.g., 1979 Imperial Valley, 1989 Loma Prieta). This observation explains why a dislocation model with average properties (i.e., slip, rise time, etc.) reproduces successfully near-fault ground motions for strike-slip faults and for dip-slip faults with intermediate-to-large earthquake magnitudes. However, for very large earthquakes, such as megathrust events on subduction zones (e.g., 1985 Michoacan), the fault region that contributes to the pulse formation encompasses individual subevents, and, consequently, crack-like slip functions (rather than dislocation models) may be more appropriate for the simulation of the near-fault ground motions. The interested reader may find a detailed discussion in Mavroeidis and Papageorgiou (2010).

Time Histories of Near-Fault Ground Motions

The advent of performance-based earthquake engineering, the growth of computer processing power, and the associated increased availability

Seismic Actions Due to Near-Fault Ground Motion, Fig. 5 Recorded (*black trace*) and synthetic (*gray trace*) near-fault ground motion time histories and *S*-wave isochrones for selected stations of the 1979 Imperial Valley, California, earthquake. Tomographic images of the static slip offset, rupture time, static stress drop, and strength excess are also illustrated (Reprinted from Mavroeidis and Papageorgiou (2010). Copyright © 2010 Seismological Society of America)

of structural analysis software have made possible the performance of sophisticated nonlinear structural analysis on a routine basis. However, the overall seismic performance-based assessment of a given structure hinges on the use of realistic earthquake ground motions that reflect the seismic hazard at the site of the structure, as well as the local site conditions (Halldorsson et al. 2011).

In general, earthquake engineers have the following options for selecting ground motion input when performing nonlinear structural analysis in the near-fault region: (1) use actual records of near-fault ground motion, (2) generate synthetic records of near-fault ground motion using physical models of the seismic source, and (3) generate synthetic records of near-fault ground motion using phenomenological models.

Recorded Near-Fault Ground Motions

The gradually increasing number of recorded near-fault ground motions has recently enabled strong motion seismologists to compile these records in publicly available ground motion databases (e.g., Pacific Earthquake Engineering Research Center Ground Motion Database, Center for Engineering Strong Motion Data, European Strong Motion Database, among others). Even though the number of near-fault records is still limited, they have served as an invaluable resource to earthquake engineers. Researchers have also proposed methodologies for identifying and extracting pulse-like motions from actual near-fault records using wavelet analysis (e.g., Baker 2007; Vassiliou and Makris 2011).

However, the selection of strong motion records for nonlinear structural analysis is not always a straightforward process. For instance, the available records may not reflect the appropriate earthquake magnitude, source mechanism, site conditions, or source-site configuration. While this could be a problem for far-field sites, it is even a greater challenge for sites in the immediate vicinity of the fault. For the above reasons, it is of paramount importance to earthquake engineers to have the ability to generate suites of realistic broadband ground motion time

histories, both in the far-field and near-fault regions (Halldorsson et al. 2011).

Synthetic Near-Fault Ground Motions Using Physical Models of the Seismic Source

Strong motion seismologists have utilized various schemes of deterministic and stochastic simulation techniques to generate broadband ground motion time histories at specific locations in the vicinity of the fault (see, e.g., Papageorgiou 1997 and references provided therein). These simulation methods are based on source mechanics principles and wave propagation theory. Site effects are also frequently taken into account. Regardless of the degree of sophistication of the various ground motion simulation methods, the Earth crustal structure and the seismic source should sufficiently be characterized and quantified.

For regions of intense seismic activity, the crustal structure is frequently defined in terms of one-dimensional velocity models. Detailed three-dimensional crustal models have also become available for specific regions in the benefit of three-dimensional wave propagation codes that may effectively take into account basin effects and complex fault geometries at the cost of increased computational demands. The characterization of the seismic source is a more complicated issue. For kinematic descriptions of the earthquake source, source parameters such as slip, rise time, rupture velocity, and slip function should properly be quantified and a priori defined. On the other hand, for dynamic descriptions of the earthquake source, the source parameters may vary as long as the elastodynamics equation with a prescribed fracture criterion on a predetermined fault plane is satisfied. The selected initial conditions and failure criterion determine the time and space evolution of the fault rupture in a dynamic source model.

Once the seismic source and Earth crustal model have been adequately described, near-fault ground motion simulations in the low-frequency range (e.g., below 1 Hz) can be performed using deterministic modeling techniques [e.g., discrete wavenumber method (DWN), finite difference method (FDM), finite element method (FEM), boundary element

S

method (BEM), spectral element method (SEM), or hybrids of them] that involve calculations of synthetic Green's functions. In order to generate broadband synthetic ground motions, the low-frequency waveforms should be combined with high-frequency ground motions (e.g., above 1 Hz) simulated using: (1) the empirical or semiempirical Green's function method or (2) a stochastic modeling technique utilizing a source model that provides an unambiguous way to distribute the seismic moment of the simulated event on the fault plane. This matter is of great importance for near-fault ground motion simulations due to the proximity of the point of observation to the source. It should be mentioned that high-frequency ground motion simulations can be carried out using synthetic Green's functions as well, excluding site effects and small-scale heterogeneities.

Synthetic Near-Fault Ground Motions Using Phenomenological Models

Ground motion simulation techniques based on kinematic or dynamic source models are not always appealing to earthquake engineers because specialized seismological knowledge and, quite frequently, demanding computational resources are required. Therefore, in practice, earthquake engineers utilize actual near-fault records of past earthquakes to investigate the dynamic response of engineering structures to pulse-like seismic excitations and rely on strong motion seismologists only for generating site-specific near-fault ground motions for the design of special structures.

To overcome this deficiency, earthquake engineers have introduced idealized waveforms, intending to represent typical ground motion pulses observed in the near-fault region, in an effort to investigate the dynamic response of engineering structures to near-fault ground motions (e.g., Makris 1997; Sasani and Bertero 2000; Alavi and Krawinkler 2001; Mavroeidis and Papageorgiou 2003, among others). These idealized waveforms should successfully capture the impulsive character of the near-fault records both qualitatively and quantitatively. In addition, their input parameters should have a clear physical

interpretation and scale, to the extent possible, with physical parameters of the faulting process.

Mathematical Representation of Near-Fault Ground Motion Pulses Proposed by Mavroeidis and Papageorgiou (2003)

The mathematical formulation for the representation of the near-fault ground velocity pulses proposed by Mavroeidis and Papageorgiou (2003) is the product of a harmonic oscillation with a bell-shaped function. That is:

$$v(t) = \begin{cases} A\dfrac{1}{2}\left[1+\cos\left(\dfrac{2\pi f_P}{\gamma}(t-t_0)\right)\right]\cos\left[2\pi f_P(t-t_0)+v\right], \\ t_0-\dfrac{\gamma}{2f_P}\le t \le t_0+\dfrac{\gamma}{2f_P} \text{ with } \gamma>1 \quad 0, \text{ otherwise} \end{cases} \quad (1)$$

where A controls the amplitude of the signal, f_P is the frequency of the amplitude-modulated harmonic (or the prevailing frequency of the signal), v is the phase of the amplitude-modulated harmonic, γ is a parameter that defines the oscillatory character of the signal, and t_0 specifies the epoch of the envelope's peak. The pulse period (T_P) is defined as the inverse of the prevailing frequency (f_P) of the signal, thus providing an "objective" assessment of this important parameter. That is:

$$T_P = \frac{1}{f_P} \quad (2)$$

The model input parameters have a clear physical meaning as they coincide with the key features that determine the waveform characteristics of the near-fault pulses (i.e., amplitude, duration, phase and number of half cycles).

The mathematical model proposed by Mavroeidis and Papageorgiou (2003) was calibrated using a large number of actual near-fault records. It successfully replicated a large set of displacement, velocity, and, in many cases, acceleration time histories, as well as the corresponding elastic response spectra. A sample of the quality of fitting of the synthetic waveforms to actual near-fault records is illustrated in Fig. 6.

The scaling characteristics of the model input parameters were also investigated through
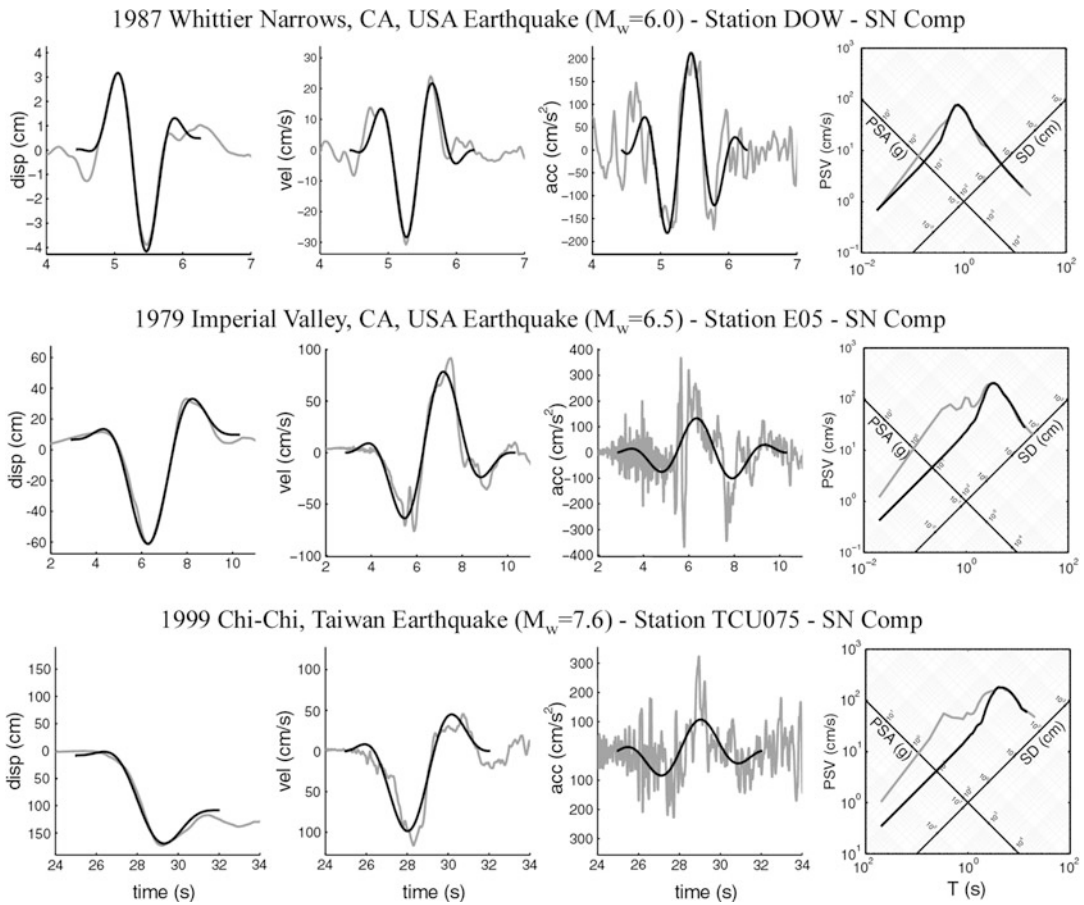
Seismic Actions Due to Near-Fault Ground Motion, Fig. 6 Sample of synthetic waveforms (*black trace*) fitted to actual near-fault records (*gray trace*). Ground motion time histories (displacement, velocity, and acceleration),

as well as the 5% damped elastic response spectra are illustrated (Reprinted from Mavroeidis et al. (2004). Copyright © 2004 John Wiley & Sons, Inc.)

regression analysis, and simple empirical relationships were proposed. By performing least-squares analysis (Fig. 7a), the following relationship was obtained between the pulse period and the earthquake magnitude:

$$\log T_P = -2.2 + 0.4 M_w \qquad (3a)$$

If the empirical relationship is required to satisfy the self-similarity condition, the following equation can be obtained:

$$\log T_P = -2.9 + 0.5 M_w \qquad (3b)$$

Equation (3) was derived by Mavroeidis and Papageorgiou (2003) using near-fault ground

motion records affected by forward directivity. Similar scaling equations have been proposed by other investigators (e.g., Somerville 2003; Bray and Rodriguez-Marek 2004; Baker 2007, among others). However, it should be pointed out that the definition of the pulse period in these studies is not the same as the definition provided by Mavroeidis and Papageorgiou (2003), and therefore caution should be utilized when the mathematical model of Eq. 1 is used in conjunction with scaling laws for the pulse period proposed by other investigators.

Mavroeidis et al. (2004) also derived an equation that relates the pulse period (T_P) with the rise time (τ) (i.e., the time it takes for a representative point on the fault plane to reach its final displacement). The rise time is an important physical
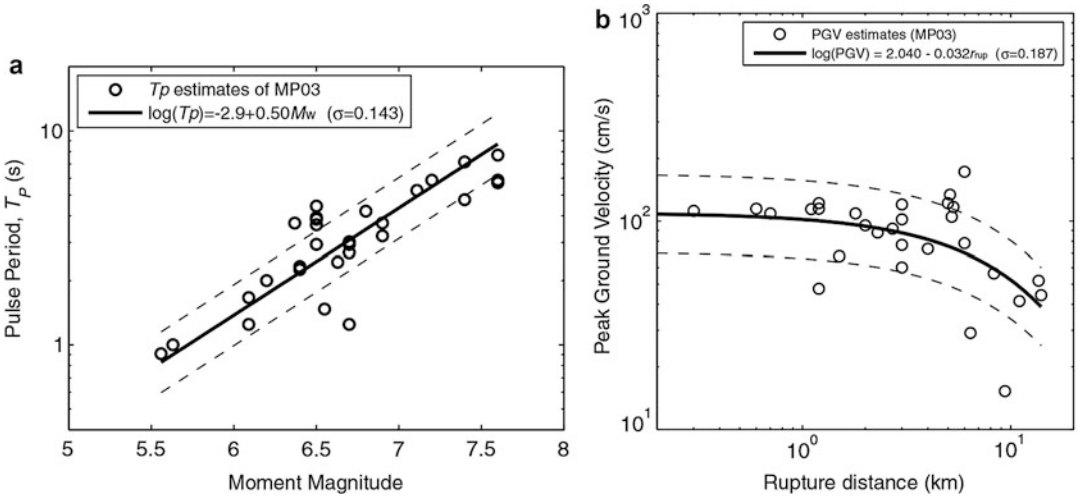
Seismic Actions Due to Near-Fault Ground Motion, Fig. 7 (**a**) Scaling of the pulse period with earthquake magnitude, according to Mavroeidis and Papageorgiou (2003), and (**b**) attenuation of peak ground velocity with rupture distance, according to Halldorsson et al. (2011) (Reprinted from Halldorsson et al. (2011). Copyright © 2011 American Society of Civil Engineers)

parameter of the fault rupture process that greatly affects strong ground motion characteristics. In fact, the rise time (and therefore the pulse period) is related to one of the characteristic corner frequencies of the source spectrum (i.e., the spectrum of seismic waves radiated by the earthquake source before these are modified by propagation path and site effects).

On the other hand, the peak value of the near-fault velocity records appears to be a fairly stable parameter. A value of 100 cm/s effectively represents peak ground velocities within a few kilometers from the causative fault regardless of the earthquake magnitude (Mavroeidis and Papageorgiou 2003). This observation is in good agreement with the typical slip velocity value of 100 cm/s frequently considered by seismologists. As indicated by Mavroeidis et al. (2004), there are solid physical reasons that explain the stability of the velocity amplitude close to the fault. More recently, Halldorsson et al. (2011) proposed the following attenuation relationship for peak ground velocity (PGV) with rupture distance (R) (Fig. 7b):

$$\log PGV = 2.040 - 0.032 \; R \qquad (4)$$

Once PGV has been determined using Eq. 4, parameter A that controls the amplitude of the

synthetic velocity pulse can be defined by considering that $A \sim (0.85–1.00)$ PGV. Other investigators have also proposed attenuation relationships for PGV (e.g., Bray and Rodriguez-Marek 2004, among others).

Finally, parameter γ varies from a value slightly larger than 1 up to a maximum value of 3, while the phase angle (v) varies from 0° to 360°. Halldorsson et al. (2011) have provided the probability density functions of γ and v, assuming that these two parameters are normally distributed.

Simplified Methodology for Generating Broadband Near-Fault Ground Motions Proposed by Mavroeidis and Papageorgiou (2003)

As explained by Mavroeidis and Papageorgiou (2003), the mathematical expression of Eq. 1 replicates accurately the intermediate-to-long-period ("coherent") features of the near-fault ground motions. The high-frequency components that are "incoherent" cannot be simulated using simplified mathematical models (see, e.g., acceleration time histories and the short-period range of response spectra in Fig. 6).

A simplified methodology for generating realistic, broadband, near-fault ground motions that are adequate for engineering analysis and design was proposed by Mavroeidis and Papageorgiou (2003).

Based on this technique, the coherent (long-period) ground motion component is simulated using the mathematical model of Eq. 1, while the incoherent (high-frequency) seismic radiation is synthesized using the specific barrier model (Papageorgiou and Aki 1983a, b) in the context of the stochastic modeling approach. The specific barrier model is a physical model of the seismic source that applies both to the "near-fault" and "far-field" regions, allowing for consistent ground motion simulations over the entire frequency range and for all distances of engineering interest. The specific barrier model has been calibrated to shallow crustal earthquakes of three different tectonic regions: interplate, intraplate, and extensional regimes (Halldorsson and Papageorgiou 2005).

This simplified methodology has been applied to hypothetical and actual earthquakes (e.g., Mavroeidis and Papageorgiou 2003; Halldorsson et al. 2011) and is demonstrated in this entry through the case study of the 1971 San Fernando earthquake (Mavroeidis 2004). According to Heaton (1982), the 1971 San Fernando earthquake with M_W 6.6 may have been a double seismic event that occurred on two subparallel thrust faults, the Sierra Madre and San Fernando Faults, as indicated in Fig. 8a. The slip distribution on the causative faults, inferred by inversion of teleseismic, strong motion, and geodetic data, is illustrated in Fig. 8b. The damage that the Olive View Hospital sustained during the earthquake has been attributed to the destructive potential of near-fault ground motions on flexible structures (Bertero et al. 1978). No strong motion instruments were installed in the immediate vicinity of the hospital building. However, there are indications that the ground motion that the Olive View Hospital sustained was equivalent or greater than the ground motion recorded at the nearby Pacoima Dam (PCD) station.

Broadband synthetic time histories are generated at the location of the Olive View Hospital for the fault-station geometry of Fig. 8a. The mathematical model of Eq. 1 is first employed to generate the coherent component of the ground motion at the Olive View Hospital. The values for the input parameters A, f_P, γ, and v are those inferred by fitting the mathematical model of Eq. 1 to the ground motion time histories and response spectra of the PCD record (see Mavroeidis and Papageorgiou 2003). For the synthesis of the incoherent seismic radiation at the location of the Olive View Hospital, the specific barrier model of Papageorgiou and Aki (1983a) is utilized. The selected parameters for the specific barrier model are consistent with the values inferred by Papageorgiou and Aki (1983b) for the 1971 San Fernando earthquake. The model consists of two subevents (Fig. 8c), in agreement with the two distinct slip patches of similar size inferred for this event by Heaton (1982) (Fig. 8b). The site characterization at the Olive View Hospital is assumed to be NEHRP site class D consistent with available information.

Figure 9a illustrates the synthetic ground motions (strike-normal component) at the Olive View Hospital. The top and middle panels display the incoherent and coherent ground motion components, while the bottom panels show the superposition of the above two components. For comparison purposes, the ground motion recorded at the nearby PCD station due to the 1971 San Fernando earthquake is also shown in Fig. 9b. The overall agreement between the synthetic ground motions at the Olive View Hospital and the recorded ground motions at PCD station is very good. It is evident that acceleration amplitudes larger than those recorded at the PCD station characterize the synthetic accelerogram at the Olive View Hospital. On the other hand, the corresponding velocity and displacement time histories are very similar. These differences in acceleration amplitudes may be attributed to the different site conditions at the locations of the Olive View Hospital (NEHRP site class D) and Pacoima Dam (NEHRP site class B; rock good enough to serve as the foundation of a concrete dam).
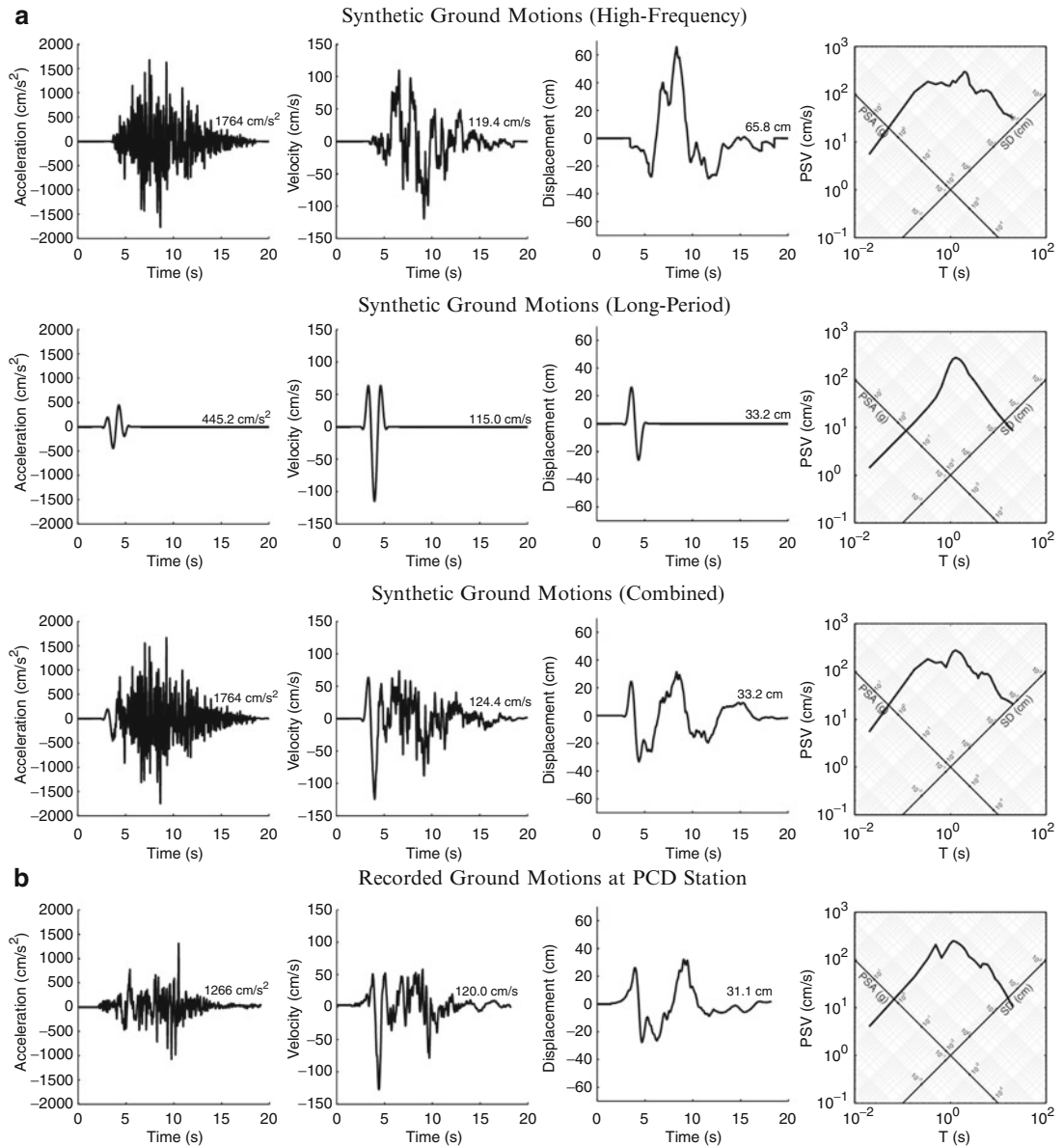
Response Spectra, Strength Reduction Factors, and Damping Coefficients for Near-Fault Ground Motions

In this section, the primary characteristics of near-fault ground motion response spectra are discussed, and recommendations are made for

S
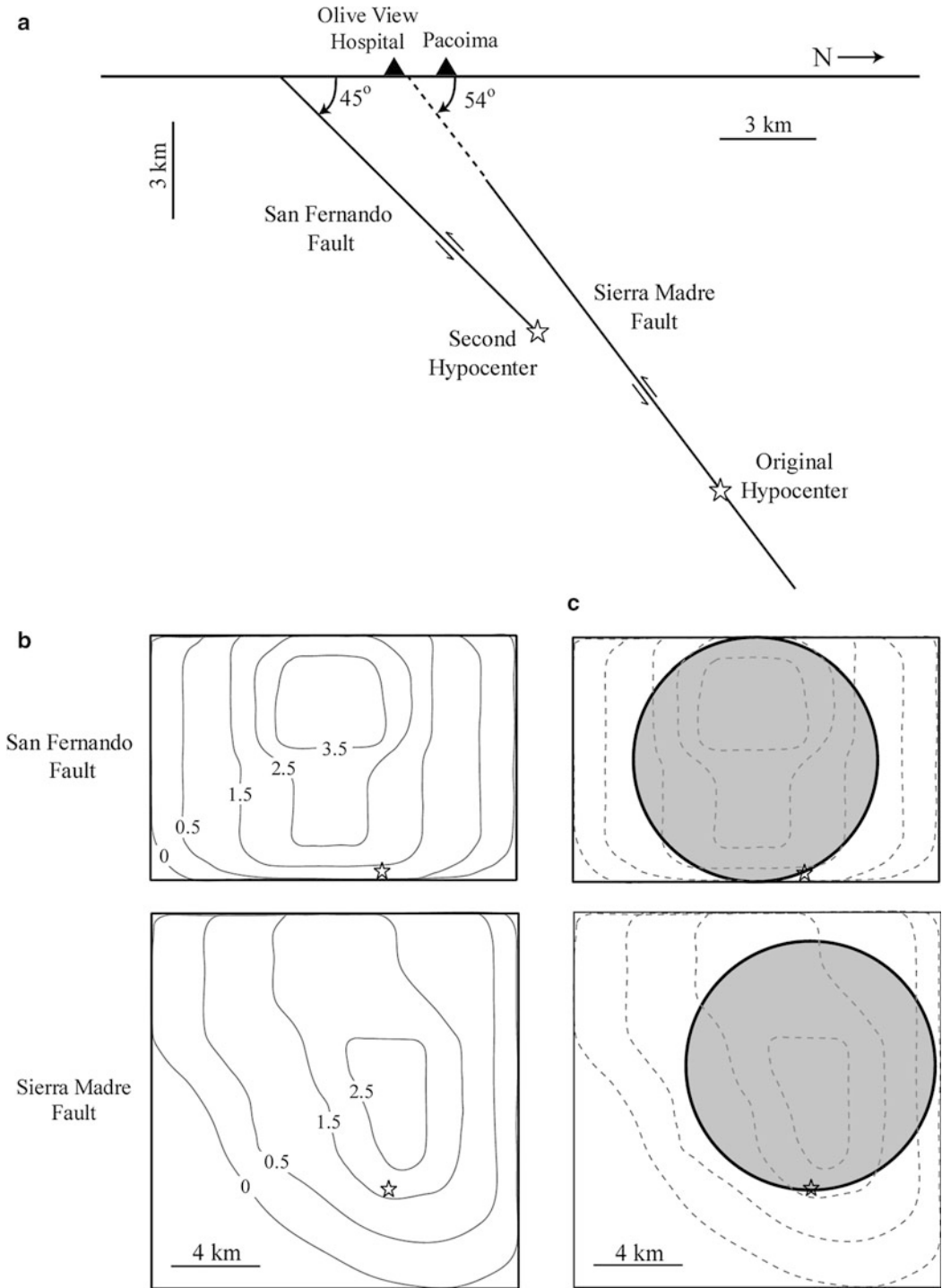
Seismic Actions Due to Near-Fault Ground Motion, Fig. 8 (**a**) Cross-sectional view of the causative faults of the 1971 San Fernando, California, earthquake (Heaton 1982), (**b**) slip distribution in meters for the two subparallel thrust faults of Fig. 8a (Heaton 1982), and (**c**) subevents of the specific barrier model represented by two circular cracks (ω^2–model) (Reprinted from Mavroeidis (2004). Copyright © 2004 G. P. Mavroeidis)
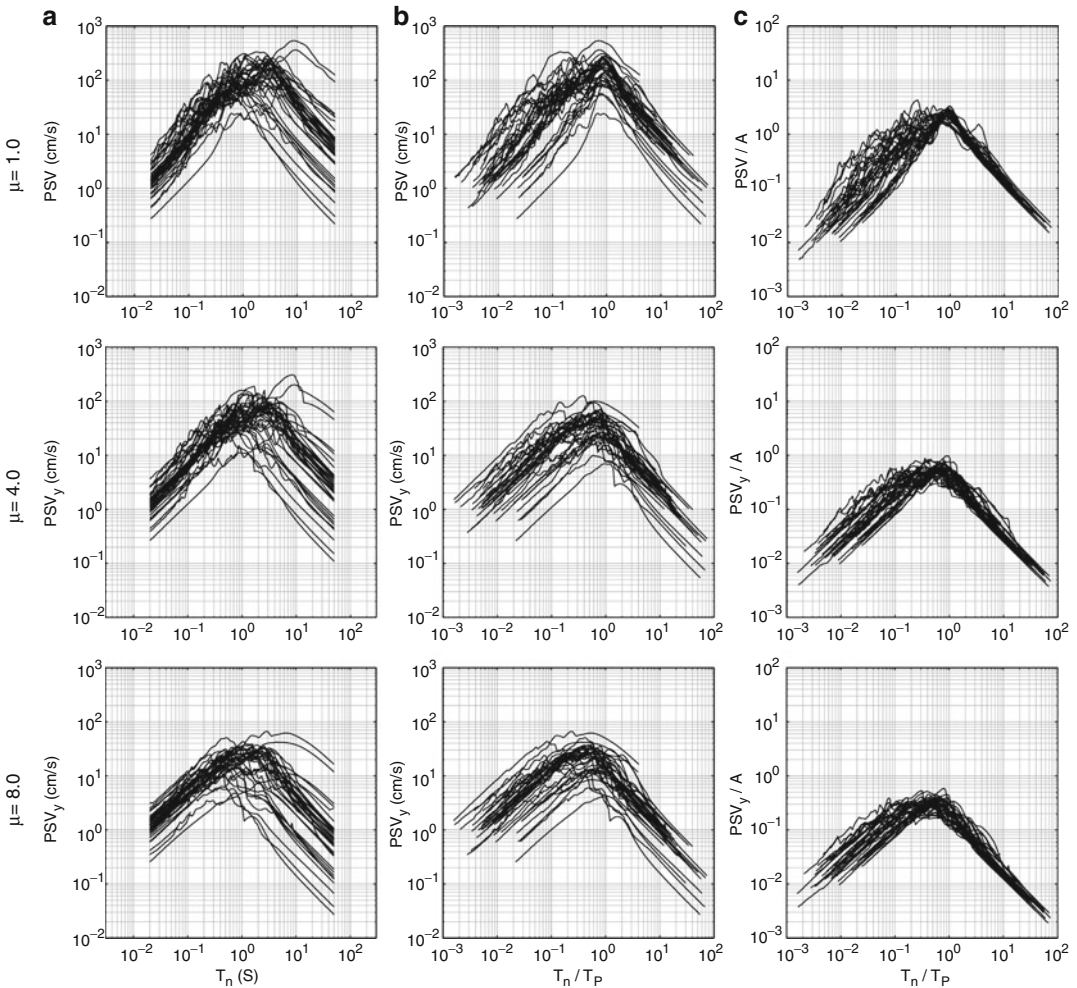
Seismic Actions Due to Near-Fault Ground Motion, Fig. 9 (**a**) Synthesis of near-fault ground motions at the location of the Olive View Hospital for the fault-station geometry illustrated in Fig. 8a; the 5% damped elastic response spectra are also shown. (**b**) Actual ground motions recorded at the nearby Pacoima Dam (PCD) station (Reprinted from Mavroeidis (2004). Copyright © 2004 G. P. Mavroeidis)

design spectra, strength reduction factors and damping coefficients for analysis, and design in the near-fault region. The interested reader may find additional information in Mavroeidis et al. (2004) and Hubbard and Mavroeidis (2011).

Response Spectra

Figure 10a displays the 5% damped equal-ductility pseudo-velocity response spectra of elastic-perfectly plastic single-degree-of-freedom (SDOF) systems subjected to a large

Seismic Actions Due to Near-Fault Ground Motion, Fig. 10 Standard and normalized 5% damped equal-ductility ($\mu = 1.0$, 4.0, 8.0) pseudo-velocity response spectra of elastic-perfectly plastic SDOF systems subjected to actual near-fault ground motion records: (**a**) *PSV* versus T_n, (**b**) *PSV* versus T_n/T_P, and (**c**) *PSV/A* versus T_n/T_P (Reprinted from Mavroeidis et al. (2004). Copyright © 2004 John Wiley & Sons, Inc.)

number of actual near-fault ground motion records (Mavroeidis et al. 2004). Inspection of this figure reveals that peak spectral amplitudes of near-fault records vary significantly, especially for smaller values of the ductility factor (μ). Furthermore, the periods that correspond to peak spectral amplitudes are characterized by significant dispersion.

Figure 10b illustrates the equal-ductility pseudo-velocity response spectra of Fig. 10a with the period axis normalized with respect to the corresponding T_P values estimated by Mavroeidis and Papageorgiou (2003). This abscissa normalization yields response spectra characterized by peak spectral amplitudes that lie within a very narrow range of the normalized period (i.e., $T_{peak}/T_P \approx 0.7$–1.0 for elastic spectra). If the ordinates of the equal-ductility pseudo-velocity response spectra are further normalized with respect to A, the normalized response spectra of Fig. 10c are obtained; these spectra are characterized by small dispersion of normalized peak spectral amplitudes. In addition, they exhibit smaller dispersion in the normalized
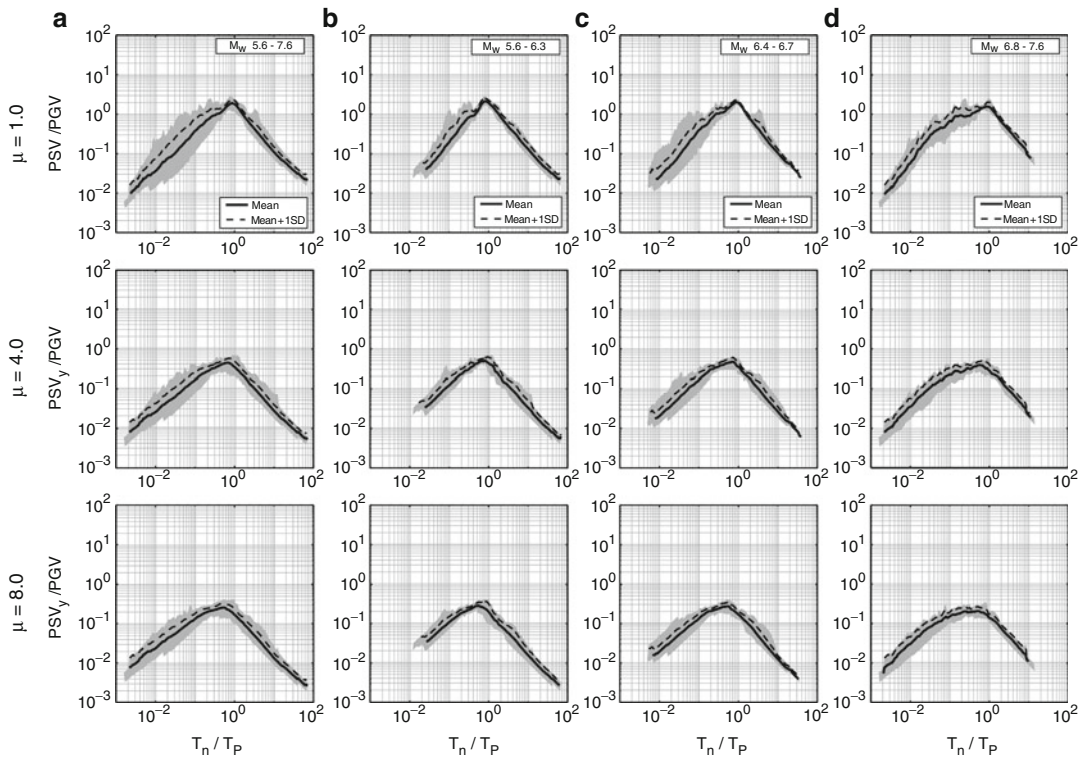
Seismic Actions Due to Near-Fault Ground Motion, Fig. 11 Normalized 5% damped equal-ductility pseudo-velocity response spectra of elastic-perfectly plastic SDOF systems: (**a**) all earthquakes (M_W 5.6–7.6), (**b**) moderate earthquakes (M_W 5.6–6.3), (**c**) moderate-to-large earthquakes (M_W 6.4–6.7), and (**d**) large earthquakes

(M_W 6.8–7.6). The *solid* and *dashed lines* represent the mean and mean-plus-one-standard-deviation pseudo-velocity response spectra, respectively. The *gray* region represents the range of variation of the spectral amplitudes (Reprinted from Mavroeidis et al. (2004). Copyright © 2004 John Wiley & Sons, Inc.)

long-period range and larger dispersion in the normalized high-frequency range (controlled by the coherent and incoherent ground motion components, respectively).

Thus, it may be concluded that parameters A and T_P can be used to normalize the response spectra of actual near-fault records. This development facilitates the systematic investigation of the response spectrum characteristics of the SDOF system subjected to near-fault ground motions. It should be noted that the normalization of the ordinates of the equal-ductility pseudo-velocity response spectra of Fig. 10a with respect to PGV values yields normalized response spectra very similar to those illustrated in Fig. 10c. This is anticipated because parameter A effectively approximates PGV.

In order to investigate the effect of the earthquake magnitude on the normalized response spectra of Fig. 11a, the seismic events have been grouped into three categories: moderate (M_W 5.6–6.3), moderate-to-large (M_W 6.4–6.7), and large (M_W 6.8–7.6) earthquakes. The normalized equal-ductility pseudo-velocity response spectra of these three categories are displayed in Fig. 11b, c, d, respectively. It becomes evident that, for smaller values of T_n/T_P, the normalized spectral amplitudes increase with earthquake magnitude. However, for larger T_n/T_P values, the normalized response spectra appear to exhibit a uniform behavior regardless of the variation in earthquake magnitude. As a consequence, the normalized response spectra of large earthquakes exhibit flatter shapes around their peaks than the
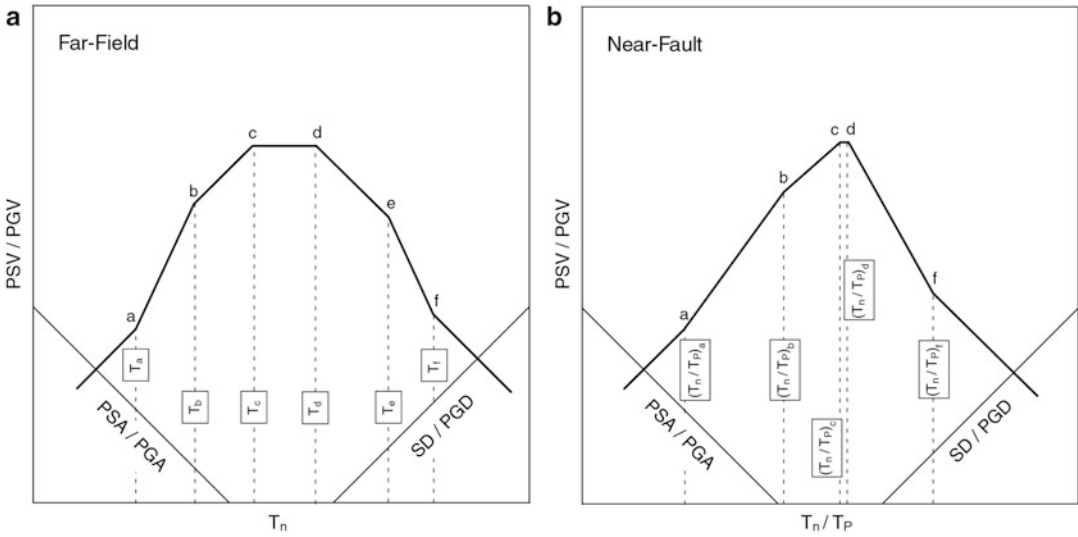
Seismic Actions Due to Near-Fault Ground Motion, Fig. 12 Schematic illustrations of idealized response spectra in four-way logarithmic plots for far-field and near-fault ground motion records (Reprinted from Mavroeidis et al. (2004). Copyright © 2004 John Wiley & Sons, Inc.)

normalized response spectra of moderate earthquakes.

The ensemble of the normalized elastic response spectra illustrated in the first panel of Fig. 11 can be utilized to derive normalized elastic design spectra for moderate, moderate-to-large, and large earthquakes, as well as for the entire set of seismic events considered by Mavroeidis et al. (2004). The solid and dashed lines in the top panel of Fig. 11 represent the mean and mean-plus-one-standard-deviation 5% damped normalized elastic response spectra. These average elastic response spectra can be used to derive normalized elastic design spectra for two different nonexceedance probability levels.

Figure 12a displays a sketch of the standard idealized elastic design spectrum derived from far-field ground motion records. This standard design spectrum has been used for many decades since it was first introduced in engineering practice. The acceleration-, velocity-, and displacement-sensitive regions of this design spectrum can readily be identified in Fig. 12a. On the other hand, the normalized response spectra of the near-fault ground motion records (see Fig. 11) can be approximated by the sequence of linear segments displayed in Fig. 12b. The values of the characteristic normalized periods

$[(T_n/T_P)_a, \quad (T_n/T_P)_b, \quad (T_n/T_P)_c, \quad (T_n/T_P)_d,$ and $(T_n/T_P)_f]$ are provided in Mavroeidis et al. (2004).

Strength Reduction Factors

The earliest and perhaps the simplest recommendation of a procedure to construct inelastic spectra from elastic spectra using ductility-dependent strength reduction factors (R_y) is based on the work of Veletsos and Newmark (1960). These results were further developed by Newmark and Hall (1982) based on a suite of far-field ground motion records. Mavroeidis et al. (2004) checked the validity of the reduction factors proposed by Newmark and Hall (1982) for the response spectra of near-fault ground motion records by normalizing the period intervals of the R_y design equations as follows:

$$R_y = \begin{cases} 1, & \dfrac{T_n}{T_P} < \left(\dfrac{T_n}{T_P}\right)_a \\[2ex] \sqrt{2\mu - 1}, & \left(\dfrac{T_n}{T_P}\right)_b < \dfrac{T_n}{T_P} < \left(\dfrac{T_n}{T_P}\right)_{c'} \\[2ex] \mu, & \dfrac{T_n}{T_P} > \left(\dfrac{T_n}{T_P}\right)_c \end{cases} \tag{5}$$

where $\left(\frac{T_n}{T_P}\right)_{c'} = \frac{\sqrt{2\mu-1}}{\mu} \left(\frac{T_n}{T_P}\right)_c$. The characteristic values of $(T_n/T_P)_a$, $(T_n/T_P)_b$, and $(T_n/T_P)_c$ are

associated with the normalized elastic design spectrum for near-fault ground motions (see Fig. 12b) proposed by Mavroeidis et al. (2004).

Figure 13 compares the computed values of R_y from the mean elastic and inelastic 5% damped normalized response spectra of Fig. 11 with the R_y values obtained from Eq. 5. The agreement between the two sets of curves is very good over the entire period range, for all specified ductility factors, and for all earthquake magnitude categories. Figure 13 demonstrates that the Veletsos-Newmark-Hall design equations can be used for near-fault ground motions as well, provided that normalized response spectra are utilized and appropriate values of $(T_n/T_P)_a$, $(T_n/T_P)_b$, and $(T_n/T_P)_c$ are selected.

Damping Coefficients

Damping coefficients are frequently used in earthquake engineering as a simple way to adjust the pseudo-acceleration or displacement response spectra associated with a viscous damping ratio of 5% to the higher values of viscous damping needed for design of structures equipped with base isolation and/or supplemental energy dissipation devices. Damping coefficients are also frequently used for predicting the maximum displacement demands of an inelastic structure from the maximum displacement demands of its equivalent linear system characterized by a longer natural period and a higher viscous damping ratio.

Damping coefficients (B) are defined as $B(T,\beta) = PSA(T,\beta = 5\%)/PSA(T,\beta)$, where T is the elastic period of vibration of the structure, β is the viscous damping ratio, and PSA are the ordinates of the pseudo-acceleration response spectra for particular values of T and β. Damping coefficients (B) are also known as "damping adjustment factors." The reciprocal of B is often used in the literature and referred to as "damping correction factor," "damping reduction factor," "spectral scaling factor," or "damping modification factor."

Hubbard and Mavroeidis (2011) calculated damping coefficients for the SDOF system subjected to near-fault pulse-like ground motions for a large range of periods and damping levels.

The results indicated that damping coefficients proposed in design codes and previous studies, based primarily on far-field ground motion records, tend to not be conservative for near-fault pulse-like seismic excitations. Figure 14a illustrates the relationships between damping coefficient and period that were established for viscous damping ratios in the range of 5–100%. These damping coefficient curves were generated using the definition of B factors and the median pseudo-acceleration response spectra for various levels of damping.

In order to investigate the effect of earthquake magnitude on damping coefficients using near-fault records, the seismic events were again grouped into three categories labeled as moderate (M_W 5.6–6.3), moderate-to-large (M_W 6.4–6.7), and large (M_W 6.8–7.6) earthquakes. The variation of damping coefficient with period for these three categories is displayed in Figs. 14b, c, d for viscous damping ratios in the range of 5–100%. While the B factors for all three earthquake magnitude categories attain approximately the same peak values for a given damping ratio, the period range over which these peak values occur clearly depends on earthquake magnitude. In addition, the damping coefficient curves of Fig. 14a derived from the entire ground motion ensemble smooth out the effect of earthquake magnitude and therefore do not capture the particular features of the damping coefficient plots illustrated in Figs. 14b, c, d.

Figure 15 indicates that the normalization of the period axis of the B plots with respect to T_P yields damping coefficient curves that show a much stronger resemblance to each other. More specifically, the normalized damping coefficient curves for all groups of records attain comparable peak values for a given damping ratio. These maximum values tend to be slightly closer than the peaks observed on the non-normalized damping coefficient curves displayed in Fig. 14. In addition, the normalized periods over which these peak values occur coincide at a value slightly lower than 1.0 on the normalized period axis, a statistic that had previously varied greatly by earthquake magnitude. The B curves illustrated in Fig. 15 may allow
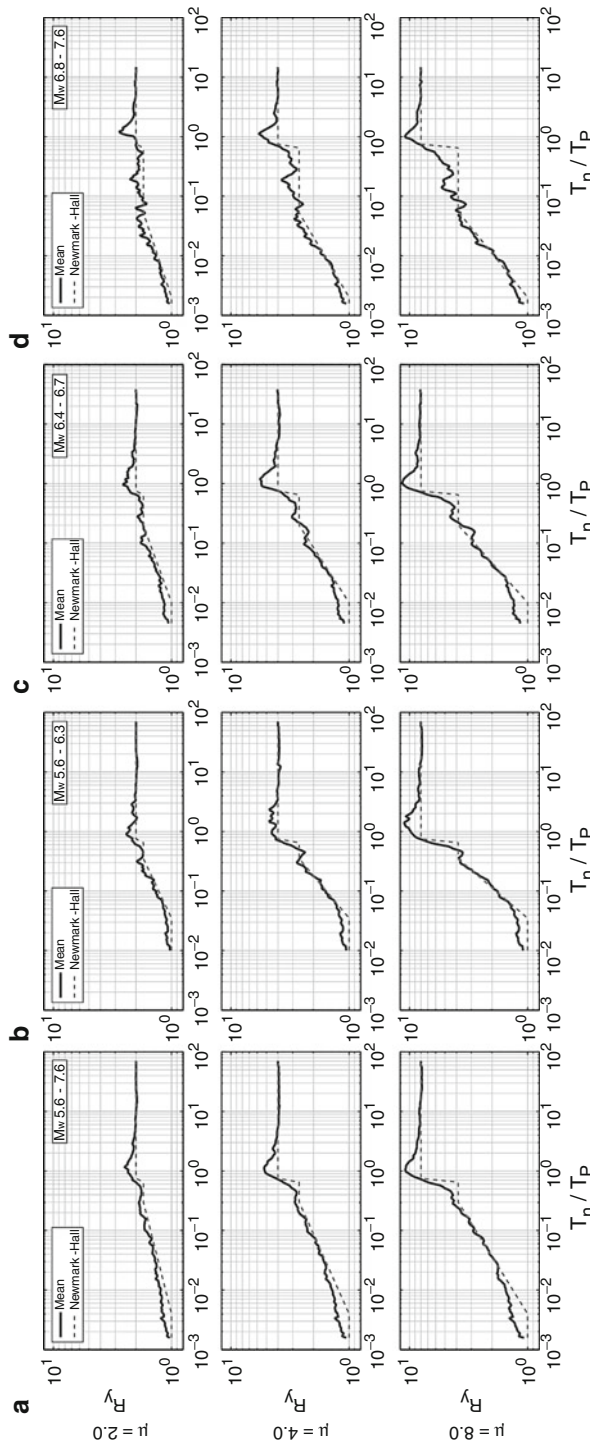
Seismic Actions Due to Near-Fault Ground Motion, Fig. 13 Comparison of mean values of the strength reduction factor with the Veletsos-Newmark-Hall design equations for: (**a**) all earthquakes (M_W 5.6–7.6), (**b**) moderate earthquakes (M_W 5.6–6.3), (**c**) moderate-to-large earthquakes (M_W 6.8–7.6); $\mu = 2.0$, 4.0, 8.0, and $\zeta = 5\%$ (Reprinted from Mavroeidis et al. (2004). Copyright © 2004 John Wiley & Sons, Inc.)
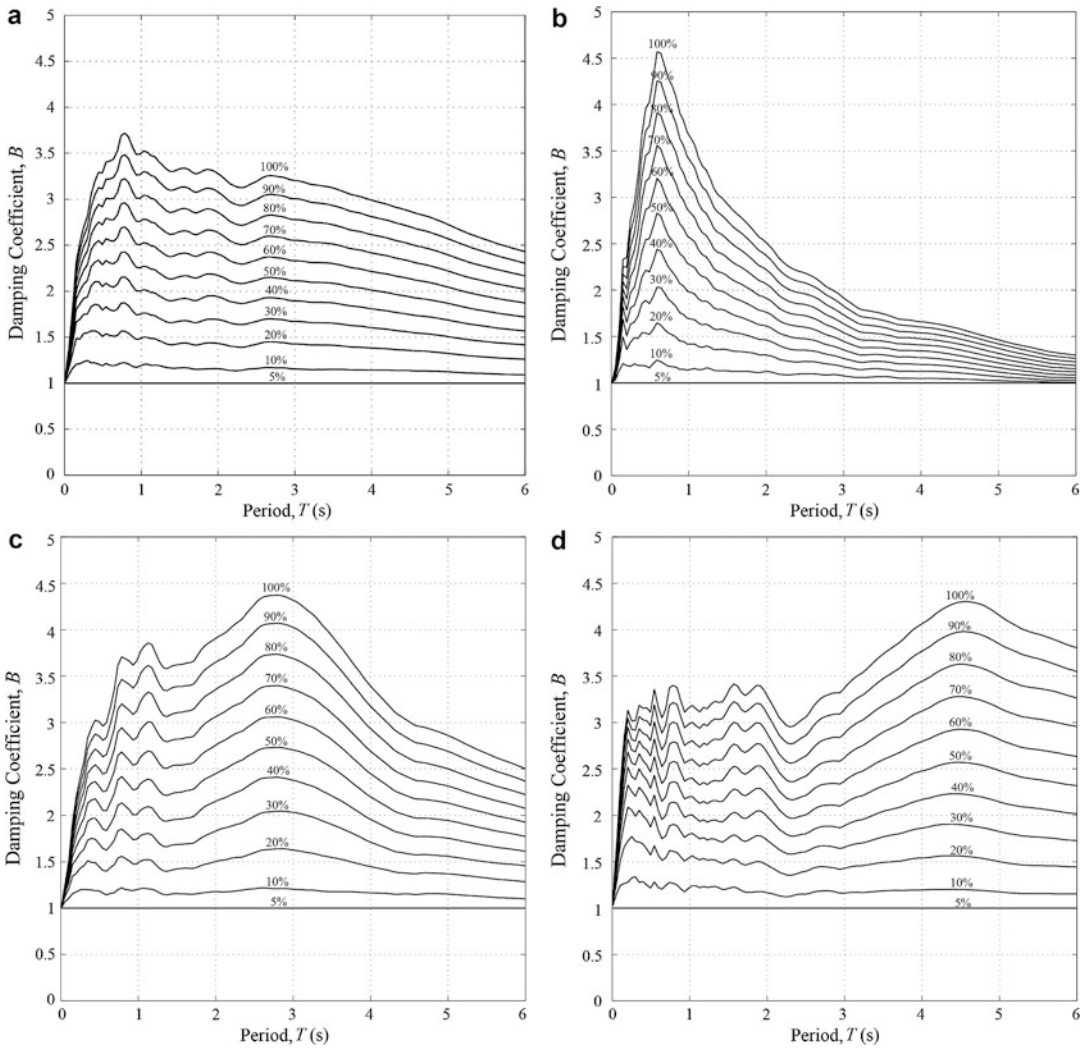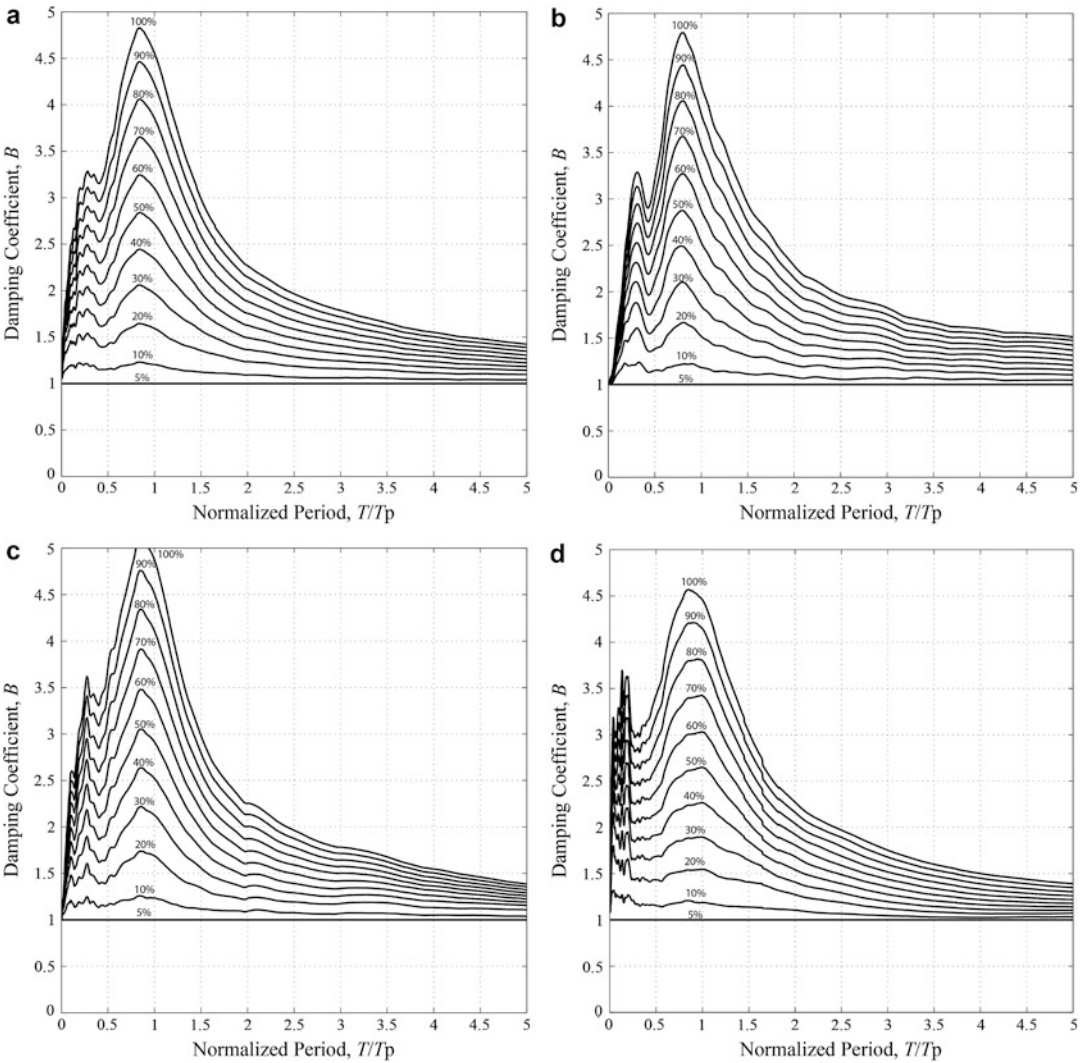
Seismic Actions Due to Near-Fault Ground Motion, Fig. 14 Calculated damping coefficients for near-fault ground motion records: (**a**) all earthquakes (M_W 5.6–7.6), (**b**) moderate earthquakes (M_W 5.6–6.3), (**c**) moderate-to-large earthquakes (M_W 6.4–6.7), and (**d**) large earthquakes (M_W 6.8–7.6) (Reprinted from Hubbard and Mavroeidis (2011). Copyright © 2011 Elsevier B.V.)

for a single set of empirical equations to represent near-fault damping coefficients that are normalized by T_P.

An empirical equation was developed by Hubbard and Mavroeidis (2011) to fit the main behavior of damping coefficients observed in the set containing all records (Fig. 15a). In order to effectively model these B curves, two equations were needed to describe different ranges of the damping ratio. Equation 6 is designed to be representative of the damping coefficients at normalized periods greater than ~0.83, a normalized period where the peak in value seems to occur. Below this characteristic normalized period, the damping coefficients linearly reduce to one at a normalized period of zero. Damping ratios are represented as fractions instead of whole numbers within the context of Eq. 6:

Seismic Actions Due to Near-Fault Ground Motion, Fig. 15 Damping coefficients for near-fault ground motion records with the period axis normalized with T_P: (**a**) all earthquakes (M_W 5.6–7.6), (**b**) moderate earthquakes (M_W 5.6–6.3), (**c**) moderate-to-large earthquakes (M_W 6.4–6.7), and (**d**) large earthquakes (M_W 6.8–7.6) (Reprinted from Hubbard and Mavroeidis (2011). Copyright © 2011 Elsevier B.V.)

$$B = 3.4 \frac{\beta^{1.3}}{(T/T_P)^{1.3}} + 1 \quad \text{for} \quad 0.10 \leq \beta \leq 0.50 \tag{6a}$$

$$B = 2 \frac{(\beta + 0.3)^{1.5}}{(T/T_P)^{1.3}} + 1 \quad \text{for} \quad 0.50 < \beta \leq 1.00 \tag{6b}$$

The empirical expression that was developed through this method is shown in Fig. 16 as a direct comparison to the computed damping coefficients for the set containing all records. For the considered ranges of damping ratio and normalized period, Eq. 6 provides a model that is conservative without exception. It also does an adequate job of capturing the shape of the curves, ensuring that there are no damping coefficients that are greatly overconservative. In addition, the model remains conservative and captures the behavior of the damping coefficient plots of the different earthquake magnitude groupings
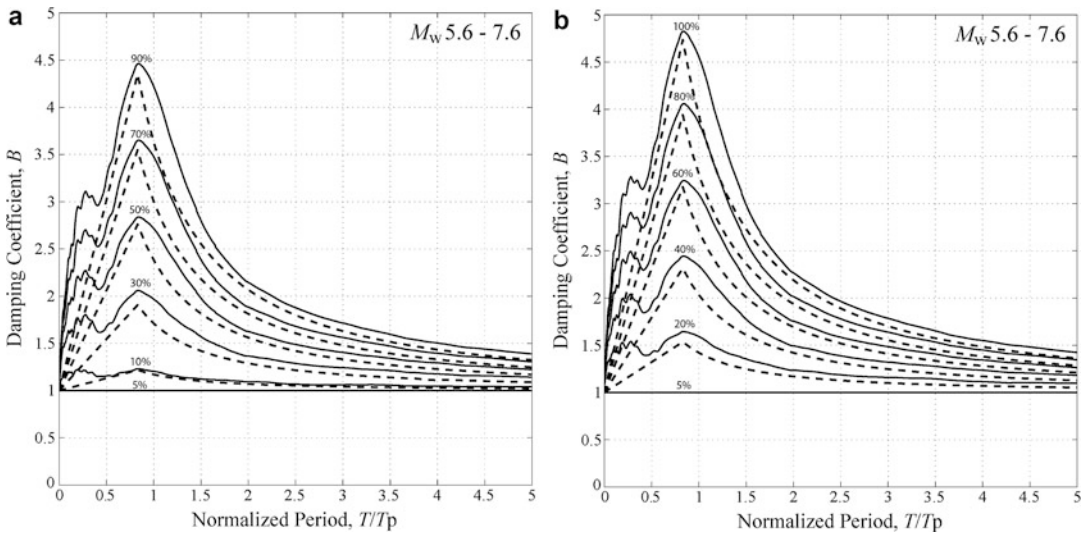
Seismic Actions Due to Near-Fault Ground Motion, Fig. 16 A comparison between the proposed method and the calculated damping coefficients with the period axis normalized with T_P (Reprinted from Hubbard and Mavroeidis (2011). Copyright © 2011 Elsevier B.V.)

of the dataset (Figs. 15b, c, d) as shown by Hubbard and Mavroeidis (2011).

Summary

This entry focuses on the description of seismic actions due to near-fault ground motions. Particular emphasis was given on synthesizing broadband near-fault ground motion time histories for earthquake engineering applications, using a simple mathematical model for the representation of the coherent ground motion component and a physical model of the seismic source for the description of the incoherent seismic radiation. In addition, recommendations on design spectra, strength reduction factors, and damping coefficients were made for engineering analysis and design in the near-fault region. This included the normalization of the period axis with respect to the period of the ground velocity pulses. The pulse period is controlled by the rise time on the fault plane and scales directly with earthquake magnitude.

Cross-References

▶ Earthquake Response Spectra and Design Spectra
▶ Engineering Characterization of Earthquake Ground Motions
▶ Physics-Based Ground-Motion Simulation
▶ Response Spectrum Analysis of Structures Subjected to Seismic Actions
▶ Response-Spectrum-Compatible Ground Motion Processes
▶ Selection of Ground Motions for Response History Analysis
▶ Stochastic Ground Motion Simulation
▶ Time History Seismic Analysis

S

References

Alavi P, Krawinkler H (2001) Effects of near-fault ground motion on building structures (Report). CUREE-Kajima Joint Research Program, Richmond

Archuleta RJ (1984) A faulting model for the 1979 Imperial Valley earthquake. J Geophys Res 89:4559–4585

Baker JW (2007) Quantitative classification of near-fault ground motions using wavelet analysis. Bull Seismol Soc Am 97:1486–1501

Bertero VV, Mahin SA, Herrera RA (1978) Aseismic design implications of near fault San Fernando earthquake records. Earthq Eng Struct Dyn 6:31–42

Bouchon M (1997) The state of stress on some faults of the San Andreas system as inferred from near-field strong motion data. J Geophys Res 102:11731–11744

Bray JD, Rodriguez-Marek A (2004) Characterization of forward-directivity ground motions in the near-fault region. Soil Dyn Earthq Eng 24:815–828

Halldorsson B, Papageorgiou AS (2005) Calibration of the specific barrier model to earthquakes of different tectonic regions. Bull Seismol Soc Am 95:1276–1300

Halldorsson B, Mavroeidis GP, Papageorgiou AS (2011) Near-fault and far-field strong ground motion simulation for earthquake engineering applications using the specific barrier model. J Struct Eng ASCE 137:433–444

Heaton TH (1982) The 1971 San Fernando earthquake: a double event? Bull Seismol Soc Am 72:2037–2062

Housner GW, Trifunac MD (1967) Analysis of accelerograms – Parkfield earthquake. Bull Seismol Soc Am 57:1193–1220

Hubbard DT, Mavroeidis GP (2011) Damping coefficients for near-fault ground motion response spectra. Soil Dyn Earthq Eng 31:401–417

Makris N (1997) Rigidity-plasticity-viscosity: can electrorheological dampers protect base-isolated structures from near-source ground motions? Earthq Eng Struct Dyn 26:571–591

Mavroeidis GP (2004) Modeling and simulation of near-fault strong ground motions for earthquake engineering applications. PhD dissertation, Department of Civil, Structural and Environmental Engineering, State University of New York at Buffalo, Buffalo

Mavroeidis GP, Papageorgiou AS (2002) Near-source strong ground motion: characteristics and design issues. In: Proceedings of the seventh U.S. national conference on earthquake engineering, Boston

Mavroeidis GP, Papageorgiou AS (2003) A mathematical representation of near-fault ground motions. Bull Seismol Soc Am 93:1099–1131

Mavroeidis GP, Papageorgiou AS (2010) Effect of fault rupture characteristics on near-fault strong ground motions. Bull Seismol Soc Am 100:37–58

Mavroeidis GP, Dong G, Papageorgiou AS (2004) Near-fault ground motions, and the response of elastic and inelastic single-degree-of-freedom (SDOF) systems. Earthq Eng Struct Dyn 33:1023–1049

Newmark NM, Hall WJ (1982) Earthquake spectra and design. Earthquake Engineering Research Institute, Berkeley

Papageorgiou AS (1997) Engineering seismology. In: Beskos DE, Anagnostopoulos SA (eds) Computer analysis and design of earthquake resistant structures. Computational Mechanics Publications, Southampton, pp 153–200

Papageorgiou AS, Aki K (1983a) A specific barrier model for the quantitative description of inhomogeneous faulting and the prediction of strong ground motion. I. Description of the model. Bull Seismol Soc Am 73:693–722

Papageorgiou AS, Aki K (1983b) A specific barrier model for the quantitative description of inhomogeneous faulting and the prediction of strong ground motion. II. Application of the model. Bull Seismol Soc Am 73:953–978

Sasani M, Bertero VV (2000) Importance of severe pulse-type ground motions in performance-based engineering: historical and critical review. In: Proceedings of the twelfth world conference on earthquake engineering, Auckland

Somerville P (2003) Magnitude scaling of the near fault rupture directivity pulse. Phys Earth Planet In 137:201–212

Vassiliou MF, Makris N (2011) Estimating time scales and length scales in pulse-like earthquake acceleration records with wavelet analysis. Bull Seismol Soc Am 101:596–618

Veletsos AS, Newmark NM (1960) Effect of inelastic behavior on the response of simple systems to earthquake motions. In: Proceedings of the second world conference on earthquake engineering, Tokyo

Seismic Analysis of Concrete Bridges: Numerical Modeling

Andreas Kappos
Department of Civil Engineering, City University London, London, UK

Synonyms

Concrete bridges; Modal analysis; Nonlinear analysis; Seismic loading

Introduction

Bridges are deceptively simple systems, since they are typically single-storey structures

wherein the horizontal members (the deck) can often be modeled either as a continuous beam or as a series of simply supported beams. In fact, the continuous beam might be a valid approximation in seismic analysis of bridges if a "spine" model is adopted; in such a model the bridge deck is simulated using 3D beam elements with 6 degrees of freedom (DOFs) at each node, located at the centroid of the cross section. It is worth noting that in gravity load analysis the geometric complexity of the deck is usually represented in the computer model to greater detail, as compared to that used for estimating the seismic response. On the other hand, bridges present peculiarities that are not commonly encountered (or are far less important) in buildings, such as the modeling of bearings, shear keys, and expansion joints, as well as the modeling of soil-structure interaction at all bridge supports, including those at the abutments which can play a major role in some cases.

In the sections that follow, modeling of the various bridge components is first addressed (section "Modeling of Bridge Components"), followed by an overview of methods currently used for seismic analysis of bridges (section "Bridge Analysis Methods"), which is a topic that is also addressed in other articles of the encyclopedia; section "Bridge Analysis Methods" also includes the presentation of a case study that illustrates the modeling and analysis procedures described in the previous sections, while both sections "Modeling of Bridge Components" and "Bridge Analysis Methods" include specific modeling examples and selected results. Finally some concluding remarks are provided in section "Summary and Concluding Remarks." Methods of seismic analysis are presented in various parts of the encyclopedia (e.g., article by Vayas and Iliopoulos (2014) focusing on modeling of steel and composite bridges). Hence, the focus herein is on analysis of concrete bridges, wherein some specific issues arise as discussed in the following; of course, several of the described models and techniques are also applicable to bridges made of other materials.

Modeling concrete bridges for seismic design purposes should always take into account the intended plastic mechanism, which, contrary to buildings, involves primarily yielding in the vertical members, i.e., the piers, while seismic energy dissipation can also take place in the bearings; in a small number of (important) bridges, supplementary damping devices (such fluid or friction dampers) are provided and have to be accounted for in the analysis. In practical design applications bridges are analyzed in the elastic range, and any inelasticity (material nonlinearity) effects, wherever entering of some members in the inelastic range under the design earthquake is allowed, are accounted for by simply reducing the design response spectrum by a "behavior" (or force reduction) factor. However, special types of bridges have been analyzed using advanced inelastic analysis tools (typically as a verification of an initial design based on the results of equivalent elastic analysis), and in general, inelastic analysis methods are gaining ground in recent years. Therefore, an attempt is made herein to present some basic concepts and models suitable for inelastic analysis, with emphasis on those that are better suited for practical application.

Modeling of Bridge Components

Deck

In concrete bridges the deck can have various forms, i.e., solid slab, voided slab, beams (usually precast I-beams) with cast in situ top slab, box girder (single-cell or multicell), and other, less common, ones. The material used is typically prestressed concrete, except for some short span bridges with slab-type deck, where ordinary reinforced concrete can be used.

Very important, both for the seismic behavior of the bridge and for its modeling, is the type of pier to deck connection. There are three basic options in this respect:

- Monolithic connection: Very common in slab bridges and box-girder bridges, especially when the cantilever method of construction is used in the latter.

S
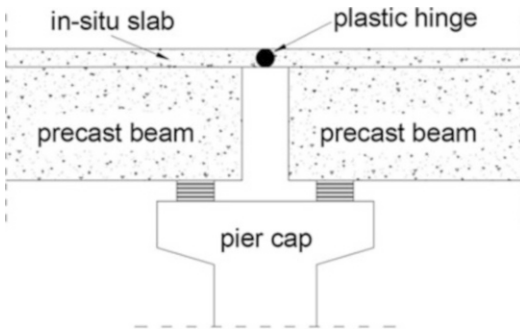
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 1 Pier region of a bridge with deck consisting of precast posttensioned beams and cast in situ slab continuous over the piers

- Bearing connection: The deck rests on the piers through two or more bearings, which in the case of modern concrete bridges are typically elastomeric bearings; this is the usual type of connection in the case of beam-type decks and also in box-girder bridges when the incremental launching method of construction is used.
- Mixed type of connection: Some piers (typically the taller ones) are monolithically connected to the deck, whereas others (typically the squat ones) have bearing connections. This is a fairly new solution, which has advantages when bridge configurations with substantially unequal pier heights have to be used (e.g., in ravine bridges).

In the case of bearing connections (very common in older concrete bridges), it is clear that the deck does not carry moments due to lateral (seismic) loading nor forms part of the energy dissipation mechanism of the bridge. An exception is the case of continuity slabs, i.e., the parts of a top slab in beam-type decks that continue over the piers (whereas beams terminate on each side of the pier), hence providing a continuous deck (Fig. 1). For these regions modern codes like Eurocode 8-2 (CEN 2005a) allow formation of plastic hinges (in bending about the transverse axis, see Fig. 1). When the focus of the analysis is on the response to the design earthquake and inelastic behavior is allowed and anticipated,

actual hinges can be introduced beforehand at these locations.

In the usual case that the deck remains elastic under the design seismic action, it can be modeled with elastic elements of any appropriate type: beam-column elements forming frames or grillages or even shell elements. For seismic analysis the recommended approach is the use of 3D beam columns (spine model), which is the simplest option and generally an adequate one for the purposes of this analysis. Relevant guidelines like those of ACI 341 (2014) and FHWA (2006) recommend four to five elements per span, but usually more elements are used in practice, since these are elastic members and do not noticeably affect the computational demands. Besides, prestressed concrete members like beams and box girders often have cross sections that vary along the span (thicker webs of beams and box girders toward the piers, where shear forces are maximum); hence, the 3D beam elements should be arranged in such a way that they properly reproduce this gradual change in geometry. Moreover, since masses required for the dynamic analysis of the bridge are typically lumped at the nodes of the model (even when uniformly distributed masses are automatically calculated from the geometry of the elements), use of a sufficient number of elements leads to a more accurate representation of the mass distribution in the bridge and hence of its dynamic characteristics.

In the case of *monolithic* connections, deck-to-pier joints carry bending moments due to seismic loading, and in principle, parts of the deck may become inelastic. However, in most cases the strength of the deck that is governed by the substantial gravity loading on the bridge plus traffic loads (very high in railway bridges) is clearly higher than that of the pier that is typically governed by seismic moments in medium to high seismicity areas; hence, the deck remains in the elastic range and the previous comments apply. In this case, accurate modeling of box girders requires accounting for shear lag (nonuniform distribution of the longitudinal stress across the flange width due to the shear deformations within the flange); the FHWA (2006) manual specifies that the flexural stiffness of the superstructure taken
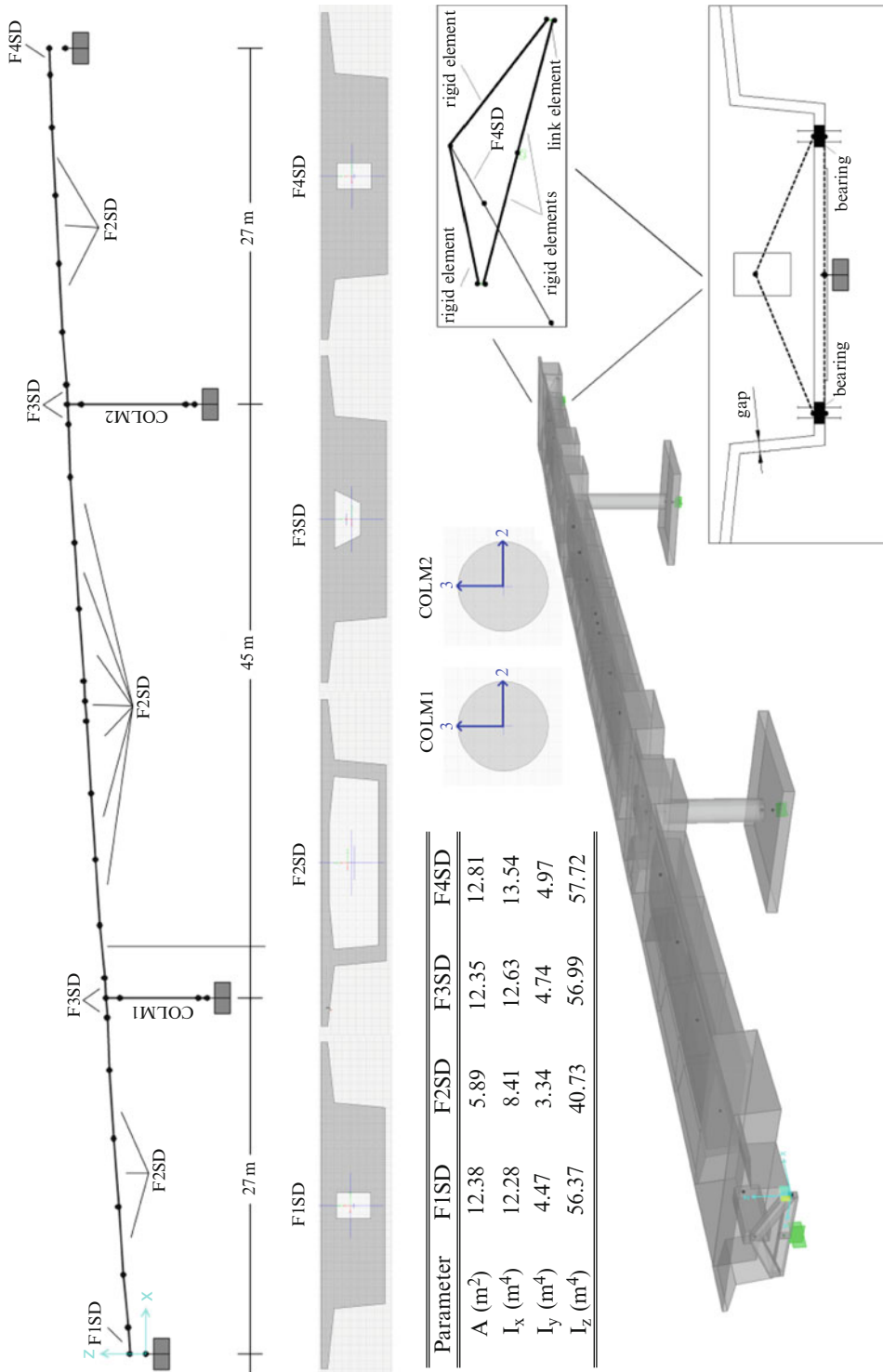
about a transverse axis should be reduced near piers when there is moment transfer between the superstructure (deck) and the pier, without providing values for this reduction. According to the ATC-32 report (ATC 1996), stiffness in these regions is based on an effective width that should be no greater than the width of the column plus twice the cap beam depth. If this width is practically equal to the entire width of the superstructure, no reduction in stiffness due to shear lag is required in the model.

An example is shown in Fig. 2 depicting the spine model of an overpass bridge (Kappos et al. 2013a) with monolithic pier to deck connections and free sliding connections at the abutments. It is seen that, depending on the length of the span, up to 11 3D beams have been used. The thickening of the webs of the box-girder section toward the end of the spans is properly modeled in the software used, SAP 2000 (CSI 2011). Figure 2 also conveys an idea of what the model would look like in case shell elements were used. As shown in a study by Kappos et al. (2002), it is possible to achieve a good match of the dynamic characteristics of a bridge modeled using shell elements, using a simple spine model; the match was particularly good (differences in significant natural periods between 4 % and 11 %) when reduced flexural rigidity was specified for the 3D beam elements close to the monolithic connection. That study also confirmed that it is not necessary to use too many elements for the deck; in fact a model involving 488 3D beam elements predicted almost identical dynamic characteristics (periods and mode shapes) as another one using only 77 elements (8 per span).

The value of *flexural rigidity* (EI) to be assigned to the elements used for modeling the deck depends on the material used. In the usual case of prestressed concrete, it is allowable to assume negligible cracking and use the value for the gross cross section (EI_g). In ordinary (non-prestressed) reinforced concrete decks, EI should account for cracking effects. Recommended values in Caltrans (2013) are 50–75 % of EI_g; the lower bound represents lightly reinforced sections and the upper bound represents heavily reinforced sections.

The results of seismic analysis in the transverse direction of a bridge with box-girder deck will be influenced by the assumption made regarding the *torsional* stiffness of the deck (which is substantial, as opposed to that of "open-type" orthotropic decks like beams with top slab). It is recommended to assume 20 % of the uncracked value, based on the ratios (10 ÷ 30 %) of cracked-to-uncracked torsional stiffness estimated by Katsaras et al. (2009).

In between the aforementioned simple (spine) and complex (shell) modeling approaches for the deck is the **grillage** model, i.e., a horizontal planar system of longitudinal and transverse 3D beams, as shown in the example of Fig. 3, referring to a ravine bridge with a top slab on posttensioned I-beams supported through laminated elastomeric bearings, studied by Ntotsios et al. (2009). The longitudinal members of the grillage have the properties of the I-beams and the tributary part of the slab, the transverse elements above the piers have the properties of the actual transverse beams in the actual bridge, while the intermediate four transverse elements of the grillage represent the coupling of the longitudinal beams in the transverse direction due to the presence of the deck. Grillage models are particularly suited for slabs or orthotropic decks like that of Fig. 3 and provide a good balance between accuracy and practicability. They are not very easy to set up nor offer particular advantages over the spine model in the case of box-girder sections. An important issue here is the modeling of the torsional stiffness of the deck which cannot be estimated from the properties of the individual elements of the grillage but rather should be derived for the entire box girder and then distributed among the longitudinal members. In case the grillage model is used for solid slab decks, it is recommended to add diagonal braces to account for the interaction between longitudinal and transverse deck action due to Poisson effects, which cannot be neglected (as in the case of orthotropic decks with primary girders, see Vayas and Iliopoulos (2014) for orthotropic composite decks). More details on the input parameters required for deck modeling with grillages (flexural and torsional rigidities, shear areas)

S

Parameter	F1SD	F2SD	F3SD	F4SD
A (m²)	12.38	5.89	12.35	12.81
I_x (m⁴)	12.28	8.41	12.63	13.54
I_y (m⁴)	4.47	3.34	4.74	4.97
I_z (m⁴)	56.37	40.73	56.99	57.72

Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 2 Finite element modeling of an overpass bridge with a box-girder superstructure monolithically connected to the piers (Kappos et al. 2013)
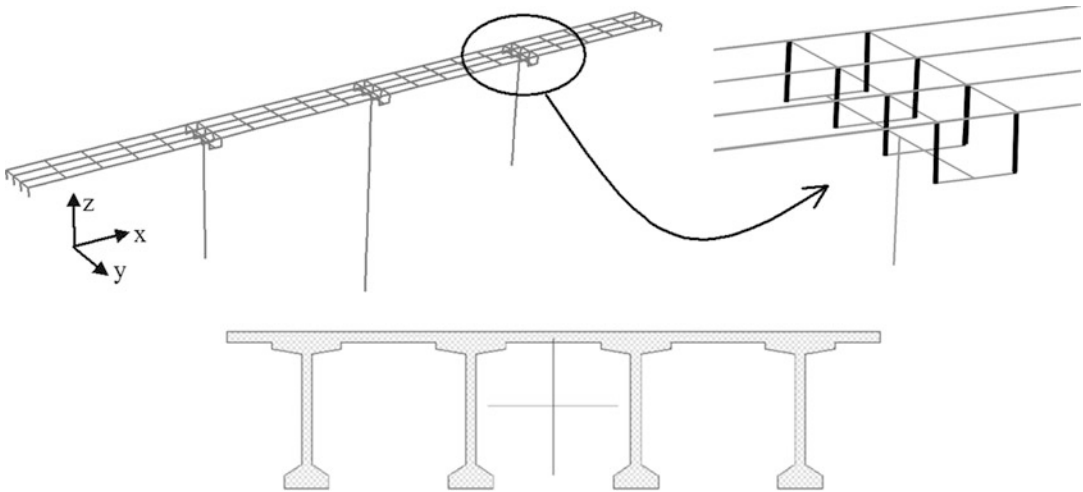
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 3 Grillage model and detail of the connection between the deck and the pier (*above*) for a ravine bridge with beam with top slab deck (*below*)

can be found in Chapter 2, of the recent book by Kappos et al. (2012). In that chapter a brief discussion of some advanced topics like the effects of skewness and curvature in plan and/or in elevation on the seismic response, and the verification of deck deformation demands when this is necessary for seismic assessment, can also be found.

Piers and Their Foundations

Piers commonly used in concrete bridges are of the following types:

- Single columns with solid circular or (less frequently) rectangular section
- Single columns with hollow circular or rectangular section
- Multicolumn bents (frame-type piers, running in the transverse direction of the bridge)
- Wall-type piers (usually of large dimensions, especially in the transverse direction)
- Braced piers (usually V shaped)

It is noted that hollow rectangular piers are typically of large dimensions (could reach 7 m or more in the transverse direction), and their structural behavior is closer to that of walls, rather than of hollow circular columns.

Geometric Considerations

Single-column piers, whether solid or hollow, are generally modeled as "sticks" using beam-column elements; four to five elements are usually enough, unless complex geometries (e.g., flared columns) are involved. As shown in Fig. 4, the top element of the column is connected via a rigid link to the centroid G of the deck section (which is the location of the corresponding horizontal element); in most software packages there is no need for specifying a different element, but rather the top element of the column extends up to G, and its end portion is specified as a rigid offset. In the common case of columns monolithically connected to the box girder of the deck, there is no need for additional elements in the column, but when the box girder is supported on bearings (which will be the case at the abutments), it is necessary to introduce transverse elements extending from the first to the last bearing and link these to the end node of the deck model using rigid links (see bottom right of Fig. 2).

As shown in Fig. 4, the moment distribution along the height of the pier is influenced by the rotational restraint at the top, which in turn depends on the torsional rigidity of the deck. This is substantial in the case of box girders
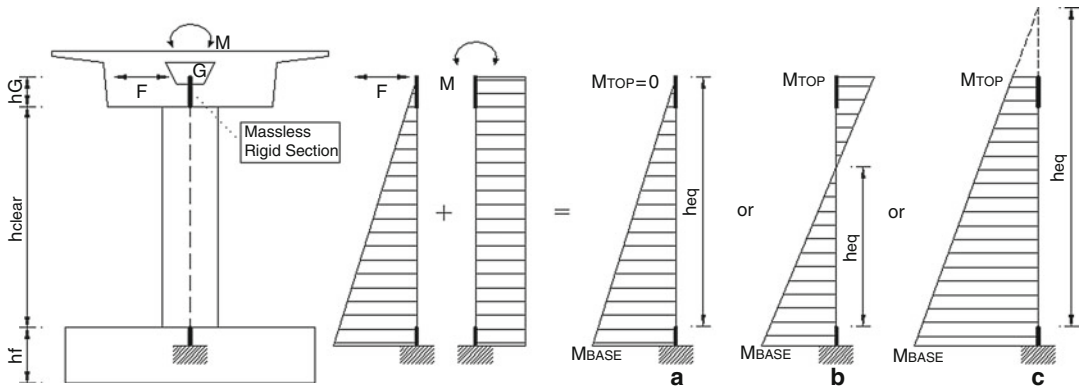
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 4 Pier modeling and transverse response accounting for the torsional stiffness of the deck

(as in Fig. 4) and low in the case of precast beams with top slab and other similar, "open-type," sections.

Multicolumn bents are naturally modeled as 2D frames, again with beam-column elements, each of which has the properties of the corresponding member (column or cap beam). It is worth noting that although columns usually have circular sections, cap beams are rectangular. Rigid offsets at the element ends properly capture the effect of the finite size of the beam-column joints that are quite massive members in bridges. In case the cap beam is monolithically cast with the deck, the torsional resistance of the top of the bent is substantially higher than that of the cap beam alone; Aviram et al. (2008) recommend multiplying the torsional resistance of the cap beam by 10^2. Clearly this does not apply in the case that the deck is bearing-supported on the cap beam. *V-shaped* piers are modeled in a similar way, but of course, vertical elements are inclined rather than upright.

There are several options available for modeling *wall-type* piers and the aspect ratio is an important parameter in this case. The simpler model is clearly the stick one, previously mentioned for the case of single-column piers. Decks are usually bearing-supported on wall-type piers (which are the preferred solution in seismic isolation designs wherein seismic energy dissipation takes place in the bearings and, whenever present, the dampers); hence, it is essential in this case to add the aforementioned horizontal rigid element at the top of the stick model. A more refined model could consist of a vertical grillage (see notes on grillage models in section "Deck"). The most refined model, feasible only in elastic analysis, is the use of shell elements; an example (not necessarily a recommended one, unless for research purposes) is shown in Fig. 5; a mesh of 18,885 shells was set up for the 3-span bridge, keeping the aspect ratio of the shells rather low at 1.2, since these elements perform best when their shape is close to square. As noted in section "Deck," the main normal modes of that bridge were captured with reasonable accuracy using a simple spine model with only 77 beam-column elements (Kappos et al. 2002).

Stiffness Considerations

Even in elastic analysis of reinforced concrete (R/C) piers for seismic loading, it is essential to account for the effect of cracking, to make sure that displacements are not underpredicted. Practically all existing codes adopt approximate values of the pier stiffness, corresponding to yield conditions, and this stiffness is assumed as known when design seismic actions (e.g., modal forces) are estimated. These approximate values are either very rough estimates, like the $0.5\,EI_g$ (50 % of uncracked section rigidity) adopted by both Eurocode 8-1 (CEN 2004) and AASHTO (2010), or slightly more sophisticated ones taking into account the level of axial loading on the pier

Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 5 Finite element (*shell*) mesh in the pier to deck connection area of a bridge with box-girder superstructure monolithically connected to hollow rectangular piers (Kappos et al. 2002)
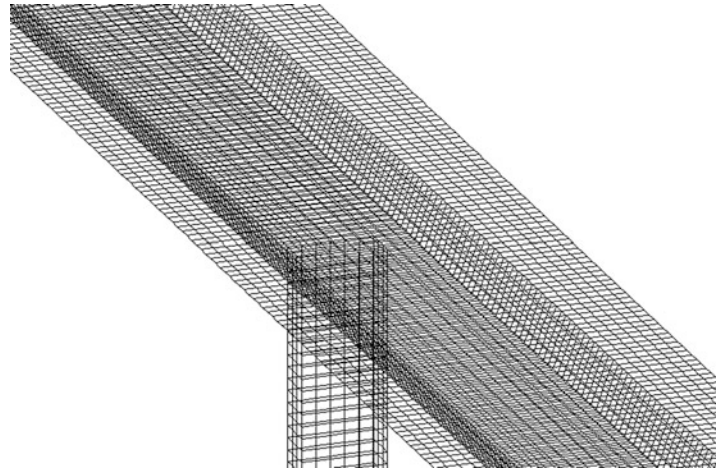
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 6 Effective stiffness of cracked reinforced concrete circular sections (AASHTO 2010)
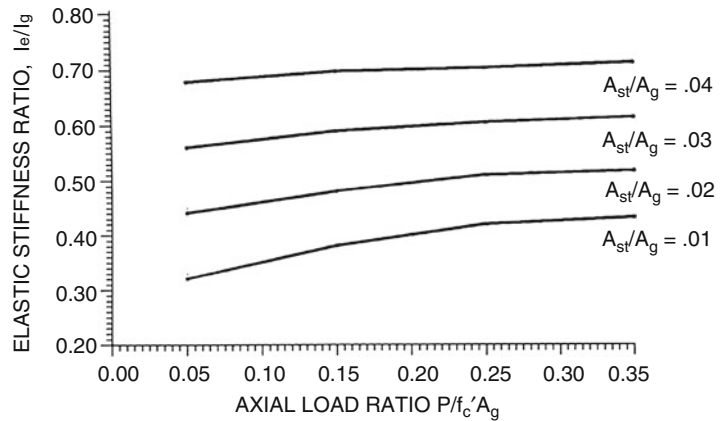
(which, in general, is not significantly affected by seismic actions) and/or the reinforcement ratio.

Eurocode 8-2 for Seismic Design of Bridges (CEN 2005a) in its (informative) Annex C suggests the following relationship for the effective moment of inertia of R/C ductile columns:

$$I_{eff} = 0.08\,I_g + I_{cr} \qquad (1)$$

where the cracked section inertia can be calculated as the secant value at yield (M_y is the yield moment and φ_y the yield curvature, E_c the concrete modulus)

$$I_{cr} = M_y / \left(E_c \cdot \varphi_y \right) \qquad (2)$$

Obviously, I_{cr} can only be estimated from Eq. 2 when the pier has been designed, so that both strength and yield curvature can be calculated; hence, use of these relationships is feasible only when iterative elastic analyses, or inelastic analysis, are used.

The Caltrans Seismic Design Criteria (2013) adopt the same concept as EC8-2 (secant value at yield), the only exception being that the 0.08 I_g term (accounting for tension stiffening effects) is not included in Eq. 1. As an alternative, the Caltrans Criteria allow the calculation of effective stiffness as a function of the axial load ratio and the pier reinforcement ratio from graphs provided by Priestley et al. (1996); diagrams like that of Fig. 6 can be directly implemented for carrying

S
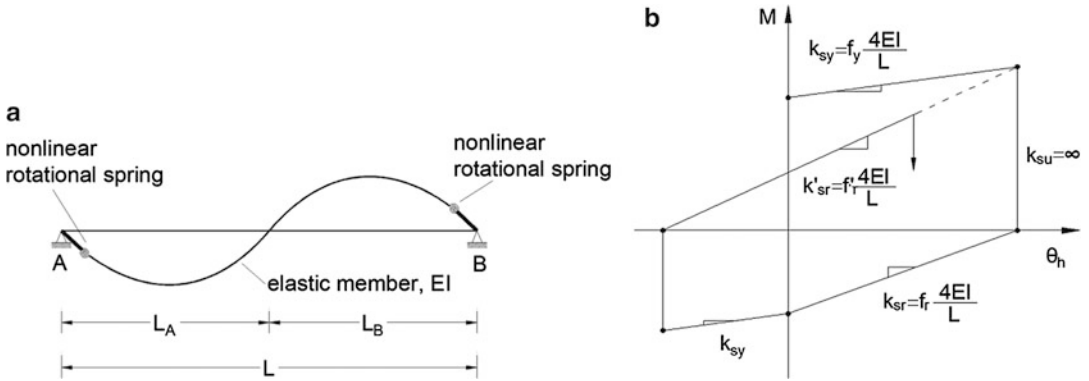
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 7 Lumped plasticity model (**a**) basic configuration; (**b**) hysteresis law for springs

out elastic analysis, assuming a reasonable reinforcement ratio (e.g., $A_{st}/A_g = 0.01$, which is the usual minimum reinforcement ratio), while, in principle, analysis should be repeated if the resulting reinforcement is substantially different.

Nonlinear Models for Piers

As mentioned in section "Introduction," the most common plastic mechanism on which seismic design is based is that involving inelastic response of the piers. Therefore, inelastic (material nonlinear) modeling of piers is important, not only for research but also for practical assessment purposes (which includes assessing an existing, probably substandard, bridge as well as a newly designed bridge that is important enough to warrant this additional design effort). Space limitations do not allow a detailed treatment of this important issue, and only a brief overview of the main available models will be provided herein; more detailed information on nonlinear modeling of bridge piers and several case studies can be found i.a. in the recent book by Kappos et al. (2012).

Nonlinear models for piers can be classified into three categories:

- Lumped plasticity models
- Distributed plasticity models
- Continuum models

Several subcategories can be defined for each of the above, as briefly discussed in the following.

Lumped plasticity models, also known as "point-hinge" models, are based on the simplifying assumption that all inelastic behavior takes place at the plastic hinge points that are typically located at the member ends. This concept can be materialized in different ways, the most efficient one consisting in inserting two nonlinear rotational springs at the element ends, as shown in Fig. 7a; more accurately, the springs are inserted at the ends of the rigid offsets located at the element ends to model the finite width of joints (e.g., between the cap beam and the column, in a multicolumn bent). All post-yield flexural deformation takes place in these springs, whereas the remainder of the beam-column element remains elastic throughout; it is emphasized that the flexural rigidity EI of the "elastic" element should account for cracking effects, as discussed in section "Stiffness Considerations." The (local) stiffness matrix for this lumped plasticity element, relating chord rotations at the ends to the corresponding bending moments, can be readily set up assuming a series connection between the springs and the beam, hence adding the flexibility matrices of each component, i.e.,

$$[F] = [F_e] + [F_S] = \begin{bmatrix} \dfrac{L}{3EI} + \dfrac{1}{K_{S_i}} & -\dfrac{L}{6EI} \\[2ex] -\dfrac{L}{6EI} & \dfrac{L}{3EI} + \dfrac{1}{K_{S_j}} \end{bmatrix}$$

(3)

where K_{si} and K_{sj} are the stiffnesses of the springs at ends i and j, which are assumed for simplicity to be uncoupled (no off-diagonal terms K_{sij}) and can be different, e.g., when one end has yielded while the other is still in the pre-yield range. The main advantage of this simple model is that K_s values can be defined on the basis of any constitutive law, whether simple or complex. Figure 7b shows a typical moment versus rotation hysteresis law (with stiffness degradation) that can be applied to either spring; note that prior to exceeding the yield moment the springs are rigid ($K_s = \infty$), hence all elastic deformation takes place in the elastic member. The 2×2 flexural stiffness matrix relating end moments to chord rotations can be easily derived by inverting [F] and then transforming to the 6×6 matrix including the rigid body modes and the axial deformations (axial stiffness EA is usually assumed to remain unaffected by flexural yielding).

Bridge piers, in particular those of the wall type, develop significant shear deformations subsequent to shear cracking, which may occur before or after flexural yielding. In the lumped plasticity context, shear deformations can be treated either in a simplistic way by using the stiffness matrix of a Timoshenko beam (involving GA' terms in addition to EI ones) and a rough reduction factor for GA', or, more rigorously, by introducing additional springs at the ends representing the relationship between shear force and shear deformation; the issue of inelastic shear is discussed later on in relation to another model. Another important source of deformation in R/C piers, especially those with substandard detailing with respect to earthquake requirements, is *bond slip*, which can give rise to substantial local rotations at the member ends ("fixed end rotations"). This effect can either be modeled indirectly by decreasing the stiffness of the M – θ law (Fig. 7b) at the member ends, something that requires proper calibration, or be modeled directly by introducing additional rotational springs to the model of Fig. 7a.

Despite their crudeness, lumped plasticity models (first develop in the 1960s) remain quite popular due to their simplicity, as well as the fact

that they can relatively easily account for the effects of shear, as well as of bond slip, and are easier to calibrate against experimental results than other, more sophisticated (and complex), models.

Distributed (or spread) plasticity models, still of the beam-column type, drop the assumption of point hinges and directly account for the spread of inelasticity along the bridge member, hence leading, in principle, to more accurate results. There are several different approaches in this respect, i.e., inelastic response can be monitored at several predetermined sections of the element and the stiffness matrix be synthesized on the basis of the tangent stiffness of each such section, or variable length plastification zones be defined, typically at the member ends, assuming the rest of the element is quasi-elastic (as in lumped plasticity models). The latter option retains some of the simplicity of point-hinge models while being more rational and, in principle, accurate and will be described in the following.

A recent spread plasticity model (Mergos and Kappos 2012) accounting for inelastic response in all mechanisms (flexure, shear, bond slip) is shown in Fig. 8. The length of the plastified zones at the ends (respective rigidities EI_A and EI_B) is defined on the basis of the moment diagram of the element and the corresponding yield moments, e.g., the left zone has a length $\alpha_A \cdot L$, where

$$\alpha_A = \frac{M_A - M_{yA}}{M_A - M_B} \leq 1 \qquad (4)$$

The flexural stiffness matrix can be set up using the principle of virtual work for the case of members with variable cross section. The current rigidities are calculated from the moment versus curvature relationship at each member end. Models with more than three parts have been proposed but are not deemed appropriate for practical application.

A similar procedure can be followed in the case of the shear sub-element (Fig. 8d), which represents the hysteretic shear behavior of the R/C member prior and subsequent to shear cracking, flexural yielding, and yielding of the shear reinforcement. In this case, the current shear

Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 8 Distributed plasticity model: (**a**) geometry of R/C member; (**b**) beam-column finite element with rigid arms; (**c**) flexural sub-element; (**d**) shear sub-element; (**e**) anchorage slip sub-element
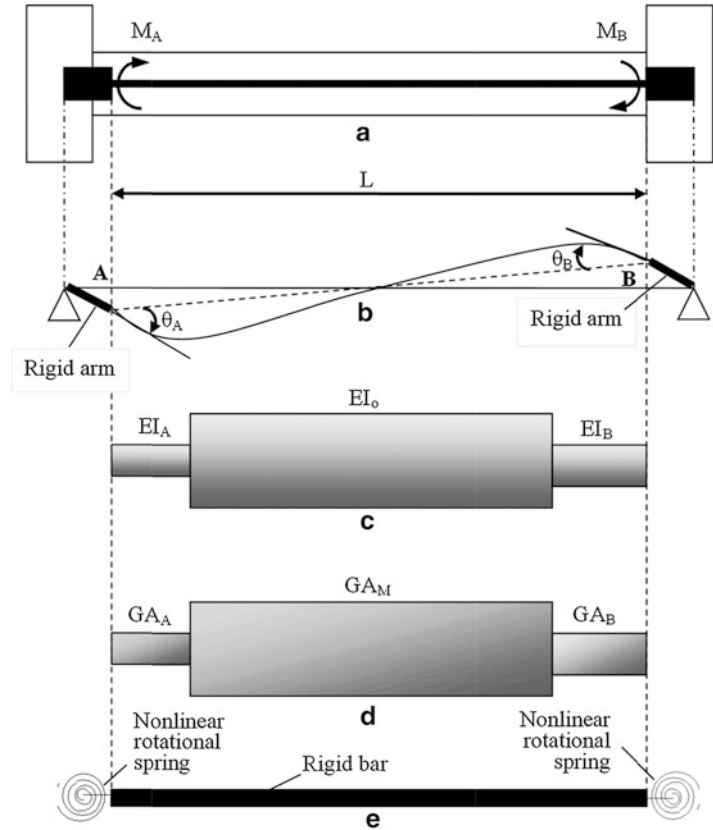
rigidities (GA_A, GA_B) are calculated from the $V - \gamma$ (shear versus shear deformation) curves at each end, details of which are given in Mergos and Kappos (2012). After determining the distribution of GA along the R/C member at each step of the analysis and by applying the principle of virtual work, the coefficients of the flexibility matrix of the shear sub-element are given by the following equation:

$$f_{ij}^{sh} = \frac{a_{As}}{GA_A \cdot L} + \frac{1 - a_{As} - a_{Bs}}{GA_M \cdot L} + \frac{a_{Bs}}{GA_B \cdot L} \quad (i, j = A, B)$$

(5)

Finally, rotations due to slip at the end anchorages are captured with the simple slip sub-element of Fig. 8e that consists of a rigid bar with two uncoupled nonlinear springs at the ends. The M-θ_{slip} skeleton curve is derived assuming uniform bond stress along different segments of the anchored reinforcement bar (details in Mergos and Kappos 2012).

In the previously presented models, the element stiffness matrix in the post-yield range is set up on the basis of EI values that are estimated from predefined constitutive laws relating bending moment to either end rotation or end curvature; when axial (EA) and shear (GA) rigidities are not taken as constant, their values are estimated from similar predefined laws (e.g., $V - \gamma$). Another option is the *fiber model*, wherein the stiffness parameters are not estimated from predefined laws, but rather from moment – curvature analysis (and, far less often, shear force versus deformation analysis) of a number of "monitoring" or "control" sections, which are divided in a number of "fibers" (in the general, biaxial, case, these are squares or rectangles rather than horizontal fibers), using Bernoulli's principle and the stress-strain constitutive laws of the pertinent materials, i.e., confined and unconfined concrete (for the core and the cover, respectively) and steel bars, as shown
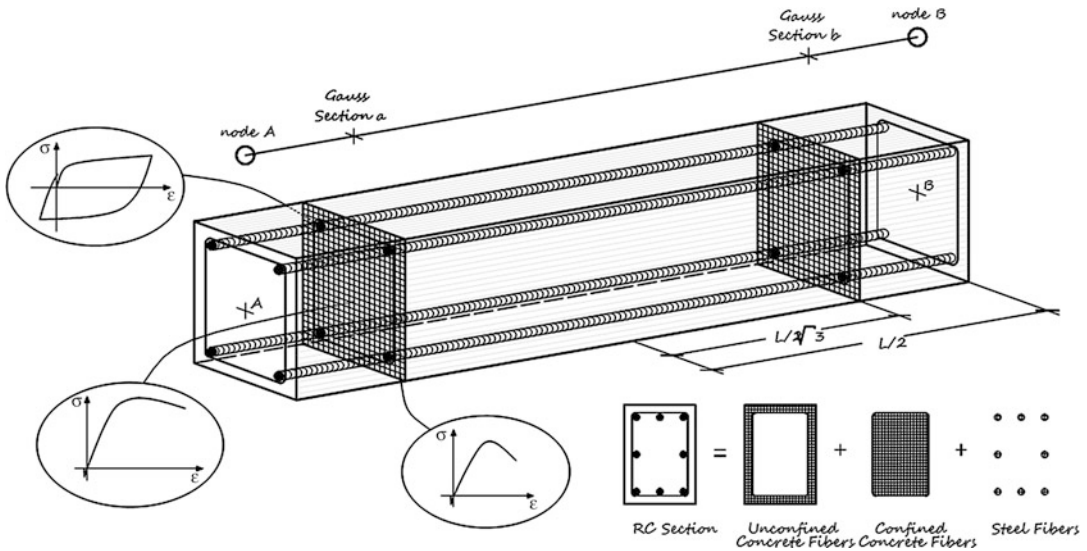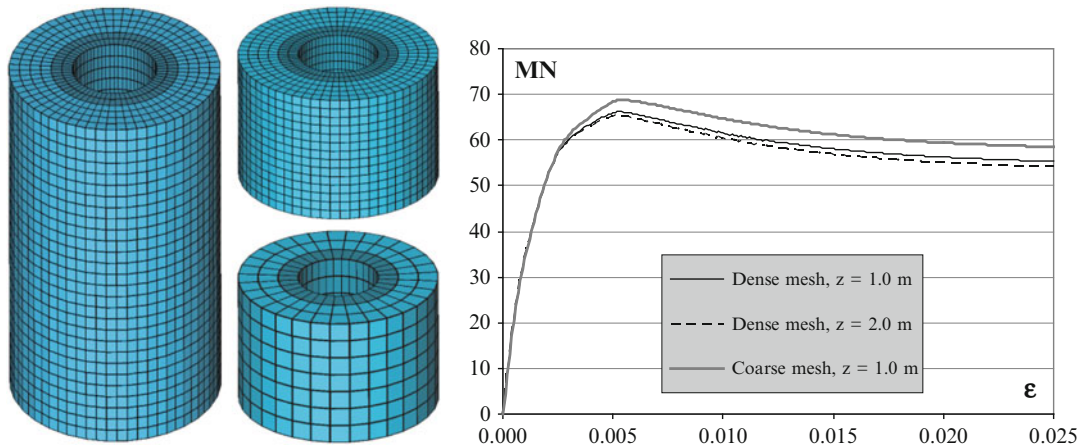
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 9 Discretization of R/C member using the fiber approach

in Fig. 9 (Kappos et al. 2012). The model has been implemented in point-hinge models but has found its main use in distributed plasticity models, wherein four (Fig. 9) or more sections are "monitored" and the element stiffness matrix is set up assuming linear variation of stiffness (or flexibility) between monitoring sections; the latter are typically taken as the end sections and the Gauss points used in the integration required for deriving the element stiffness matrix (e.g., using the Gauss-Lobatto quadrature scheme).

A rigorous application of the fiber model involves a number of difficulties, and different approaches have been put forward, some of them based on the stiffness approach (and involving displacement shape functions) and others on the flexibility approach (involving force shape functions, which do not change in the inelastic range); hybrid procedures have also been used. The flexibility approach is numerically more advantageous but computationally more demanding. Details of all these procedures fall beyond the scope of this entry and can be found in the literature (e.g., Fardis 1991; Kappos et al. 2012). However, it has to be emphasized here that although the fiber approach is more rigorous than the "phenomenological" approaches based on predefined force – deformation relationships,

it is not necessarily more accurate than the latter, except in the case of R/C members with negligible effect of shear and bond-slip deformations, which is not very common, even in well-designed bridge members. For instance, a pier usually has such an aspect ratio that shear deformations cannot be ignored. There are versions of the fiber model wherein shear deformations are included (Kappos et al. 2012), but the computational cost involved is particularly high. Finally, because most of the available fiber model-based software packages (like SeismoStruct) include constitutive laws assuming that concrete is initially uncracked, they overestimate the actual stiffness of real bridges (and other structures) that are cracked prior to being subjected to seismic loading (due to shrinkage, traffic and ambient vibrations, and possible previous small earthquakes).

Nonlinear *continuum models* are typically used for research purposes, whose aim is to study in detail the response of critical regions in piers. Material nonlinearity is taken into account using either a "standard" plasticity model or a combination of plasticity and damage models, the latter being able to affect the elastic component of the deformation as (seismic) damage propagates. Geometric nonlinearity is less critical in R/C piers, except for very tall ones. The type of

Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 10 Modeling of hollow circular piers with solid elements using different heights and mesh densities

finite elements used can be quire "heavy," i.e., shell and 3D solid ("brick") elements have been used for concrete piers. Figure 5 shows an example of using shell elements for both the deck and the piers, but nonlinear behavior was not taken into account in that model.

Figure 10 is taken from a study (Papanikolaou and Kappos 2009) focusing on the effect of confinement on the strength and ductility of solid and hollow piers; solid elements were used for concrete, whose behavior was governed by a sophisticated plasticity model accounting for confinement effects, while line elements embedded to the solid elements were used for the reinforcement (transverse reinforcement consisted of spiral or hoop circular reinforcement, with or without transverse links). It is seen that the axial load versus axial deformation curves resulting from the "coarse" mesh are not substantially different from those from the dense mesh; it is worth pointing out that an upper limit of about 6,000 solid elements was found, beyond which the computational cost and volume of results were excessive, which is a good indicator of the type of models that can be analyzed in such a context. Further examples of applications of continuum models to concrete pier components can be found in Kappos et al. (2012).

Soil-Structure Interaction Effects

A key aspect in proper modeling of piers (whether inelastic or equivalent elastic) is capturing the effect of foundation compliance, i.e., of the fact that pier foundations, especially the shallow but also the deep ones, like piles, do not provide full fixity to the base of the pier, but displace (horizontally), rotate, and even settle (vertically), as the bridge is subjected to seismic loading. Large-size foundations have also the effect of modifying the seismic input to the structure ("kinematic" interaction or "wave scattering" effect), but this issue is not further addressed herein. The interaction of the foundation ground with the bridge substructure (the piers and their foundations) does not only modify the fixity conditions but also increases the damping ("radiation" damping at the ground-foundation interface); in simplified analysis this can be safely ignored, but there are ways to explicitly include it, such as the addition of dashpot elements at the base of the piers.

A commonly adopted practical approach for calculating the pseudo-static interaction between the bridge foundation and the soil is the *Winkler spring* model, wherein the soil reaction to the foundation movement is represented by independent (linear or nonlinear) unidirectional translational spring elements. In the case of surface foundations, the vertical springs are distributed below the surface of the footing, while in pile foundations horizontal springs are distributed along the pile shaft. Although approximate, Winkler formulations are widely used not only

because their predictions are in good agreement with results from more rigorous solutions but also because the variation of soil properties with depth can be relatively easily incorporated. Moreover, they are efficient in terms of computational time required, thus allowing for easier numerical handling of the structural inelastic response, wherever this is deemed necessary.

In the case of *surface foundations*, a simple system of three translational (two horizontal, x and y, one vertical, z) and two (less often three) rotational springs can be used at the base of the footing; the spring constants for the x, y, and z springs can be estimated from relationships (ASCE 2007) derived from the solution of the problem of a rigid plate resting on the surface of a homogeneous half-space:

$$
\begin{aligned}
K_x &= \frac{GB}{2-\nu}\left[3.4\left(\frac{L}{B}\right)^{0.65} + 1.2\right] \\
K_y &= \frac{GB}{2-\nu}\left[3.4\left(\frac{L}{B}\right)^{0.65} + 0.4\frac{L}{B} + 0.8\right] \quad (6) \\
K_z &= \frac{GB}{2-\nu}\left[1.55\left(\frac{L}{B}\right)^{0.75} + 0.8\right]
\end{aligned}
$$

where G is the shear modulus of the ground, ν the Poisson ratio (0.35 for unsaturated soils and 0.5 for saturated soils), L the larger dimension of the (rectangular) footing, and the smaller one; similar relationships are given in ASCE (2007) and FHWA (2006) for the rotational springs. These relationships are particularly convenient to use since they only include very fundamental properties of the soil, which can always be estimated (e.g., the initial modulus G_0 can be estimated from the shear wave velocity and the specific weight of the soil). However, an upper and lower bound approach to defining stiffness and (in nonlinear models) capacity is recommended because of the uncertainties in the soil properties and the static loads on the foundations of existing bridges. The large-strain effective shear modulus, G, can be roughly estimated on the basis of the anticipated peak ground acceleration (PGA); for regions of low-to-moderate seismicity, a value of $G = 0.5\,G_0$ is recommended, while

for regions of moderate-to-high seismicity, $G = 0.25\,G_0$ is suggested (FHWA 2006).

In *pile foundations* the mechanical parameters for the springs are frequently obtained from experimental results (leading to P-y curves for lateral and N-z curves for axial loading) as well as from very simplified models. A commonly used P-y curve is the lateral soil resistance versus deflection relationship proposed by the American Petroleum Institute (API):

$$
P = 0.9p_u\tanh\left[\frac{kH}{0.9p_u}y\right] \quad (7)
$$

where p_u is the ultimate bearing capacity at depth H, y is the lateral deflection, and k is the initial modulus of subgrade reaction. The tip of the pile can either be modeled using a vertical spring (with a carefully selected axial stiffness) or assumed to be vertically fixed but free to rotate.

Figure 11 shows an example (Kappos and Sextos 2001) of modeling a pile group using Winkler springs with initial properties calculated from Eq. 7. This is a fully inelastic model wherein both the piles and the pier (modeled using the lumped plasticity approach, but with several elements for each pile) can yield; in practical applications equivalent linear properties (based on secant stiffness at the estimated maximum displacement) are often used, especially for the soil. The limitations of the P-y approached are discussed in detail in Kappos et al. (2012), where an overview of more advanced soil-bridge interaction (kinematic and inertial) models is also provided. It will only be mentioned here that despite the abundance of models and software, nonlinear analysis wherein both the bridge and the foundation ground are modeled with nonlinear models is not only cumbersome but often leads to convergence problems; hence, explicit treatment of nonlinearity should be confined where the main interest lies, typically in the piers and perhaps in the piles (cf. Fig. 11).

Abutments and Backfills

There are two common types of bridge abutments:
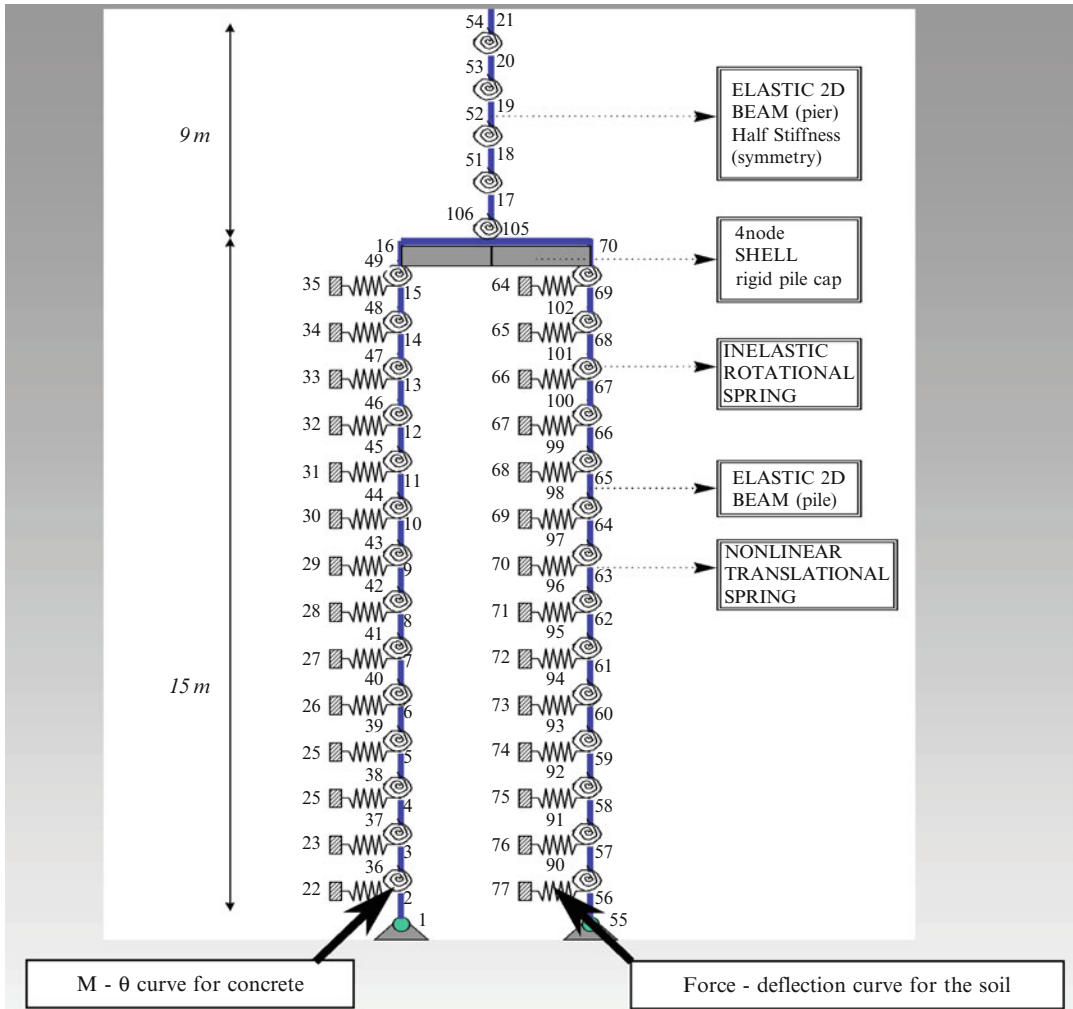
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 11 Modeling of soil-pile foundation-pier system using Winkler springs

- Seat type (the deck is bearing-supported on the horizontal seat of the abutment; see Fig. 2 bottom right)
- Integral or diaphragm type (abutment mono-lithically connected to the end of the deck)

The modeling of the abutment system, which also includes the wing walls and the foundation (footing or piles) can vary from very simple to very complex, depending on the situation. Some common cases are briefly discussed in the following.

Integral- or *diaphragm*-type abutments are always included in the model of the bridge, as a continuation of the deck model; in addition to the body of the abutment, they also include the foundation, which typically consists of piles (often relatively flexible ones). It is essential that the flexibility of the abutment foundation be modeled; otherwise the displacements of the integral bridge are seriously underestimated; besides problems with earthquake analysis, failure to capture the flexible end supports of the bridge also results in unrealistic stresses from
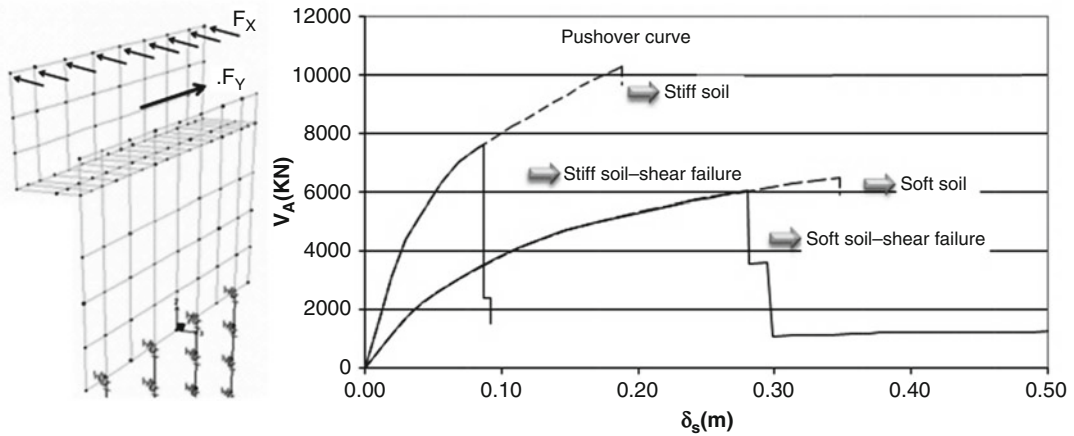
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 12 Modeling of abutment system and resulting pushover curves

temperature variations and shrinkage of concrete. In addition to the foundation, the backfill and the embankment have to be accounted for in the model, especially in short bridges with stiff deck; a concept usually adopted in these cases is the "effective" (or "critical") length of the embankment, whose properties are calculated and then introduced as springs (discussed later in this section) at the end of the elements representing the integral abutment.

Seat-type abutments can be modeled in a simple way, as shown in Fig. 2, i.e., by just including in the model the bearings (as springs, see section "Bearings" for property definition) of the abutment seat; so long as the end connection of the bridge is dominated by the properties of the bearings (in modern concrete bridges these are either elastomeric or pot bearings that can slide in all directions, see section "Bearings"), this simple model is quite adequate. If pot bearings are used, even a simple sliding connection (roller support) can be defined in the model; this ignores friction forces at the pot bearings, which are small if monolithic connections or fixed bearings are used elsewhere in the bridge. However, when a longitudinal joint closes, or a shear key blocks the transverse movement of the deck, or the gap between the deck and the abutment stem wall closes (see Fig. 2 bottom right), the abutment-backfill system is activated and significant forces

can develop at the bridge ends; further movement (in either direction) can be captured by the model only if the flexibility of the system, which includes both concrete members and the backfill soil, is modeled.

In abutments the important aspects of soil-structure interaction (section "Soil-Structure Interaction Effects") are modeled in practice-oriented applications through a system of linear or nonlinear *springs* at the ends of the bridge. The properties of the springs can best be defined by an analysis of the abutment-backfill system, preferably accounting for nonlinear effects directly, or at least by proper selection of reduced properties (e.g., estimates of G of the ground consistent with the expected deformations). Figure 12 shows the modeling of an abutment system (Kappos and Sextos 2009); the abutment wall is modeled with 2D shell elements, while the piles with frame elements supported on (depth dependent) nonlinear horizontal springs. In the vertical direction, friction springs were used along the piles and an appropriate vertical stiffness was introduced with the use of a (compression only) spring at the tip of the piles; a simpler model was also analyzed wherein infinite vertical stiffness was assumed (tip displacements restrained). On the right of Fig. 12 are shown the *pushover curves* (i.e., seismic force versus monitoring point displacement) derived for the transverse direction

(force F_y), wherein the behavior of the abutment system is dominated by the nonlinear response of the piles. For investigation purposes, the analysis was performed both for soft soil conditions (which was the actual case in the specific abutment) and for the case of a significantly stiffer supporting soil. These pushover curves reveal the sensitivity of the abutment response to the soil conditions, as well as the importance of accounting for all possible failure modes in the analysis (shear failure of the piles limits the deformation capacity, i.e., the ductility, of the entire system).

Equivalent linear or nonlinear springs based on the stiffness defined by curves like those in Fig. 12 can be used for modeling abutments to which the deck forces are directly transferred (due to the previously mentioned reasons). A full-range model, covering all stages of the response, should include two (nonlinear) springs in series, one representing the stiffness of the bearings and one the stiffness of the abutment system; a gap element (with the width of the longitudinal joint of the bridge) should be added if analysis is carried out in a single run; otherwise separate analyses with open and closed joint should be carried out (a usual practice in the USA).

In lieu of carrying out a proper analysis of the abutment, simplified procedures can be used for estimating a reasonable stiffness (and, in nonlinear analysis, strength) for the abutment. Arguably the most popular procedure is that prescribed in Caltrans (2013), wherein the *longitudinal* stiffness can be calculated from the initial embankment fill stiffness $K_i \approx 28.7$ kN/mm/(m width of the wall), and this has to be adjusted proportionally to the backwall (or diaphragm) height (h):

$$K_{abut} = K_i \times w \times (h/1.7) \qquad (8)$$

where w is the projected width of the backwall or diaphragm, for seat and diaphragm abutments, respectively, and (h/1.7) is a proportionality factor based on the 1.7 m height of the diaphragm abutment specimen tested at UC Davis (the actual relationship is not linear, but so far there is no sufficient data to develop a more sophisticated

one). The aforementioned K_i applies to well-compacted backfills as required by Caltrans (2013); otherwise it should be reduced by 50 %. The ultimate abutment load was assumed to be limited by a maximum static soil passive pressure of 240 kPa; the latter is multiplied by the corresponding surface, e.g., the product of the backwall width and height in seat-type abutments as well as the proportionality factor (h/1.7). The stiffness value for Eq. 8 applies when the elastic response of the bridge is dominated by the abutments; when this is not the case, Caltrans prescribes reductions depending on the ratio of the longitudinal displacement demand at the abutment (from elastic analysis) to the effective longitudinal abutment displacement at idealized yield (ratio of strength to stiffness).

In the *transverse* direction, a nominal abutment stiffness equal to 50 % of the elastic transverse stiffness of the adjacent bent can be used; this nominal stiffness has no direct correlation or relevance to the actual residual stiffness (if any) provided by the failed shear key but is meant to suppress unrealistic response modes associated with a completely released end condition.

Clearly, a full model for the abutment-backfill system should also include the backfill soil, as well as part of the embankment that is activated during the seismic excitation of the bridge; as mentioned previously, the latter is important in the case of short bridges (like two-span overpasses) wherein the embankment plays a key role in the seismic response of the bridge. In particular, consideration of the abutment-soil system participating mass has a critical effect on the mode shapes and consequently the dynamic response of the bridge. The critical length L_c of the embankment to be considered in the analysis can be estimated from the relationship (Zhang and Makris 2002):

$$L_c \approx 0.7\sqrt{S B_c H} \qquad (9)$$

where S is the slope of the embankment, H its height, and B_c its width at the crest. It should be borne in mind that L_c actually changes with the level of the seismic action, but this is difficult to capture in practical analysis.

Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 13 Modeling of abutment and backfill system using 3D finite elements
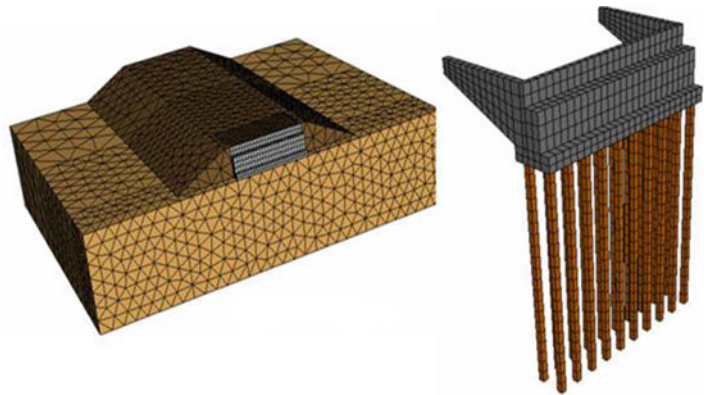
An example of a "heavy" finite element model of the entire system is shown in Fig. 13 from Kappos et al. (2012) wherein further examples and details can be found. Results of analysis of a number of typical abutment and backfills using sophisticated models such as that of Fig. 13, which considered the soil (backfill, embankment, and foundation) as the nonlinear material mechanism, have shown that both the stiffness and the strength estimated according to the pre-2013 Caltrans provisions (adopting K_i = 11.5 kN/mm/m, i.e., 60 % lower than the new value) underestimated the values found from the 3D FE models.

So long as the soil behind the abutment has been analyzed and its (macroscopic) stiffness reduced to a spring constant, the entire bridge can be modeled by combining in series the aforementioned translational springs (one for the backfill-embankment system and one for the abutment and its foundation) in each direction of the bridge. More details on modeling abutments and backfills using a system of nonlinear springs can be found in Aviram et al. (2008).

Bearings, Joints, and Shear Keys

Bridge furnishings include a number of components, i.e.,

- Bearings
- Joints
- Parapets – rails
- Waterproofing system

Among these, critical components of the bridge, particularly in an earthquake resistance context, are the bearings and the joints. Modeling of these critical components for seismic analysis is discussed in the remainder of this section, which also covers some other components, normally located close to, or even within, the bearings, i.e., shear keys and damping devices.

Bearings

Bearings are mechanical systems which:

- Transmit loads from the superstructure (deck) to the substructure (piers, abutments)
- Accommodate relative displacements between them

In the past, steel bearings of the pin, roller, rocker, or sliding type have been used, and they are still used in some bridges, in particular steel ones. In modern concrete bridge construction, bearings typically belong to one of the following categories:

- Pot bearings
- Elastomeric (common) bearings
- Elastomeric (special) bearings

Pot bearings (Fig. 14) allow sliding and rotation and consist of a shallow steel cylinder (or "pot") on a vertical axis with a neoprene disk which is slightly thinner than the cylinder and fitted tightly inside. A steel piston fits inside the cylinder and bears on the neoprene, while flat
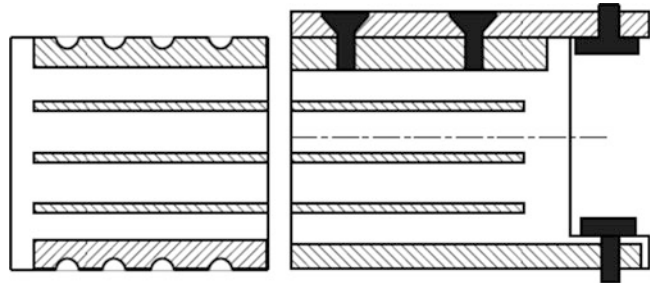
S

**Seismic Analysis of
Concrete Bridges:
Numerical Modeling,
Fig. 14** Pot bearings
(www.agom.it): free-
sliding (*left*) and
transversely guided sliding
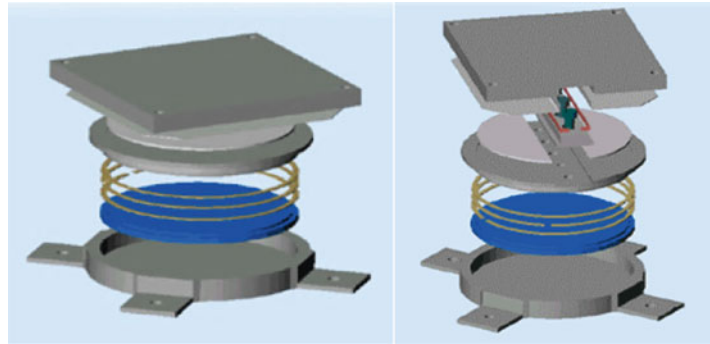
**Seismic Analysis of
Concrete Bridges:
Numerical Modeling,
Fig. 15** Laminated
bearing with outer steel
plates: profiled (*left*) or
allowing fixing (EN1337,
CEN 2005b)

brass rings are used to seal the rubber between the piston and the pot; the rubber behaves like a viscous fluid, flowing as rotation occurs. Sliding can either take place in any direction or be guided (through a groove and sliding bar system; see Fig. 14-right) in a specific direction. Use of pot bearings is very common, especially at the seats of the abutments.

In general, pot bearings do not have to be explicitly modeled; it suffices to release the corresponding degrees of freedom at the support of the deck.

Common elastomeric bearings are made of elastomer, i.e., either natural or synthetic rubber (e.g., neoprene), which is flexible in shear (low GA) but very stiff against volumetric change. To avoid bulging (lateral expansion that adversely affects the properties of the elastomer) different types of reinforcement are used in the elastomer (fiberglass, cotton, steel). The most common and efficient (and also the most expensive) type of reinforced elastomeric bearings is that reinforced with thin steel plates as shown in Fig. 15; these are constructed by vulcanizing elastomer to these steel plates.

The design of elastomeric bearings is carried out (in European and some other countries) according to the European Standard EN1337 (CEN 2005b). This standard prescribes maximum strains due to vertical load, rotations and horizontal actions, such as loads or displacements, and minimum thickness of the internal and external steel plates. It also prescribes a number of ultimate limit state verifications (limitation of distortion and rotation, tension in the steel plates, bearing stability (buckling), and slip prevention). The procedure for designing a bridge so that the seismic action is resisted entirely by elastomeric bearings on all supports ("seismic isolation") is prescribed in Chap. 7 of Eurocode 8-3 (CEN 2005a).

Proper modeling of elastomeric bearings is essential in the framework of seismic design. In all cases at least the horizontal shear stiffness (K_h) should be captured, but in more refined models the flexural (K_b) and the axial (K_v) stiffness of the bearings are also introduced in the model. These three stiffness values can be calculated from the following relationships:
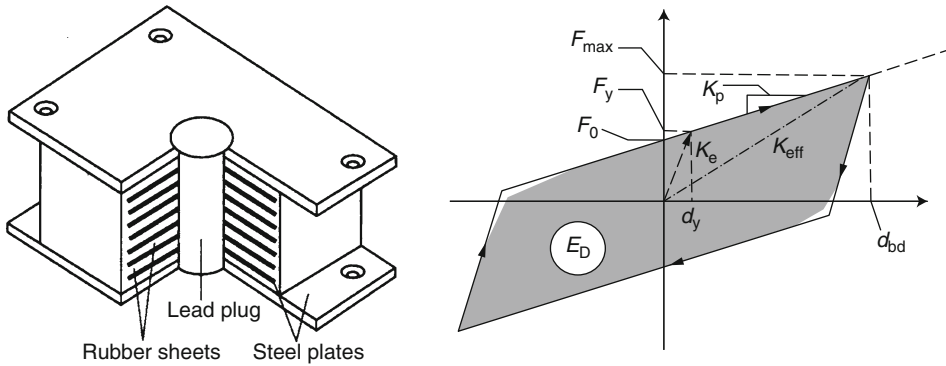
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 16 Lead rubber bearing (rectangular section) and corresponding hysteresis law

$$K_h = GA/t_r \qquad (10a)$$

$$K_b = 0.329 E_c I/t_r \qquad (10b)$$

$$K_v = E_c A/t_r \qquad (10c)$$

where t_r is the thickness of the elastomer (not the total height of the bearing) and the other symbols have their usual meaning. The shear modulus G can be taken as 0.9 MPa for static loading and 1.8 MPa for dynamic loading (rubber is a visco-elastic material, i.e., its resistance increases with the loading rate).

Special elastomeric bearings are those that provide the high energy dissipation needed to resist strong earthquakes. The most common types used today are:

- High damping elastomeric bearings
- Lead-rubber bearings (laminated bearings with lead core)

Bearings of the first category can provide fairly high values of effective damping ratio ($\xi_{eff} \approx 10\text{–}25\ \%$). The main concern about them is durability, as rubber properties are known to deteriorate with time; of course, companies manufacturing them claim design lives appropriate for normal uses. Lead-rubber bearings (LRB) are laminated bearings with a cylindrical lead core, as shown in Fig. 16; they are either circular or rectangular. The selection of lead as the core

material is due to its high hysteretic energy dissipation ("fat" elastoplastic hysteresis loops under cyclic loading; see Fig. 16 right). In LRBs the equivalent viscous damping ratio can reach quite high values (in the range of 30 %). In addition to damping, the lead core also provides resistance to service lateral loads.

Modeling of special elastomeric bearings is similar to that of common bearings as far as stiffness is concerned (relationships 10), but the additional damping has to be properly introduced in the model. In all types of elastic analysis, a viscous damping ratio higher than that of reinforced concrete (5 %) is introduced, and the corresponding spectra are used to derive the seismic actions; in Eurocode 8 the ordinates of the elastic spectrum for $\xi \neq 5\ \%$ are estimated by multiplying the reference spectrum by

$$\eta = \sqrt{{}^{10}/_{(5+\xi)}} \geq 0.40 \qquad (11)$$

The value of effective damping ratio to be used can be estimated from (CEN 2005a)

$$\xi_{eff} = \frac{1}{2\pi} \left[\frac{\Sigma E_{D,i}}{K_{eff} d_{cd}^2} \right] \qquad (12)$$

where $\Sigma E_{D,i}$ is the sum of dissipated energies of all special bearings i in a full deformation cycle at the design displacement d_{cd} and $K_{eff} = \Sigma K_{eff,i}$, i.e., the sum of the composite stiffnesses of the

isolator unit and the corresponding substructure (pier) i. For the specific case of LRBs, the dissipated energy $E_{D,i}$ is calculated from the pertinent elastoplastic hysteresis loop (Fig. 16 right).

An interesting type of special bearing, used exclusively for seismic isolation, is the *friction pendulum* shown in Fig. 17, wherein the sliding surface of the bearing is concave; hence, the restoring force is provided by the horizontal component of the structure itself. Sliding on the concave surface is resisted by friction of the contact material which is PTFE (polytetrafluoroethylene, most common commercial name Teflon); the friction coefficient is high initially (hence, no swaying of the superstructure takes place under normal loading conditions) but substantially higher under high velocities induced by earthquake.

The *articulation* of the bridge, i.e., the arrangement of the different types of bearings, is a critical aspect of the design of the bridge, in particular the seismic one. Figure 18 shows an

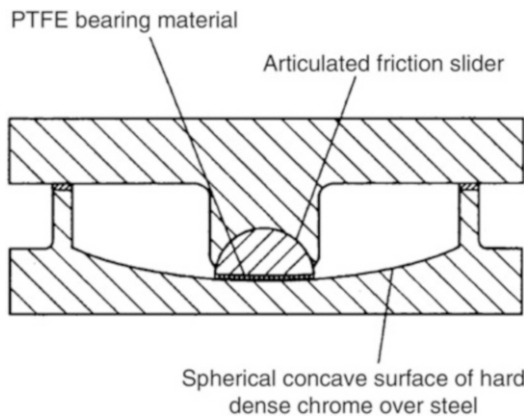**Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 17** Friction pendulum bearing

example of an actual railway bridge wherein a seismic isolation approach was adopted, involving a combination of lead rubber bearings and viscous fluid dampers; it is seen that the LRBs located toward the end of the bridge are movable horizontally, while free sliding pot bearings are used at the abutments.
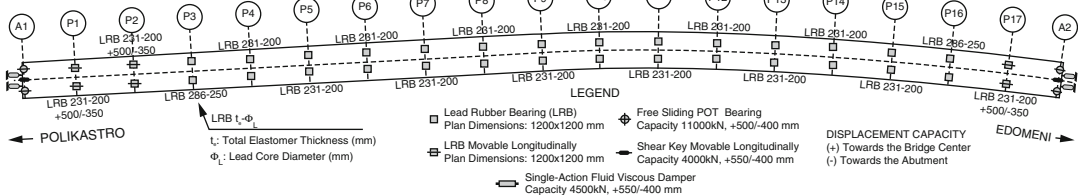
Supplemental Damping Devices
On several occasions involving large bridges, the amount of damping that can be provided by realistic arrangements of special bearings is not sufficient for limiting displacements to the required limits (recall that long-span bridges are long-period structures and when isolation is used their fundamental period can become very long, more than 3 s). In such cases a more efficient solution can be the use of special damping devices (separate from the bearings) that will supplement the energy dissipation provided by the bearings (high damping or LRB); alternatively, common elastomeric bearings or friction pendulum bearings can be used in combination with the damping devices. The most commonly used devices are

• Viscous fluid dampers
• Steel yielding devices

Viscous fluid dampers are based on the concept (long used in the automotive industry) of forcing through a piston a viscous fluid (usually silicon oil) through an orifice. Another, more recent, alternative are shear panels containing high-viscosity fluids. The constitutive law of such dampers is not restricted to the well-known linear dependence on velocity (through C, the

Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 18 Arrangement of bearings in a seismically isolated railway bridge (Kappos et al. 2014)

damping coefficient) but is often nonlinear, of the form

$$F_D = C \cdot |\dot{u}|^\alpha \cdot \text{sgn}(\dot{u}) \tag{13}$$

where F_D is the damping force and α is an exponent between 0 and 1; $\alpha = 1$ corresponds to standard viscous damping, but in practical applications lower values are used since they lead to higher F_D at lower velocities (e.g., in the bridge shown in Fig. 18, $\alpha = 0.15$ was selected for the viscous dampers, which had $C = 5,440$ kN·s/m).

Modeling of viscous fluid dampers can be made using two approaches: in the simpler approach, appropriate for practice-oriented elastic analysis, the dampers are substituted by an effective value which is the sum of the basic damping ξ_0 (typically 5 %) and the contribution of the fluid dampers (having exponent α and damping coefficients C_i). Considering the fundamental mode of the bridge (modal displacements φ_i at each mass m_i) and calculating the energy dissipated by the nonlinear dampers in a cycle of sinusoidal motion, it can be shown (Hwang 2002) that

$$\xi_{\text{eff}} = \xi_0 + \frac{\sum\limits_{j} \lambda C_j \phi_{rj}^{1+\alpha} \cos^{1+\alpha}\theta_j}{2\pi A^{1-\alpha}\omega^{2-\alpha}\sum\limits_{i} m_i\phi_i^2} \tag{14}$$

where ω is the circular frequency, θ_j is the angle of inclination of the damping device j, λ is a function of the exponent α (tabulated values of λ are given in FEMA 1997), and $u_i = A\varphi_i$ (A is the amplitude) are the actual displacements of the masses m_i.

In a more involved, nonlinear, analysis, the viscous dampers can be directly introduced in the model at the particular locations of the bridge wherein they are installed; such elements are available in some software packages but are rarely used in practical design.

Steel yielding devices provide an almost elastoplastic hysteresis loop, and it is possible to get a great variety of damping ratio values by properly selecting the yield displacement and post-yield stiffness ratio (hardening) of the isolation system. They can be modeled either by expressing the hysteretic energy dissipation as an equivalent damping (by equating the area of the hysteresis loop to that of the ellipse representing viscous damping energy) or by directly including yielding elements at the pertinent positions of the bridge (such elements are available in most programs).

Shear Keys

Shear keys serve the purpose of preventing the displacement of the bridge deck in a certain direction and can be located at several positions in a bridge; a typical one is at the abutments for blocking the transverse movement of the deck. They can be either external as shown in Fig. 19 left (i.e., forming part of the seat) or as interior short cantilevers interlocking with corresponding grooves in the deck (Fig. 19 right); exterior shear keys are preferable because they are easy to inspect and repair. Shear keys might directly bear on the surrounding part of the deck, immediately blocking its movement, or be located at a selected distance, forming a local joint (see section "Joints" for modeling of joints).
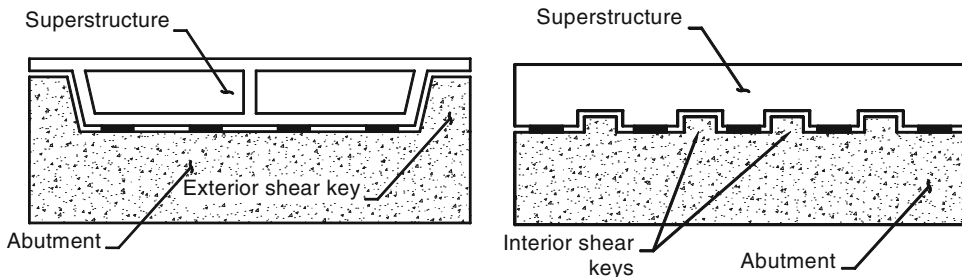
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 19 Exterior and interior shear keys in bridges (Kappos et al. 2012)

Modeling of shear keys should be consistent with their design "philosophy," i.e., whether the shear key is meant to remain essentially elastic for the entire range of seismic response of the bridge considered in design, or a "sacrificial" element whose capacity should be limited with a view to protecting other more critical and/or more difficult to repair components of the bridge; a popular design concept in California is that the capacity of the shear keys should not exceed or be close to the shear capacity of the piles (on which the abutment is supported). Eurocode 8-2 (CEN 2005a) vaguely specifies that the design actions for the seismic links (one type of which is shear keys) should be derived as capacity design effects, with the horizontal resistance of the bearings assumed equal to zero, without explaining how these effects should be derived (clearly one possibility could be to relate the strength of shear keys to that of piles, as mentioned previously).

The strength of shear keys can be derived from sophisticated models such as strut and tie ones or simplified ones, usually based on the shear friction concept. Bozorgzadeh et al. (2006) proposed the following relationship for the nominal capacity of a shear key:

$$V_n = \frac{\mu_f \cdot \cos a + \sin a}{1 - \mu_f \cdot \tan \beta} \cdot A_{uf} \cdot f_{su} \qquad (15)$$

where μ_f is a kinematic friction coefficient, β is the angle of inclined face of the shear key (Fig. 19 left), α is the angle of kinking of the vertical bars with respect to the vertical axis (recommended value from test results 37°), and f_{su} is the (ultimate) tensile strength of the vertical reinforcement that has an area A_{vf}. Such models strongly depend on the friction coefficient that varies substantially depending on the detailing of the shear key; for sacrificial keys with smooth finishing of the concrete interface, $\mu_f = 0.36$ is recommended (Bozorgzadeh et al. 2006), but for properly detailed joints of adequate roughness, much higher values apply (up to 1.4 for keys monolithically cast with the abutment seat).

Having established a proper value for the shear key strength, the key can be modeled simply as a rigid-plastic spring (or "link") with a displacement capacity of around 100 mm or using more sophisticated multilinear constitutive laws with ascending and descending branches (see more details in Kappos et al. 2012). In linear, "code-type," analysis, one usual option in the USA is to carry out two analyses, one with and one without thee shear keys (essentially one assuming displacement is blocked in the pertinent direction and one with the deck allowed to displace freely).

Joints

Joints (also called *expansion joints* which is inappropriate when they are also designed for seismic actions) are required to accommodate (with negligible resistance and noise) movements of the deck due to:

- Thermal expansion/contraction
- Shrinkage and creep of concrete
- Earthquake-induced horizontal movement

There are several types of joints (compression seal, strip seal, finger plate, sliding plate, modular), each of them appropriate for up to a certain design movement of the bridge. In bridges designed for high seismic actions, the joint gap might result as quite substantial (over 200 mm). It is noted that current seismic codes like Eurocode 8-2 require the joint gap to satisfy

$$d_{ED} = d_E + d_G + \psi_2 d_T \qquad (16)$$

where d_G is the long-term horizontal displacement due to permanent and quasi-permanent actions (posttensioning, shrinkage, and creep), d_T the displacement due to thermal actions ($\psi_2 = 0.5$ for road bridges), and d_E is the seismic displacement calculated as $\pm \eta \mu_d d_{Ee}$ where d_{Ee} is the displacement derived from the analysis for the seismic loading combination (with the design spectrum reduced by the behavior factor q), μ_d is the design ductility ($\mu_d = q$ in the common case that the fundamental period of the bridge $T \geq T_o = 1.25 T_c$ where T_c is the corner period of the design spectrum) and η is the damping correction factor for the design spectrum ($\eta = 1$ for $\xi = 5\%$).
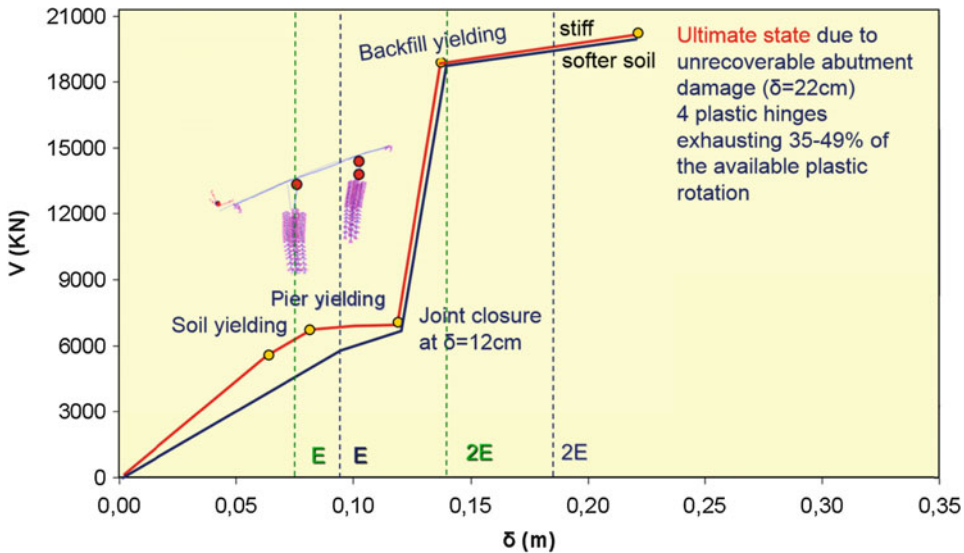
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 20 Pushover curves for the longitudinal direction of an overpass bridge (Kappos and Sextos 2009); "E" denotes the design earthquake level

When large movements ($d_{Ed} > 100$ mm) have to be accommodated, *modular* joints are used, wherein sealing elements and rail elements are coupled; in this case intermediate elements (rails), supporting elements, and linkage elements (e.g., folding trellis) causing equal gap widths are needed, and the total movement accommodated by the joint is the sum of the distances of the rails.

It is common practice in Europe to provide substantial joint gaps (as well as ample seat lengths) and make sure that the joints remain open during the design earthquake, without carrying out specific verifications for the case that the gap closes. In California, "dual" analysis is the recommended (by Caltrans) practice, wherein the bridge is analyzed assuming either free movement or full restraint at the "compression end" of the bridge (movement is always free at the "tension end" where the deck moves away from the abutment) and taking the most unfavorable response quantities from either set of analyses; clearly this is conservative and might result in increased costs. Another interesting difference in US and European practice is that *restrainers* (either cables or bolt linkages, the bolts passing through holes in the deck) are often used at the

joints of US (and New Zealand) bridges, with a view to preventing *unseating* during earthquakes stronger than the design one. In Europe this is seen as an option mainly in *retrofitting* of existing bridges with insufficient seating lengths.

Modeling of joints in seismic analysis is essential, since the bridge boundary conditions are drastically different when joints are open or closed. For instance, in the longitudinal direction of the bridge, there are at least two joints at the ends (over the seat-type abutments) that are essential for accommodating "non-seismic" displacements (d_G and d_T); when either of these joints closes during the earthquake, substantial forces are transferred to the abutment-backfill system (see section "Abutments and Backfills" and Fig. 20 later in this section), and the horizontal forces developed in the bridge-abutment-backfill system can be much higher than those developed when longitudinal movement was free.

Proper modeling of joints requires *nonlinear* analysis using the gap (or other special link) elements currently available in most software packages. This element is often combined in series with the spring (or link) elements modeling the bearings and/or shear keys at the abutment.

It is noted that although the gap element has an elastic behavior subsequent to gap closure, the analysis of the bridge is nonlinear, as the boundary conditions change during the analysis. The importance of capturing the effect of joint closure on the seismic behavior of a bridge can be seen in Fig. 20, where pushover curves are given for the longitudinal direction of a typical overpass (whose abutment system is shown in Fig. 12). Recall that two different soil conditions were studied; this has a noticeable effect on the initial stiffness of the bridge, but as soon as the longitudinal gap (120 mm in this case) closes, a drastic increase in both stiffness and strength is noticed, as the abutment-backfill system is now activated. The final failure of the bridge is estimated to take place during this second stage of the response (at a displacement of almost twice the gap length) and is attributed to unrecoverable damage to the soil behind the abutment (50 % loss in strength), while the piers are still well within their rotational capacity (35–49 %). A different failure mechanism (exceedance of available ductility of piers) would have been predicted had the end support been simulated as longitudinal restraint (as per the Caltrans simplified approach).

Bridge Analysis Methods

Methods of analysis can be classified as:

- Elastic (equivalent) static
- Elastic dynamic (response spectrum)
- Inelastic static (pushover)
- Inelastic dynamic (response history)

The basics of all these methods are presented in a series of articles in the encyclopedia and will not be repeated herein. Instead, some aspects of analysis (other than member modeling that was presented in section "Modeling of Bridge Components") specific to concrete bridges will be briefly described, and the current trends in nonlinear static (pushover) analysis of bridges (not covered elsewhere here) will also be

presented. These will be followed by a case study involving application of different analysis method to an actual concrete bridge.

Code-Prescribed Analysis of Concrete Bridges

Among the important documents (codes-guidelines) for the *analysis* of concrete bridges are those regularly published by Caltrans (2013, latest version) and the ACI (2014, latest version) in the USA and Eurocode 8-2 (CEN 2005a) in Europe. Concrete bridges are the sole type covered in the ACI Report 341.2, while the Caltrans Criteria and Eurocode 8-2 also cover bridges made of other materials (steel and composite). For *retrofitting* of bridges a comprehensive document is the FHWA (2006) Manual. Some key aspects of these documents specific to the analysis of concrete bridges are discussed in the following. Other important sections relevant to the design of concrete bridges are those prescribing the procedures for resistance verification of reinforced concrete sections and the detailing of R/C members (piers, abutments, and retaining walls), which fall beyond the scope of this entry that focuses on modeling for structural analysis. Of course, calculation of R/C member strength and ductility (especially for piers) is essential for nonlinear analysis of concrete bridges; aspects of this issue are covered in the case study presented later in section "Comparative Case Study."

- All types of analysis are permitted for concrete bridges, the *equivalent static procedure* being subject to a number of limitations regarding the effect of higher modes; this type of analysis is usually suitable for the longitudinal direction of straight bridges, which is dominated by a single mode (that is often the fundamental mode of the bridge). Three versions of the method (Rigid Deck Model, Flexible Deck Model, Individual Pier Model) are prescribed by Eurocode 8-2 (see basic aspects of these methods in Vayas-Iliopoulos (2014)). The reference method in practical design is the (elastic) *dynamic response spectrum analysis*, while *nonlinear* methods are only used in practice for the verification of the design of

some important bridges that have initially been designed using response spectrum analysis. Nonlinear methods are much more common in the case of *assessment* of existing bridges. The FHWA (2006) Retrofit Manual provides sufficient guidance (especially with respect to estimating strength and deformation capacity) for the application of both the pushover and response-history analysis methods, which are covered more briefly in Eurocode 8-2; notably, assessment of existing bridges is not currently covered by the Eurocode package (it is one of the issues that will be added at the next stage of development).

- Application of elastic methods to bridges is the same as for other structures, but there are two aspects specific to concrete bridges that have to be properly addressed:
    - In most concrete bridges both prestressed and ordinary (non-prestressed) concrete are used, the former for the deck, the latter for the piers and abutments. The *damping* ratio ξ is different for these two materials (primarily due to the different degrees of cracking in each), i.e., 5 % for reinforced concrete and 2 % for prestressed concrete. In EC8-2 this is accounted for by considering the response spectrum for an equivalent damping ratio:

$$\xi_{ef} = \frac{\Sigma \xi_i E_{di}}{\Sigma E_{di}} \quad (17)$$

    where E_{di} is the deformation energy induced in member i by the seismic action. This quantity is not a standard output of common structural analysis programs, and in order to avoid ad hoc spreadsheet calculations, designers often prefer to simply use an average value of 3.5 %. It is worth noting here that E_{di} is *not* the energy *dissipation* through yielding mechanisms, but rather refers to the pre-yield state; hence, it is not appropriate to consider a spectrum for $\xi = 5$ % on the basis that only reinforced concrete members yield and dissipate energy (prestressed concrete decks remain quasi-elastic as already mentioned in section "Deck").

    - The most important factor in reducing the elastic response spectrum to the design one (which is, in fact, an inelastic spectrum except when elastic response is foreseen for the design seismic action) is the force reduction factor, called *behavior factor* (q) in Eurocode 8 and *response modification factor* (R) in the US codes. Values for q for concrete bridges depend on whether the bridge is designed as ductile or "limited ductile." For *ductile* concrete bridges with *vertical* piers (working in prevailing bending) $q = 3.5\lambda(\alpha_s)$, where $\lambda(\alpha_s) = 1.0$ for $(\alpha_s) \geq 3$ ($\alpha_s = L_s/h$ is the shear span ratio of the pier, where L_s is the distance from the plastic hinge to the point of zero moment and h is the depth of the cross section in the direction of flexure of the plastic hinge) and $\lambda(\alpha_s) = \sqrt{(\alpha_s/3)}$ for $3 > (\alpha_s) \geq 1$. For ductile concrete bridges with piers consisting of *inclined* struts (e.g., V shaped), $q = 2.1\lambda(\alpha_s)$. For *limited ductile* concrete bridges, the corresponding values are 1.5 (vertical piers) and 1.2 (inclined struts). For abutments rigidly connected to the deck (integral bridges) $q = 1.5$, except in "locked-in" structures, i.e., bridge structures whose mass essentially follows the horizontal seismic motion of the ground (hence, they do not experience significant amplification of the horizontal ground acceleration), in which case $q = 1$. These values, as well as similar ones specified by the American Code (AASHTO 2010), are in many cases conservative. In a recent study evaluating the actual force reduction factors for existing bridges in Europe, Kappos et al. (2013) found that in all bridges studied the available q values were higher than those used for design in both the longitudinal and transverse directions. In fact, in many cases the code-specified values (in particular those of AASHTO for single-column bents) seem to significantly underestimate the actual energy dissipation capacity of concrete bridges. Seen from another perspective,

S

this is a clear indication that modern bridges possess adequate margins of safety and are able to withstand seismic actions that are often substantially higher than those used for their design. This high performance is due to their ductility, as well as their overstrength; previous studies that have ignored the latter led to unrealistically low estimation of q-factor values.

Nonlinear Static Analysis of Concrete Bridges

Interesting and useful work has been carried out in the last decade on nonlinear (inelastic) static, also known as pushover, analysis of bridges; nevertheless this is clearly less than that for buildings. A recent book presenting all available methods for pushover analysis of bridges is that by Kappos et al. (2012), which also includes a substantial number of case studies involving the comparative application of several methods. Due to space limitations, only one approach will be presented herein which, in the writer's opinion, combines sufficient accuracy with relatively limited effort and the possibility to be applied using available (commercial) software tools, with very limited need for additional spreadsheet calculations; in fact, software for "single-run" application of the method is currently at an advanced level of development. The method is usually referred to as *(multi-)modal pushover analysis (MPA)*; it was presented in a comprehensive form for buildings by Chopra and Goel (2002) and was extended to bridges by Paraskeva et al. (2006). The key idea is to perform multiple pushover analyses of the structure, one for each significant mode, and combine statistically the resulting displacements and rotations. The steps involved in the latest version of the method (Paraskeva and Kappos 2010), which includes a number of improvements, are summarized in the following.

Step 1: Compute the natural periods, T_n, and mode shapes, $\boldsymbol{\phi}_n$, for linearly elastic vibration of the structure.

Step 2: Carry out separate pushover analyses for force distribution $\mathbf{s}_n^* = \mathbf{m} \cdot \boldsymbol{\phi}_n$ for each significant mode of the bridge and construct the pushover curve (base shear V_{bn} vs. displacement u_{cn} of the "control" or "monitoring" point) for each mode; $\mathbf{s}_n^*$ are loading *patterns*; hence, the relative significance of each mode is not accounted for at this stage; this will be done at Step 5, through the target displacement for each modal pushover analysis. Gravity loads are applied before each MPA and P-Δ effects are included, if significant (e.g., in bridges with tall piers).

Step 3: The pushover curve must be idealized as a bilinear curve so that a yield point and ductility factor can be defined and subsequently used to appropriately reduce the elastic response spectra representing the seismic action considered for assessment. This idealization can be done in a number of ways, some more involved than others; it is suggested to do this once using the full pushover curve (i.e., analysis up to "failure" of the structure, defined by a drop in peak strength of about 20 %) and the equal energy absorption rule (equal areas under the actual and the bilinear curve). It is noted that the remaining steps of the methodology can be applied even if a different method for producing a bilinear curve is used.

Step 4: Several procedures are available (FEMA 1997; Chopra and Goel 2002; CEN 2004, all referring to buildings) for defining the earthquake displacement demand associated with each of the pushover curves derived in Step 3. Paraskeva et al. (2006) adopted the capacity and demand spectra procedure based on inelastic demand spectra (Fajfar 1999); hence, Step 4 consists in converting the idealized $V_{bn} - u_{cn}$ pushover curve of the multi-degree-of freedom (MDOF) system to a "capacity diagram" (Fig. 21). The base shear forces and the corresponding displacements in each pushover curve are converted to spectral acceleration (S_a) and spectral displacements (S_d), respectively, of an equivalent single-degree-of-freedom (SDOF) system, using the relationships (Chopra and Goel 2002):
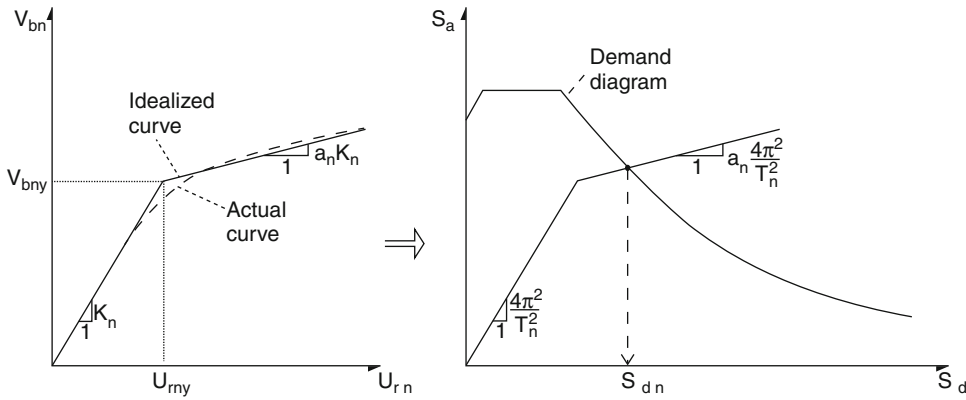
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 21 Idealized pushover curve of the nth mode of the MDOF system and corresponding capacity curve for the nth mode of the equivalent inelastic SDOF system

$$S_a = \frac{V_{bn}}{M_n^*} \qquad (18a)$$

$$S_d = \frac{u_{cn}}{\Gamma_n \cdot \phi_{cn}} \qquad (18b)$$

wherein ϕ_{cn} is the value of ϕ_n at the control (or "monitoring") point, $M_n^* = L_n \cdot \Gamma_n$ is the effective modal mass, with $L_n = \boldsymbol{\phi}_n^T \mathbf{m} \cdot \mathbf{1}$, $\Gamma_n = L_n / M_n$ is a mass participation factor, and $M_n = \phi_n^T \mathbf{m} \cdot \boldsymbol{\phi}_n$ is the generalized mass, for the nth natural mode. For *inelastic* behavior, estimation of the displacement demand at the monitoring point is made with the aid of inelastic spectra.

Step 5: Since the displacement demand calculated in Step 4 (for each mode) refers to SDOF systems with periods equal to those of the corresponding modes, the next step is to correlate these displacements to those of the actual bridge. Hence, Step 5 consists in converting the displacement demand of the nth mode inelastic SDOF system to the peak displacement of the monitoring point, u_{cn} of the bridge, using Eq. 18b. The selection of this point is a critical issue for MPA of bridges, and as discussed by Paraskeva et al. (2006), several choices of monitoring point are acceptable as long as the derived pushover curve has a reasonable shape, but they do not

lead to equally good results as far as the final response quantities are concerned. For practical purposes, a good selection is the deck point above the most critical support (pier or abutment) of the bridge.

Step 6: In this step, a correction is made of the displacement of the monitoring point of the bridge, which was calculated at the previous step. The correction is necessary only for cases that significant inelasticity develops in the structure. If the structure remains elastic or close to the yield point, the MPA procedure suggested by Paraskeva et al. (2006) is used to estimate seismic demands for the bridge. The response displacements of the structure are evaluated by extracting from the database of the individual pushover analyses the values of the desired responses at which the displacement at the control point is equal to u_{cn} (see Eq. 18b). These displacements are then applied to derive a new vector $\boldsymbol{\phi}_n'$, which is the deformed shape (affected by inelastic effects) of the bridge subjected to the given modal load pattern. The target displacement at the monitoring point for each pushover analysis is calculated again with the use of $\boldsymbol{\phi}_n'$, solving Eq. 18b for u_{cn}', and recalculated Γ_n using $\boldsymbol{\phi}_n'$.

Step 7: The response quantities of interest (displacements, plastic hinge rotations, forces in the piers) are evaluated by extracting from

the database of the individual pushover analyses the values of the desired responses r_n, due to the combined effects of gravity and lateral loads for the analysis step at which the displacement at the control point is equal to u_{cn} (or u_{cn}').

Step 8: Steps 3–7 are repeated for as many modes as required for sufficient accuracy; there is little merit in adding modes whose participation factor is very low (say less than 1 %), and application of the method to a number of bridges shows that it is not necessary to assure that the considered modes contribute to 90 % of the total mass.

Step 9: The total value for any desired response quantity (and each level of earthquake intensity considered) can be determined by combining the peak "modal" responses r_{no} using an appropriate modal combination rule, e.g., SRSS or CQC. This simple procedure is used for both displacements and plastic hinge rotations, which are the main quantities commonly used for seismic assessment of bridges. If member forces (e.g., pier shears) have to be determined accurately, a more involved procedure of combining modal responses should be used, consisting in correcting the bending moments at member ends (whenever yield values were exceeded) on the basis of the relevant moment versus rotation (M – θ) diagram and the value of the calculated plastic hinge rotation; this procedure blends well with the capabilities of currently available software.

Comparative Case Study

The overpass shown in Fig. 2 (some aspects of its modeling were discussed in section "Modeling of Bridge Components") has three spans and total length equal to 100 m, typical in modern motorway construction. Piers have a cylindrical cross section, while the pier heights are 8 m and 10 m. The deck is monolithically connected to the piers, while it rests on its two abutments through elastomeric bearings; movement in both the longitudinal and transverse directions is initially allowed at the abutments, but transverse displacements are restrained whenever the 150 mm gap shown

at the bottom of Fig. 2 is closed. The Greek Seismic Code design spectrum (similar to that of EC8) scaled to a PGA of 0.16 g was used for seismic design. The design spectrum corresponded to ground category "B" (close to ground "C" in the final version of EC8 (CEN 2004)). The bridge was designed as a ductile structure (plastic hinges expected in the piers) for a behavior factor q = 2.4.
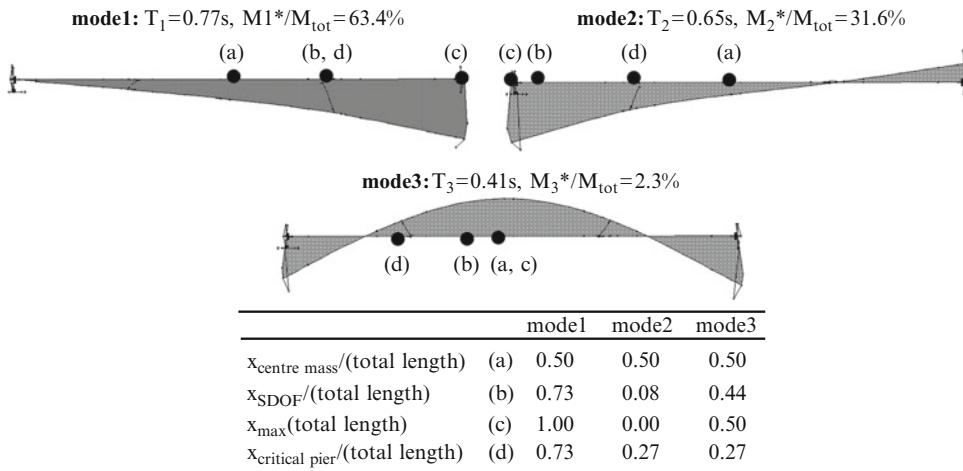
The bridge was analyzed applying a number of popular assessment procedures, i.e.:

- Modal analysis
- "Standard" pushover analysis (SPA) (first mode loading)
- Pushover analysis for a "uniform" loading pattern (as required by Eurocode 8 (CEN 2004) and by the ASCE Standard 41–06 (ASCE/SEI 2007))
- Modal pushover analysis (MPA) as proposed in Paraskeva et al. (2006)
- Improved modal pushover analysis as proposed by Paraskeva and Kappos (2010)
- Nonlinear response-history analysis (NRHA), for artificial records closely matching the demand spectrum (see Paraskeva et al. 2006)

All inelastic analyses were carried out using the SAP2000 software package (CSI 2011). Plastic hinging in the piers had to be modeled slightly differently in the NRHA and the pushover analysis, due to limitations of the software used. More specifically, nonlinear rotational spring elements were used in the finite element models used in NRHA, while the built-in beam hinge feature of SAP2000 was implemented in the models set up for pushover analysis. In both cases, though, the same moment versus rotation (M-θ) relationship was used (i.e., bilinear with 2–6 % hardening, depending on the calculated ultimate moment), with input parameters defined from fiber analysis performed for each pier section, utilizing the in-house developed computer program RCCOLA-90.

Nonlinear Static Analysis

The dynamic characteristics of the bridge were determined using standard modal (eigenvalue)

mode1: $T_1 = 0.77s$, $M_1^*/M_{tot} = 63.4\%$ **mode2:** $T_2 = 0.65s$, $M_2^*/M_{tot} = 31.6\%$

   (a)   (b, d)   (c)   (c) (b)   (d)   (a)

mode3: $T_3 = 0.41s$, $M_3^*/M_{tot} = 2.3\%$

   (d)   (b) (a, c)

		mode1	mode2	mode3
$x_{centre\ mass}$/(total length)	(a)	0.50	0.50	0.50
x_{SDOF}/(total length)	(b)	0.73	0.08	0.44
x_{max}(total length)	(c)	1.00	0.00	0.50
$x_{critical\ pier}$/(total length)	(d)	0.73	0.27	0.27

Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 22 Modal force distribution, location of the equivalent SDOF systems, and modal parameters for the main transverse modes of the overpass bridge

analysis. Figure 22 illustrates the first three transverse mode shapes of the overpass bridge, together with the corresponding participation factors and mass ratios, as well as the locations of alternative monitoring points for each mode. Consideration of the modes shown in Fig. 22 assures that more than 90 % of the total mass in the transverse direction is considered. For MPA, applying the modal load pattern of the nth mode in the transverse direction of the bridge, the corresponding pushover curve was constructed and then idealized as a bilinear curve (Fig. 21). As noted under Step 4 of the MPA procedure (see section "Nonlinear Static Analysis of Concrete Bridges"), the inelastic demand spectra method was used for defining the displacement demand for a given earthquake intensity.

Figure 23 illustrates the deck displacements of the bridge derived using pushover analysis for each mode independently, as well as the MPA procedure initially proposed in Paraskeva et al. (2006). If the structure remains elastic for the given earthquake intensity, both spectral displacement S_d and the product $\Gamma_n \cdot \phi_n$ will be independent of the selection of the control (monitoring) point; this means that deck displacements are independent of the location of the monitoring point. On the contrary, it was found that deck displacements derived with respect to different control points, for *inelastic* behavior of the

structure, are not identical but rather the estimated deformed shape of the bridge depends on the monitoring point selected for drawing the pushover curve for each mode.

For inelastic behavior, it appears that the estimated values of u_{cn} are different not only because of the deviation of the elastic mode shape ϕ_n from the actual deformed shape of the structure but also due to the fact that the spectral displacement S_d is dependent on the selection of monitoring point if the structure exhibits inelastic behavior (due to the bilinearization of the capacity curve). An improved target displacement of the monitoring point is calculated (from Eq. 18b) using ϕ_n', the actual deformed shape of the structure (see Fig. 23), while the spectral displacement remains the same. The response quantities of interest are evaluated by extracting from the "database" the values of the desired responses, r_n, for the analysis step at which the displacement at the control point is equal to u_{cn}' (the improved estimate of u_{cn} derived on the basis of ϕ_n').

Figure 24 illustrates the deck displacements of the overpass bridge, calculated from MPA using u_{cn}' as target displacement for each mode. It is noted that, due to the approximations involved in the capacity and demand spectra procedure, deck displacements derived with respect to different control points are not the same, but differences are significantly reduced and results are deemed
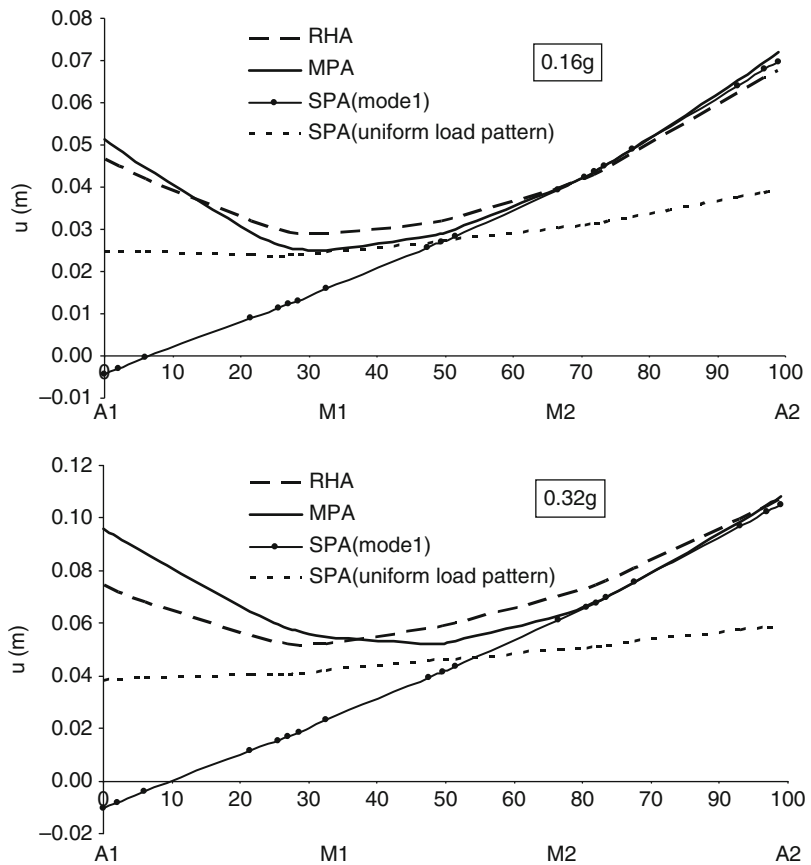
S

Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 23 "Modal" deck displacements derived with respect to different control points – *inelastic* behavior of the overpass bridge ($A_g = 0.16$ g)
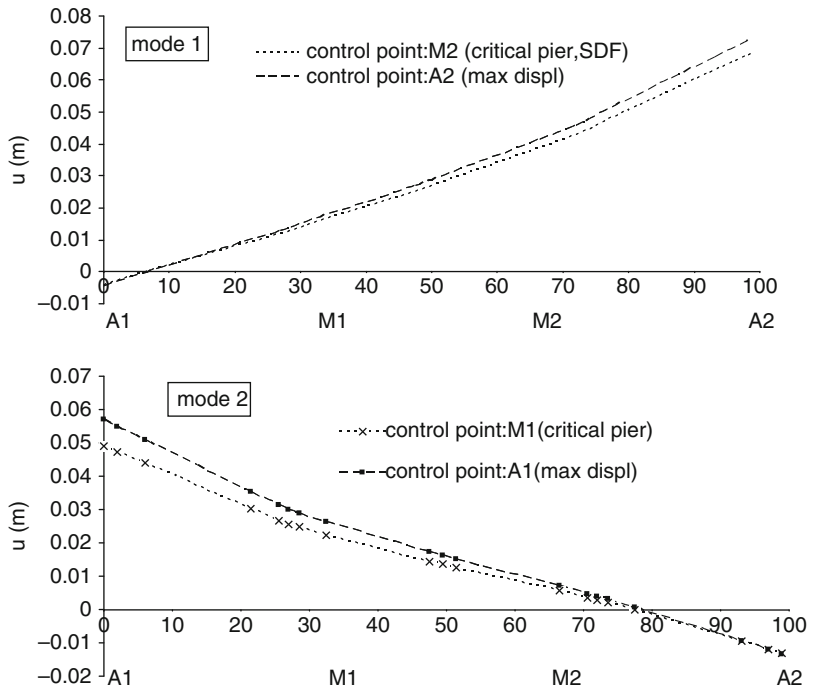
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 24 "Modal" deck displacements derived with respect to different control points using u_{rn}' as target displacement according to the *improved* MPA procedure – overpass bridge ($A_g = 0.16$ g)
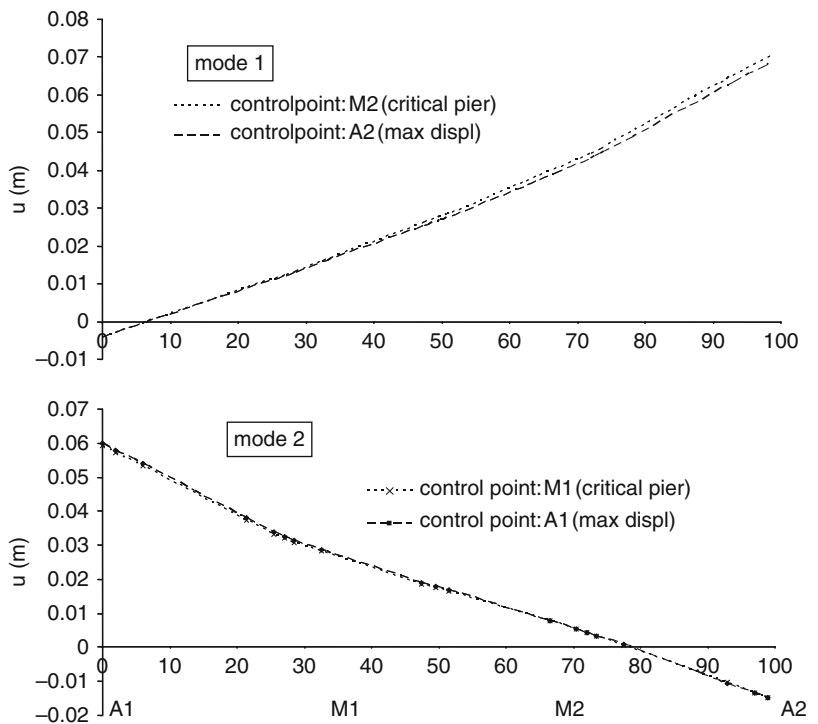
Seismic Analysis of Concrete Bridges: Numerical Modeling, Fig. 25 Response to the design earthquake ($A_g = 0.16$ g) and to twice the design earthquake ($A_g = 0.32$ g) calculated from SPA, MPA, and NRHA: deck displacements of the overpass bridge

acceptable for all practical purposes. Paraskeva and Kappos (2010) also studied other bridges with different configurations and noted that the differences between deck displacements derived with respect to different control points, as well as the improvement in the prediction of deck displacements using the procedure described in section "Nonlinear Static Analysis of Concrete Bridges," are more significant in bridges longer than the overpass of Fig. 2, especially if the increased length is combined with significant curvature in plan, which amplifies the complexity of its dynamic behavior and results in more significant contribution of the higher modes.

Reliability of Static Analysis Procedures

Results of the standard and modal pushover approaches were evaluated by comparing them with those from nonlinear response-history analysis, the latter considered as the most rigorous procedure to estimate seismic demand. To this effect, a series of NRHAs was performed using five artificial records compatible with the design elastic spectrum. The Newmark $\gamma = 1/2$, $\beta = 1/4$ integration method was used, with time step $\Delta t = 0.0025$ s and a total of 10,000 steps (25 s of input). A uniform damping value of 5 % was assumed for all modes of vibration, while hysteretic damping was accounted for through the elastoplastic behavior of the structural members.

The displacements determined by the SPA and MPA procedures were checked against those from NRHA for increasing levels of earthquake excitation, as shown in Fig. 25. It is noted that the deck displacements shown in the figures as the NRHA case are the average of the peak displacements recorded in the structure during the five response-history analyses. In this study the

displacement demand was estimated independently in static and dynamic (response-history) inelastic analysis, whereas in some previous studies comparisons of displacement profiles were made assuming the same maximum displacement in both cases; the choice adopted here is deemed as more relevant for practical applications, as it permits an evaluation of all aspects of the proposed procedure (including the uncertainty in estimating displacement demand in pushover analysis).

As shown in Fig. 25, the MPA procedure predicts well (i.e., matches closely the values from the NRHA approach) the maximum transverse displacement. On the other hand, the SPA procedure underestimates the displacements of the deck at the location of the abutment A1 and the first pier of the bridge, compared to the more refined NRHA approach. This is not surprising if one notes the differences between the first two mode shapes in the transverse direction (Fig. 22), which are strongly affected by torsion (they contribute more than 90 % of the torsional response, as well as over 90 % of the transverse response of the bridge) due to the unrestrained transverse displacement at the abutments (until the 150 mm gap closes), combined with the different stiffness of the two piers caused by their different height. What is essentially achieved by the MPA is the combination of these first two modes (the 3rd transverse mode is not important in this particular bridge), each of which dominates the response in the region of the corresponding abutment. In the case of applying ground motions with twice the design earthquake intensity (also shown in Fig. 25), where the structure enters deeper into the inelastic range and higher mode contributions become more significant (without substantial alteration of the mode shapes), it is noted that the displacement profile derived by the MPA method tends to match that obtained by the NRHA, whereas SPA predictions remain poor. Note that, regardless of earthquake intensity, the uniform loading pattern (also shown in Fig. 25) fails to capture the increased displacements toward the abutments; nevertheless its overall prediction of the displacement profile could be

deemed better than that resulting from using a single modal load pattern.

Additional case studies reported by Paraskeva and Kappos (2010) confirm that SPA predicts well (i.e., matches closely the values from the NRHA approach) the maximum transverse displacement, when applied to bridges of regular configuration, where the higher mode contribution is not significant. In such cases the improvement of the displacements derived by the MPA procedure is not significant even for high levels of earthquake excitation.

Summary and Concluding Remarks

It is clear that today the bridge engineer has at his/her disposal a set of powerful analysis tools that can be used for the seismic design or assessment of any bridge type. The potential of these tools, when properly utilized, was revealed by their success in predicting the response of bridges tested under high levels of earthquake actions that caused substantial amounts of inelasticity; an example was presented in section "Comparative Case Study." The information on bridge modeling presented in section "Modeling of Bridge Components" offers to researchers and designers the necessary information regarding the available models for the various parts of the bridge (deck, bearings and shear keys, isolation and energy dissipation devices, piers, foundation members), as well as tools for modeling the dynamic interaction between piers, foundation, and soil, as well as the abutment-embankment-superstructure system. It also provides information on important parameters that help ensuring that inelastic analysis of bridge earthquake response is conducted properly. It has to be emphasized in this respect that the power and versatility of the analysis tools also makes the results particularly sensitive to improper application.

Special emphasis was given to modeling of piers (section "Piers and Their Foundations"), as these members are both the ones wherein energy dissipation through plastic hinging is intended to

occur (unless a seismic isolation system is used) and those whose inelastic response is relatively easier to model in inelastic analysis (compared, for instance, to the abutment-backfill system or some foundation types or, indeed, some types of joints). Having said this, it is also clear from the material presented in this entry that proper modeling of the other components of the bridge, even those that are typically assumed to remain elastic during the seismic excitation (such as prestressed concrete decks), is also important, since, through their stiffness characteristics, they affect the dynamic characteristics of the bridge and the way seismic actions are transferred to the dissipating zones. Of great importance is also the modeling of the various connections in the bridge system, i.e., those between piers and deck, abutments and deck, and, in the common case (especially in the transverse direction) that the movement of the deck is restrained at the location of the abutment, the proper modeling of the response of the abutment-backfill system. As noted in section "Abutments and Backfills," in a practical context and when the main objective of the analysis is the response of the bridge itself (rather than that of the surrounding ground), the recommended solution is to carry out an independent analysis of the abutment-embankment system, determine its resistance curves (in all relevant directions), and use them to describe the nonlinear response of the equivalent springs to which the bridge model will be connected. If such an analysis cannot be afforded, the properties of these springs can be defined on the basis of simplified guidelines from the literature. For pier-foundation-soil interaction, the existing literature is more mature and often it is not necessary to carry out separate analysis of the system to derive the nonlinear properties of the soil-foundation dynamic impedance to be introduced in the bridge model, especially when surface foundations are used; in these cases information from the literature can be used to account approximately for the interaction with the surrounding ground.

Regarding the feasibility and reliability of different methods used for the analysis of concrete bridges, these depend primarily on the configuration of the bridge analyzed. As a rule, the longitudinal direction of the bridge is the easier one to analyze, and even simple, equivalent static elastic methods can lead to a reasonable design. In most other cases dynamic analysis is required to properly capture the higher mode effects that are important, especially in the areas close to the abutments. When the expected plastic mechanism does not involve more-or-less uniform yielding in the energy dissipation zones of the bridge (this is the case of irregular pier configurations), inelastic analysis is strongly recommended for verifying the design initially carried out using standard modal (response spectrum) analysis. Nonlinear analysis is clearly the preferred choice in the case of assessing existing bridges not properly detailed for seismic performance; for such bridges the simplest choice is standard pushover analysis, but whenever more than one mode affects the response (this is very often the case in the transverse direction of the bridge) more sophisticated tools like the multimodal pushover method presented in sections "Nonlinear Static Analysis of Concrete Bridges" and "Comparative Case Study" have to be used. The use of such analytical tools is expected to increase when the software required for applying them in a single run becomes widely available.

Cross-References

▶ Buildings and Bridges Equipped with Passive Dampers Under Seismic Actions: Modeling and Analysis
▶ Seismic Analysis of Steel and Composite Bridges: Numerical Modeling

References

AASHTO (American Association of State Highway and Transportation Officials) (2010) LRFD bridge design specifications, 5th edn. American Association of State Highway and Transportation Officials, Washington, DC

ACI (American Concrete Institute) Committee 341 (2014) Seismic analysis and design of concrete bridge systems (ACI341.2R-14). American Concrete Institute, Detroit

ASCE (American Society of Civil Engineers) (2007) Seismic rehabilitation of existing buildings – ASCE standard 41-06. American Society of Civil Engineers, Reston

ATC (Applied Technology Council, USA) (1996) Improved seismic design criteria for California bridges: provisional recommendations. ATC report no ATC-32. Applied Technology Council, Redwood City

Aviram A, Mackie KR, Stojadinović B (2008) Guidelines for nonlinear analysis of bridge structures in California. PEER report 2008-03. University of California, Berkeley

Bozorgzadeh A, Megally S, Restrepo JI, Ashford SA (2006) Capacity evaluation of exterior sacrificial shear keys of bridge abutments. J Bridg Eng ASCE 11(5):555–565

Caltrans (California Department of Transportation) (2013) Seismic design criteria version 1.7. Caltrans Division of Engineering Services, Sacramento, California

CEN (2005a) Eurocode 8: design of structures for earthquake resistance – part 2: bridges (EN 1998-2). CEN, Brussels

CEN (2005b) Structural bearings – part 3: elastomeric bearings (EN 1337-3). CEN, Brussels

CEN (Comité Européen de Normalisation) (2004) Eurocode 8: design of structures for earthquake resistance – part 1: general rules, seismic actions and rules for buildings (EN1998-1-1). CEN, Brussels; Corrigenda: July 2009 and Jan 2011

Chopra AK, Goel RK (2002) A modal pushover analysis procedure for estimating seismic demands for buildings. Earthq Eng Struct Dyn 31(3):561–582

Computers & Structures Inc. (CSI) (2011) SAP2000 – version 15.0.1: linear and non linear static and dynamic analysis and design of three-dimensional structures. Computers & Structures Inc. (CSI), Berkeley

Fajfar P (1999) Capacity spectrum method based on inelastic demand spectra. Earthq Eng Struct Dyn 28(9):979–993

Fardis MN (1991) Member-type models for the non-linear seismic response analysis of reinforced concrete structures. In: Donea J, Jones PM (eds) Experimental and numerical methods in earthquake engineering. CEC, JRC, Ispra

FEMA (1997) NEHRP guidelines for the seismic rehabilitation of buildings, FEMA-273. FEMA, Washington, DC

FHWA (Federal Highway Administration, USA) (2006) Seismic retrofitting manual for highway bridges part 1 – bridges (FHWA- HRT-06-032). Turner-Fairbank Highway Research Center, McLean

Hwang JS (2002) Seismic design of structures with viscous dampers. In: International training program for seismic design of building structures. NCREE, Taiwan

Kappos AJ, Sextos AG (2001) Effect of foundation type and compliance on the lateral load response of R/C bridges. J Bridg Eng ASCE 6(2):120–130

Kappos AJ, Sextos AG (2009) Seismic assessment of bridges accounting for nonlinear material and soil response, and varying boundary conditions. In: Coupled site and soil-structure interaction effects with application to seismic risk mitigation, NATO science for peace and security, series-C. Springer, Dordrecht, pp 195–208

Kappos AJ, Manolis GD, Moschonas IF (2002) Seismic assessment and design of R/C bridges with irregular configuration, including SSI effects. Eng Struct 24(10):1337–1348

Kappos AJ, Saiidi M, Aydinoglu N, Isakovic T (2012) Seismic design and assessment of bridges: inelastic methods of analysis and case studies. Springer, Dordrecht

Kappos AJ, Gkatzogias KI, Gidaris I (2013a) Extension of direct displacement-based design methodology for bridges to account for higher mode effects. Earthq Eng Struct Dyn 42(4):581–602

Kappos AJ, Paraskeva TS, Moschonas IF (2013b) Response modification factors for concrete bridges in Europe. J Bridg Eng ASCE 18(12):1328–1335

Kappos AJ, Papadopoulos I, Tokatlidis A (2014) Design of a seismically isolated railway viaduct over axios river in Northern Greece. In: Railways 2014: the second international conference on railway technology, Ajaccio, Corsica, 8–11 Apr, Civil-Comp Press, paper no 80

Katsaras CP, Panagiotakos TB, Kolias B (2009) Effect of torsional stiffness of prestressed concrete box girders and uplift of abutment bearings on seismic performance of bridges. Bull Earthq Eng 7:363–375

Mergos PE, Kappos AJ (2012) A gradual spread inelasticity model for R/C beam-columns, accounting for flexure, shear and anchorage slip. Eng Struct 44:94–106

Ntotsios E, Karakostas C, Lekidis V, Panetsos P, Nikolaou I, Papadimitriou C, Salonikios T (2009) Structural identification of Egnatia Odos bridges based on ambient and earthquake induced vibrations. Bull Earthq Eng 7(2):485–501

Papanikolaou VK, Kappos AJ (2009) Numerical study of confinement effectiveness in solid and hollow reinforced concrete bridge piers: Methodology. Comput Struct 87(21–22):1427–1439

Paraskeva TS, Kappos AJ (2010) Further development of a multimodal pushover analysis procedure for seismic assessment of bridges. Earthq Eng Struct Dyn 39(2):211–222

Paraskeva TS, Kappos AJ, Sextos AG (2006) Extension of modal pushover analysis to seismic assessment of bridges. Earthq Eng Struct Dyn 35(10):1269–1293

Priestley MJN, Seible F, Calvi GM (1996) Seismic design and retrofit of bridges. Wiley, New York

Vayas I, Iliopoulos A (2014) Numerical modeling guidelines for seismic analysis of steel and composite

bridges. In: Encyclopedia of earthquake engineering. Springer, Berlin

Zhang J, Makris N (2002) Kinematic response functions and dynamic stiffnesses of bridge embankments. Earthq Eng Struct Dyn 31(11):1933–1966

Seismic Analysis of Masonry Buildings: Numerical Modeling

Christos A. Zeris and Elizabeth Vintzileou
Department of Structural Engineering, Faculty of Civil Engineering, National Technical University of Athens, Athens, Greece

Synonyms

Adobe; Confined masonry; Finite elements; Load-bearing masonry; Masonry building reinforced concrete; Reinforced masonry; Seismic analysis; Unreinforced masonry

Introduction

Masonry building construction encompasses the large inventory of structures built worldwide solely or partially of individually laid masonry units bonded or not together with some type of mortar, possibly with the incorporation of reinforcement; depending on the constituent materials, stacking, and bonding techniques, different technologies are adopted for building the masonry elements that comprise the entire building; thus, materials and techniques adopted worldwide vary with local customs, socioeconomic conditions, and available technology. A comprehensive continuously updated description of the different materials and technologies of construction around the world is given in the World Housing Encyclopedia (EERI/IAEE).

Because of their widespread use, masonry structures house not only the population but also important social and economic operations such as hospitals and schools, high congregation areas, business, small to medium industry, and civil administration (UNIDO 1983); furthermore, being historically one of the earliest forms of construction, they also house the people's cultural heritage, most often being such by themselves (Figs. 1, 2, 3, and 4).

Due to their geographic spread, masonry structures are exposed to different levels of seismic hazard; in fact, a significant portion of cultural heritage structures are located in earthquake prone-areas of Europe, Asia, and South America. Recent devastating earthquakes (e.g., in New Zealand, Italy, Chile, India, Pakistan, and elsewhere) have shown that existing masonry structures are quite vulnerable to seismic actions, as also recorded in many seismic damage reconnaissance reports published after major earthquakes (e.g., among others, Hughes et al. 1990; Rossetto et al. 2009; DesRoches and Comerio 2011).

As a consequence of this fact, social and economic requirements for human safety and operability as well as the need for preservation of cultural heritage require more and more often that these structures be analyzed for seismic actions, in order to be designed (new construction), or their seismic vulnerability can be assessed and evaluated (existing construction) for the purpose of repair, rehabilitation, and/or strengthening to current seismic standards. The seismic performance assessment of existing masonry buildings also follows performance-based design (PBD) and analysis procedures, similar to other types of structures (e.g., concrete and steel buildings): hence, in the design of new construction, normative regulations, practices, and experience of good seismic performance have been encompassed (e.g., EC6 2005 for gravity load design and EC8 2004 for seismic design). On the other hand, codes and guidelines are being drafted for the assessment and retrofit of existing masonry buildings (e.g., FEMA-356 2000) or for cultural heritage structures (Moro 2007; ICOMOS; ISCARSAH).

Compared to the more recently evolved types of construction like steel and reinforced concrete, however, masonry buildings have certain inherent idiosyncrasies, which their modeling for seismic analysis should account for:

S

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 1 The stone masonry castle of Mycenae, Peloponnese, Greece, 2nd millennium BC (Source http://en.wikipedia.org/wiki/Mycenae)

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 3 The Alcantara bridge (104 AD), Spain (Source: http://en.wikipedia.org/wiki/Roman_bridge)

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 2 The Byzantine church of Hagia Sophia (360 AD), Istanbul, Turkey

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 4 The Aoos river bridge (eighteenth century), Hepirus, Greece

(i) Masonry buildings are both non-engineered, namely, structures built with traditional means and experience, and engineered, namely, structures designed and constructed following engineering principles and, more recently, code regulations (e.g., EC6 2005; EC8 2004). In fact, earlier engineered masonry buildings have been designed for gravity loads only; only some types of recent masonry construction, following the evolution of seismic regulations, have also been designed for earthquake (IAEE 2008; ASCE 2013; EC6 2005). In addition to the possible lack of proper design, they are also characterized by the usual problems in the load-bearing system similar to other structures, such as: irregularities in plan and/or elevation, improper foundation conditions, a history of (possibly undocumented) modifications in plan and elevation, and the decay of the material properties under environmental exposure.

(ii) Depending on the prevailing socioeconomic conditions, masonry buildings are and have

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 5 Residential masonry buildings in Greece, built in the nineteenth and early twentieth century, reflecting different levels of socioeconomic status

been constructed using a variety of locally available technologies and materials (Fig. 5). Often, currently acceptable levels of quality control and supervision were not enforced. New masonry buildings in developed countries may use units manufactured industrially and with quality control; this is not the case for existing buildings in all countries or new masonry construction in underdeveloped countries (see typical examples offered in the World Housing Encyclopedia, EERI/IAEE).

(iii) Masonry, unless suitably reinforced or confined, as in the case of new construction, cracks and fails in a brittle manner. Seismic analysis methods for existing or historical unreinforced masonry buildings need to take this fact into account.

(iv) Masonry buildings suffer from structural inadequacies in the load-bearing system inherent in this type of construction, such as: the presence of diaphragms which are poorly connected to the masonry for lateral load transfer, the relatively high flexibility of the floor diaphragms, and the way the masonry wall elements are constructed and tied together through the thickness and/or at their intersections. Seismic analysis modeling techniques should not overlook these particularities, if the analysis results are to be reliable.

Depending therefore on the problem at hand, reliable modeling for seismic analysis of masonry structures will have to account of these building characteristics, in order to reliably predict the damages expected and to identify the methods and extent of intervention required for strengthening these structures. Following a brief review of masonry building characteristics, the methods and limitations of different methodologies adopted for the seismic analysis of masonry construction are subsequently considered.

Modeling Techniques for Seismic Analysis of Masonry Buildings: Classification and Definitions

Masonry Materials Used

According to the type of masonry units employed, masonry is classified as stone masonry (Fig. 5), industrially manufactured clay brick (solid, hollow), site-produced mud brick or plinth (also known as adobe masonry), or industrially manufactured hollow or solid concrete masonry units (CMU) in fully, partially grouted, or ungrouted construction. Other industrially manufactured types of block are also used worldwide such as lightweight concrete, cinder, fly ash, or autoclaved concrete, among others. Depending on the type of construction, reinforcement is also used to enhance the bearing capacity of masonry elements and to provide ductility to the structure. Another construction system (covered by modern codes, e.g., EC6 2005) is that of confined masonry: horizontal and vertical reinforced concrete elements are provided, at distances depending on the dimensions of the building and on the seismicity of the region. The ties, constructed during the construction of the masonry, function as linear tensioned members. They assist masonry in taking shear and bending (in and out of plane) while they also contribute to the confinement of plain masonry, thus enhancing the ductility of the system.

In cultural heritage buildings in Europe and elsewhere, most frequent construction types of stone masonry are double-leaf masonry, made of two leaves, either independent or connected between them with sporadic header stones (Fig. 6), and three-leaf masonry, made of two independent leaves, with the space between them being filled with a more or less loose material of poor mechanical properties. Other types of masonry, like cavity masonry or timber reinforced one, are also quite frequent in historic structures.

Finally, the material that bonds the masonry units together may be none (also called dry construction), mud, or different types of mortar such as lime, lime–pozzolan, lime mortar reinforced with animal hair, or, in case of modern masonry construction, cement or lime–cement.

Load-Bearing Function of the Masonry

The type of load path of the vertical and transverse loads down to the foundation defines two basic types of masonry:
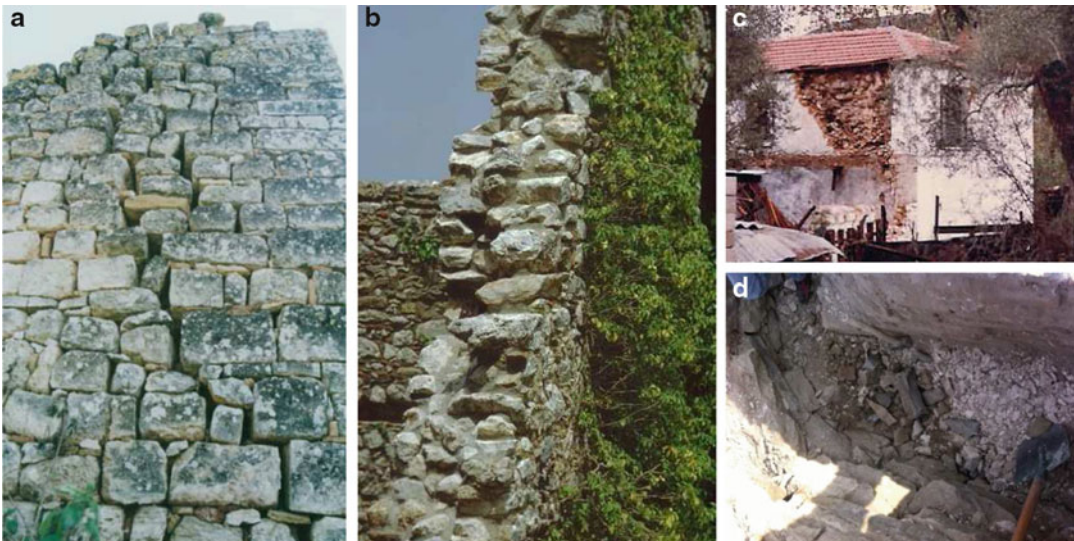
Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 6 Unreinforced masonry construction technologies. (**a**) Dry masonry construction (Tiryns, Argos, -Peloponnese, thirteenth to fourteenth century BC). (**b**) Double-leaf masonry wall. (**c**) Three-leaf rubble stone masonry. (**d**) Cavity wall (Van, Turkey)

(i) **Load-bearing (LB) masonry**, bearing all the gravity and lateral loads from floors and roof to the foundation. Thus, the masonry elements themselves are responsible to provide overall lateral stability and to transfer the seismic inertia loads under in-plane and out-of-plane actions, down to the foundation, through combined flexural, axial, shear, or rigid body rocking mechanisms. Seismic analysis of this type of masonry involves adequate modeling of the entire load-bearing structural system response at the individual element level, including the diaphragms and the foundation, accounting for both in-plane and out-of-plane response, as discussed in the next section. As a special case of LB masonry construction, masonry arches are constructed with keystone elements at the apex, in order to provide the path of the line of thrust to the foundation.

(ii) **Non-load-bearing (NLB) masonry**, in which the bearing function is provided by a structural skeleton made of other structural materials such as steel and timber of reinforced concrete (RC). Such is the case of the widely adopted masonry panel infilled RC building construction, a form of confined masonry, whose modeling is covered in section "Modeling of NLBM Infilled Frame Buildings for Seismic Analysis."

If anything else, NLBM infill panels have to support with adequate resistance capacity and in a stable manner their own inertia forces as these are materialized at the building elevation relative to ground in which they are supported. Furthermore, through the deformation compatibility between the infill and the confining frame structure, they are forced to resist seismic load through a friction contact separation mechanism with the frame under lateral load response. Therefore, unless specifically isolated by adequate details from the bearing frame (over the entire lateral inelastic deformation expected), NLBM panels are seismically bearing, and consequently, even NLBM elements need to be included in seismic analysis of the entire building in which they are constructed.

Because of this contact mechanism, practical seismic analysis of this type of masonry construction involves macromodeling using axial load-bearing struts, as briefly discussed in section "Modeling of NLBM Infilled Frame Buildings for Seismic Analysis."

Masonry Wall Construction Types and Resisting Mechanisms

Masonry walls, whether LBM or NLBM, range from single-leaf or multiple-leaf walls (Fig. 6) with interior cavities among the leaves, which are either void (e.g., the common double-leaf walls with space for sliding window frames) or are filled with rubble (typical for old masonry buildings, cultural heritage buildings included) or concrete (in modern masonry construction of reinforced cavity walls according to EC6 2005). Multiple-leaf walls may be tied to each other in modern construction (e.g., modern veneer brick masonry walls) or, as often the case in existing structures, they are untied.

The laying of the masonry units varies according to the local techniques, the material, and the production form of the unit. Industrial units are laid in courses (use of header courses or random laying construction and keystones at the intersections) and the provision of collar joints filled with mortar (and possibly steel) and/or the inclusion of bond beam elements. The vertical elements are traditionally built on a stone or brick or concrete footings, on which they rest or are tied to with reinforcement. Often, interior walls are in fact lighter construction partitions of brick or wood not tied to the load-bearing system, or the wall intersections were poorly connected.

Irrespective of material and load-bearing type, masonry is classified according to its resisting mechanisms as unreinforced masonry, reinforced masonry (prestressing possibly included), and confined masonry, depending on whether reinforcement or additional confining elements within the masonry are used. The use of each type in new constructions in seismic regions depends on the seismicity of the region, on the number of stories, etc. (see, e.g., EC8 2004).

(i) **Unreinforced masonry (URM)**, namely, masonry without any or very small amounts
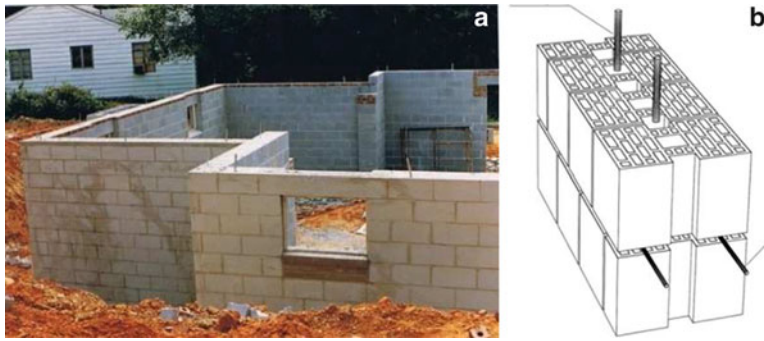
Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 7 Reinforced masonry construction technology and detail. (**a**) Modern CMU reinforced masonry residential building (Source: http://en. wikipedia.org/wiki/Masonry_structure). (**b**) Reinforced masonry construction detail using clay bricks with improved thermal insulation properties

of reinforcement included and without any additional confining members within the masonry body. This form of masonry construction is the most frequently encountered type of construction in existing or historical buildings and is associated with the largest amount of modeling problems, due to the variety of techniques and materials adopted and the variability of their properties. Seismic analysis methods for URM buildings need to account for the mechanical characteristics of the material, obtained from in situ evaluation using flat jacks or through testing (e.g., Clough et al. 1979; Magenes et al. 2008; Vintzileou and Miltiadou-Fezans 2008; Ruiz-Garcia and Negrete 2009; and the TCCMaR test series, Kingsley 1994); furthermore, modeling should account for the brittle nature of its response. Only elastic methods of analysis are therefore meaningful in their seismic performance assessment, whereas a low behavior factor ($\sim$1.50) is applicable for evaluation of the seismic forces.

(ii) **Reinforced masonry (RM)** is masonry with horizontal reinforcement in mortar bed joints and vertical reinforcement positioned in a cavity or in holes of the vertically perforated masonry units or in grooves between adjacent blocks (Fig. 7a, b).

In modern construction, reinforcement comprises steel reinforcing in the form of bars, trusses or cut wire mesh, dovetails, and other special shaped proprietary ties, placed vertically and horizontally. Other reinforcing materials include prestressing strands (prestressed masonry) and polymers, used for external strengthening existing masonry. In addition to steel reinforcement used in modern construction, timber elements have traditionally been used in historical masonry structures and are still used in new buildings constructed in seismic regions in the developing world, with very good seismic performance characteristics (also called timber-laced buildings, Figs. 8 and 9).

Generally, reinforcement serves to connect the leaves of the masonry wall, to allow for uniform distribution of vertical loads at floors and roof levels, to act as shear and flexural reinforcement for in-plane and out-of-plane seismic load transfer, and also to provide ductility, when reinforcement yielding precedes brittle failure of the masonry; therefore, its beneficial role should be included in seismic analysis; furthermore, for new construction, when ductility is enforced by design, both elastic and inelastic methods of analysis are meaningful in seismic performance assessment, with a suitable response reduction coefficient applied in the expected seismic loads; in this context, behavior factors up to 3.0, comparable to RC wall construction, can be used in modern RM building design (EC6 2005).

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 8 Traditional timber-laced stone masonry buildings in earthquake-prone areas. (**a**) Antalya, Turkey. (**b**) Kastoria, Greece

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 9 Bhatar construction: traditional unreinforced load-bearing timber-laced masonry structure in Pakistan in modern construction (Source: http://www.holcimfoundation.org/Projects/advocacy-of-traditional-earthquake-resistant-construction-north)

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 10 Model of the Gaiola pombalina masonry confinement with timber, developed in Lisbon after the 1755 earthquake (http://en.wikipedia.org/wiki/Gaiola_(construction))

(iii) **Confined masonry (CM)** is masonry which encompasses during construction horizontal and/or vertical confining RC or RM elements, monolithically bonded into the masonry structure (e.g., RC lintel beams and pilasters). In fact, modern clay brick walls in seismic regions make use of such horizontal and vertical RC elements. In the case of RC confining elements, similar modeling conventions for seismic analysis are adopted as for RM buildings, using however higher seismic forces; in this case, a behavior factor of 2.0 is adopted in EC8 (2004).

A special form of URM and LBM construction originating in older (historic) buildings (a well-known case being the Pombalinas, constructed during the rebuilding of Lisbon following the devastation of the city on 1755) and still adopted in less developed countries (e.g., the *dhajji* in Pakistan) makes use of timber reinforcing elements in the form of tension–compression braces; they are built within the masonry walls at the floor levels or between floors, in order to confine the masonry infilling the voids (Fig. 10), thereby enhancing its in-plane shear resistance and providing damping under seismic response. This type of LBM construction is also behaving as CM and partially RM, due to the load-carrying mechanism provided by the timber elements and

S

the in-plane stabilization and stiffening role provided by the masonry infill, confined in turn by the timber structure; seismic modeling of such systems follows the macromodel approach (see section "FE Modeling of LBM Structures for Seismic Analysis"), whereby the confining timber elements are included in the model together with the masonry and are verified accordingly.

Seismic Analysis Methods of Masonry Buildings

Depending on the problem at hand, both linear and nonlinear analysis methods, as also employed for seismic analysis of RC and steel buildings, are being applied to LBM buildings according to the limit state objective of the verification and the ductility capacity of the masonry elements (if any).

Linear Elastic Analyses
For serviceability limit state (cracking, service load deformation) verifications, linear analysis methods are adequate, up to the onset of cracking of the masonry. The use of linear elastic analysis for the ultimate limit state (strength) design verification of URM buildings under seismic load combinations provides meaningful verification results only under seismic load levels obtained using a response reduction coefficient equal to 1.0; for the design of new RM buildings, it is possible to use linear elastic methods with higher behavior factors (up to 3.0), due to the inherent ductility capacity of the masonry. It should be noted, however, that in case of historic buildings, lower seismic actions than those prescribed by modern codes are frequently adopted. Actually, the application of the requirements of current codes may lead to invasive interventions that are against the internationally accepted charters for the preservation of the built cultural heritage (e.g., ICOMOS; ISCARSAH).

Linear elastic analysis methods can also be adopted for ultimate limit state verifications of URM buildings, provided the structure does not have excessive irregularities (EC6 2005) and lacks significant torsional effects, thereby exhibiting clearly separated orthogonal modes of vibration. Linear elastic modeling neglects cracking of the masonry elements and its influence in the kinematics and the redistribution of forces. Even if cracking is neglected, however, linear elastic models provide useful information about the structure and the model adopted since: (i) they give an indication of the areas of increased tensile or compressive stress in the masonry, which potentially need to be strengthened or rehabilitated; (ii) they can be easily compared with a visual or in situ measurement of cracking and stresses in the masonry, thus allowing to verify the reliability of the model and the reasons for the existing condition of the structure; and iii) possible interacting factors of overstress can be established (e.g., earthquake following a long-term preexisting foundation settlement).

Static and Modal Elastic Analyses
Linear elastic analysis includes both equivalent static and modal analysis:

(i) Static analysis can be used in buildings in which higher modes are not dominant in the response, the building is orthogonal in plan, and it does not exhibit major irregularities: two different lateral load distributions with height should be considered, representative of different modes of lateral deformation, namely, (a) triangular distribution and (b) uniform distribution. In the case of micromodels (see section "FE Modeling of LBM Structures for Seismic Analysis"), possible refinements on the linear elastic model can be considered through local modification of the finite element (FE) stiffness characteristics in order to account for cracking of the masonry, whether such is predicted from an elastic analysis or it is obtained from field inspection of the condition of the structure.

(ii) Modal analyses, when used, should include all the modes with a modal mass that is greater than 5 % of the total mass of the building and enough modes whose modal masses sum to at least 75 % of the total building mass.

For the evaluation of the modal response characteristics – deformations and internal forces – using macromodels (see section "FE Modeling of LBM Structures for Seismic Analysis"), the square root of the sum of squares combination rule of modal quantities can be used if the modes differ with each other by as much as 90 %, or, better, using the complete quadratic combination rule, giving accurate maxima for closely coupled mode combinations. For seismic modal analyses using micromodels (e.g., shell finite elements), the peak response characteristics (deformations and internal stresses) E_{max} should be evaluated following:

$$E_{max} = \max_t \left\{ \sum_{j=1}^{n} E_j(t) \right\} \qquad (1)$$

where $E_j(t)$ is the time history of the corresponding parameter due to response in the jth eigenmode.

Inelastic Analyses

In currently accepted methodologies for the assessment of existing LBM buildings (both RM or URM) in the context of PBD, nonlinear static analysis methods and corresponding modeling conventions are employed with different levels of modeling detail, depending on whether cracking, post-ultimate, and cyclic hysteretic characteristics are included in the model (Fig. 11). The purpose of these methodologies is to evaluate the inelastic lateral load deformation of the building, taking into account the actual inelastic characteristics of the elements, brittle or ductile; evaluation of this capacity curve yields the expected target deformation demands, under different seismic excitation levels (performance levels), at which point the onset, distribution, and extent of structural damages are obtained and compared to code damage levels (FEMA-356 2000), either in terms of element resistance (brittle element response) or in terms of inelastic deformations (ductile element response). For the evaluation of these generalized force–deformation characteristics, different failure mechanisms – accounting also for out-of-plane effects – can be considered, and the weakest governing mechanism should be adopted as governing the failure response.

Modeling of LBM Buildings for Seismic Analysis

The primary load-bearing elements of LBM construction are the vertical load-supporting elements, namely, the floor and roof structure, as well as the perimeter and interior bearing masonry walls. All these elements carry the vertical loads (including self-weight) and the lateral forces to the foundation. Secondary elements, not part of the lateral resisting system (such as light partitions), are not included in the model as earthquake resisting elements. Only their self-weight alone is taken into account in analysis.
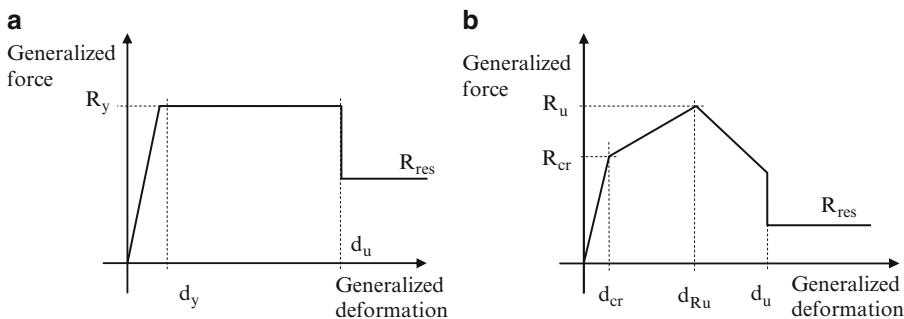
Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 11 Inelastic analysis constitutive models of masonry building elements for in-plane only action (bending or shear). (**a**) Simplified. (**b**) Refined

Floor and Roof Elements and Diaphragmatic Action

Floor structures and the roof structure transfer the vertical loads and in-plane inertia seismic loads to the rest of the bearing elements. While in modern masonry buildings floor elements are typically stiff RC diaphragms, in typical existing or historical masonry construction, light wooden floors on wood or steel purlins have been used traditionally. Brick or stone masonry vaults are also typical in the lower story of several historic structural systems; other variants include arched brick constructions between joists or two-way Zoellner diaphragms made of brick infilled voids. Purlins usually span one way and are simply resting or encased in the bearing walls at each end. A peculiarity of masonry buildings is that masses are not concentrated at floor levels; they are distributed over the height of the building.

Unlike modern construction concrete slabs which provide diaphragmatic action that distributes the inertia loads in plan and tie all the vertical masonry elements at the floor level, existing masonry structures were constructed with flexible diaphragms which deform in plane and operate differently during the earthquake (Fig. 12a). Therefore, the presence or absence of diaphragmatic action and the way the diaphragm is tied to the vertical elements are two important aspects to consider in seismic modeling, since the diaphragm stiffness will affect both the dynamic characteristics of the masonry building and the transfer of forces among the stiff vertical wall elements.

Along the same context, the function of the diaphragm (one-way flexible or two-way stiff action) shall also define the distribution of the floor plan masses to the walls: one-way joist diaphragms will only distribute inertia reaction loads across the walls at which their wooden joists are inframed, making the usual uniform mass distribution assumption in the building model, namely, a lumped rotational/translational mass at the center of mass, incorrect.
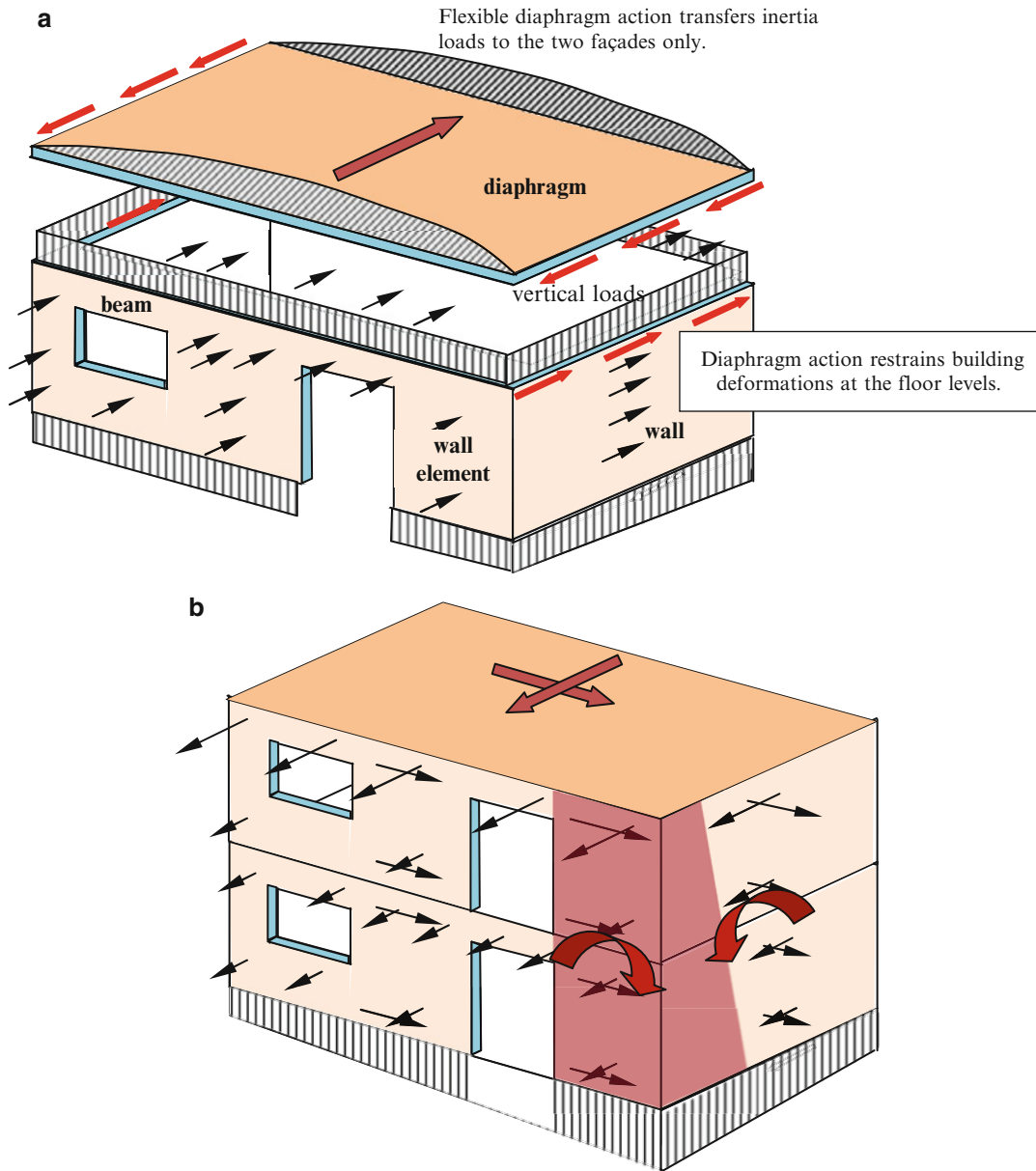
In addition to load distribution, failure of the diaphragm is also important to model: apart from failure of arch-supported diaphragms that tend to be sensitive to relative in-plane deformations of the supporting walls, failure or collapse of flexible diaphragms takes the form of failure (usually pullout) of the diaphragm-to-wall connection, due to inertia force transfer or excessive out-of-plane deformation of the wall and loss of support of the timber elements to the wall. This type of behavior depends mostly on the vertical support system of the diaphragm and the detailing of the connections, which all need to be part of the building model (see, e.g., Vintzileou et al. 2007, for a description of the system used in the traditional masonry structures in Lefkada, Greece).

Seismic Load-Bearing (Primary) Vertical Elements

The primary lateral load-bearing elements of LBM buildings are the perimeter and interior bearing walls, which are typically perforated with openings, forming wall elements, spandrel beams (also arched lintel beams), and pilasters (Fig. 12b). The in-plane lateral load transfer of the wall elements depends on their aspect ratio (the height to width ratio): walls with relatively longer width compared to the element height (height to width ratio less than 2–3) tend after initial cracking to transfer the lateral force to the lower level directly through an inclined strut (including also the vertical load), whose horizontal component equals the lateral load; more slender walls or multiple wall elements (wallets) created in a wall with openings may opt for a more flexure-dominated behavior (similar considerations apply also to the horizontal spandrel beams between openings), while failure of these elements is brittle.

Since the walls are the primary lateral load resisting elements, their distribution in plan and their stiffness (namely, geometric size and percentage of opening area) determine the eccentricity between the center of application of the floor inertia forces (the center of mass for stiff diaphragms) and the center of rigidity of the building in plan. Consequently, irregularities that may arise in plan and also in elevation due to the wall distribution and geometry will influence the distribution of the lateral loads among the different resisting elements, something which is crucial to model in seismic analysis. It should

a

Flexible diaphragm action transfers inertia loads to the two façades only.

diaphragm

vertical loads

beam

Diaphragm action restrains building deformations at the floor levels.

wall

wall
element

b

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 12 Good practices in the seismic analysis of masonry buildings. (**a**) Contributions of the seismic forces to the masonry façade in a typical URM building through diaphragmatic action (where this exists). (**b**) Contributions of the three-dimensional analysis to the in-plane and out-of-plane actions on the facades and to the corner element forces

further be overemphasized that these structural characteristics are dependent on the intensity of the earthquake, since masonry (like concrete) cracks in tension or softens under extreme compression stresses, and therefore the relative stiffness of the bearing elements changes beyond what has been assumed in an elastic analysis, redistributing forces as well as increasing the lateral deformation demands during the an earthquake. The analysis sophistication therefore and the modeling capabilities should reflect the level of response prediction.

S

In-plane response failure of the walls or other structural elements (where present) will take place due to inadequate resistance or excessive interstory drift. Masonry element failure includes wall pier, pilaster, or spandrel beam failure under in-plane actions. Depending on the aspect ratio of these elements and the existence of diaphragmatic action, failure of the wall elements affects individual masonry elements and will jeopardize the entire vertical load path to the lower floors, the building dynamic characteristics, as well as the redistribution of forces in and out of plane. Element failure takes the following forms:

(i) Flexure-dominated failure including cracking and/or rocking of the wall, compressive toe failure of the wall

(ii) Shear failure in plan, including sliding at mortar bed joints, diagonal cracking or diagonal crushing of the wall between the cracks

(iii) Entire wall movement

(iv) Apart from these in-plane response failure modes, walls may also fail under concurrent in-plane and out-of-plane action. In this case failure is closely associated with the existence or not of adequate diaphragmatic action and a suitable tensile diaphragm connection to the façade walls that will mobilize the entire building façade from the footing line (below grade) to the roof (Fig. 12a). Special cases of out-of-plane bucking failure under in-plane compressive action are also included in this combined failure mode, primarily for slender wall elements though, e.g., unsupported multiple-leaf walls.

(v) Failure under combined biaxial effects (combined axial load) characterizes the corner walls and pilasters, due to the three-dimensional nature of the building response. Bidirectional rocking of the building (Fig. 12b) induces axial loads that are considerably higher than uniaxial predictions, together with biaxial bending and shear forces. In this case, a three-dimensional model of the building is needed to predict such overstress in the corners, with the results of plane analyses only being unconservative.

Secondary Elements

Such elements are typically the interior masonry (or other) column elements that support the diaphragm or narrow, slender elements on the façade that function as columns. The failure of these elements does not alter significantly the response of the building, and therefore, they are ignored in the seismic analysis model. However, their ability to bear vertical loads under the seismic deformation of the building should still be verified.

Foundation

The foundation of the masonry walls is not necessarily responding as rigid and non-deforming, as typically assumed in seismic modeling. Foundations may settle in the long term under the weight of the walls, inducing tensile cracks in the building that change the distribution of forces and the way the vertical elements respond (rocking rather than flexing). Furthermore, due to the lack of reinforcement, contact type of connection develops between the URM footing, the masonry wall, and the soil, with possible uplift and rocking under tensile or rocking response. These mechanisms should be captured in the seismic model, if the vertical and lateral loads are such as to allow for this kind of motion. Consequently, if preliminary analyses indicate this to be the case, the soil resistance to the footing stresses should be modeled using an elastic but tensionless type of behavior (e.g., a Winkler model with uplift), in order to obtain the proper footing flexibility as is the case in situ. An in situ geotechnical study and adequate knowledge of the foundation shall provide the soil constitutive characteristics and the foundation conditions (e.g., the possible existence of a well or a septic tank near the footing).

Principle Modeling Requirements for Seismic Analysis of LBM Buildings

In summary, the building seismic model used should identify the basic force transfer mechanisms of the structural system, irrespective of the method of analysis used:

(a) *Plane or three-dimensional model*. Plane frames are often used as an approximation for modeling the building. It is important to

note that the proper idealization of seismic behavior is through modeling of the three-dimensional response. Two-dimensional models are unable to monitor the three-dimensional response of the entire structure; consequently their use will not predict the spatial response effects, namely, (i) the simultaneous action of seismic axial forces and biaxial bending effects in the corner piers, (ii) the corner element vertical deformation compatibility under concurrent actions in the two orthogonal directions, (iii) the influence of out-of-plane bending in the resistance of in-plane actions of the vertical elements, and (iv) the possible in plan torsional effects which will enter into the response in the case of a relatively rigid diaphragm and eccentric distributions of the mass (e.g., an opening) or the stiffness (e.g., asymmetric facade opening distributions and/or interior masonry walls).

(b.1) *Modeling of the diaphragm.* The presence or lack of a diaphragm in the model should be in accordance with the function of the diaphragm in the structural system at hand. The in-plane rigidity of the diaphragm is an important consideration to account for in the model, particularly because of the fact that masonry buildings tend to be very stiff and the relative diaphragm-to-wall element in-plane stiffness will define the distribution of inertia loads from the floor to the vertical seismic load resisting elements.

(b.2) *Modeling of the diaphragm connection with the walls.* Furthermore, as far as the modeling of the diaphragm is concerned, it is important to consider in the model whether its connections with the vertical masonry elements justify the use of a deformation fully compatible FE nodal connection, transferring load. The same is true for the roof structure, which is normally simply supported on the masonry walls, and sliding of the roof trusses on the walls is often possible under differential lateral seismic movements across walls.

(c) *Modeling of the mass.* The conventional frame analysis assumption of a lumped mass idealization is not justified in masonry modeling, and, instead, the model should incorporate distributed masses; due to the large masonry element size (compared to the normal operational loads of the building) and often the lack of a heavy concrete diaphragm, loads are not distributed according to the lateral stiffness of the elements. For this reason, three-dimensional modal dynamic analysis should at best be adopted, reflecting more accurately the system deformation and the load path of the inertia forces from the diaphragm to the foundation. Similarly, out-of-plane effects need also to be taken into account in a "distributed with height" sense for assessing the in- and out-of-plane interaction of forces to the walls, the dynamic connection forces at the wall intersection, and the seismic deformations and forces at the roof-to-wall connections.

(d) *Modeling of the building foundation.* Masonry building response and past deformation history are affected by the deforming foundation at the base of the building. Consequently, full foundation fixity is often an unrealistic assumption, given that the stresses under the masonry wall may be relatively high. Evidence of cracking around openings or unsymmetrical distribution of cracks in an existing masonry building is often the effect of differential settlements due to variable ground conditions in plan, such as the presence of an abandoned well or a septic tank at one end, improper ground preparation at the time of construction, or a partial plan basement. For modeling the foundation, an acceptable practical modeling approach makes use of tensionless elastic springs, providing vertical, lateral, and bending restraint in the embedded footings.

FE Modeling of LBM Structures for Seismic Analysis

From the analysis of the complexity of the response, the failure modes, and the fact that different technologies and materials comprise

masonry construction, practical modeling of masonry buildings for seismic analysis relies on the use of FE models and follows two different techniques (Fig. 13):

(i) The less refined "global" (phenomenological) macromodels, suitable for the analysis of entire LBM buildings. In this case, the walls, piers, and spandrel beams of the masonry structure are modeled using one-dimensional line elements with nonlinear characteristics. These macromodel elements (Fig. 13d) are characterized by equivalent axial load, bending, and shear interaction response characteristics. Such models have been used both for elastic and inelastic three-dimensional seismic analysis for design, assessment of seismic vulnerability, and fragility studies of masonry buildings under monotonic or cyclic loading.

(ii) The refined "local" FE micromodels where the entire masonry building or a plane portion of the building, such as a masonry façade under planar response, is modeled using two- or three-dimensional FE approximations, with associated material and loading description (Fig. 13c). In terms of geometric representation, thick shell or across the length and through thickness brick FE approximations are adopted, possibly coupled with beam FE for the reinforcement, if any. Even more refined micromodel approximations have been adopted in research studies, in which the actual masonry unit and mortar have been separately modeled using brick and plate elements, respectively.

In terms of material approximation, phenomena such as cracking and compression nonlinearity or the presence of steel are smeared within the element integration area through equivalent stiffness and resistance modifications at the FE integration point. Models of this type have been used in the parametric investigation of conventional or historical masonry construction and for the validation of test results. Complexity and computational cost and resources are the

primary issue in this case as well as in certain aspects, the actual ability to model the material behavior under cyclic loading conditions.

Macromodels for Entire Building Analysis

Since the use of micromodels is too expensive and complicated for entire building analysis under seismic excitations and due to the fact that available software capabilities are often limited in handling such FE micromodels, equivalent FE macromodels are being used for practical seismic analysis of LB or NLB masonry construction.

For load-bearing masonry, the usual modeling conventions adopted in conventional frame building analysis are also adopted for masonry buildings as well: roof and diaphragm elements are modeled using line FE, or plain diaphragmatic action is enforced (if it exists). Similarly, wall piers, columns (pilasters), and spandrels are modeled at their centerline using equivalent or actual property line FE with linear or nonlinear characteristics, while the element joint regions are modeled using infinitely stiff elements and/or rigid zone transformation models depending on the analysis software conventions.

For inelastic analysis, phenomenological axial, shear, and flexural constitutive relations should be specified, possibly with cracking and post-failure modeling capabilities and, if possible, interaction of axial/flexural and shear stresses. Where this is not possible and given that the variation of axial loads is not high for low-rise buildings, uncoupled values may be assumed based on initial state vertical load levels.

Masonry building macromodels for inelastic analysis evolved from: (i) the simplified weak spandrel strong pier model, whereby spandrels crack early and are neglected (therefore the piers are considered to act as uncoupled cantilevers joint by hinged rigid link beams at the floor levels) (Fig. 14b); (ii) the strong spandrel weak pier model (the shear frame analogy), whereby the piers crack first and are therefore assumed to deform with their inelastic characteristics, the spandrels remaining relatively rigid (Fig. 14c); and (iii) the equivalent frame model, where the masonry structure is modeled as an assembly of
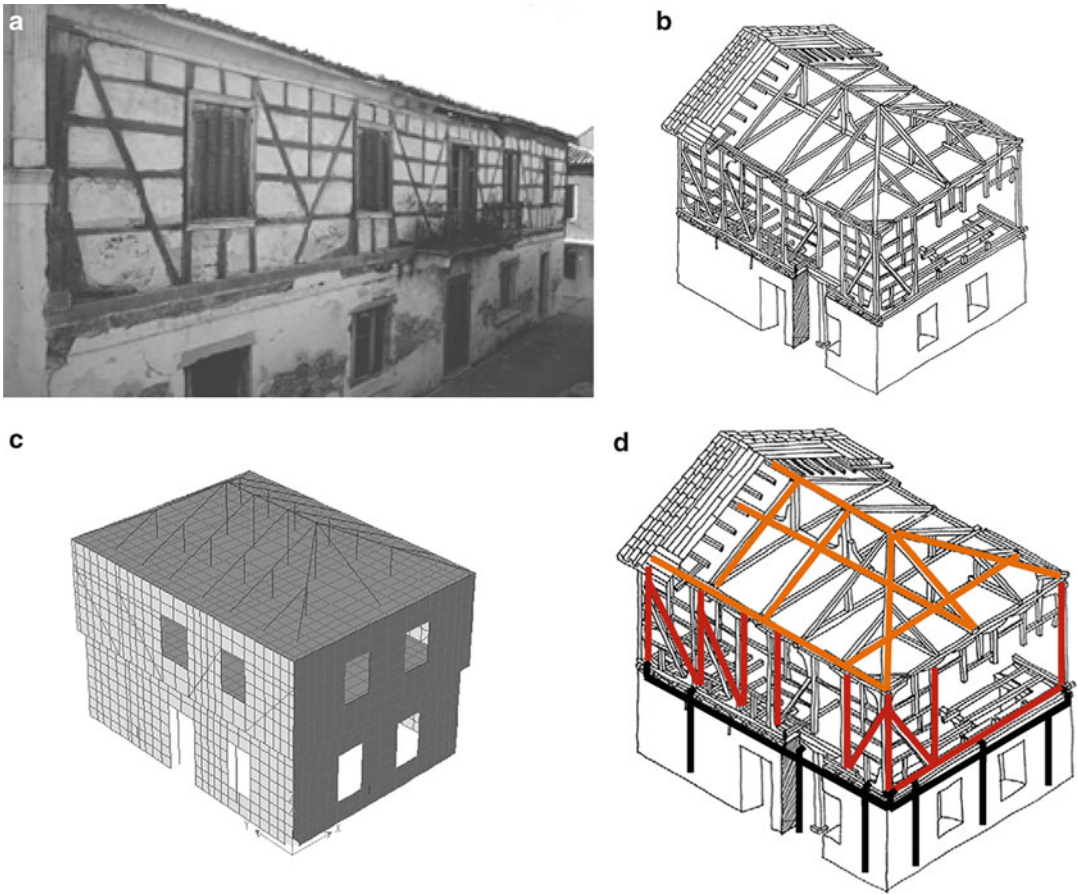
Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 13 Modeling example of a historical masonry building in Lefkada, Greece, built of timber confined plinth masonry walls at the upper floor on a stone masonry ground floor. (**a**) Photo). (**b**) Bearing structure. (**c**) Micromodel. (**d**) Macromodel (Vintzileou et al. 2007). Note that several load-bearing elements in the model are omitted for clarity

vertical pier and horizontal spandrel line elements interconnected by rigid joint regions (Fig. 14d): the geometry of the joints is obtained from the geometry of openings and an equivalent pier height, which is defined by the extent of cracking observed in the vertical elements following an earthquake or, if uncracked, assuming a crack inclination at about 30° that extends from the opening toward the joint (Lagomarsino et al. 2013).

For the constitutive modeling of inelastic FE macromodels under static seismic-type load, the multilinear (simplified or more refined) shear force–interstory drift or bending moment–pier rotation diagrams of Fig. 11 have been proposed

by several investigators for modeling both the in-plane shear and the flexural response of masonry wall piers, incorporating the different failure mechanisms of these elements (Fig. 15).

Magenes and Calvi (1997) proposed that the peak shear resistance R_u (Fig. 15) of a rocking masonry wall pier bearing an axial load P is given by:

$$R_u = \frac{D^2 t}{H_0} \frac{p}{2} \left(1 - \frac{p}{k f_u} \right) \qquad (2)$$

where D and t are the length and thickness of the pier; H_0 is the effective height equal to the shear span, namely, the height to zero moment, taken as

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 14 Equivalent frame model for load-bearing masonry construction. (**a**) The masonry facade to be modeled. (**b**) Weak spandrel and strong cantilever pier model. (**c**) The weak pier and rigid spandrel shear frame model. (**d**) The equivalent frame model using spandrels, piers, and rigid joint regions (Lagomarsino et al. 2013)
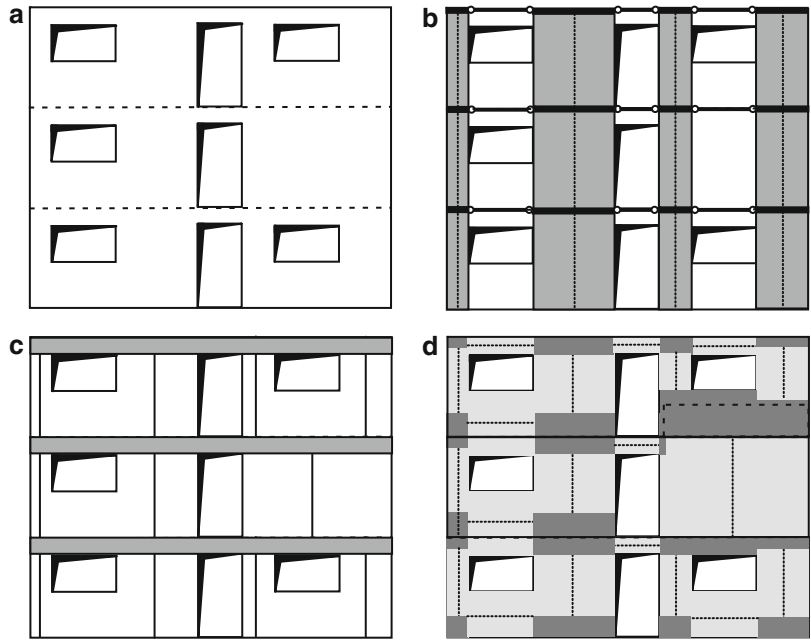
Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 15 Model of the wall pier and lateral load deformation constitutive relation under axial force
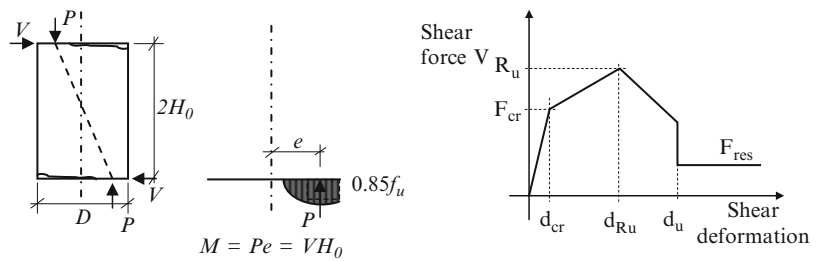
being equal to the pier height, for cantilever piers, and half the pier height, for piers in contraflexure; p is the average vertical load pressure, equal to the axial load P divided by the wall area ($p = P/Dt$); f_u is the compressive strength of the masonry; and k is a coefficient that depends on the stress distribution at the toe of the wall ($k = 0.85$ for an equivalent rectangular stress block; see Fig. 15).

For shear failure mechanisms of brick masonry walls, when mortar bed and head stone failure is involved, they proposed to adopt a Mohr–Coulomb friction model for the wall, namely:

$$R_u = Dt\tau_u = Dt(c + \mu p) = Dt\left(c + \mu\frac{P}{Dt}\right) \quad (3)$$

where, in addition to the parameters defined above, τ_u is the average masonry shear strength; μ is the Coulomb friction coefficient; and c the cohesion of the wall (namely, the average frictional resistance at $P = 0$). Both of the latter parameters are global material constants for the wall and are obtained from testing of masonry elements. Following micromodel analysis at failure and comparison with test results, their model accounted for two types of shear wall failure, namely:

(i) Failure at the cracked wall ends, with the peak shear strength τ_u being resisted in the compression area of the cracked section D' by t, where D' (the compressed portion of

the wall length D) is obtained by equilibrium of normal force and moment (Eq. 4a):

$$D' = \beta D = \left(1.5 - \frac{3V}{P}\frac{H_0}{D}\right)D \qquad (4a)$$

(ii) Failure at the mid-height of the wall, with the maximum shear strength τ_u being resisted by the entire wall thickness (area D by t) decreasing in an inverse linear manner with the wall shear span ratio H_0/D:

$$R_u = Dt\frac{\tau_u}{1+\frac{H_0}{D}} = Dt\left(\frac{c+\mu p}{1+\frac{H_0}{D}}\right) \qquad (4b)$$

The shear strength of the wall in diagonal tensile failure is therefore given by the minimum strength of the two resisting mechanisms – failure modes:

$$R_u = Dt\tau_u, \quad \tau_u = \min\left\{\left(\frac{1.5\bar{c}+\bar{\mu}}{1+\frac{3\bar{c}H_0}{pD}}\right), \left(\frac{\bar{c}+\bar{\mu}p}{1+\frac{H_0}{D}}\right)\right\} \qquad (5)$$

It is further noted in the above expressions that the coefficients of friction and cohesion μ and c may be modified so as to obtain effective values $\bar{\mu}$ and $\bar{c}$, respectively, corrected for the geometry of the masonry unit, in accordance with the fact that the expressions above sometimes overestimated the experimental value of strength, due to the influence of the masonry headjoints.

(iii) In addition to shear failure by Coulomb friction at the joints, shear failure due to in-plane cracking of the bricks was also experimentally observed for weak brick and strong mortar, in the presence of high axial stresses. It was proposed that the shear strength be estimated in this case in terms of the tensile strength of the bricks f_{bt}, following Eq. 6:

$$R_u = Dt\tau_b = Dt\frac{f_{bt}}{2.3\left(1+\frac{H_0}{D}\right)}\sqrt{1+\frac{p}{f_{bt}}} \qquad (6)$$

and the minimum of the values (Eqs. 5 and 6) used for resistance. In terms of deformation, the wall pier deformation corresponding to the ultimate strength (d_{Ru}, Fig. 15) was found to be close to 0.5 % of the wall height H in most test results they performed or evaluated (Magenes and Calvi 1997).

Detailed Micromodels for Seismic Analysis

In addition to the macromodels above, micromodels have been proposed and are employed for equivalent static linear or nonlinear seismic analysis of masonry buildings and (primarily) historical structures; furthermore, micromodels have also been used to calibrate macromodel topology and the masonry wall failure and constitutive response. Micromodels idealize masonry in detail using:

(i) Two-dimensional thick plate and shell elements that account for both in-plane and out-of-plane stiffness and resistance characteristics (for out-of-plane bending effects of the walls). The use of shell element elastic models is quite common in the seismic analysis of masonry structures, since they do not require excessive computational resources and provide the basic load path and demand concentrations within the building, accounting for complex geometries and multitude of materials, such as timber, masonry, etc. (Fig. 13c). These are therefore suitable to use in entire building seismic analysis models.

(ii) Three-dimensional (brick) FE models of the entire masonry structure, taking into account in the model both complex geometric idealizations of the structure and the foundation and the material complexities associated with the presence and interaction of several different materials with complex constitutive characteristics, such as: stress-strain

nonlinearity due to cracking or crushing; tri-axial capacity interaction and volumetric dilatancy under loading for the mortar, concrete, stone, or brick; yielding or pullout phenomena of the reinforcement where it exists; interface failure between brick, mortar, or steel (where it exists); and inelasticity of timber.

Depending also on the capabilities of the software adopted for the seismic evaluation, micromodels adopted for modeling of inelastic seismic response of LBM buildings include:

- The smeared representation models. These model nonlinearity as spatially averaged, by considering the distributed cracking of masonry in the vicinity of the FE integration points and/or the average constitutive response of masonry over the entire FE integration volume (area) using suitable two- or three-dimensional inelastic constitutive behavior and a strength interaction surface. Plane stress or three-dimensional yield surface characteristics have been proposed for CMU, brick masonry, or stone, possible candidates being, among others, the models proposed by Gambarotta and Lagomarsino (1997) and Stavridis and Shing (2010).
- The discrete representation models. These model nonlinearity discretely through detailed FE modeling of all different material regions involved, namely, individual modeling of the mortar as a brick, shell, or zero-length contact FE, the masonry units and the concrete (where present) as a brick or plain stress shell element (primarily for industrially manufactured units of constant geometry), and the steel or timber reinforcement (where present) as truss or beam elements (for CM or RM buildings).
- As a special case of these are the discrete crack representation models, which further monitor crack formation using fracture energy criteria and the evolution of cracking within the masonry element through mesh redefinition (average material representation) and/or through or along predefined mortar beds (where these exist in physical and model

space), using suitable contact friction elements.
- Models based on the discrete element idealization that fall into this category have been promoted for modeling primarily historical monuments; in this case, the brick units (or stone building blocks) are modeled as individual deformable or undeformable volume elements, and their interface is described with Coulomb frictional contact-separation characteristics.

Modeling of NLBM Infilled Frame Buildings for Seismic Analysis

Due to the abundant use of CM infilled panels in steel and RC frame structures, the seismic analysis modeling of entire frame buildings with NLBM infills is also briefly examined herein for completeness. Only the modeling of the infills is considered herein, since the modeling of the entire frame is beyond the scope of this text. It should be noted, however, that in the design of RC frame buildings, the infills are (and have been) neglected in the structural model assuming that these contribute only to the inertia mass. Only recently modern seismic codes (EC8 2004) provide structural forming and detailing guidelines for taking into account possible adverse response effects due to the presence of the infill; furthermore, PBD methodologies for the assessment and retrofit of existing RC frames require that these be fully accounted for in the seismic model, in their as-built configurations and properties (FEMA-356 2000).

Infill panels provide a large increase in the lateral stiffness of the confining frame with a disproportionate increase in its mass; consequently, masonry infilled frames exhibit short fundamental periods compared to the bare frame structure and therefore attract higher inertia forces at shorter drifts. Properly engineered infill panels, constructed of good quality modules and mortar and adequately wedged into the panel, without excessive openings and a regular in plan and in height configuration, will provide the bare frame structure with considerable overstrength and stiffness enhancement.

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 16 Damages to infilled RC frames following the 1985 Mexico City earthquake. (**a**) Damage of the confining frame elements. (**b**) Out-of-plane failure of the infills. Note that the panels were confined by RC lintels and pilasters tied to the RC frame
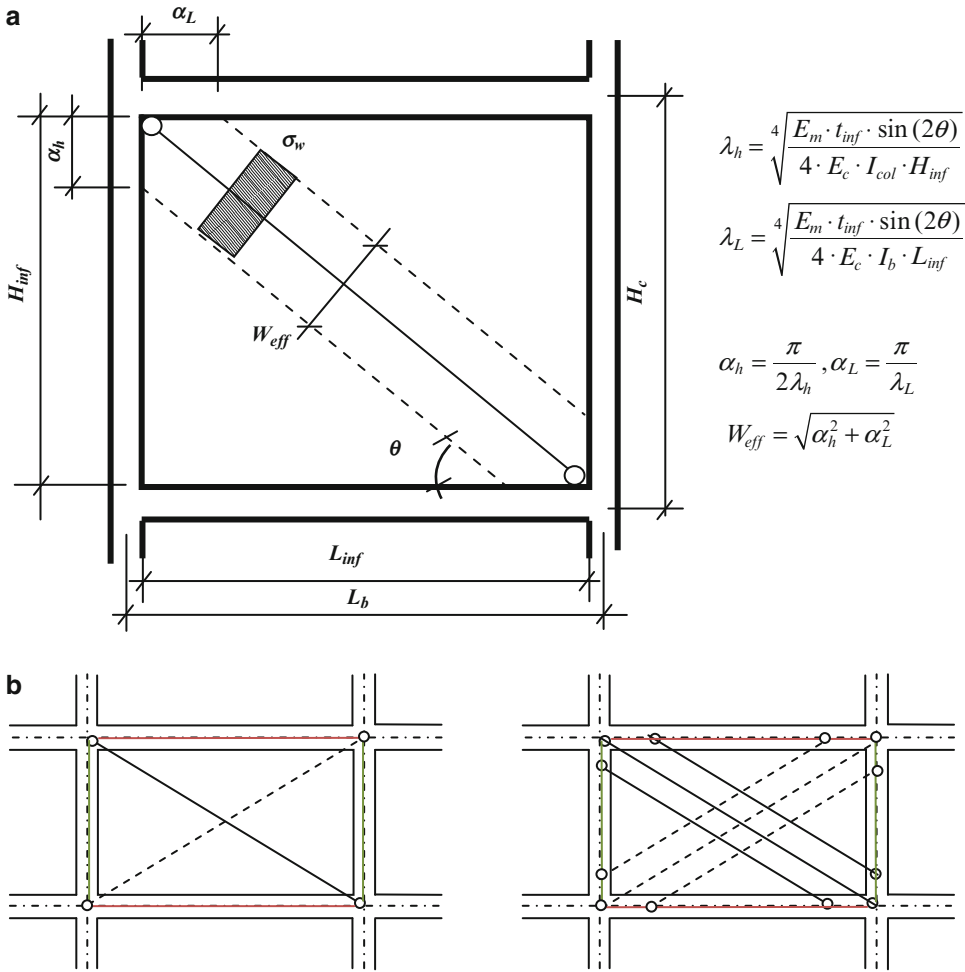

Masonry infills, like other masonry elements already discussed, tend to fail in plane or out of plane (Fig. 16) following similar mechanisms as other masonry elements (diagonal thrust or tension, corner compression, sliding at joints through brick or mortar, out-of-plane collapse, etc.). Such a full or partial failure of the infills will lead to local failures of the confining elements due to the formation of unintentional short column effects or due to a shear failure of the top of the column or the beam–column joint. For these reasons, in the PBD assessment approach of existing structures, the infills need to be accounted for in seismic analysis.

In modeling infilled frame buildings in practical seismic analysis of entire structures, macromodels are used for the panels, as shown in Fig. 17. Micromodels have also been adopted in the literature, however primarily for research applications and in order to calibrate the macromodel parameters. The use of macromodels stems from the observation that, due to the partial contact separation of the panel from the surrounding frame elements under lateral response of the infilled frame, the function of the infill can be modeled as an equivalent axial load-bearing diagonal strut element resisting compression only in the direction of the lateral force. The strut has a thickness t_{inf} equal to the panel thickness and an equivalent width W_{eff} based on the frame to panel contact length; proposed strut widths adopt contact lengths that account for the relative stiffness characteristics of the confining frame and the infill panel in terms of $\lambda_{h,L}$ (Stafford Smith and Carter 1969), where $\lambda_{h,L}$ expresses the relative stiffness of the infill and the concrete elements assuming elastic contact (Fig. 17a):

where E_m and E_c are the Young's moduli of masonry and concrete; I_{col} is the uncracked moment of inertia of the confining column elements; θ is the geometric inclination of the infill strut (Fig. 17a); L_{inf} and H_{inf} are the clear length and height of the infill; and H_c is the centerline distance between the beams. For seismic applications cross strut configurations are used with compression only properties (Fig. 17b). In order to model the effect of short column formation in the infilled frame response, multiple strut configurations as shown in Fig. 17b have also been proposed (Crisafulli et al. 2000). Their use is recommended in the case of asymmetric infill configurations within a floor, such as infilled frame bays next to open bays or end bays of the infilled frame.

S

$$\lambda_h = \sqrt[4]{\frac{E_m \cdot t_{inf} \cdot \sin(2\theta)}{4 \cdot E_c \cdot I_{col} \cdot H_{inf}}}$$

$$\lambda_L = \sqrt[4]{\frac{E_m \cdot t_{inf} \cdot \sin(2\theta)}{4 \cdot E_c \cdot I_b \cdot L_{inf}}}$$

$$\alpha_h = \frac{\pi}{2\lambda_h}, \alpha_L = \frac{\pi}{\lambda_L}$$

$$W_{eff} = \sqrt{\alpha_h^2 + \alpha_L^2}$$

Seismic Analysis of Masonry Buildings: Numerical Modeling, Fig. 17 Equivalent strut macromodel of infilled RC frames and effective strut width equation. (**a**) Model proposed by Stafford Smith and Carter (1969). (**b**) Single and multiple strut macromodels of infilled RC frames for static lateral load and cyclic dynamic analysis

Summary

The seismic response and forms of failure for different types of masonry construction have been presented and discussed. It is commonly believed that masonry buildings, being the most common form of past construction in all seismic affected areas, as well as a financially viable alternative for new low- to mid-rise construction, they need to be evaluated for seismic loading, due to the seismic vulnerability of such structures as it has been proven in numerous seismic events, past and recent.

Several modeling conventions have been proposed so far and are still under rigorous research investigations, for the seismic analysis of masonry buildings. Overall, these range, depending on the capabilities of the FE code at hand and also on the reliability of the structural information available, from simple strut-and-tie models (the equivalent frame analysis models, which are also referred to as macromodels) to the refined micromodels, using plane or three-dimensional FE analysis tools, both linear and nonlinear material-wise and geometry-wise.

Not all methods are suitable for all cases, and often analysis "overkill" for a problem that is quite complex to model will give a false sense of security: due to the complexity of the problem, involving different masonry materials (brick, mortar, stone, rubble, and also RC or steel and timber), different types of construction (LB or NLB, CM), and different structural topologies (low or mid rise, with or without diaphragms, irregular in plan or elevation, with openings, with flexible foundation), equally complex methods should be used, justifying the accuracy of the input information, namely, the material properties for all materials involved in the construction and an adequate knowledge of their interaction.

For entire structural models, macromodels or area micromodels with averaged properties over the FE region are adequate to capture global quantities (forces and deformations) and seismic performance. For detailed damage prediction (cracking, crushing, region disintegration under overload), three-dimensional FE models are adopted, using either deformable elements with smeared or discrete cracking representation or discrete FE models with contact friction interfaces, where crack spreading is not feasible (e.g., URM construction and historic building analyses). These models, although they provide realistic damage predictions compared to observation and test, are too detailed and complex to be applied to ordinary construction, and they need to be substantiated by adequate testing in order to establish the material behavior as input to the FE model.

Cross-References

References

ASCE (2013) Building code requirements and specification for masonry structures. Standards ASCE/SEI 5-13, 6-13. American Society of Civil Engineers, Reston

Clough R, Mayes R, Gülkan P (1979) Shaking table study of single-story masonry houses, vol 3, Summary, conclusions and recommendations, UCB/EERC-79/25. Earthquake Engineering Research Center, University of California, Berkeley

Crisafulli F, Carr A, Park R (2000) Analytical modeling of infilled frame structures; a general review. Bull N Z Soc Earthq Eng 33(1):30–47

DesRoches R, Comerio M (2011) The 2010 Haiti Earthquake. Earthquake Spectra, vol 27, S1. Earthquake Engineering Research Institute, Oakland

EC6 (2005) Eurocode No. 6, Design of masonry structures – Part 1-1: General rules for reinforced and unreinforced masonry structures, EN-1996-1. European Committee for Standardization, Brussels

EC8 (2004) Eurocode No. 8, Design of structures for earthquake resistance – Part 1: General rules, seismic actions and rules for buildings, EN-1998-1. European Committee for Standardization, Brussels

EERI/IAEE. World Housing Encyclopedia. http://www.world-housing.net/

FEMA-356 (2000) Prestandard and commentary for the seismic rehabilitation of buildings. FEMA, Washington, DC

Gambarotta L, Lagomarsino S (1997) Damage models for the seismic response of brick masonry shear walls. Part II: The continuum model and its applications. Earthq Eng Struct Dyn 26:441–462

Hughes R, Nichols R, Paul T, Riding V (1990) The Augusta, Sicily earthquake of 13 December 1990, a field report. EEFIT, Institution of Structural Engineers, London

IAEE (2008) Regulations for seismic design: a world list. http://www.iaee.or.jp/worldlist.html

ICOMOS. International Council on Monuments and Sites. http://www.icomos.org/en/

ISCARSAH. The International Scientific Committee on the Analysis and Restoration of Structures of Architectural Heritage. http://iscarsah.icomos.org/

Kingsley G (1994) The U.S.-TCCMAR full-scale five-story masonry research building test, Part 2, Report SSRP-94/02. University of California, San Diego

Lagomarsino S, Penna A, Galasco A, Cattari S (2013) TREMURI program: an equivalent frame model for

S

the nonlinear seismic analysis of masonry buildings. Eng Struct 56:1787–1799

Magenes G, Calvi GM (1997) In-plane seismic response of brick masonry walls. Earthq Eng Struct Dyn 26(11):1091–1112

Magenes G, Morandi P, Penna A (2008) Test results on the behaviour of masonry under static cyclic in plane lateral loads. ESECMaSE project, report RS-01/08. Department of Structural Mechanics, University of Pavia, Pavia

Moro L (coord. editor) (2007) Guidelines for evaluation and mitigation of seismic risk to cultural heritage. Gangemi Editore, Rome

Rossetto T, Peiris N, Alarcon J, So E, Sargeant S, Libberton C, Verrucci E, Del Re D, Free M (2009) The L'Aquila (Italy) Earthquake of 6th April 2009, a field report. EEFIT, Institution of Structural Engineers, London

Ruiz-Garcia J, Negrete M (2009) Drift-based fragility assessment of confined masonry walls in seismic zones. Eng Struct 31:170–181

Stafford Smith B, Carter C (1969) A method of analysis for infill frames. Proc Inst Civil Eng 44:31–48

Stavridis A, Shing PB (2010) Finite element modelling of nonlinear behaviour of masonry-infilled RC frames. J Struct Eng 136(3):285–296

UNIDO (1983) Repair and strengthening of reinforced concrete, stone and brick masonry buildings. Building construction under seismic conditions in the Balkan region, vol 5. UNDP/UNIDO Project RER/79/015. United Nations Industrial Development Organization, Vienna

Vintzileou E, Miltiadou-Fezans A (2008) Mechanical properties of three-leaf stone masonry grouted with ternary or hydraulic lime based grouts. Eng Struct 30(8):2265–2276

Vintzileou E, Zagkotsis A, Repapis C, Zeris C (2007) Seismic behaviour of the historical structural system of the Island of Lefkada, Greece. Construct Build Mater 21(1):225–236

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling

Ioannis Vayas[1] and Aristidis Iliopoulos[2]
[1]Traffic Engineering Laboratory, National Technical University of Athens (N.T.U.A.), Athens, Greece
[2]Peikko Greece SA, Marousi, Athens, Greece

Synonyms

Bearings; Bridge modeling; Composite bridge; Seismic design; Seismic isolation

Introduction

Reinforced concrete slabs rigidly connected with steel girders have been used to form the basic superstructure of large numbers of deck bridges for many decades. This is due to the fact that the composite construction method offers the bridge engineers a great variety of solutions for different types of problems. A typical composite cross section of a highway bridge is shown in Fig. 1. A series of parallel steel girders are rigidly connected with a reinforced concrete slab through shear connectors. The shear connectors installed are mostly welded studs allowing use of the deck as part of the top flange (*deck plate girders*). The longitudinal bending of the composite T-girders, at sagging bending areas, results in tension in steel and compression in concrete. The simultaneous operation of both of these materials generates the composite action which is the most important feature for the formation of stiff and high-strength cross sections. At hogging moment areas, concrete is considered to be fully cracked, and only the slab reinforcement, but not concrete, contributes to bending resistance.

In pure steel bridges the reinforced concrete slab is replaced by an orthotropic steel deck. This is composed of the steel deck plate; the longitudinal, mostly trapezoidal, stiffeners; and the crossbeams (Fig. 2). Orthotropic are lighter than concrete decks. However, they require high fabrication costs due to intensive welding operations and are susceptible to fatigue. Therefore, pure steel bridges (see EN 1993-2 2006) are nowadays mostly limited to cases where it is essential to limit the deck weight, e.g., for very large spans, movable bridges, etc. The present article refers mainly to composite bridges that constitute the vast majority in modern steel bridge construction.

Modeling for analysis is required in order to determine internal forces and moments, deformations, and vibrations of bridge decks including bearings, piers, abutments, piles, etc. In addition models should include the foundation when soil-structure interaction is accounted for. A bridge analysis model should be based on the following criteria:
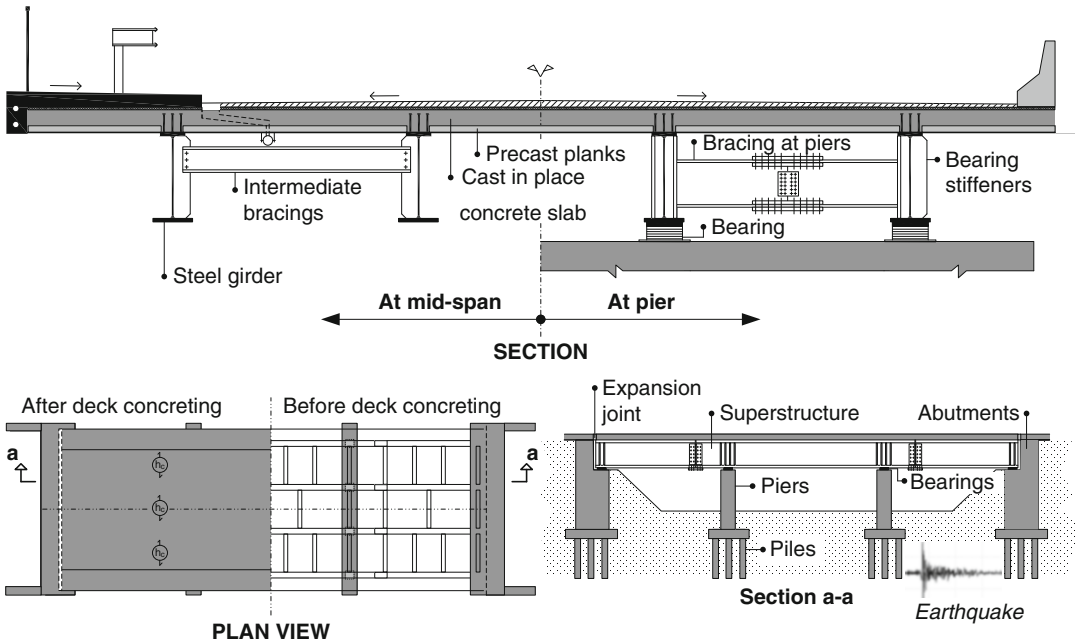
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 1 Layout of a typical composite deck bridge

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 2 Steel bridge with orthotropic deck
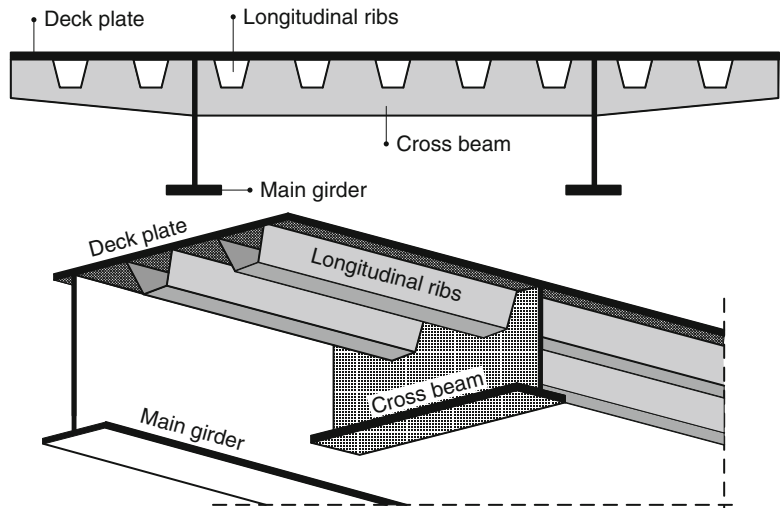
- It should reflect the structural response in terms of deformation, strength, and local and global stability.
- It should include as many as possible structural elements and parts of the bridge deck (cross frames, stiffeners, etc.) and their possibly eccentric connections.
- It should also include bearings and piers individually, piles, etc.
- It should cover all construction stages and loading cases.

- Loads should be easily introduced.
- It should allow the performance of dynamic analysis and include the most important modes of vibration.
- The resulting output should be such that it enables easily the execution of the code-prescribed verifications.
- It should be supported by commercial analysis and design software.

Models for Seismic Analysis

There exist several possible models for bridge seismic analysis that could be employed depending on the bridge configuration, the bearing types, the connection between superstructure (bridge deck) and piers or abutments, the type of foundation with possible soil-structure interaction, etc.

Figure 3 shows possible modeling levels for seismic analysis of steel and composite bridges, starting from the simple to the comprehensive ones. Simple models, like the first three in Fig. 3, can be used for preliminary seismic analysis. They may be employed for the derivation of seismic forces and displacements on bearings, piers, foundations, or the soil. Comprehensive models, like the last three of Fig. 3, are mostly used at the main design phase since they also cover analysis for construction stages and service conditions where other loads due to traffic, wind, temperature, time-dependent concrete effects, etc. must be taken into account.

In the following the various analysis models for steel and composite steel-concrete bridges are presented. Models for superstructures (decks) are linear since decks are expected to remain elastic in the seismic situation. Any inelastic activity should be restricted to bearings, piers, piles, soil, etc. For bearings and piers, nonlinear characteristics are provided. For other elements (piles, abutments, soil), reference is made to other chapters.

Rigid Deck Model

In this model the superstructure is represented as a single mass, M_{dir}, acting on a spring of stiffness

K_{dir}. M_{dir} includes the entire mass of the deck without the mass of the piers. The global spring stiffness K_{dir} represents the combined stiffness of bearings and piers and the foundation of soil and is calculated from

$$\frac{1}{K_{dir}} = \frac{1}{\sum K_{bearings}} + \frac{1}{\sum K_{piers}} + \frac{1}{\sum K_{foundation}} \quad \text{where dir} = \text{direction X or Y}$$

(1)

The fundamental period for this single mass oscillator is calculated from (see Fig. 3a)

$$T_{dir} = 2 \cdot \pi \cdot \sqrt{\frac{M_{dir}}{K_{dir}}}, \, M_{dir} = \sum m \quad (2)$$

The seismic forces acting on the entire deck $F_{EA, dir, tot}$ are determined from the relevant response spectrum. The seismic forces of bearings at the top of one pier i may be determined from

$$F_{bearing, i} = \frac{\sum \left(K_{bearing, i} + K_{pier, i} + K_{foundation, i} \right)}{K_{dir}} \cdot F_{AE, dir, tot}$$

(3)

This model may be applied for the longitudinal direction of straight bridges with continuous decks, when the mass of the piers is less than 20 % of the tributary mass of the deck. The model may be also applied for the transverse direction provided that all conditions (a) to (c) referred below apply (see EN 1998-2 2005):
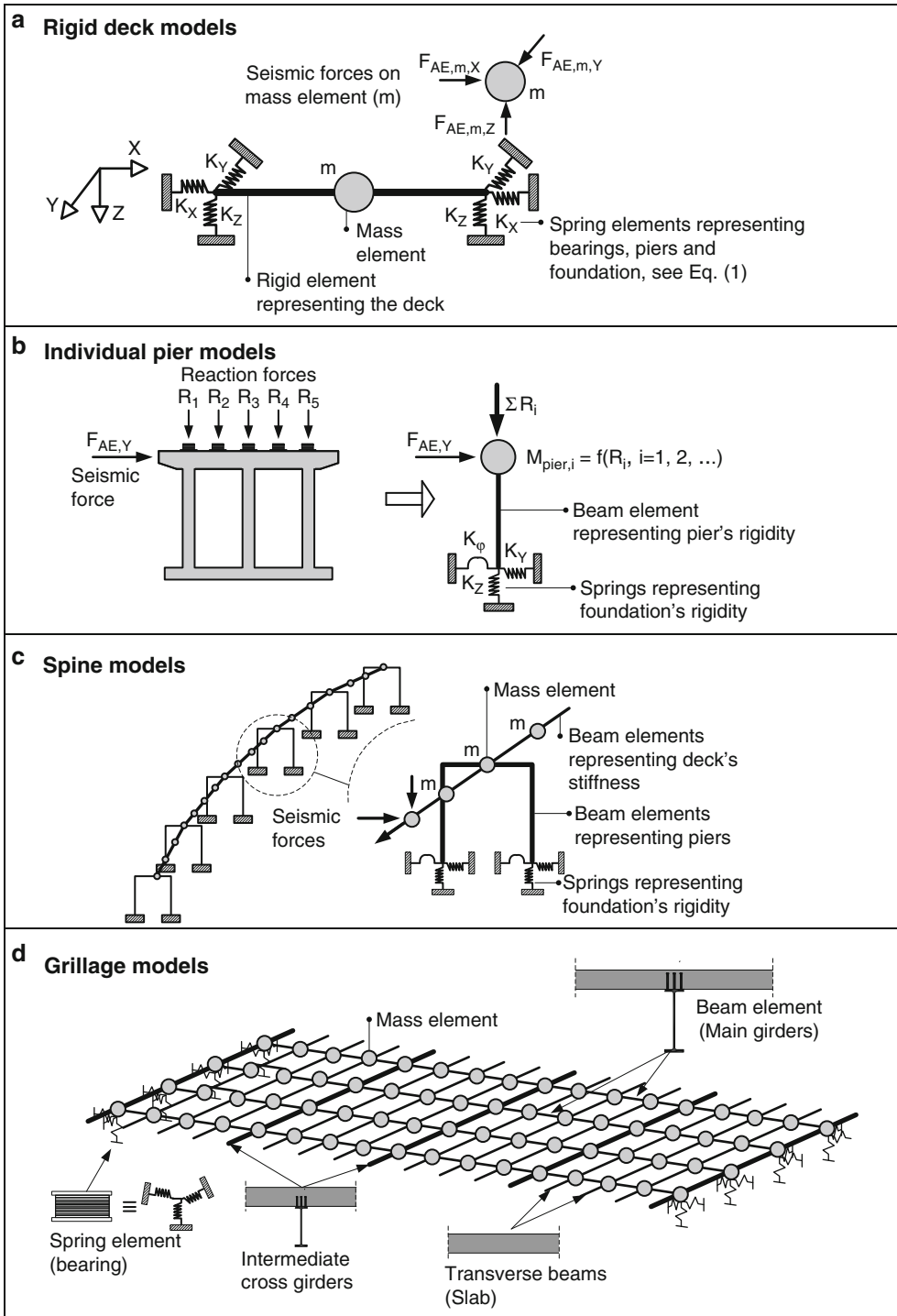
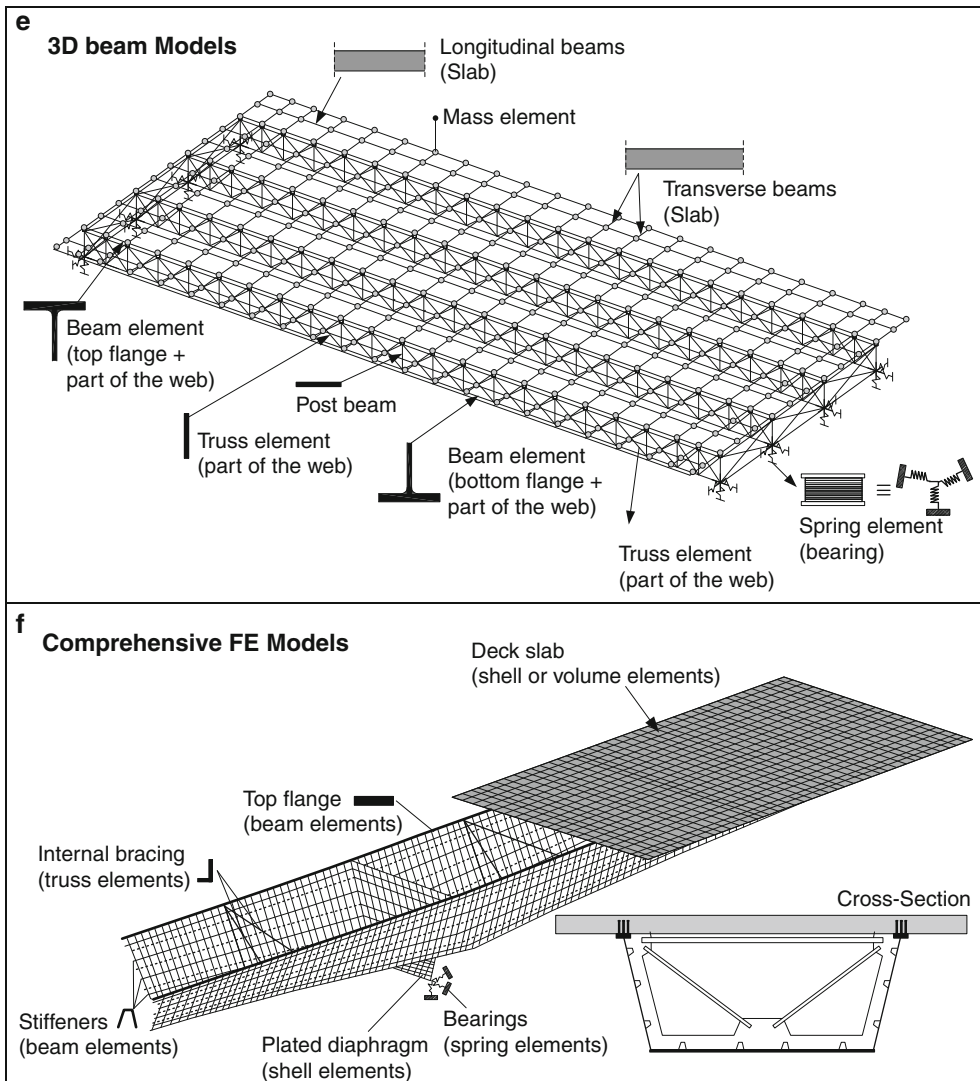(a) $L/b \le 4.0$
   where
   L is the total bridge length and
   b is the width of the deck.
(b) $\Delta d/d_a \le 0.2$
   where Δd and d_a are, respectively, the maximum difference and the average value of the displacements in the transverse direction of all pier tops under $F_{AE,Y}$.

a Rigid deck models

Seismic forces on mass element (m)

$F_{AE,m,X}$ $F_{AE,m,Y}$

m

$F_{AE,m,Z}$

X

Y   Z

K_Y K_Y

K_X K_Z m K_Z K_X

Mass element

Rigid element representing the deck

Spring elements representing bearings, piers and foundation, see Eq. (1)

b Individual pier models

Reaction forces
R_1 R_2 R_3 R_4 R_5

ΣR_i

$F_{AE,Y}$
Seismic force

$F_{AE,Y}$ $M_{pier,i} = f(R_i, i=1, 2, ...)$

Beam element representing pier's rigidity

K_φ K_Y
K_Z

Springs representing foundation's rigidity

c Spine models

Mass element

m

m

m

Seismic forces

Beam elements representing deck's stiffness

Beam elements representing piers

Springs representing foundation's rigidity

d Grillage models

Mass element

Beam element (Main girders)

Spring element (bearing)

Intermediate cross girders

Transverse beams (Slab)

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 3 (continued)

S

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 3 Analysis models for seismic design of steel and composite bridges. (**a**) Rigid deck models. (**b**) Individual pier models. (**c**) Spine models. (**d**) Grillage models. (**e**) 3D beam models. (**f**) Comprehensive models

(c) The theoretical eccentricity between the mass center of the deck and the stiffness center of the supporting members does not exceed 5 % of the deck's length.

The application of the rigid deck model may be extended to cases where bearings with damping properties are employed. The bearing stiffness $K_{bearing}$ is expressing the secant stiffness of the bearing device and is determined iteratively to correspond with the bearing's displacements. In addition higher damping values are achieved, so that a reduction factor must be employed to the resulting seismic forces. More detailed information is given in sections "High-Damping Reinforced Elastomeric Bearings" and "Lead Rubber Bearings (LRB)."

Individual Pier Model

This model may be used for seismic analysis of bridges in the transverse direction. Each pier and the associated part of the superstructure is considered separately and represented as a single mass oscillator. The mass of the oscillator is $M_{pier,\,i}$, and includes the mass of the deck between half distances of piers, while $K_{pier,\,i}$ represents the pier stiffness.

The fundamental period for pier i is calculated from

$$T_{pier,\,i} = 2 \cdot \pi \cdot \sqrt{\frac{M_{pier,\,i}}{K_{pier,\,i}}} \qquad (4)$$

Based on the fundamental period, the seismic forces acting on top of the pier $F_{AE,Y}$ are determined from the relevant response spectrum. The individual pier model is not appropriate for curved or skew bridges, bridges with varying spans, varying pier lengths, etc. It may be used for long bridges where each pier is able to act independent of the rest of the bridge. These requirements are met when following conditions apply:

(a) The seismic action in transverse direction is mainly resisted by piers.
(b) There is no significant interaction between adjacent piers.
(c) $0.90 \leq \frac{T_{pier,\,i}}{T_{pier,\,i+1}} \leq 1.10$.

Spine Models

Spine models are appropriate for performing multimodal dynamic analysis on bridges and let the seismic forces resisted by each pier determined in a "natural" way in accordance with their relative stiffness. They may be employed for "normal" bridges, normal meaning more or less straight, low skew, narrow deck, and cross section with limited distortion, i.e., with rigid closely spaced transverse frames or crossbeams.

In spine models the bridge deck is represented by beam elements that are positioned at the centroid of the cross section and have six (6) degrees of freedom (DOFs) at end nodes (Fig. 4). The degrees of freedom are the translations (u, v, w) along the principal axes coordinate system (x, y, z) and the corresponding rotations (θ_x, θ_y, θ_z). The resulting internal forces (N, M_y, M_z) act at the gravity center, while the shear forces and torsion moments (V_y, V_z, M_x) at the shear center. Cross-sectional warping may be taken into account by introduction of an additional seventh DOF per node as independent variable ($\theta_x' = d\theta_x/dx$), which results in the bimoment M_w as additional internal moment (see Kindmann and Kraus 2011).

The initial beam cross section is the gross cross section. However, for wide flanges, either of concrete or steel, effective widths must be introduced to consider shear lag effects (Fig. 5). Although effective widths vary along the length of the bridge, they are smaller at internal supports
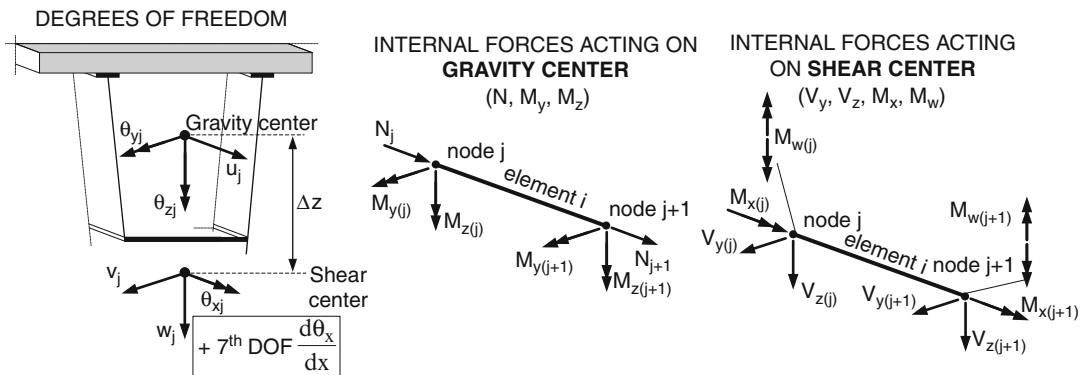
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 4 Representation of the bridge deck by 7 DOF beam elements

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 5 Gross and effective cross section for wide concrete or steel flanges

Bending stress distribution in concrete flange

Bending stresses $[\sigma_x]$

Effective part

Effective part

Bending stress distribution in bottom steel flange

Uncracked zone   Cracked zone   Uncracked zone

$0.85 \cdot L_1$ $0.15 \cdot L_1$ $0.15 \cdot L_2$ $0.85 \cdot L_2$
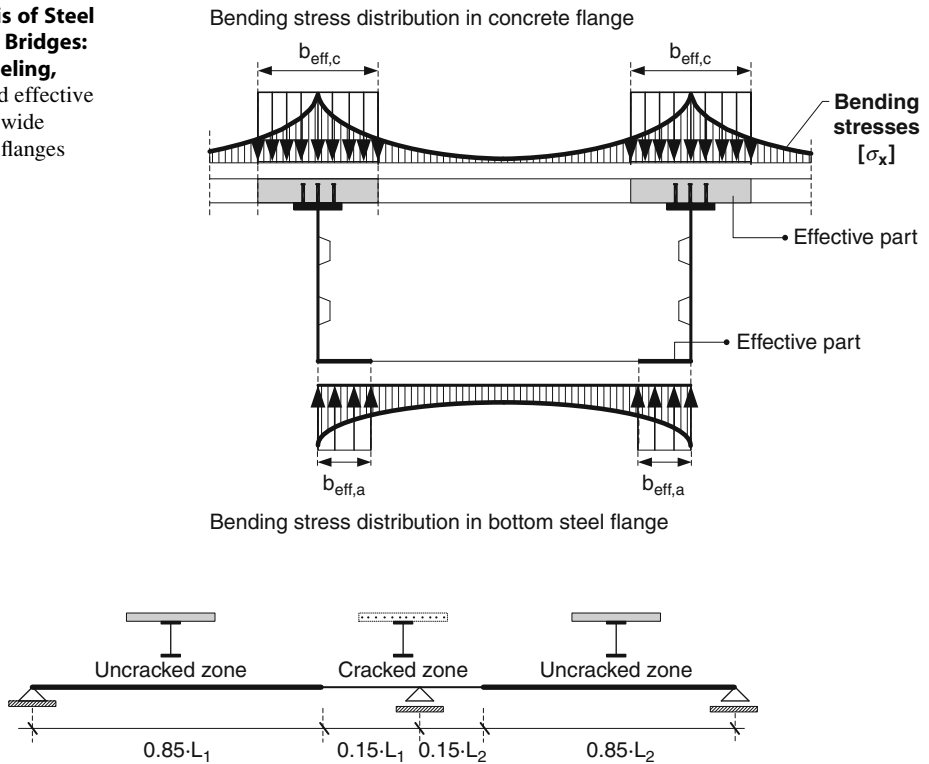
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 6 Determination of the cracked regions for composite bridges

than at spans, fixed values, equal to those at spans, are considered for global analysis. It should be mentioned that effective widths are calculated differently for concrete flanges than for steel flanges (see EN 1993-2 2006 and EN 1994-2 2005).

For composite bridges the flexural stiffness, denoted as $E \cdot I_1$, is calculated for the uncracked section in regions where concrete is in compression. For seismic analysis, the short-term modulus of elasticity of concrete is considered. At hogging moment areas, concrete is in tension for beam-type bridges. Cracking of concrete in those regions is considered by introducing the flexural stiffness $E \cdot I_2$ of the "cracked" section in which the contribution of the concrete slab is neglected. For continuous bridges the cracked region may be considered to be adjacent to the internal supports in a length equal to 15 % of the corresponding span length (Fig. 6). The 15 % rule constitutes a rough

approximation of the true behavior (EN 1994-2 2005), however, for the purpose of seismic analysis is considered as adequately precise.

As an alternative to the fully cracked section, the *tension stiffening effect* of concrete may be taken into account in the cross-sectional properties by increasing the area of steel reinforcement by the factor $1/[1 - (0.5 \cdot f_{ctm})/(\rho_{s,\,tot} \cdot f_{sk})]$ where $\rho_{s,\,tot}$ is the total reinforcement ratio, f_{ctm} the mean tensile strength of concrete, and f_{sk} the characteristic yield strength of the reinforcing steel (see Vayas and Iliopoulos 2013).

The torsional stiffness $G \cdot I_t$ for composite box girders may be calculated from mechanics, where the shear modulus of concrete is taken into account by introducing 0.2 or 0 as the Poisson's ratio for uncracked and cracked regions correspondingly.

Although the superstructure is represented by a single beam in spine models, bearings and piers

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 7 Two-span continuous bridge. (**a**) Physical model and (**b**) numerical spine model
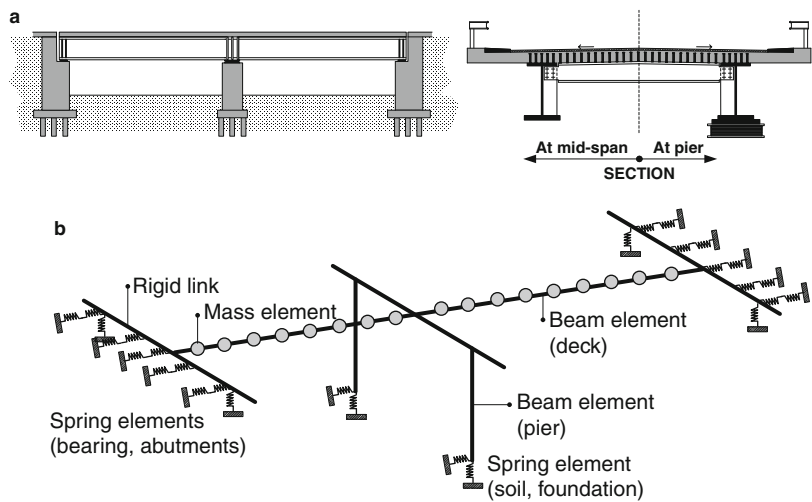
appear individually in the model. Figure 7a shows for a two-span continuous composite girder bridge that is supported by two pile bent abutments and one two-column bent. The cross section is a composite section consisting of two I-girders and a concrete deck. The girders rest individually on bearings. Figure 7b shows the spine model for this bridge. The cross section of the superstructure is represented by a single beam element, while each bearing by two horizontal springs. The beam and the springs are coupled by a rigid link that represents the crossbeam. Below the crossbeam are the two piers that are represented by beam elements that rest on translational and rotation springs representing the pile foundation and the soil. Similar conditions apply to the abutments, where each bearing is represented by two horizontal springs, coupled in series with horizontal abutment/foundation springs.

Grillage Models

General
The most popular computer-aided modeling method for the analysis of composite bridges is the simulation by means of a plane grillage system. This is due to the fact that this system is easy to apply and comprehend as well as that it has been proved to be sufficiently accurate for a wide variety of bridge decks. In this model, the structure is idealized by means of a series of longitudinal and transverse beam elements rigidly interconnected at nodes. Each element is given an equivalent bending and torsion inertia to represent the relevant portion of the deck.

Plate Girder Bridges
Figure 8 illustrates a grillage representation of a simply supported composite bridge with four main girders that may similarly be applied to continuous systems or different number of main girders. Longitudinal grillage members are arranged to represent the main girders with the inertia properties of the composite section (steel section with a part of the slab corresponding to the effective width). Transverse members represent the deck slab with thickness h_c equal to the thickness of the slab and width b equal to the distance between transverse beams; it is convenient to select b equal to the distance of the axle loads. A non-cracked flexural rigidity for the slab elements is usually applied. The torsional rigidity of the transverse slab elements can be set to zero. The total in-plane second moment of area of the slab is equally shared to the two extreme main girders (A and D), while the intermediate girders (B and C) are given $I_Z = 0$. This is because wind loads mainly act on the edge girders of the bridge. In case of intermediate cross girders whose stiffness may influence the transverse distribution of
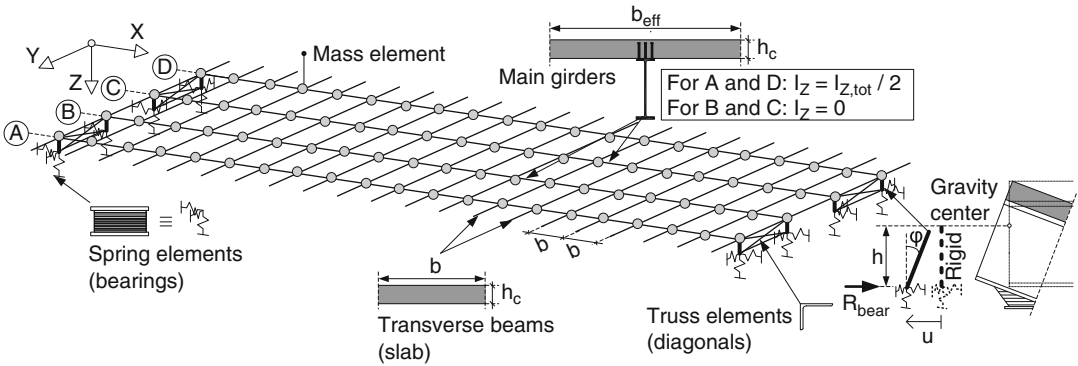
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 8 Grillage model of a plate girder bridge

the vertical loads, these girders are taken into account by appropriate modification of the transverse member properties. In case of long distances between adjacent transverse beams, X-bracing concrete elements may need to be set in the deck's plane in order to simulate the diaphragmatic behavior of the deck slab.

Reinforced elastomeric bearings, usually implemented in bridges, are represented by three axial springs of equivalent stiffness corresponding to the relevant stiffness properties in horizontal and vertical directions; the calculation of the bearings' stiffness is presented in section "Reinforced Elastomeric Bearings." The axes of the main beams coincide with the center of gravity of their cross sections. However, the bearings are positioned beneath the lower flange. Accordingly, rotations of the main girders result in horizontal deformations u of the bearings and additional support reactions R_{bear}. The support nodes are therefore put at a lower level from the grillage members and are connected to the longitudinal beams by rigid vertical bars whose height h is equal to the distance between the center of gravity of the main composite beams and the bottom flange; for better accuracy, the shear center of the cross section should be used which is assumed to be the "real" center of rotation. In case of intermediate cross girders whose stiffness may influence the transverse distribution of the vertical loads, these girders are taken into account with beam elements of an appropriate stiffness.

At piers truss elements are used for the representation of the cross braces. Due to the height h of the rigid elements, the geometry of the bracings in model may not follow the exact geometry of the bracings in real structure. A height adjustment for the rigid elements may then be necessary. This should be done only for the purpose of estimating the forces of the bracing members due to horizontal loadings, i.e., wind or earthquake. It has to be stated that in most bridges the gravity center of the composite cross sections is located near the top flange. For such cases a height adjustment has little influence on final results.

Skew Bridges

In skew bridges the support abutments or piers are placed at angles other than 90° from the longitudinal centerlines of the girders (Fig. 9). The skew angle is usually defined as the angle between the longitudinal axis of the bridge and a line square to the supports. The presence of skew affects the geometry and the behavior of the structure. Special phenomena, like twisting and out-of-plane rotation of the main girders during concreting, uplifting forces at bearings, and fatigue problems due to out-of-plane web distortion, make the analysis and design of skewed bridges intricate. The transverse elements representing the slab are usually oriented perpendicular to the main girders (orthogonal mesh); this is the most usual grillage model used by the designers. Alternatively, the transverse members
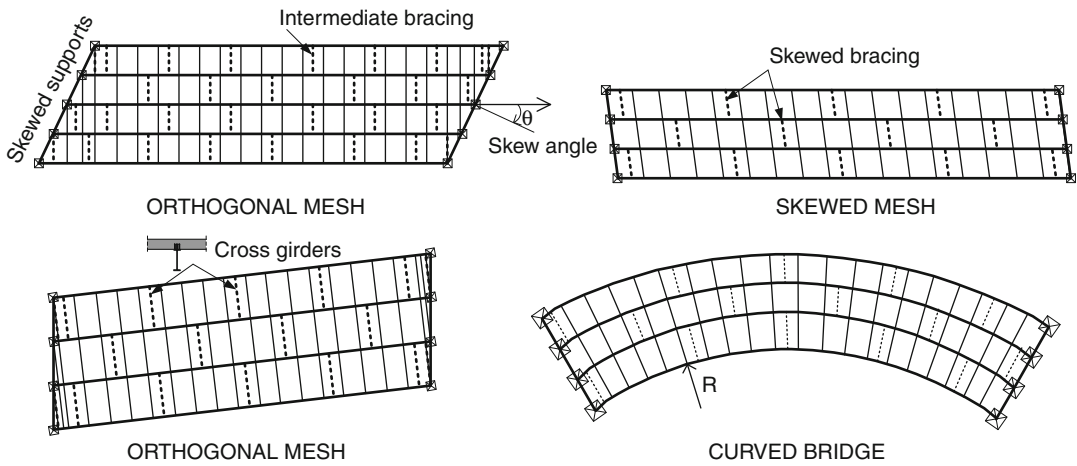
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 9 Grillage models for skew and curved plate girder bridges
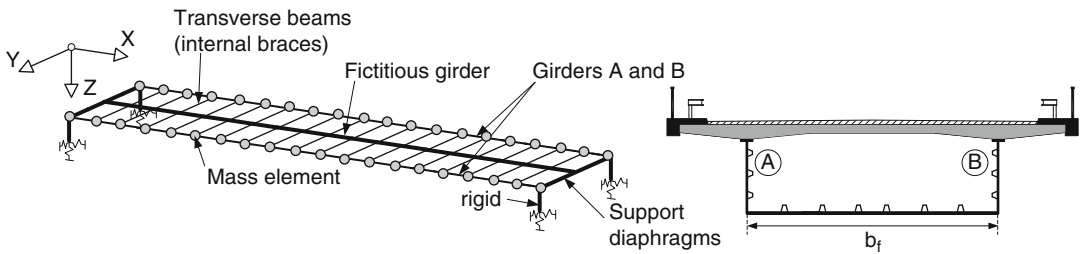
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 10 Grillage model of a simply supported box girder bridge

can be placed parallel to the line of supports (skewed mesh). Generally the skewed mesh is convenient for low skew angles ($\theta < 20°$) or when the intermediate bracing is not arranged square to the main girders.

Curved Bridges

Curved decks pose no particular problem for grillage modeling (Fig. 9). A curved bridge deck can be represented by a grillage of curved members or of straight members. Some computer programs support curved members but others do not. Generally, a grillage of straight beams with a very fine mesh is for small values of curvature sufficiently accurate. For highly curved bridges, 3D – or FE – models should be used (Adamakos et al. 2011).

Box Girder Bridges

A grillage model can be implemented also for single-box girder bridges (Fig. 10). The box girder is divided in two opened composite cross sections in which the shear lag effect in the deck slab and the lower flange is considered through the effective widths. The grillage is thus composed of two main composite girders A and B transversely connected with beams representing the internal braces or diaphragms, not the slab. The torsional rigidity of the composite box girder is represented in the model by a fictitious girder located between the main composite girders. The central girder comprises also the whole bending ($I_{Z,\ tot}$) – and shear stiffness of the deck slab ($A_Y \approx$ slab area). The flanges for girders A and B extend over their effective
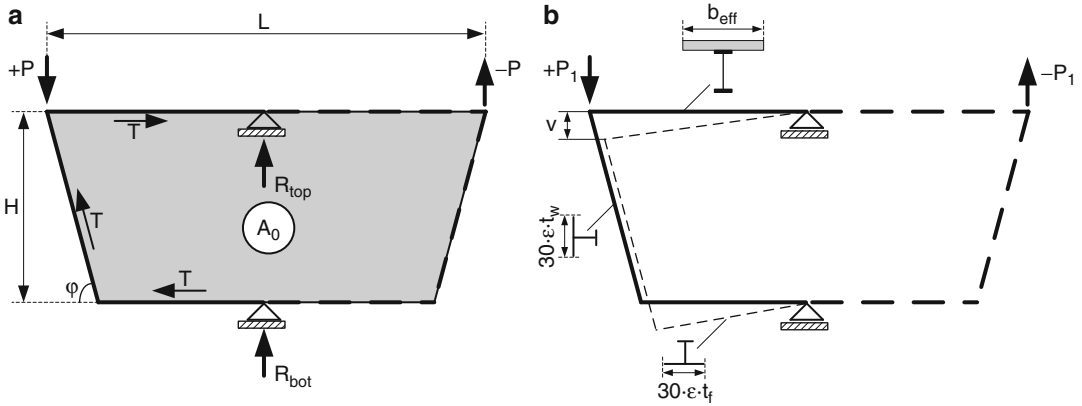
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 11 (**a**) Shear flow of the box section due to *St. Venant* torsion. (**b**) Deformation of cross frames

widths. Care should be given that these widths are different for concrete and for steel flanges and also different for analysis and for design.

Longitudinal stiffeners in the bottom flange are considered by "smearing" them over the flange width. The bottom flange is then considered having a total thickness of $t_{tot} = t + t_{add}$, where t is the thickness of the panel and $t_{add} = \sum A_{stiffeners}/ b_f$. The same procedure is followed for bridges with orthotropic steel deck, where the top flange panel is taken into account with increased thickness due to "smearing" of the longitudinal stiffeners. However, the increase in thickness does not apply to the web panel due to the fact that existing longitudinal stiffeners are usually placed only to enhance its buckling resistance and are not necessarily continuous.

Transverse beams in the global model represent the flexibility of the cross frames or of cross braces. Since a single beam shall represent an entire frame, the definition of its properties requires some preliminary analysis. The beams are considered as rigid in bending but flexible in shear deformations. Accordingly they are assigned an infinite in-plane second moment of area and a shear area A_s. Cross frames or cross braces, and accordingly the transverse beam, resist part of the torsion while they are not strained due to global bending. A global torsion moment M_t is partly resisted by the *St. Venant* action and partly by the cross frames or cross braces. This torsion is split into antisymmetric loading $\pm P = M_t/ L$ (see Fig. 11). The *St. Venant* shear flow is given by

$$T = \frac{M_t}{2 \cdot A_0} = \frac{P \cdot L}{2 \cdot A_0} \tag{5}$$

where A_0 is the shaded area in Fig. 11a.

Therefore, the forces resisted by the cross elements are equal to

$$P_1 = P - T \cdot H_{inclined} \cdot \sin \varphi = P - T \cdot H \tag{6}$$

where $P = M_t /L$. It may be seen that the cross elements are resisting part of the global torsion. If the webs are not inclined ($\varphi = 90^0$), the force resisted by the cross elements is equal to:

$$P_1 = P - \frac{P \cdot L}{2 \cdot H \cdot L} \cdot H$$
$$= \frac{P}{2}, \quad \text{i.e., } 50\% \text{ of the acting forces}$$
$$\tag{7}$$

Subsequently the vertical deformation v of the cross frame due to antisymmetric loading $\pm P_1$ as determined before is numerically calculated (Fig. 12). Due to symmetry, only half of the cross section is analyzed and hinges are placed in the middle of the flanges. The two flanges and the web are represented by beam elements with cross

sections composed of the transverse stiffeners or transverse girders and an associated flange width. The effective width for the web and the bottom flange panels is equal to $30 \cdot \varepsilon \cdot t, \varepsilon = \sqrt{235 \cdot f_y}$, where f_y = yield stress in MPa and t = thickness of the web or bottom flange.

Orthotropic steel decks are represented in a similar way. For composite bridges with concrete decks, the cross section of the relevant beam element is composed of the transverse girder and the effective width of the concrete slab.

The deformation of a transverse beam is equal to $v = P \cdot L/(A_s \cdot G)$ where G is the shear modulus (see Fig. 12). By setting this deformation equal to the deformation of the cross frame calculated previously, the shear area of this beam is defined as equal to

$$A_s = \frac{P \cdot L}{v \cdot G} \qquad (8)$$

where v is the deformation of the system of Fig. 11 due to the load P_1 given before.

Half-Through Bridges

Half-through bridges may be also represented by plane grillage models (Fig. 13). Main and cross girders are represented by beam elements, the former with steel sections, the latter with composite section including the effective slab width. Lateral torsional buckling phenomena of the top compression flange cannot be captured by this grillage model. 3D models are then recommended.

3D Models

General

The structural representation of bridge decks with truss girders or I-shaped plate girders may be done by means of 3D models as proposed from Vayas et al. (2010, 2011) and Vayas and Iliopoulos (2013). Truss girders are represented by their chord and bracing members, while plate girders are transformed to equivalent trusses. Such models have been proven to be advantageous for modeling orthogonal, skewed, and curved bridges. Unlike grillage models, they are able to consider:

- Eccentricities among the structural elements of a bridge and therefore additional internal forces and possible load distributions
- The transversal variation in the level of the neutral axis
- Torsion and distortional warping effects
- The dispersed structural behavior of the deck slab, in which bending takes place in two directions
- Buckling phenomena of the steel girders during erection stages
- Diaphragms, bracing systems, and stiffeners – possible overload or fatigue effects are taken into account

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 12 Deformation of the transverse beam of the global model
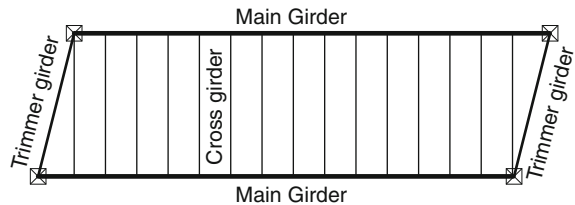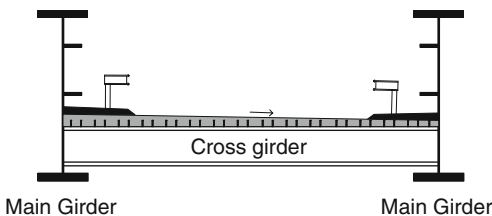
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 13 Grillage model of a half-through bridge

Girder Representation

Steel and composite I cross sections are modeled by a "hybrid" truss as shown in Fig. 14. For the steel girder, the flanges of the truss are beam elements with a cross section composed of the flange and a part of the web of the steel girder. Comparative analyses showed that one third of the web height may be associated to the flange. Therefore, the flanges of the truss are T-sections consisting of the flange of the steel girder and one third of the web and are positioned at the center of gravity of the T-section. The webs are represented by diagonal truss elements with width equal to one third of the web height and thickness equal to the web thickness. It has been also shown that the cross-sectional area $A_d = h_w \cdot t_w / 3$ for the diagonals adequately corresponds to the shear stiffness of the web. The post-beams are located at a spacing $s \approx 5\%$ of the span of the bridge. This distance is generally acceptable for small and medium span bridges because the angle between the diagonals and the flange elements usually remains between $35°$ and $45°$. Post-beams represent both the in-plane and out-of-plane stiffness of the web.

For a composite section the same procedure is followed, with the concrete slab represented by another beam element connected with the upper flange of the truss through the appropriate offset: offset = distance between the centroids C_c

and C_{fo}. The nodes of the elements that represent the slab are the same nodes of those representing the upper flange of the truss. It is recommended that a fine mesh is used for the beam elements of the concrete slab and the top flange of the steel girder so that a full shear connection is achieved. Without a fine mesh the beam elements of the slab may deflect differently than those of the top flange.

Slab Representation

Slabs are structurally continuous in both directions X and Y, and they resist applied loads by shear forces, bending moments and torques which are coupled with each other. For this reason it was previously mentioned that the transverse slab elements of the grillage models should not be used for the final design of the slab.

A grillage model which considers the dispersed bending and torsion stiffness of a solid slab is illustrated in Fig. 15. The grillage mesh should be sufficiently fine so that the grillage deflects in a smooth surface in a similar way as a real slab. A smooth deflected surface is equivalent to the requirement that the twist $\partial^2 w / \partial x \partial y$ is the same in orthogonal directions and that $m_{yx} = m_{xy}$. The spacing of the beams shouldn't be less than 2.5 times the slab depth. Transverse beams should have spacing similar to that of the longitudinal beams. It is also recommended that
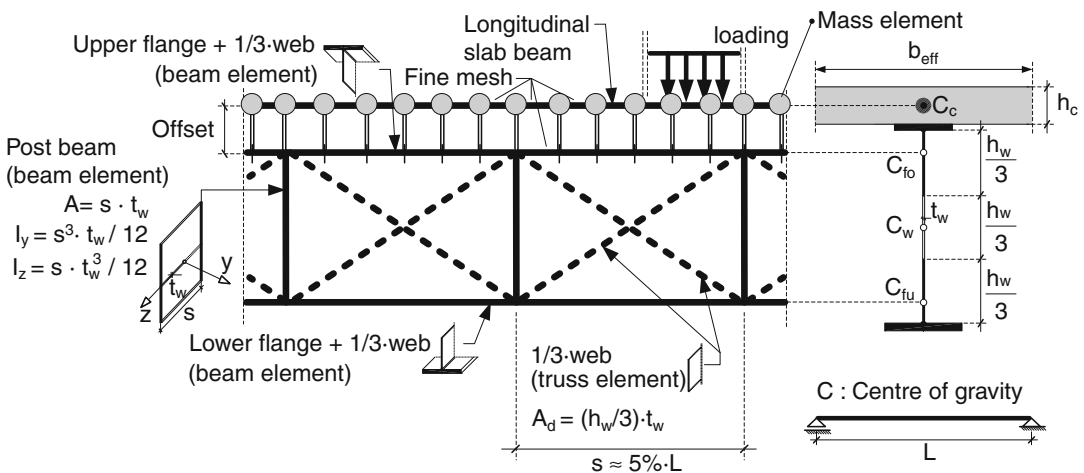
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 14 Truss idealization for a steel-concrete composite girder
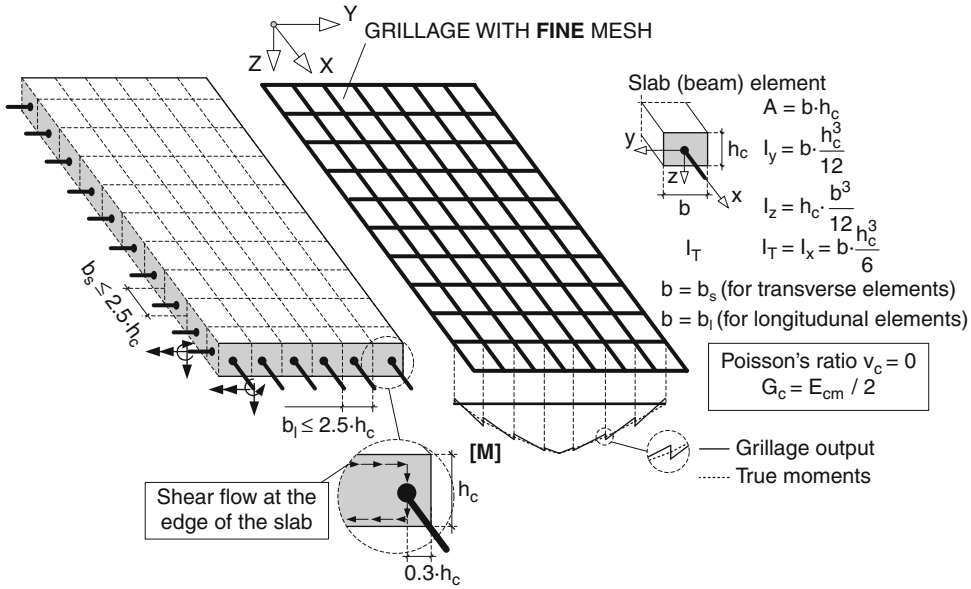
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 15 Grillage model for an isotropic solid slab

the row of longitudinal beams at each edge of the grillage should be located in a distance of $0.3 \cdot h_c$ from the edge of the slab, where h_c is the slab depth (see Hambly 1990). This is where the resultant of the shear flows is located. The width of the edge member for the calculation of I_T should be therefore reduced to $b-0.3 \cdot h_c$.

The 3D Model Implementation

The grillage model for the slab's representation in Fig. 15 can be combined with the truss model which is shown in Fig. 14. Figure 16 illustrates a 3D model which is recommended for the structural analysis both of simple and continuous composite bridges. Attention must be paid so that the grillage has its longitudinal members coincident with the center lines of the steel sections. At sagging moment areas, longitudinal slab elements are used with their uncracked properties. At hogging moment areas, concrete is considered as fully cracked and the total reinforcement is positioned due to simplicity at center of the slab. Transverse slab elements can be considered with their uncracked properties.

One can see that the model can be set up in a detailed way by taking into account all the necessary structural elements, i.e., cross-bracings, bearings, etc. Imperfections, precambering, and girders with variable cross sections can also be implemented in the model. Therefore, structural phenomena which may be difficult or impossible to investigate with plane grillages are included in the outputs of the 3D model, e.g., arch effects in integral bridges with longitudinally variable cross sections.

Modeling of Bearings

General

Bridge decks rest usually on bearings that are important elements for seismic analysis and design. Most of the bearings, especially modern bearings, provide seismic isolation to the bridge, reducing the forces on piers, foundations, and the soil. Seismic isolation is provided due to the low stiffness of such bearings that results in a shift to longer fundamental periods and lower spectral values (see Chopra 1995). An appropriate representation of bearings in the structural model is essential for seismic bridge design and will be presented in the following. Bearings may deform
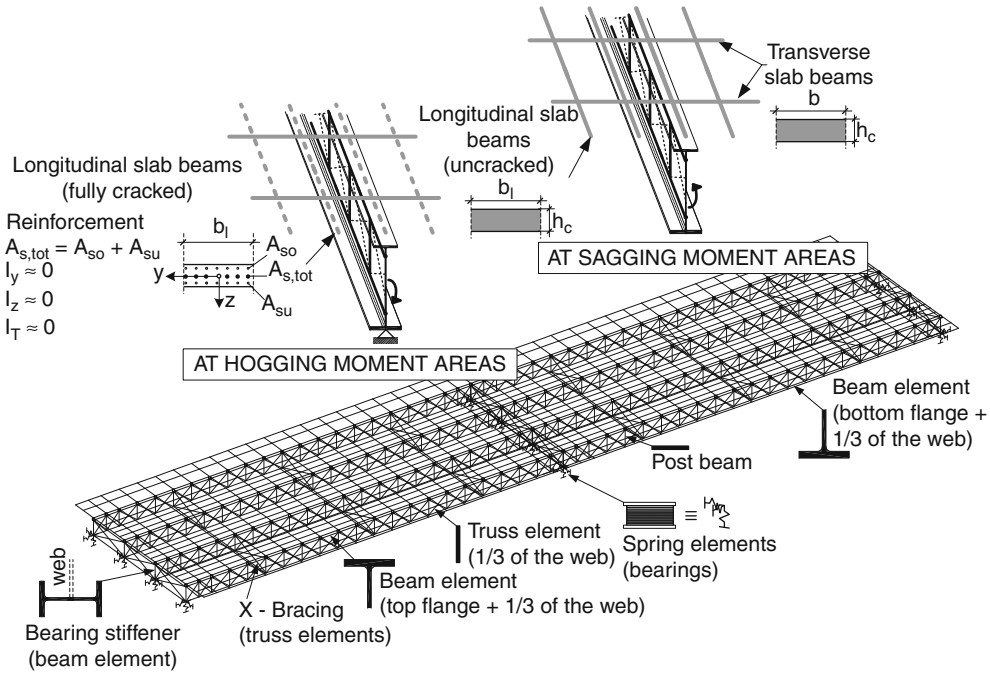
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 16 3D model of a two-span continuous bridge

during seismic action in the nonlinear range. Accordingly linear and nonlinear characteristics are provided.

Reinforced Elastomeric Bearings

Reinforced elastomeric bearings (Fig. 17) consist of uniformly spaced layers of elastomer (natural or synthetic rubber) and reinforcing steel plates that obtain their bonding through the process of vulcanization. The equivalent viscous damping ratio for common elastomers ξ is less than 6 %.

Reinforced elastomeric bearings are introduced as linear springs in global analysis. The spring stiffness in each unrestrained horizontal direction may be obtained from

$$K_{dir} = \frac{A \cdot G_b}{T_e} \qquad (9)$$

where:

$A = a \cdot b$ or $\pi \cdot D^2/4$ is the plan area of the bearings.

T_e is the nominal thickness of the elastomer layers.

dir = global direction X or Y.

The shear modulus G_b considered has an increased value compared to $G = 0.9$ MPa of elastomers to account for the speed of loading in the seismic situation so that (see EN 1998-2 2005):

$$G_b = 1.1 \cdot G \qquad (10)$$

Beyond this, upper and lower values of the shear modulus are introduced in the seismic combination with recommended values as follows:

$$G_{b, max} = 1.5 \cdot G_b = 1.65 \cdot G \qquad (11a)$$

$$G_{b, min} = 1.0 \cdot G_b = 1.1 \cdot G \qquad (11b)$$

Upper values are supposed to result in maximum forces, minimum values maximum displacements. However, this might not be always the case, depending on the resulting periods and
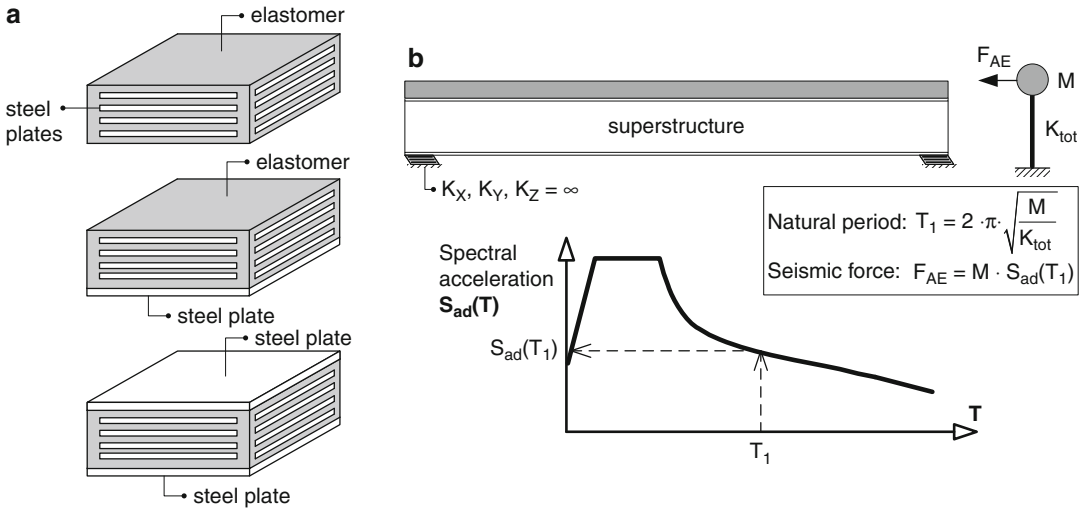
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 17 (**a**) Composition of reinforced elastomeric bearings. (**b**) Single DOF seismic analysis of a simply supported composite bridge with a rigid deck model
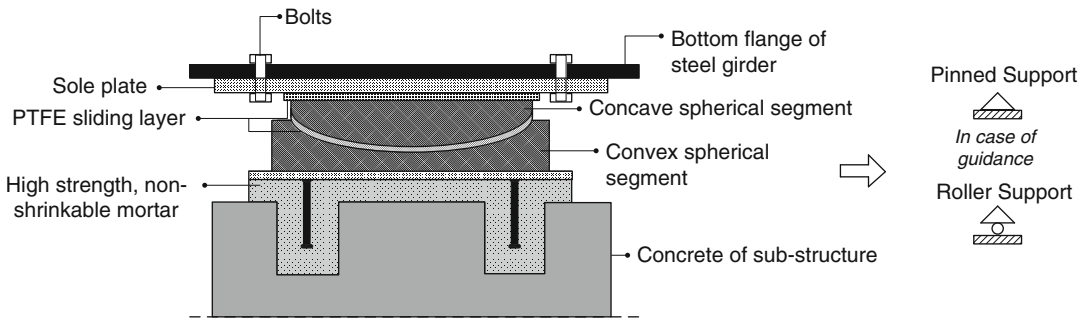
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 18 Layout and representation of spherical bearings

the spectral values. In vertical direction the bearing is practically incompressible so that the vertical displacement is considered to be restrained without use of springs.

High-Damping Reinforced Elastomeric Bearings

These are bearings in which the common elastomer is substituted by high-damping elastomer. Their equivalent viscous damping ratio ξ reaches values between 10 % and 20 %, while common elastomeric bearings have damping ratios ξ below 6 %. These bearings are modeled in global analysis as linear springs like common elastomeric bearings. The reduction in seismic forces

is taken into account by application of the reduction factor $\eta = \sqrt{10/(5 + \xi)} \geq 0.55$ on the forces determined from a response spectrum with 5 % damping.

Spherical Bearings

Spherical bearings consist of a sole plate that transfers loads from the superstructure, a concave spherical segment and a convex spherical segment that provides a mating surface for the concave segment and transfers load to the substructure (Fig. 18). PTFE (Teflon) sliding layers are provided between the three parts to allow horizontal displacements and rotation in all directions. Spherical bearings provide
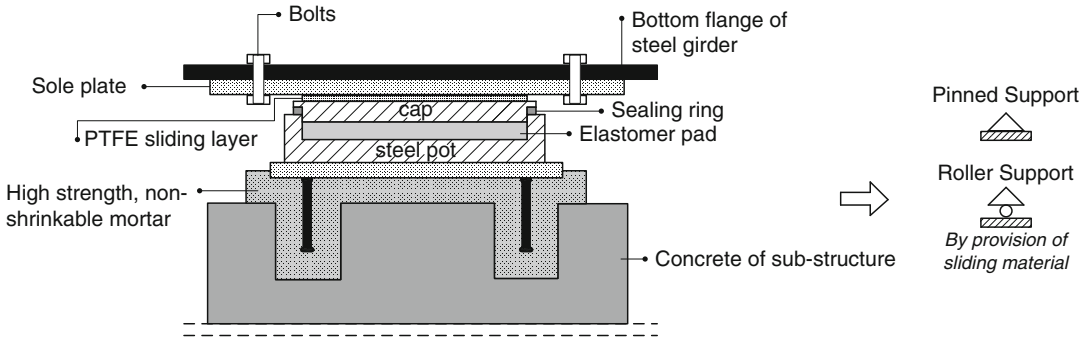
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 19 Layout and representation of pot bearings
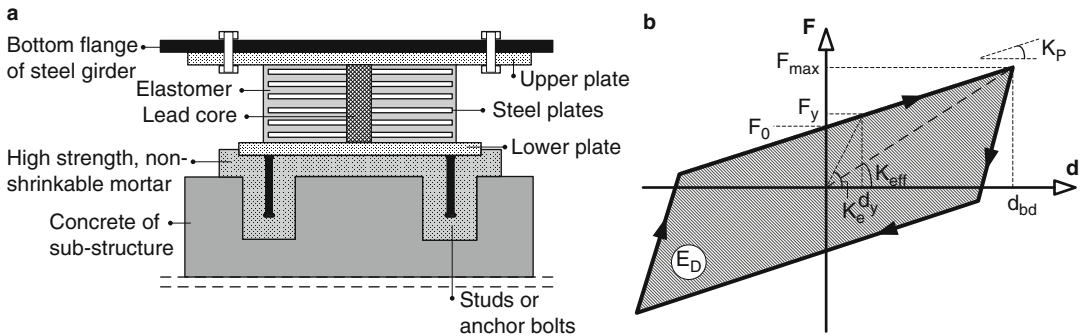
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 20 (**a**) Layout of lead rubber bearings. (**b**) Force-displacement behavior of lead rubber bearings

therefore only vertical support. They are accordingly represented by hinges allowing also horizontal displacements. Spherical bearings may be also guided to restrain displacements in one or both horizontal directions. In this case displacements in the relevant horizontal direction are restrained.

Pot Bearings

Pot bearings are based on the incompressibility of natural rubber when placed in a closed steel pot where natural rubber behaves like a fluid (Fig. 19). Pot bearings are able to transfer high compression forces in a small surface and allow rotations around all axes. They provide horizontal and vertical support while rotations are free. Provision of a sliding material allows horizontal displacements in which case the bearings provide only vertical support.

Lead Rubber Bearings (LRB)

These are common reinforced elastomeric bearings with low-damping elastomer and a cylindrical lead core that may reach damping values up to 40 % (Fig. 20a).

During cyclic loading the lead core is yielding and strain hardening so that the hysteretic response of the bearings, as illustrated in Fig. 20b, is typical for a yielding and strain hardening material. In analysis the bearings may be accordingly represented by a bilinear horizontal springs with elastic stiffness K_e for displacements up to the yield displacement and post-elastic tangent stiffness K_p for larger displacements.

The elastic stiffness may be determined from

$$K_e = K_R + K_L \tag{12}$$

where:

K_R is the shear stiffness of the elastomers.
K_L is the shear stiffness of the lead core.

The yield force is determined from

$$F_y = F_{Ly} \cdot \left(1 + \frac{K_R}{K_L}\right) \tag{13}$$

where F_{Ly} is the yield force of the lead core.

The force at zero displacement is equal to

$$F_0 = F_y - K_p \cdot d_y \tag{14}$$

The post-elastic stiffness is attributed to the lead core only and is calculated from

$$K_P = \frac{F_{max} - F_y}{d_{bd} - d_y} = K_R \tag{15}$$

However, the above bilinear representation is appropriate for nonlinear analyses. For linear analysis the spring properties are linearized by introduction of the secant stiffness. This effective stiffness is written as

$$K_{eff} = \frac{F_{max}}{d_{bd}} \tag{16}$$

The analysis must be iterated in this case by introduction of an initial value for K_{eff} and update it consecutively as a function of the resulting maximum displacement of the previous step.

Figure 20b indicates that lead rubber bearings dissipate energy when displaced beyond the yield displacement. The energy dissipation per cycle equals to the shaded area of Fig. 20b and may be determined from

$$E_D = 4 \cdot \left(F_y \cdot d_{bd} - F_{max} \cdot d_y\right) \tag{17}$$

The energy dissipation may be transformed to increased damping with an effective damping ratio ξ_{eff} determined from

$$\xi_{eff} = \frac{E_D}{F_{max} \cdot d_{bd}} = \frac{E_D}{2 \cdot \pi \cdot K_{eff} \cdot d_{bd}^2} \tag{18}$$

where all symbols are indicated in Fig. 20b. For linear analysis the value of ξ_{eff} is updated in each iteration step in order to determine the reduction factor for increased damping on the resulting seismic forces. When the rigid deck model is employed, all bearings are represented by a single spring with an equivalent stiffness and damping ratio determined from

$$K_{eff,\,tot} = \sum K_{eff,\,i} \tag{19}$$

$$E_{D,\,tot} = \sum E_{D,\,i} \tag{20}$$

where the summation refers to all bearings i.

The equivalent damping ratio for all bearings is calculated from

$$\begin{aligned}
\xi_{eff,\,tot} &= \frac{E_{D,\,tot}}{2 \cdot \pi \cdot K_{eff} \cdot d_{bd}^2} \\
&= \frac{1}{2 \cdot \pi \cdot d_{bd}^2} \cdot \frac{E_{D,\,tot}}{\sum K_{eff,\,i}}
\end{aligned} \tag{21}$$

The steps of the iterative procedure for the rigid deck model are as follows:

1. Selection of the bearings
2. Initial assumption for the displacement d_{bd}
3. Calculation of $K_{eff,\,tot}$, $\xi_{eff,\,tot}$ and $E_{D,\,tot}$ as a function of d_{bd}
4. Determination of the fundamental period of the equivalent system $T_{eff} = 2 \cdot \pi \cdot \sqrt{M_{tot}/K_{eff,\,tot}}$ where M_{tot} is the total mass
5. Determination of the seismic forces and the new displacement from Table 1
6. Repetition of steps 2–5 until converge of the displacements is achieved

Friction Pendulum Bearings (FPS)

Friction pendulum bearings constitute a variation of the spherical bearings, where the plane sliding top surface is replaced by a curved one (Fig. 21a, b). They exhibit, unlike spherical bearings, a re-centering capability and are not subjected to rotations. The force at zero displacement is provided by the friction and is given by

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Table 1 Seismic forces and design displacements (EN 1998-2 2005)

Fundamental period	Seismic forces	Design displacement
$T_C \leq T_{eff, tot} \leq T_D$	$F_{AE} = \beta \cdot \frac{T_C}{T_{eff, tot}} \cdot n \cdot A_g \cdot W$	$d_{bd} = \frac{T_{eff, tot}}{T_C} \cdot d_C$
$T_D < T_{eff, tot}$	$F_{AE} = \beta \cdot \frac{T_C \cdot T_D}{T_{eff, tot}^2} \cdot n \cdot A_g \cdot W$	$d_{bd} = \frac{T_D}{T_C} \cdot d_C$
Spectra	Acceleration spectrum	Displacement spectrum
	Spectral acceleration $S_{ad}(T_{eff})$	Displacement spectrum $d_{bd}(T_{eff})$
Notation	A_g peak ground acceleration, W total weight, and $\eta = \sqrt{10/(5 + \xi_{eff, tot})} \geq 0.55$	
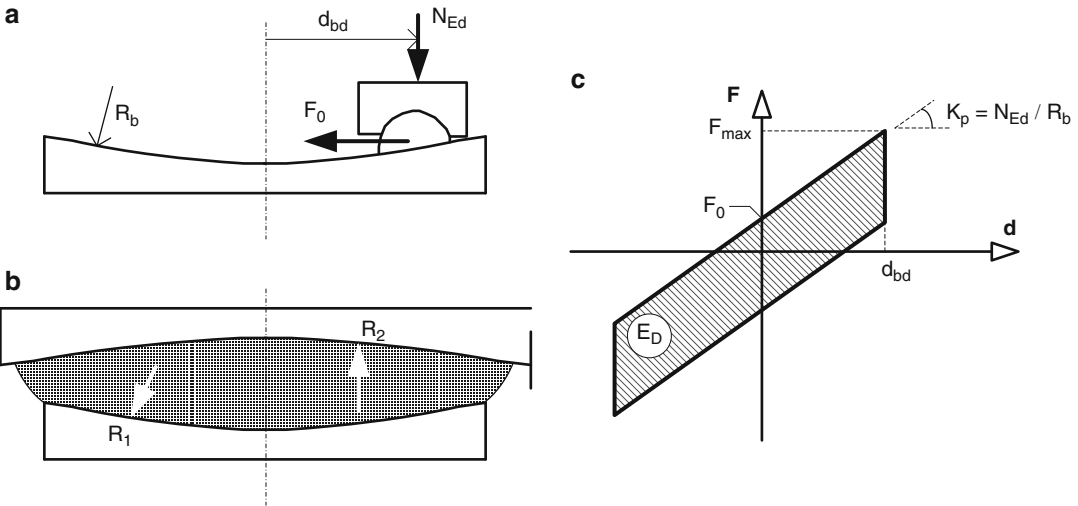
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 21 Layout of (**a**) single- and (**b**) double-friction pendulum bearings. (**c**) Force-displacement behavior of friction pendulum bearings

$$F_0 = \mu \cdot N \qquad (22)$$

The kinematics of the bearings leads to following relations:

$$\sin \varphi = d/R \qquad (23a)$$

$$\tan \varphi = F/N \qquad (23b)$$

Equating the above equations for small angles $\sin\varphi \approx \tan\varphi$ that may be considered as valid with

good approximation when $d/R \leq 0.25$ and considering the developing friction between surfaces, the following relation may be derived for the restoring force:

$$F = \frac{d \cdot N}{R} + \mu \cdot N \cdot \text{sign}(\dot{d}) \qquad (24)$$

where:

N is the vertical force.

d is the horizontal displacement.

R is the radius of the spherical surface.

μ is the friction coefficient.

$\text{sign}(\dot{d})$ is the sign of the velocity.

It may be seen that the restoring force increases linearly with the displacement and the support reaction, i.e., with the mass. For two spherical surfaces, similar relations apply, where $R = R_1 + R_2$ (Fig. 21b).

Figure 21c displays the force-displacement characteristic of friction pendulum bearings that indicates a symmetrical response. The effective linear spring stiffness is proportional to the compression force and is determined from

$$K_{eff} = \frac{F}{d} = N \cdot \left(\frac{1}{R} + \frac{\mu}{d} \right) \qquad (25)$$

The energy dissipation in a cycle is given by

$$E_D = 4 \cdot \mu \cdot N \cdot d \qquad (26)$$

and the equivalent effective damping ratio by

$$\xi_{eff} = \frac{E_D}{2 \cdot \pi \cdot F \cdot d} = \frac{2 \cdot \mu}{\pi \cdot (d/R + \mu)} \qquad (27)$$

When the rigid deck model is employed, the equivalent stiffness of all bearings is determined from

$$K_{eff, tot} = \sum K_{eff, i} = W_{tot} \cdot \left(\frac{1}{R} + \frac{\mu}{d} \right) \qquad (28)$$

where the summation refers to all bearings i and W_{tot} is the total weight of the bridge in the seismic situation.

The individual spring stiffness for each bearing i is then

$$K_{eff, i} = \frac{N_i}{W_{tot}} \cdot K_{eff, tot} \qquad (29)$$

where

N_i is the compression force of bearing i.

If bearings with identical radius R are selected, the damping ratios for each bearing

and for the overall system are equal to $\xi_{eff, tot} = \xi_{eff, i}$.

The steps of the iterative procedure for the rigid deck model are as follows:

1. Selection of the bearings, specifically the friction coefficient μ and the radius R
2. Initial assumption for the displacement d
3. Calculation of $K_{eff, tot}$ and $\xi_{eff, tot}$ as a function of d
4. Determination of the fundamental period of the equivalent system
   $T_{eff} = 2 \cdot \pi \cdot \sqrt{M_{tot}/K_{eff, tot}}$, where M_{tot} is the total mass
5. Determination of the seismic forces and the new displacement from Table 1
6. Repetition of steps 2–5 until converge of the displacements is achieved

The equivalent spring stiffness and energy dissipation may then be calculated for each individual bearing.

Fixed Bearings

Until the 1950s, steel bearings were used that consisted of four types: *pins*, *rollers*, *rockers*, and *metal sliding bearings*. Pins are fixed bearings allowing rotations. Rollers and rockers allow translation and rotation, while sliding bearings utilize one plane metal plate sliding against another, with PTFE as intermediate lubricant material, to accommodate translations. These older bearing types are modeled as hinges (see Fig. 22). Fixed bearings may also be represented by springs, their flexibility resulting from the flexibility of the bearing bars and the elongation of the anchor bolts.
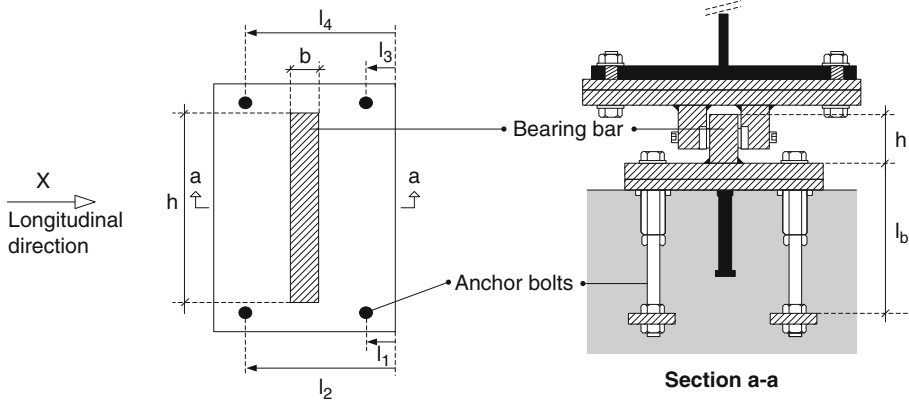
The stiffness in longitudinal direction of fixed bearings may be determined from

$$\frac{1}{K} = \frac{h^3}{3 \cdot E \cdot I} + \frac{h \cdot l_b}{E \cdot A \cdot \sum_1^4 l_i} \qquad (30)$$

where:

E is the modulus of elasticity for steel.

I is the second moment of area of the bearing bar $(I = b^3 \cdot h/12)$.

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 22 Fixed bearing

h is the height of the bearing bar.

A is the area of the anchor bolts.

l_b is the length of the anchor bolts.

l_i are the distances of the anchor bolts from the tip of the base plate.

The stiffness in transverse direction is given by a similar equation, neglecting the first term due to the high stiffness of the bearing bar in this direction and changing the direction of the lengths l_i. If a bearing set is going to be represented by one translational and one rotational spring, the relevant spring stiffness may be expressed by

$$K_{hor} = \sum_1^n K_i \qquad (31a)$$

and

$$K_{rot} = \sum_1^n K_i \cdot l_{bi}^2 \qquad (31b)$$

where:

K_i is the stiffness of each bearing as determined before.

n is the number of bearings in the set.

l_{bi} is the distance of bearing i to the centerline of the bridge deck.

Modeling of Piers

Piers are generally represented by means of 6 DOF beam elements. The pier is usually
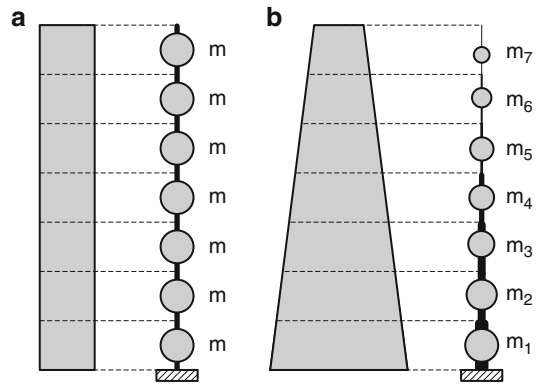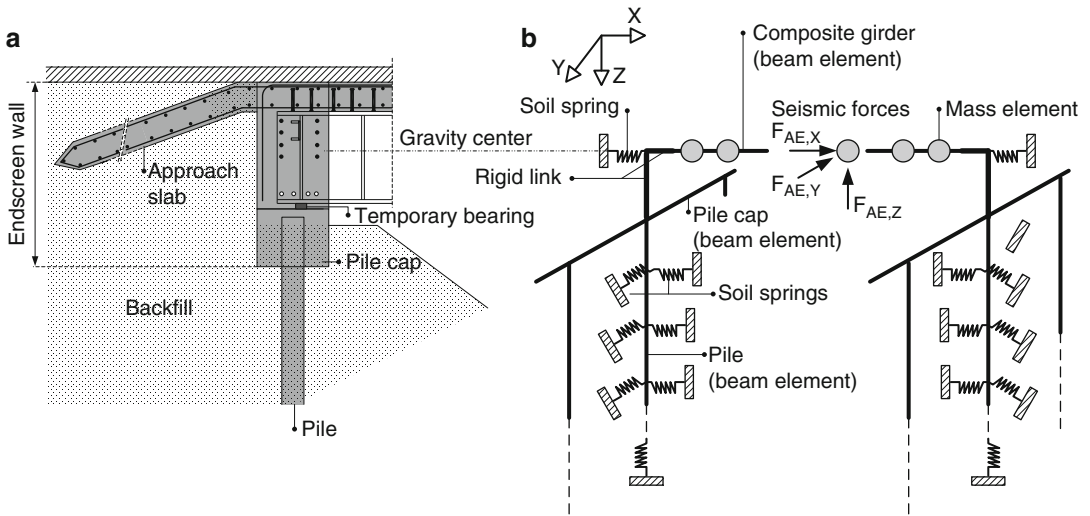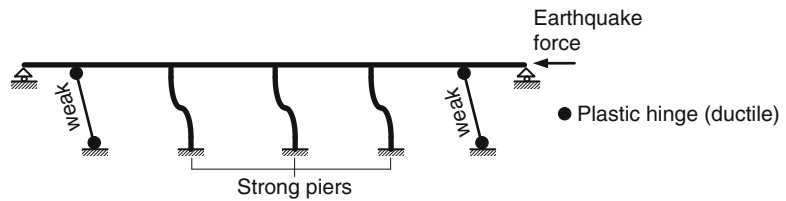

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 23 Representation of piers through discrete beam mass models – pier with (**a**) constant and (**b**) variable dimensions

subdivided in more elements, especially if its cross section varies along the height. The cross section of each element corresponds to the average pier section within its length (Fig. 23).

In frame bridges the superstructure runs continuously and is rigidly connected to the piers so that bearings and expansion joints are avoided. Such bridges are allowed to behave nonlinear in the seismic situation. Unlike in moment-resisting frame buildings where plastic hinges are developing in the beams, plastic hinges are expected to develop in bridges in the piers so that the superstructure behaves elastic (Fig. 24).

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 24 Plastic hinges in frame bridges

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 25 (**a**) End connection of superstructure to abutments in integral bridges. (**b**) Numerical grillage model of the end region of integral bridges

Accordingly for nonlinear analyses, the nonlinear properties of the piers shall be modeled. Two types of models may be used:

(a) **Plastic hinge models**: The nonlinear behavior of the pier elements is located in rotational springs at the ends of the elastic behavior part of the element. The characterization of a plastic hinge requires a moment-curvature diagram to be defined. This is obtained from the cross-sectional response to monotonic loading that is derived from a finite element analysis.

(b) **Plastic zone models**: This accounts for spread-of-plasticity effects in sections and along the beam-column element. Beam-column finite fiber elements may be used that are able to better characterize reinforced concrete elements. Consequently, higher accuracy in the structural damage estimate is attained, even for the case of high inelasticity levels.

A structural model that includes nonlinearity in a distributed fashion, using finite fiber elements, is able to characterize in higher detail the reinforced concrete elements and thought to capture more accurately response effects on such elements. Geometrical and material properties are the only required ones as input.

Modeling of Integral Bridges

Integral bridges are those where the steel girders are encased in the abutment walls so that bearings and expansion joints are avoided (Fig. 25a). Steel girders are connected to the abutment through reinforcement that passes through their web and by shear studs in the flanges or the web. Movements of the bridge due to thermal actions or time-dependent deformations of concrete are accommodated by the flexibility of piles that

Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 26 Finite element model of a filler deck bridge
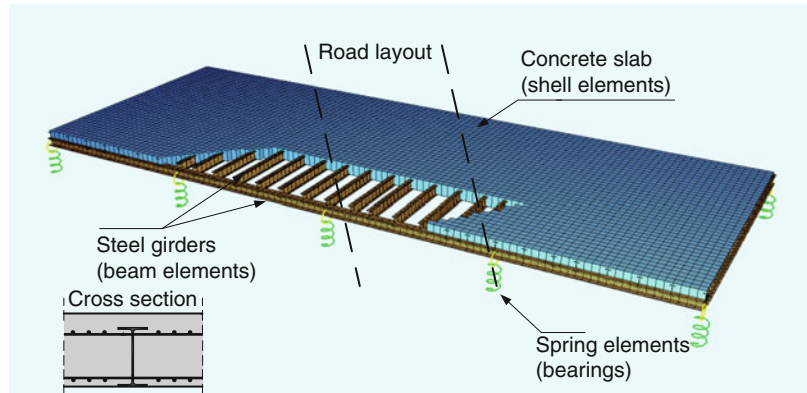
support the abutments. To increase the support flexibility, steel piles are oriented with their flanges parallel to the girders so that bending occurs around their weak axis. Integral bridges are allowed for straight and skewed bridges but not for curved bridges. Settlements of the backfill are accommodated by an approach slab. Integral bridges exhibit a good seismic performance due to increased redundancy, smaller displacements, and larger damping due to nonlinear soil-pile-structure interaction.

A grillage model for the end region of integral bridges is illustrated in Fig. 25b. Steel girders and piles are represented by 6 DOF beam elements. Beam elements representing the girders are assigned the properties of the cracked composite cross section due to the fact that negative moments develop in abutment so that the concrete slab is in tension. This section includes the section of the steel girder and the reinforcement within the effective width. Due to the small distance between piles, the pile cap is supposed to be rigid so that piles are connected with horizontal rigid elements and not with beam elements. The connection of the embedded girder and the abutment is supposed to be rigid. Accordingly the beam elements of the girders are connected by vertical rigid elements with the pile cap. Finally, horizontal rigid elements are connecting the beam elements of the girders. The vertical pile support is supposed to be hinged. Springs in two horizontal directions represent the soil that supports the piles, with stiffness varying in depth.

Horizontal springs in the longitudinal direction represent the soil-passive pressure behind the abutment.

When nonlinear methods of analysis are employed, the soil springs and the beam elements of the piles are assigned nonlinear properties. In an extension of the model, the connection between girders and abutment is considered flexible so that compression-only springs are introduced at the contact points between girders and concrete material.

Modeling of Filler Deck Bridges

Filler beam where small distanced steel beams are encased in concrete are small span bridges and are used both as simply supported (max. span ≈ 15 m) and continuous systems (max. span ≈ 30 m). Shear connection is ensured by transverse reinforcement that passes through holes at the webs of the steel beams. Figure 26 shows such a bridge that is composed of closely spaced longitudinal IPE girders resting on two transverse steel H-shaped (HEA section) girders that are supported by reinforced elastomeric bearings. The bridge is orthogonal but the road is skew.

The longitudinal elements are represented by beam elements their section being composed from the steel girder and the associated concrete embedment. The transverse H girders are also represented by beam elements with an HEA cross section. Shell finite elements represent the
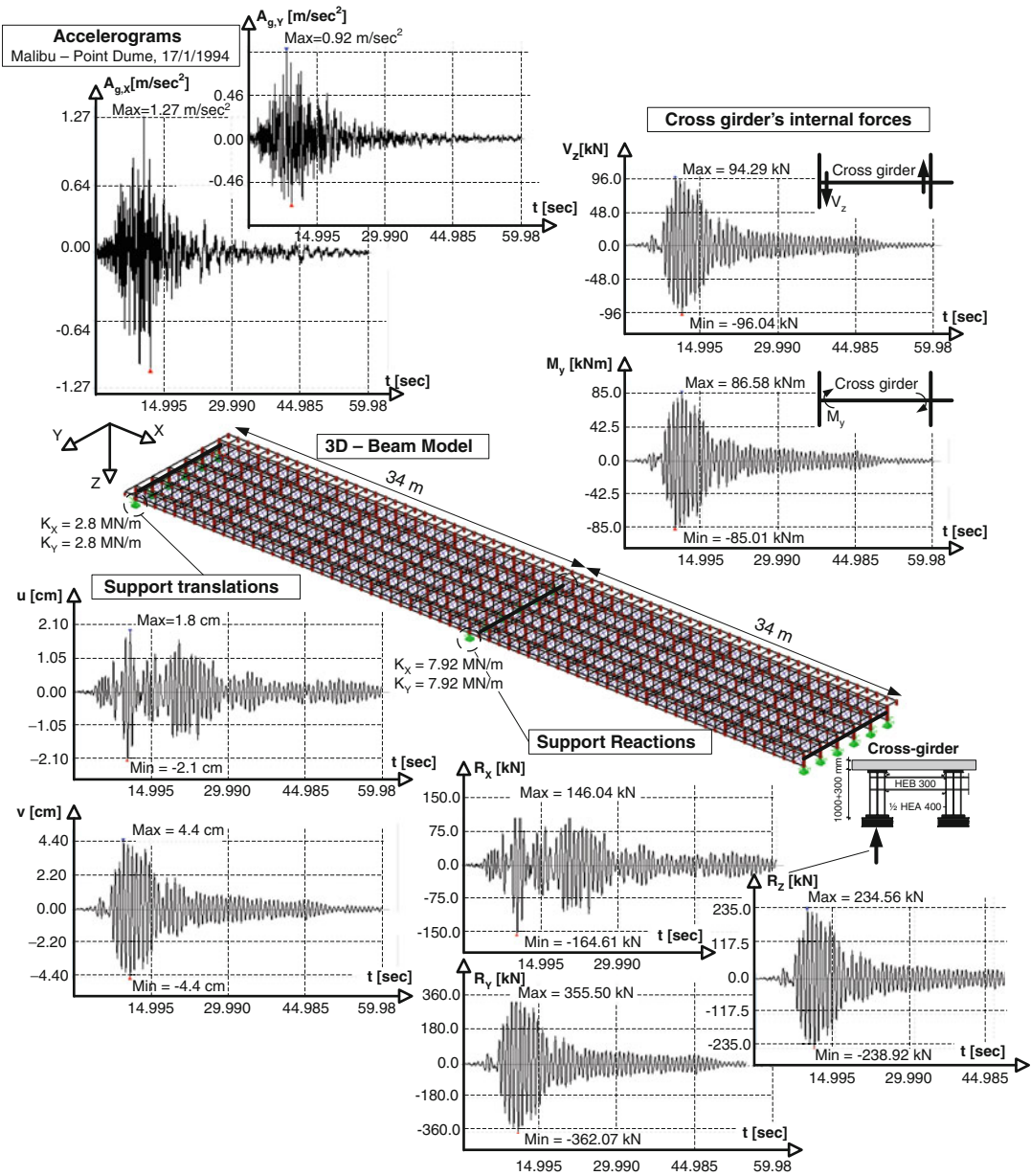
Seismic Analysis of Steel and Composite Bridges: Numerical Modeling, Fig. 27 Time history analysis of 2 × 34 m long span plate girder bridge with a 3D beam model

concrete slab that serve for an easy application of traffic loads, distribute vertical and horizontal loading between longitudinal beams, and ensure diaphragm action. In order to correctly represent the stiffness of the deck, the concrete of the longitudinal beam elements is assigned a very

low modulus of elasticity. However, concrete strength is introduced with its design value in order to allow correct design of the composite action. Deflections are calculated as average values between uncracked and cracked analysis, the latter being simulated by considering the

thickness of the concrete shell elements not as the total slab thickness but as the thickness of the compression zone. Bearings are represented by linear springs. The above-described model is a combination of a grillage and a FEM shell element model.

Application: Time History Analysis with a 3D Beam Model

A time history analysis of a two-span composite bridge with six main girders is indicatively demonstrated in Fig. 27. The bridge has been modeled according to the recommendations of the section "3D Models." The superstructure rests on low-damping elastomeric bearings whose stiffness has been estimated through Eqs. 8, 9, 10, and 11a and 11b. For the analysis a damping factor equal to 2 % for the three first natural modes has been implemented, value taken from the fabricator of the bearings. The mass of the superstructure includes all the permanent loads and 20 % of the traffic loads. One can see that with the 3D beam model, the time variation of the vertical reaction forces R_Z can be calculated; such a calculation with a grillage model would not be feasible. Moreover, the internal forces of the cross girders are calculated. Not only the grillage model but a detailed FE model as well would not give the possibility of such a calculation since shell or volume elements offer stresses and not forces or bending moments as final results. With the 3D beam model, the frame action of the cross girders (HEB 300) and the transverse stiffeners (1/2 HEA 400) could be adequately investigated.

Summary

Different types of models are used for the design and seismic analysis of steel and steel-concrete composite bridges. In this entry 1D, grillage, spine, 3D beam, and finite element models are presented. The advantages and drawbacks of each model are debated, and modeling recommendations for plate, filler deck, and box girder bridges are provided. The implementation of structural bearings in the models is also discussed, and guidance on the calculation of the bearings' stiffness and the seismic forces is given. The entry ends with an example of a time history analysis of a continuous composite beam and the calculation of reaction forces and displacements.

Cross-References

▶ Buildings and Bridges Equipped with Passive Dampers Under Seismic Actions: Modeling and Analysis

▶ Seismic Analysis of Concrete Bridges: Numerical Modeling

References

Adamakos T, Vayas I, Petridis S, Iliopoulos A (2011) Modeling of curved composite I-girder bridges using spatial systems of beam elements. J Construct Steel Res 67(3):462–470

Chopra AK (1995) Dynamics of structures. Prentice Hall, Upper Saddle River

EN 1993-2 (2006) CEN European Committee for Standardization. Design of steel structures. Part 2: Steel bridges

EN 1994-2 (2005) CEN European Committee for Standardization. Design of composite steel and concrete structures. Part 2: General rules and rules for bridges

EN 1998-2 (2005) CEN European Committee for Standardization. Design of structures for earthquake resistance. Part 2: Bridges.

Hambly EC (1990) Bridge deck behaviour, 2nd edn. E. F. & N. Spon, London

Kindmann R, Kraus M (2011) Steel structures – design using FEM. Ernst & Sohn, Berlin

Vayas I, Iliopoulos A (2013) Design of steel concrete composite bridges to Eurocodes. CPC Press, Taylor and Francis Group, Boca Raton

Vayas I, Iliopoulos A, Adamakos T (2010) Spatial systems for modelling steel-concrete composite bridges – comparison of grillage systems and FE models. Steel Construct Des Res 3:100–11

Vayas I, Adamakos T, Iliopoulos A (2011) Three dimensional modeling for steel-concrete composite bridges using systems of bar elements – modeling of skewed bridges. Int J Steel Struct 11(2):157–169

Seismic Analysis of Steel Buildings: Numerical Modeling

Charis J. Gantes
Institute of Steel Structures, School of Civil
Engineering, National Technical University of
Athens, Zografou Campus, Athens, Greece

Synonyms

Numerical simulation of steel buildings; Structural design of steel buildings in seismic regions

Introduction

The aim of this contribution is to present the state of the art in numerical modeling of steel buildings, aiming at their structural analysis and design. To that effect, the successive stages of this process will be presented, including the conceptual design, numerical modeling, application of support conditions and loads, structural analysis, and design checks. The two latter issues, structural analysis and design checks, will be discussed briefly, to the extent that they interact with the modeling process, as they are also covered in detail in other parts of this encyclopedia. The interaction with other engineering disciplines throughout the design process will be demonstrated. For the purpose of this presentation, a prototype building will be used, which is described in the next section.

Description of Prototype Building

A simple single-story laboratory building, part of an industrial project, will be used as prototype building for this presentation. The building's plan view is 28.00 m by 15.00 m (Fig. 1), and there is a small loft in part of the building (Fig. 2). The roof is double pitched with minimum required clearance equal to 4.50 m (Fig. 3). The locations of exterior doors are fixed, imposing restrictions on the layout of structural elements, as will be described in the following section. The preliminary architectural drawings include proposed column locations, indicated by their I-sections in the plan views of Figs. 1 and 2.

Conceptual Design

The initial stage of the structural design process consists of the so-called conceptual design, in which the different structural elements that are needed to support the structure are selected and are geometrically defined, taking into account the architectural requirements. In parallel, all restrictions imposed by the building's use, as well as its location and associated environmental actions have to be satisfied.

To that effect, close cooperation and interaction of the structural engineer with other involved engineering disciplines are necessary. Such disciplines include, at a minimum:

- The architectural engineering team, to determine the overall geometry; positions of openings; types of materials for floors, cladding, etc.; and associated dead loads and, in general, to ensure that structural choices do not interfere with the building's desired function and aesthetics
- The mechanical/electrical/HVAC engineering team, to determine loads that are associated with these functions of the building and to prevent conflicts with required shafts and other openings for passage of lifelines
- The geotechnical engineering team, to obtain data about the soil and its mechanical properties and to determine the type of the building's foundation

In conceptual design, all basic decisions, regarding material properties, structural system, loads, and codes to be applied, are taken. All structural engineering experts agree that this is the most important design stage and that these decisions affect significantly the project's design and construction schedule and
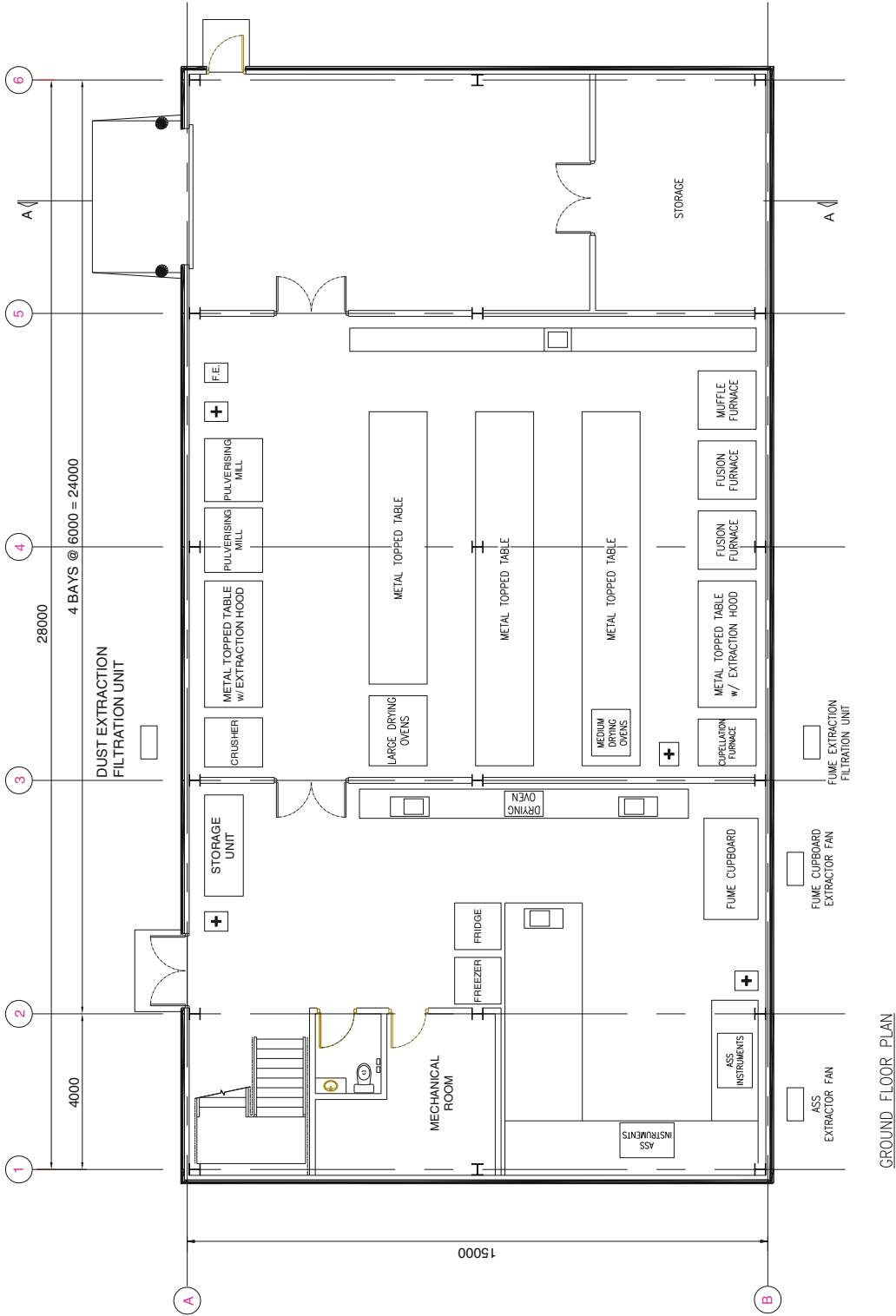
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 1 Architectural plan view of ground floor of prototype building (out of scale)
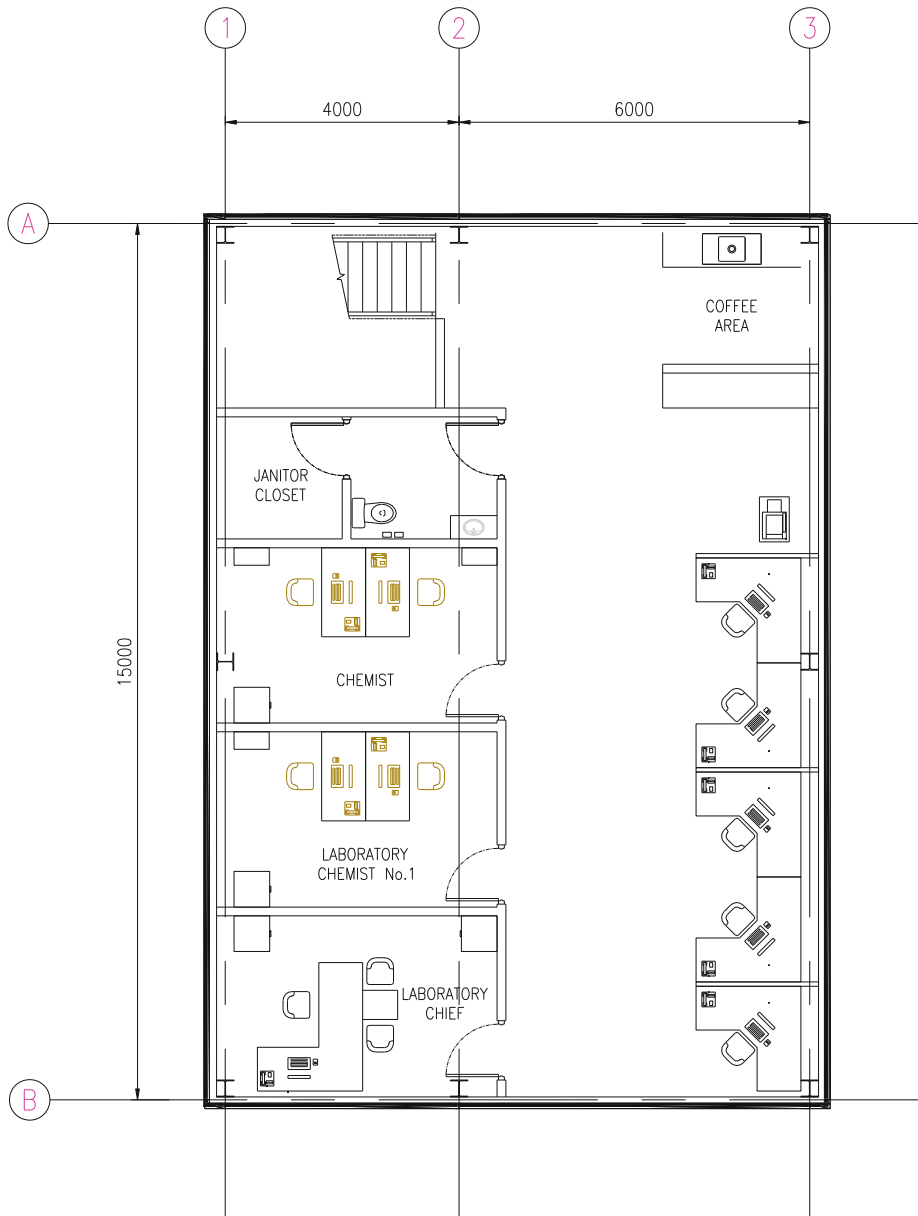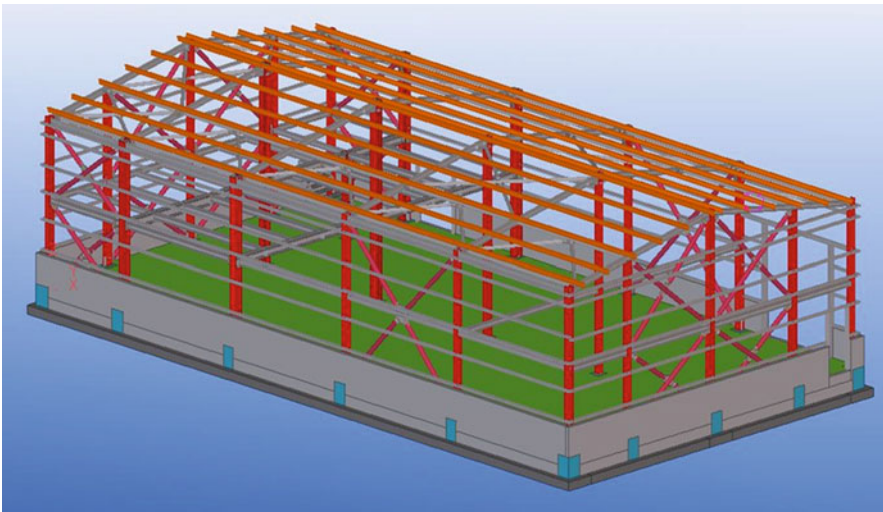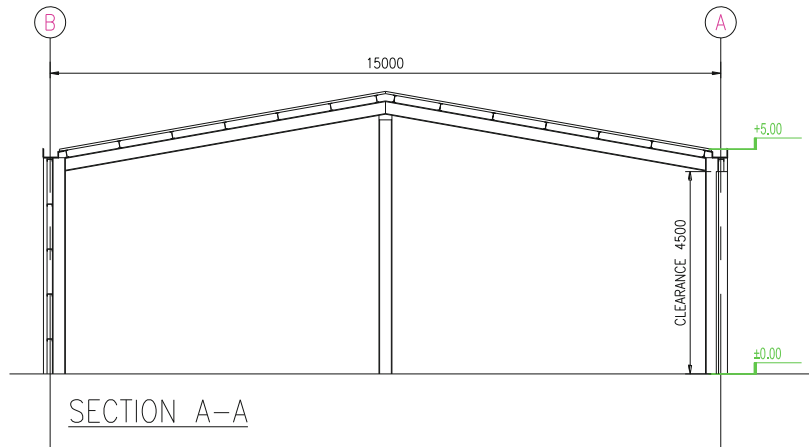
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 2 Architectural plan view of loft of prototype building (out of scale)

cost, sometimes even its feasibility. This is a process mostly dictated by the available experience of all participating design teams but can nowadays be assisted by modern software tools, both for preliminary design of the different building's facilities but also for communication between involved teams (building information modeling, BIM). However, a

Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 3 Transverse section A-A of prototype building (out of scale)

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 4 3D view of prototype building's 3D model

more detailed treatment of this design stage is beyond the scope of this presentation.

Several conceptual design choices for the prototype building described in section "Description of Prototype Building" are illustrated in Figs. 4, 5, 6, 7, 8, 9, 10, and 11. The figures have been obtained from a 3D model of the building's structure, made with a specialized modeling software, which enables integration of architectural and structural engineering and other solutions and production planning systems, working for both detailing and producing steel structures of any type.

The building's structural system consists of:

- Two-span main frames spaced at 6.00 m (with the exception of the first span, which is equal to 4.00 m) in accordance with the architectural proposal. Columns and girders have I-section, oriented so that the strong axis acts for in-plane loading. It is noted that a 6.00 m distance between main frames is optimum for avoiding loss of material in purlins and longitudinal beams, taking into account that steel members are commonly fabricated in 12.00 m lengths. Haunches are provided between
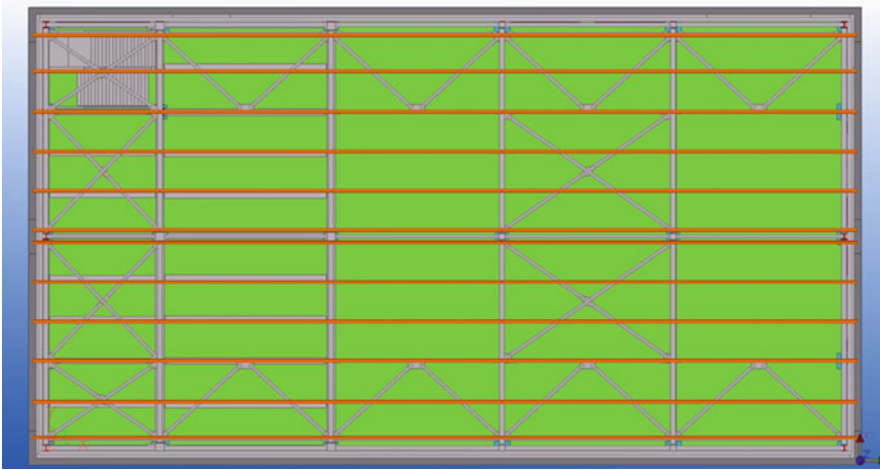
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 5 Plan view of prototype building's 3D model above foundation level
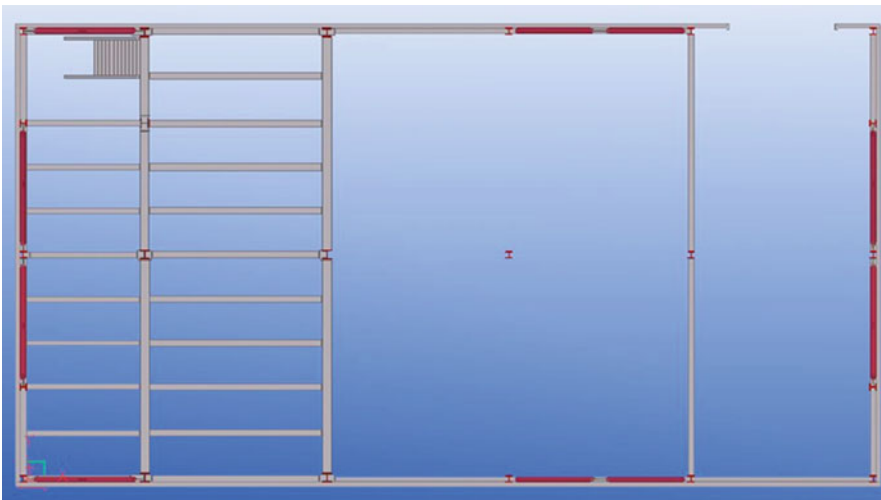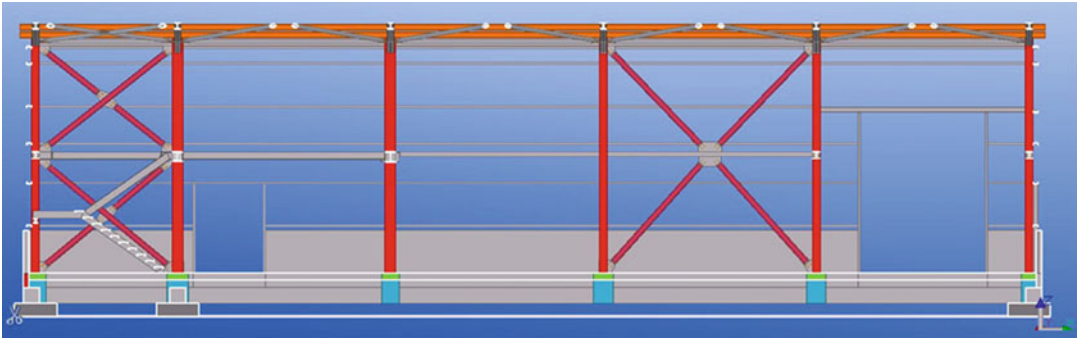

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 6 Plan view of prototype building's 3D model above loft level

columns and girders at internal frames, in order to achieve reduction of girder cross section and to accommodate the required number of bolts for a moment connection. External frames are equipped with vertical X-bracing, so that truss action dominates rather than frame action. As a result, smaller column and girder sections are possible, and there is no need for haunches between columns and girders.

- Longitudinal beams connecting the main frames at the top and at mid-height of columns. On one side of the building, the mid-height head beam is eliminated in the last span, to accommodate a required door at that location. A beam over the door is used instead. It is noted that this is not an optimum structural configuration, but it is an acceptable compromise to respect functional requirements.

S

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 7 Roof plan view of prototype building's 3D model

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 8 Side view of prototype building's 3D model
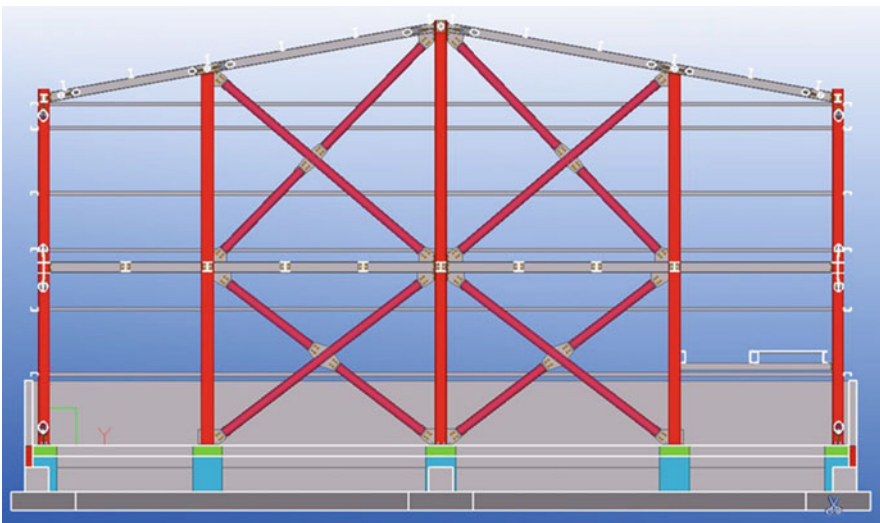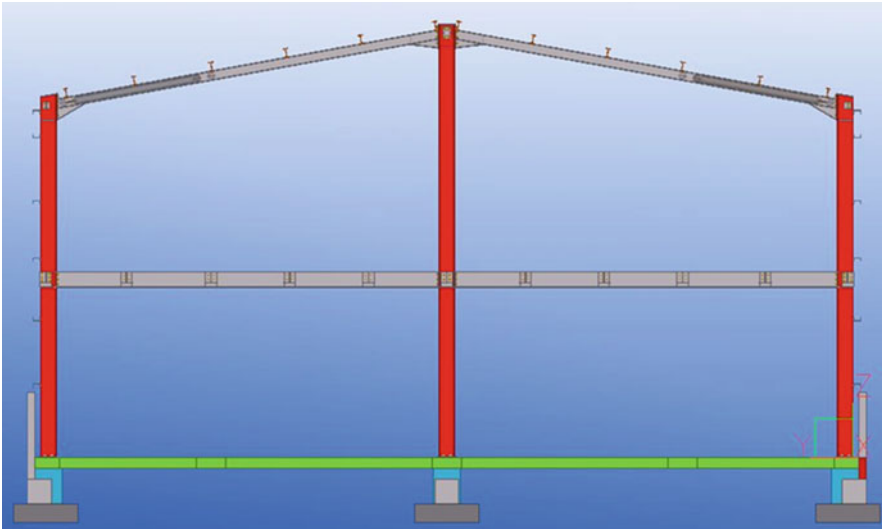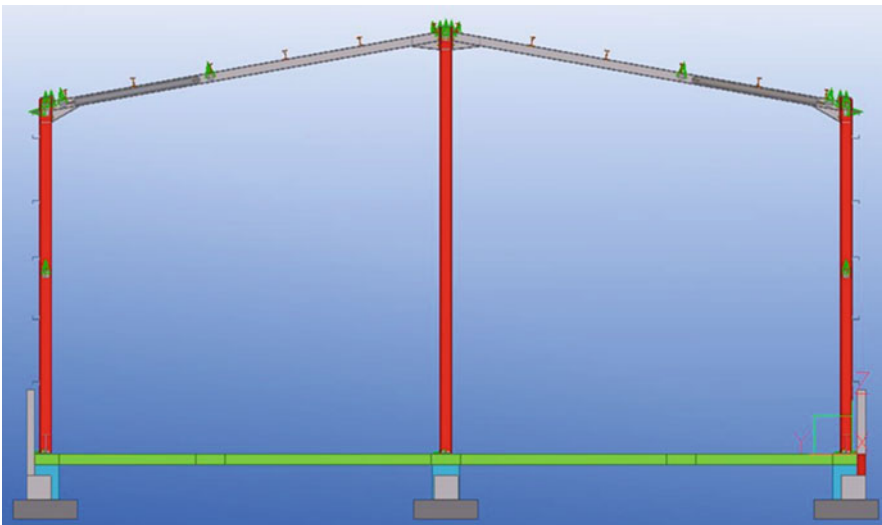

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 9 Front view of prototype building's 3D model

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 10 Transverse section of prototype building's 3D model at spans with loft

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 11 Transverse section of prototype building's 3D model at spans without loft

- Vertical X-bracing at two spans to provide lateral stability in the longitudinal direction. The type of bracing depends on panel geometry, aiming at inclination that is as close as possible to 45°.
- Horizontal roof bracing at the same spans where vertical bracing is provided, as well as at the roof's peripheral panels, to ensure sufficient diaphragm action.
- Roof and side purlins to support cladding.

A peripheral reinforced concrete wall is provided, to ensure water tightness and to protect the steel columns from eventual vehicle collisions.

Taking also into account the available geotechnical information for this site, foundation beams are provided at the periphery of the building, along the axis of intermediate columns and below the auxiliary columns supporting the loft. A steel staircase is provided to connect the ground floor with the loft.

Numerical Modeling

For the simulation, analysis and design of superstructure and foundation use is made of an appropriate finite element software. Views of the finite element model are shown in Figs. 12, 13, 14, and 15.

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 12 3D view of prototype building's finite element model
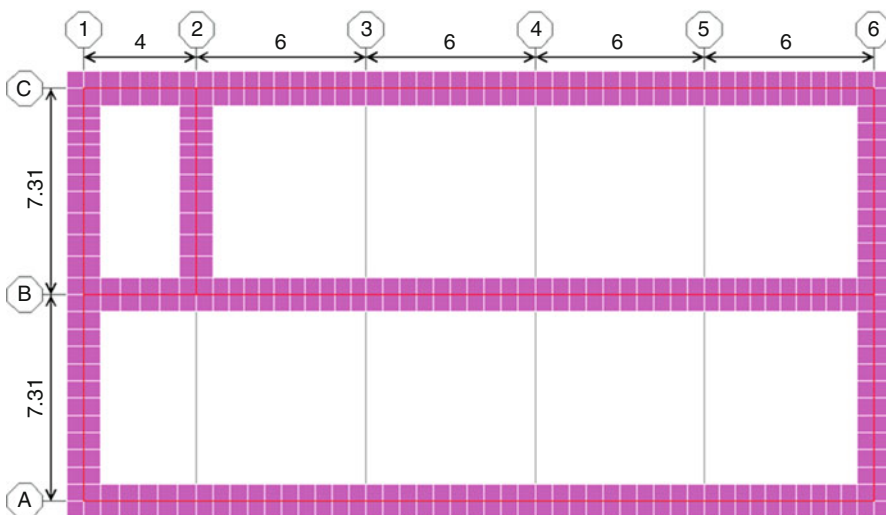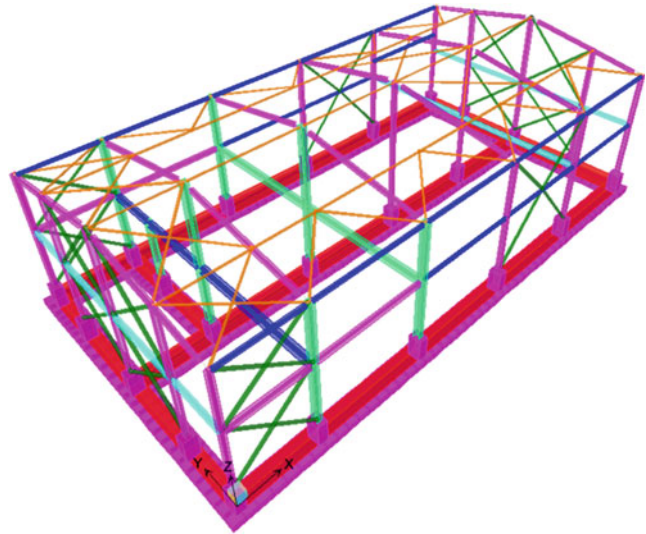
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 13 Foundation plan view from prototype building's finite element model
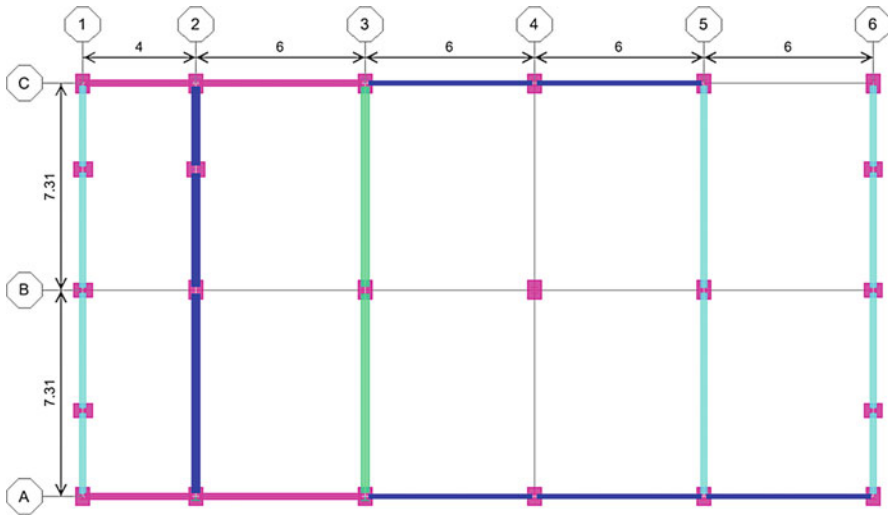
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 14 Loft's plan view from prototype building's finite element model
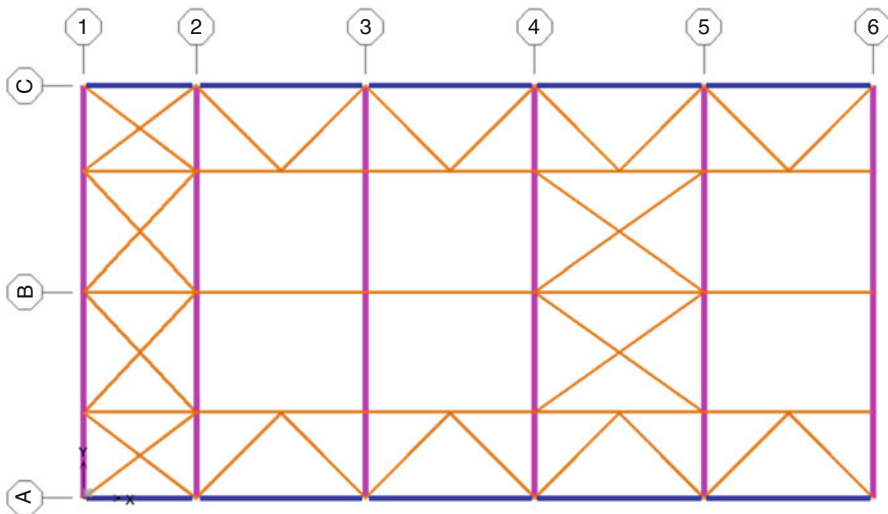
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 15 Roof's plan view from prototype building's finite element model

Some recommendations of good practice when setting up the finite element model are the following:

• It is recommended not to include roof and side purlins and other secondary members in the numerical model. The type of connection between purlins and main frame, which is a simple support so that both purlins and frame

columns and girders are continuous but do not transmit moments to each other, would require rigid links for correct modeling, which is possible in most structural programs, but would render the model unnecessarily complex. Only those members that have an active role in the building's spatial response to either vertical or horizontal actions need to be modeled. Other members can be modeled, analyzed, and

Fig. 16 Column base
connections of prototype
building, modeled as hinges
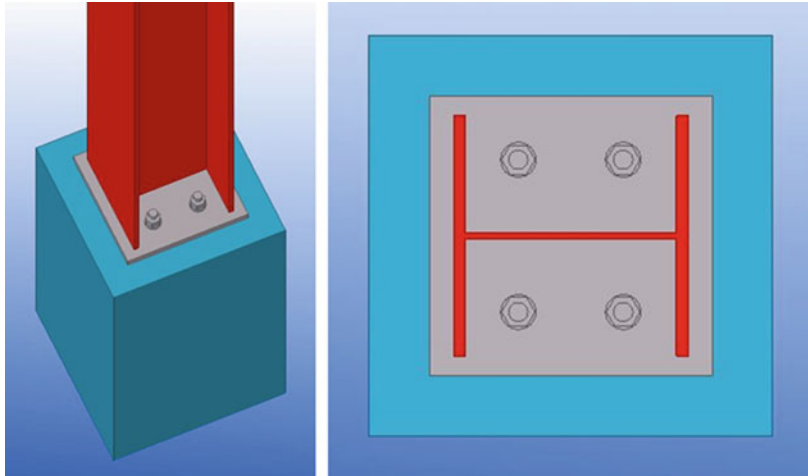
dimensioned independently, and the loads
they transfer to the frames can be applied
directly upon the frame columns and girders
in the 3D finite element model.

- Haunches can be modeled by means of uniform beam section over the entire length of the
haunch, usually the one at haunch mid-length.
This provides a satisfactory representation of
increased girder stiffness in the area of the
haunch and enables safe prediction of strength
for the design phase.

- Particular attention is needed for the appropriate modeling of supports and connections, by
releasing the appropriate member degrees of
freedom depending on the connection detail.
Even though real connections between steel
members are neither perfect hinges nor perfect
moment connections, it is common practice to
model all connections as either hinges or
moment connections, at least initially. Modern
codes provide tools for eventually classifying
connections as semirigid (e.g., European
Committee for Standardisation 2003c), in
which case either the connections should be
stiffened to become rigid or a rotational spring
should be inserted in the model to represent
the partially restricted connection between the
members in question. The decision about the
type of joint to be included in the model
depends mainly on the lever arm between
bolts in the connection. Connections of
I-beams where all bolts are between the

flanges are commonly modeled as hinged, as
is the case with the supports of all columns in
the prototype building (Fig. 16). In case bolts
are provided outside the flanges, rigid moment
connections are considered in the model, as is
the case with the column to girder connections
in the prototype building (Fig. 17). Welded
connections are usually modeled as rigid,
while connections between or at the end of
bracing members (Fig. 18) are modeled as
hinged.

For moment connections, no intervention is
required in the model, as this is the default
option for frame structures. Hinged connections are modeled by means of releases of
appropriate rotational degrees of freedom at
the members' ends. In Figs. 19 and 20,
moment releases at the ends of roof and vertical bracing as well as secondary beams are
denoted by black dots.

- If there is a diaphragm in the structure, as, for
example, is the case with the composite slab of
the loft in the prototype building, it is
recommended to avoid the more complex
option of modeling it with shell elements and
instead create a "master" node that will be
connected with all nodes of the diaphragm
(Fig. 21). In this case, in order to account for
uncertainties in the location of masses and in
the spatial variation of the seismic motion, the
calculated center of mass at each floor shall be
considered as being displaced from its

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 17 Column to girder connections of prototype building, modeled as rigid moment connections
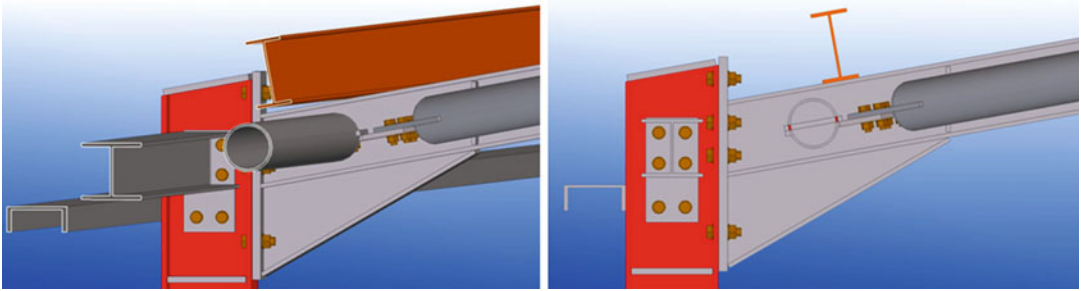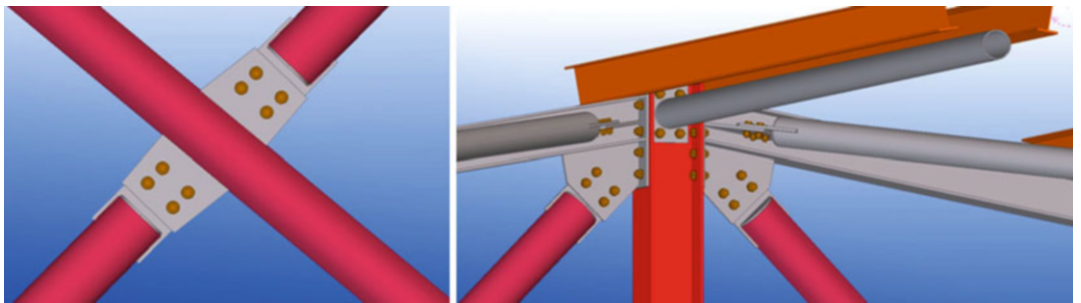

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 18 Connections between or at the end of bracing members, modeled as hinges
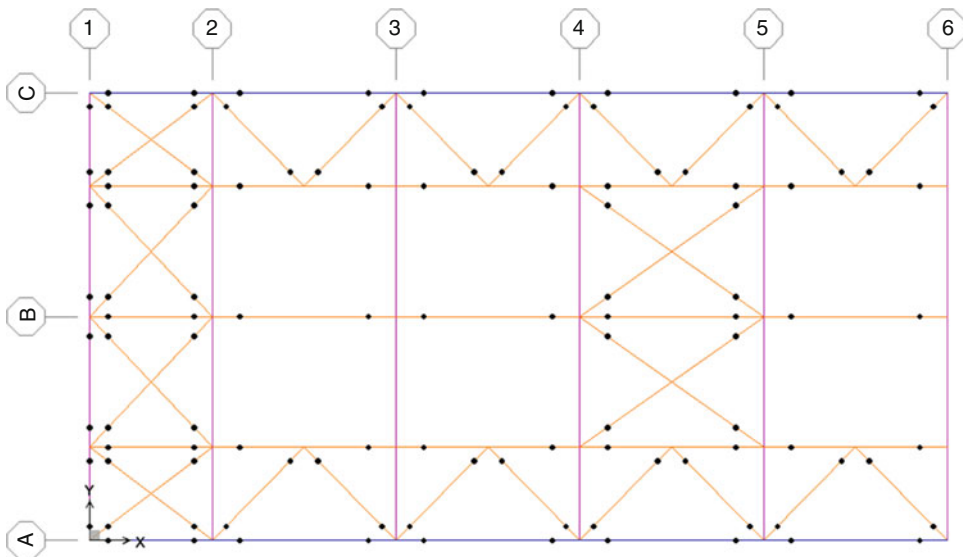
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 19 Roof's plan view from numerical model illustrating with dots the joint releases on roof's bracing and on secondary beams
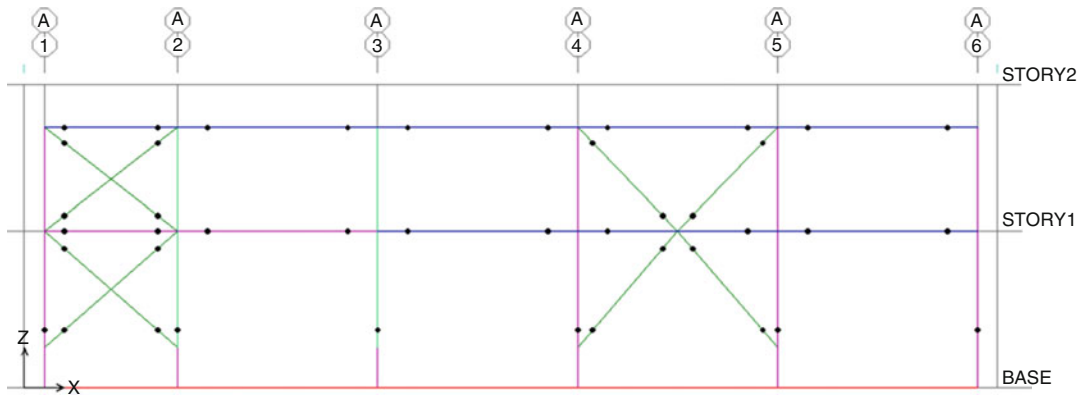
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 20 Building's side view from numerical model illustrating with dots the joint releases on column bases, on vertical bracing, and on secondary beams (reinforced concrete pedestals on which steel columns are supported are also shown)

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 21 Diaphragm definition of loft slab
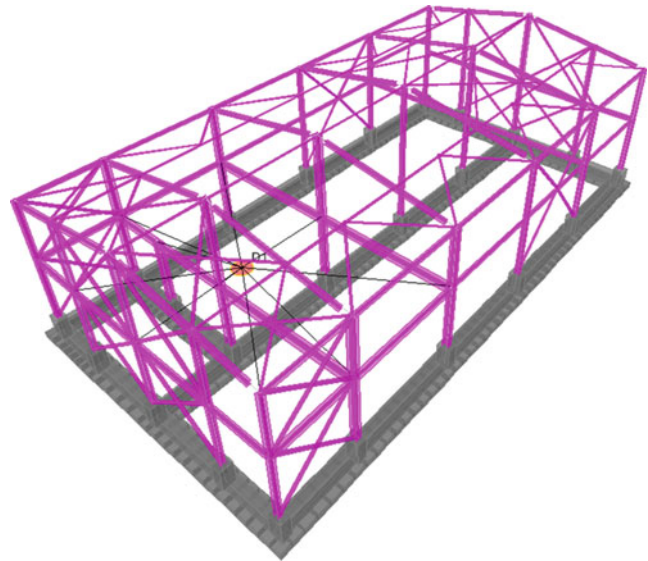
nominal location in each direction by an accidental eccentricity, commonly taken equal to $e = 0.05 \cdot L$, where L is the corresponding span length.

Loads

The next phase of the design process consists of setting up appropriate loading situations including basic loads and load combinations. Basic loads commonly include permanent loads, live loads depending on the building's intended use, and environmental loads such as those due to snow, wind, and temperature, as well as seismic actions in case the building is located in a seismic region. Characteristic nominal values of live, snow, and wind loads and seismic actions are prescribed by pertinent codes (e.g., European Committee for Standardisation 2001b, 2003a, b, d, 2004a).

For the prototype building presented here, Excel spreadsheets for the calculation of snow, wind, and seismic loads are illustrated in Figs. 22, 23, 24, and 25. Issues that may require special attention include snow accumulation due

Snow load calculation			
According EN 1991-1-3			

	$a_1=$	12.0	°		
	Roof's slope	0°<α<30°	30°<α<60°	a>60°	
For a_1	$\mu_1(\alpha_1)$	0.80	1.28	0	
	$\mu_2(\alpha_1)$	1.12	1.60	0	
For a_2	$\mu_1(\alpha_2)$	0.80	1.60	0	
	$\mu_2(\alpha_2)$	0.80	1.60	0	

$$s = \mu_i * C_e * C_t * s_k$$

117.5

where

$\mu_i=$	0.80	snow load shape coefficient
$C_e=$	1.00	exposure coefficient
$C_t=$	1.00	thermal coefficiant

$$s_k = s_{k,0}\left[1+\left(\frac{A}{917}\right)^2\right]$$

$s_{k,0}=$	0.8	design value of exceptional snow load on the ground for a given location (kN/m2)
$A =$	130	Site altitude above sea level [m]
$s_k=$	0.82	characteristic value of snow on the ground at the relevant site(kN/m^2)
$\gamma=$	1.00	weight density of snow (kN/m^3)
$s=$	0.65	snow load on the roof (kN/m^2)

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 22 Excel spreadsheet for snow load calculation for the prototype building according to Eurocode 1 – Part 1–3

to roof geometry, increased snow density in regions of low temperatures, or unusual building shapes not covered by the code guidelines for wind loads. Choices that differentiate the design spectrum of steel structures from those of structures made of different materials are the values of damping, commonly taken equal to 2 % for welded and 4 % for bolted steel structures, and the values of behavior factor, for which recommendations are provided in the pertinent seismic codes, depending also on the type of structural system.

In case part of the structure is buried (e.g., a basement), it is recommended to define accordingly the level above which the structure vibrates independently of the ground.

S

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 23 Excel spreadsheet for wind load calculation on the side walls of the prototype building according to Eurocode 1 – Part 1–4
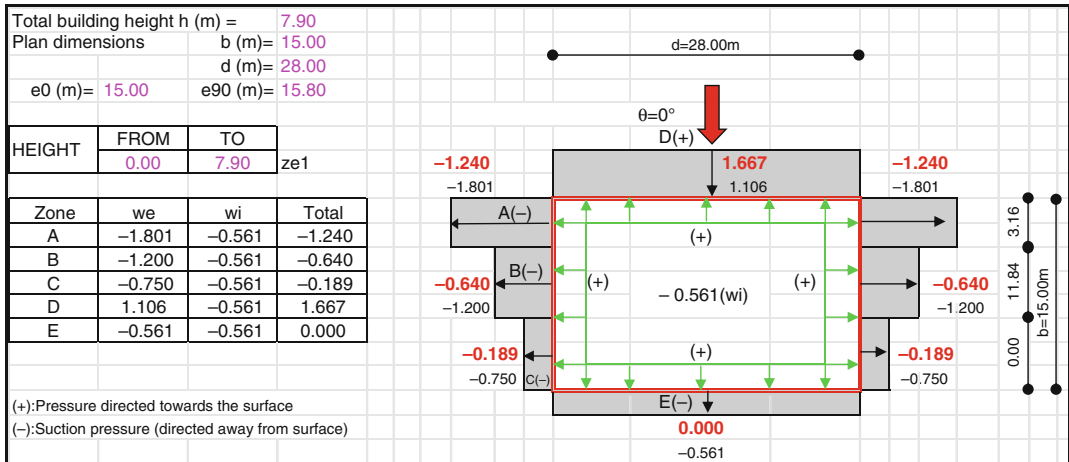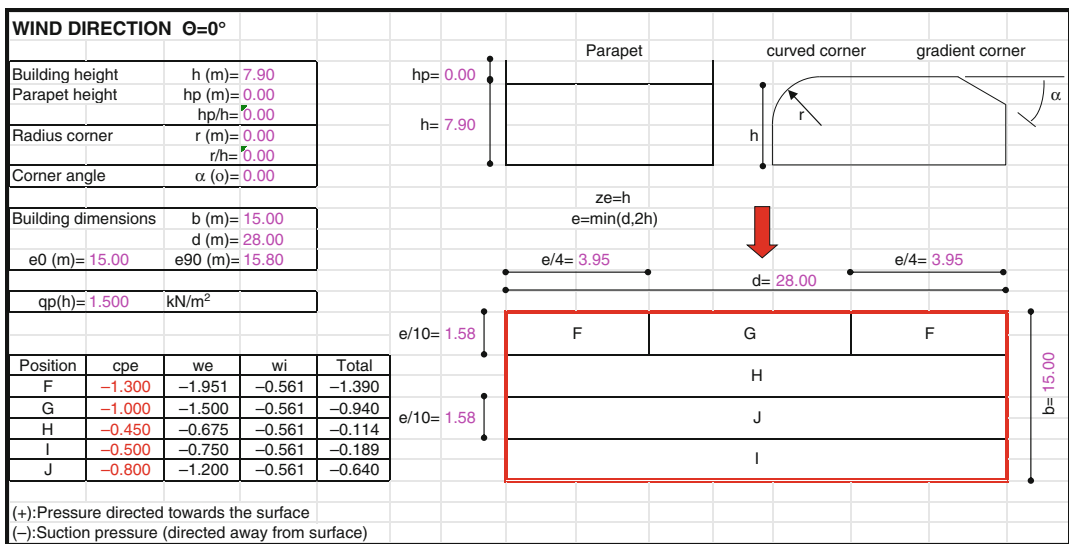

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 24 Excel spreadsheet for wind load calculation on the roof of the prototype building according to Eurocode 1 – Part 1–4

Following load calculation, basic loads are then applied on the model, as shown schematically in Figs. 26, 27, 28, and 29.

The next step is to describe pertinent load combinations in accordance with the relevant codes (e.g., European Committee for Standardisation 2001a), including combinations in the ultimate (ULS), serviceability (SLS), and seismic limit states. Common load combinations

for ordinary steel buildings are provided in Table 1.

In Table 1:

- Live load corresponds to all relevant types of live load, such as service load or load due to snow.
- Wind load corresponds to four basic individual load cases for each direction, including

Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 25 Horizontal component of elastic response spectrum of the prototype building according to Eurocode 8
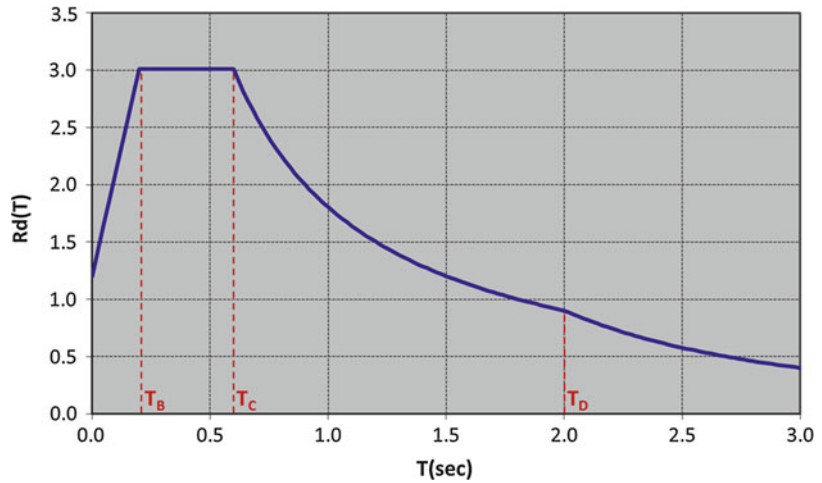
Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 26 Application of dead loads on the finite element model of the prototype building
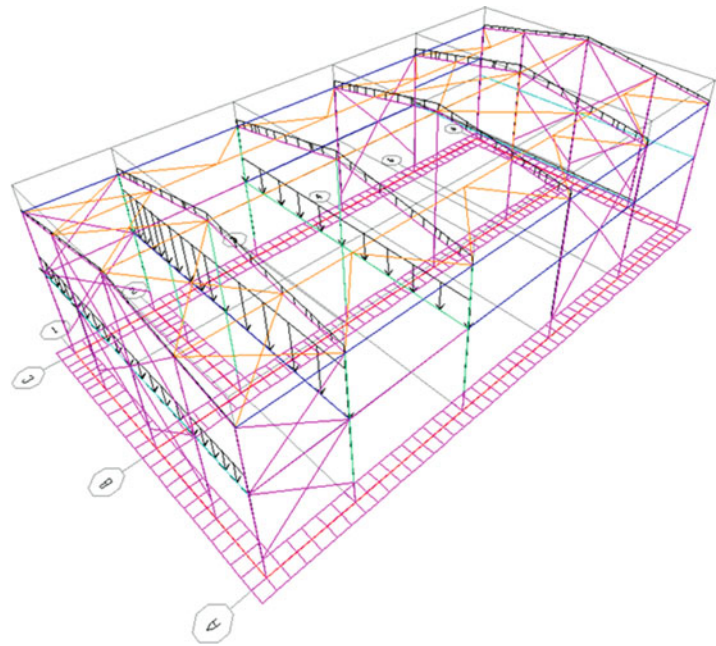
uplift or downward pressure on the roof and positive or negative pressures on the walls.

- Earthquake load corresponds to three basic individual load cases, one for each global direction.
- Temperature load corresponds to two basic individual loads (increase, decrease).

Structural Analysis

A decision that has to be taken next concerns the type of structural analysis to be performed.

A variety of linear and nonlinear (geometrically and/or material) and static or dynamic analysis algorithms are commonly available in commercial structural software (e.g., Gantes and Fragkopoulos 2010). However, in the vast majority of common steel buildings, linear static analysis is used in design practice for all loads and load combinations, with the exception of seismic actions, for which frequently both equivalent static as well as response spectrum dynamic analysis are carried out. The principle of

Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 27 Application of live loads on the finite element model of the prototype building
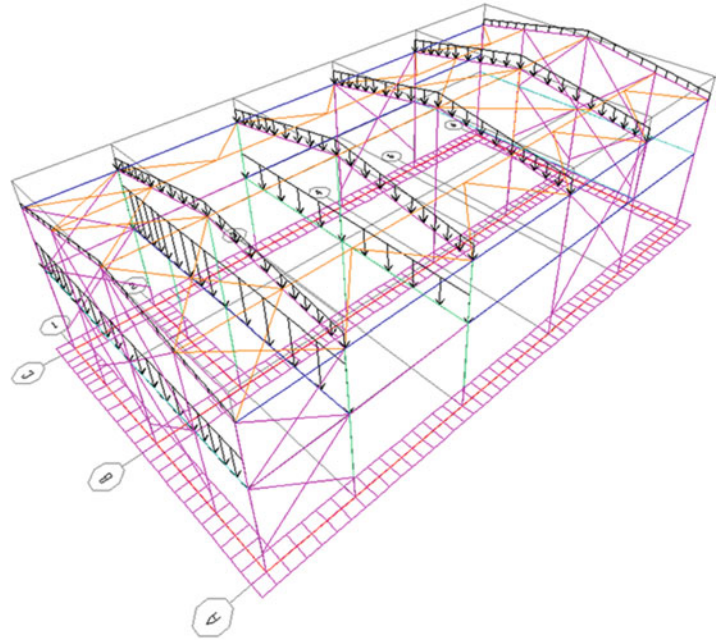
Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 28 Application of wind loads on the finite element model of the prototype building
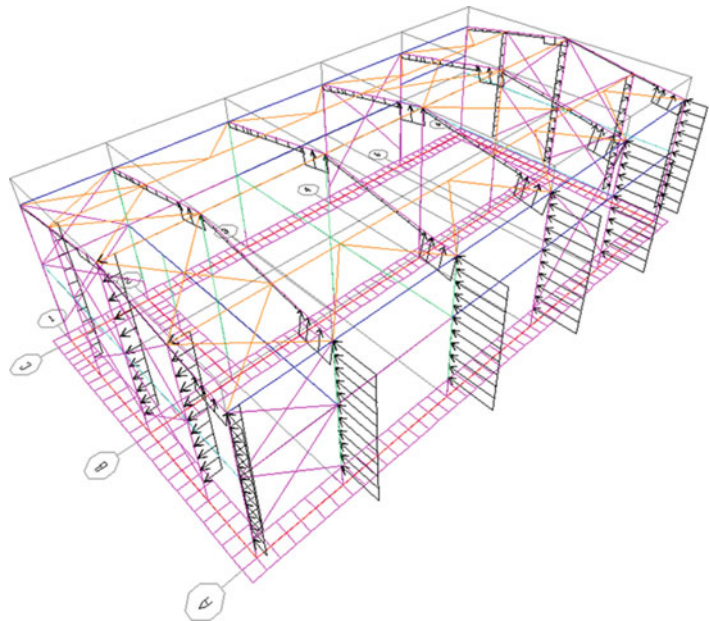
superposition is routinely employed to obtain action effects for load combinations.

In the equivalent static method, first the base shear is calculated, which is then distributed over the height of the building as specified by the pertinent seismic code. In the response spectrum analysis, for a given direction of acceleration, the maximum displacements, forces, and stresses are computed throughout the structure for each vibration mode. These modal values for a given response quantity (displacements, forces, or stresses) are then combined appropriately to produce a single, positive result for the given direction of acceleration, using, for example, the SRSS
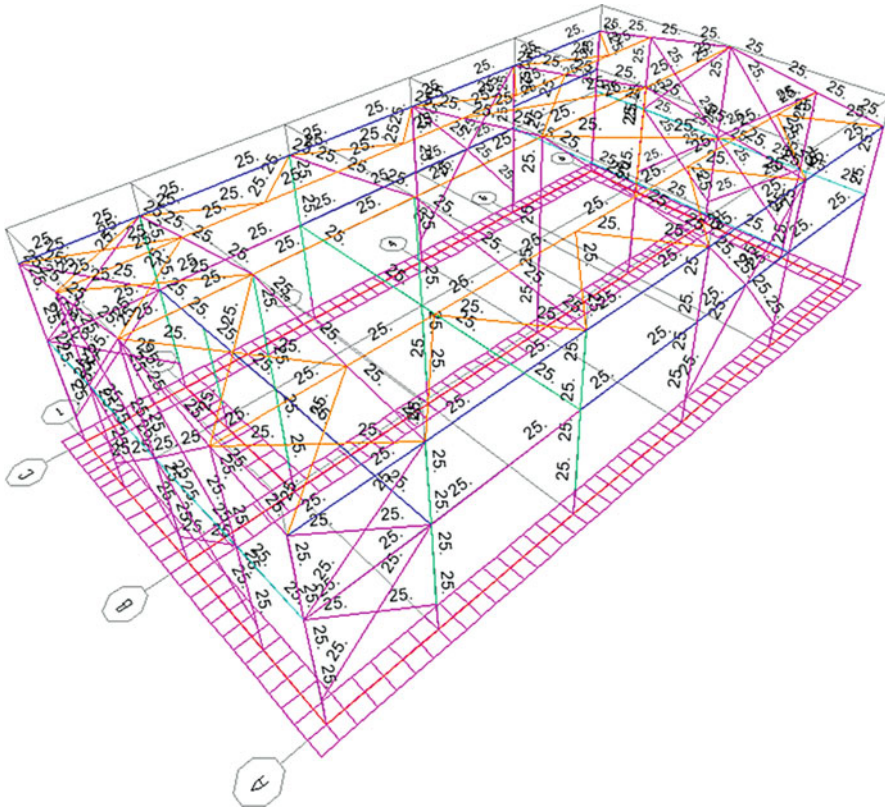
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 29 Application of temperature increase on the finite element model of the prototype building

or the CQC method. It is noted that most modern codes contain a requirement of modal participating mass ratios at least equal to 90 %. In case this is not satisfied, the response spectrum analysis must be repeated with higher number of participating modes.

The analysis results include mode shapes and natural periods of vibration as well as deformed shapes and internal force diagrams for all individual loads and load combinations. A qualitative evaluation of analysis results is always highly recommended in order to detect possible modeling errors. For that purpose, it is a good practice to start with mode shapes and natural periods of vibration, proceed with deformed shapes for individual load cases, continue with internal force diagrams for individual load cases, and finally, conclude with internal force diagrams for load combinations and envelopes, which are also used for design. Indicative results for the prototype building are presented in the following figures, including mode shapes (Figs. 30 and 31), seismic deformed shapes (Figs. 32 and 33), and internal force diagrams for individual load cases (Fig. 34) and load combination envelopes (Figs. 35, 36, 37, and 38).

Structural Design Checks

Following the qualitative evaluation of analysis results on the basis of mode shapes and corresponding vibration periods, deformed shapes and internal force diagrams, and provided that confidence is gained for the correctness of these results, design checks are performed, consisting of general checks, member checks, and connection checks. Foundation checks are

Seismic Analysis of Steel Buildings: Numerical Modeling, Table 1 Common load combinations for ordinary steel buildings in the ultimate (ULS), serviceability (SLS), and seismic limit states according to the Eurocodes

	Load combination	Load factors $\gamma_{unfavorable}$; $\gamma_{favourable}$				
		Dead	Live	Wind	Temperature	Earthquake
ULS	COMB1	1.35	1.50			
	COMB2	1.35	1.50	0.90		
	COMB3	1.35	1.50		0.90	
	COMB4	1.35	1.50	0.90	0.90	
	COMB5	1.35;1.00		1.50		
	COMB6	1.35;1.00	0.90	1.50		
	COMB7	1.35;1.00		1.50	0.90	
	COMB8	1.35;1.00	0.90	1.50	0.90	
	COMB9	1.35			1.50	
	COMB10	1.35	0.90		1.50	
	COMB11	1.35		0.90	1.50	
	COMB12	1.35	0.90	0.90	1.50	
SLS	COMB13	1.00	1.00			
	COMB14	1.00	1.00	0.60		
	COMB15	1.00	1.00		0.60	
	COMB16	1.00	1.00	0.60	0.60	
	COMB17	1.00		1.00		
	COMB18	1.00	0.60	1.00		
	COMB19	1.00		1.00	0.60	
	COMB20	1.00	0.60	1.00	0.60	
	COMB21	1.00			1.00	
	COMB22	1.00	0.60		1.00	
	COMB23	1.00		0.60	1.00	
	COMB24	1.00	0.60	0.60	1.00	
SEISMIC	COMB25	1.00	0.30			1.00

Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 30 Dominant vibration mode in transverse direction of prototype building
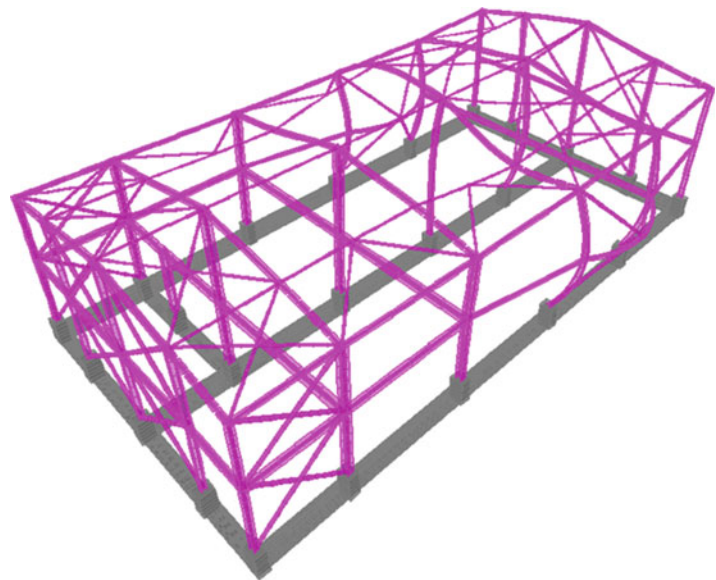
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 31 Dominant vibration mode in longitudinal direction of prototype building
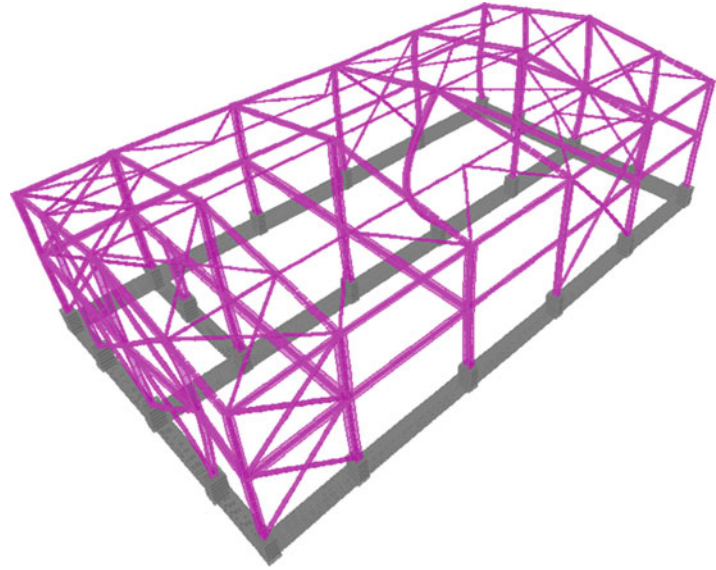
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 32 Deformed shape of prototype building from equivalent static method – transverse direction
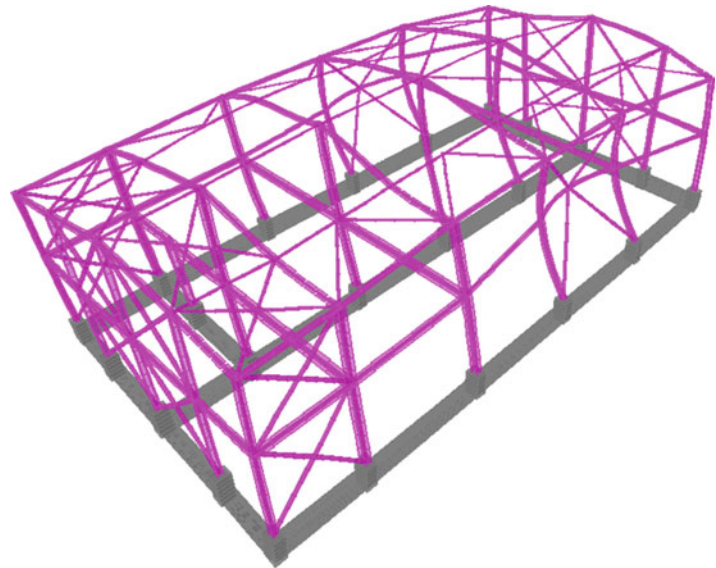
also part of this process, which are presented in other parts of this encyclopedia. This is an iterative process, and frequently, there is a need to return to the model, modify it, and run a new set of analyses and checks, as schematically illustrated in Fig. 39.

General checks (e.g., European Committee for Standardisation 2004b) consist mainly of confirming that overall structure deformations are acceptable for all load combinations. This includes interstory drifts and overall building drift for all load combinations with predominantly horizontal components, such as wind and seismic combinations. Depending on the use of the building, general checks may also include restrictions about the vertical vibration frequencies, associated with a sensation of unease of the users, as is the case in stadium grandstands. In case such checks

Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 33 Deformed shape of prototype building from equivalent static method – longitudinal direction
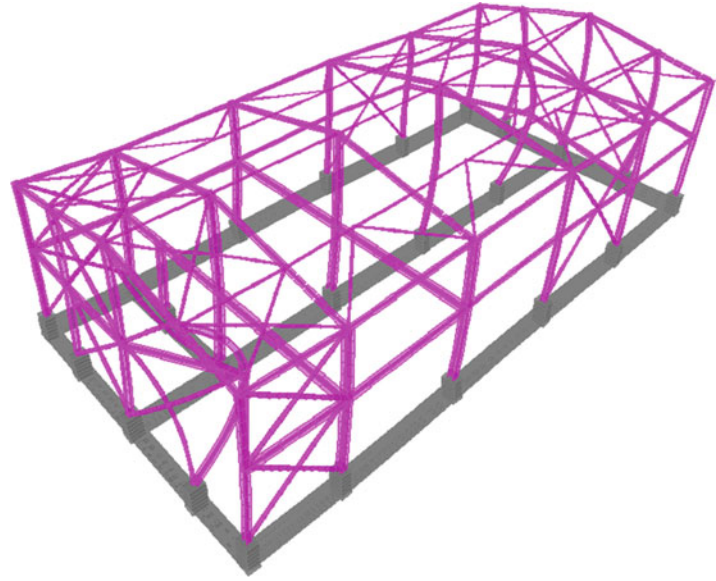
Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 34 Moment M_3 diagrams of prototype building from equivalent static method – seismic action in transverse direction
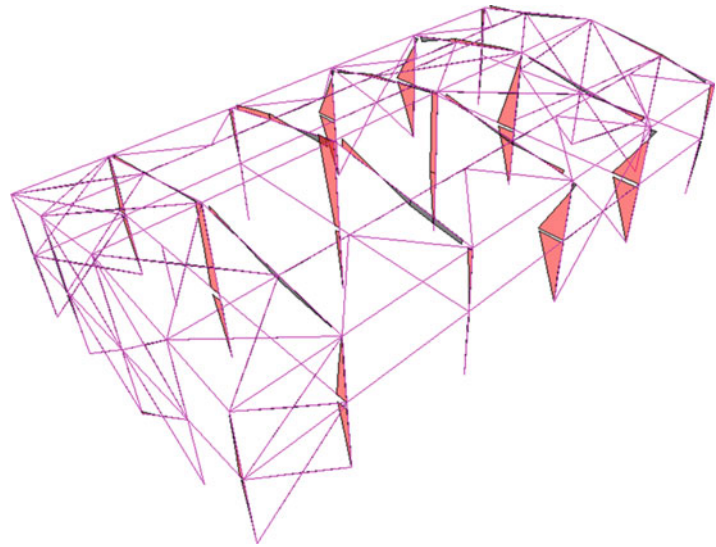
are violated by far, a modification of the structural system may be the only solution, while smaller violations may be addressed by increasing member cross sections.

Member checks (e.g., European Committee for Standardisation 2004b) consist of comparison between actions and resistances in the ultimate limit state and comparison of

maximum deflections to allowable upper bounds in the serviceability limit state. As flexural and lateral-torsional buckling are in most cases critical for steel members, it is highly recommended that the engineer reviews and modifies as needed the buckling lengths initially proposed by the software for the corresponding checks. It is also

Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 35 Envelope of moment M_3 diagrams of prototype building for ULS combinations
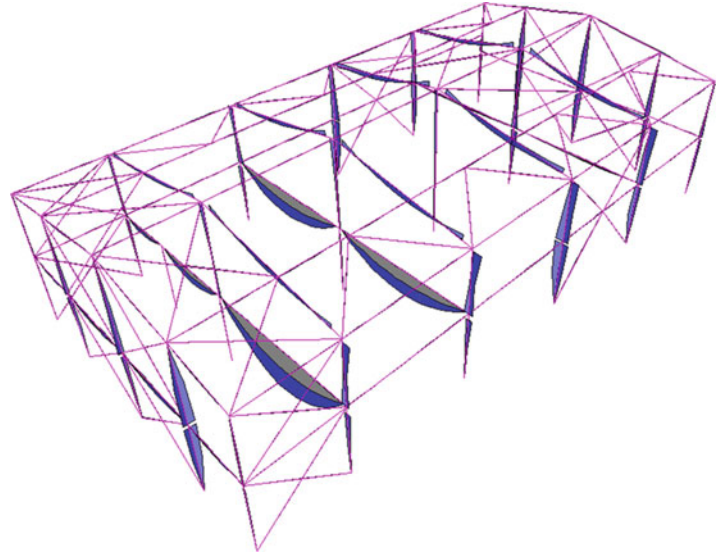
Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 36 Envelope of moment M_2 diagrams of prototype building for ULS combinations
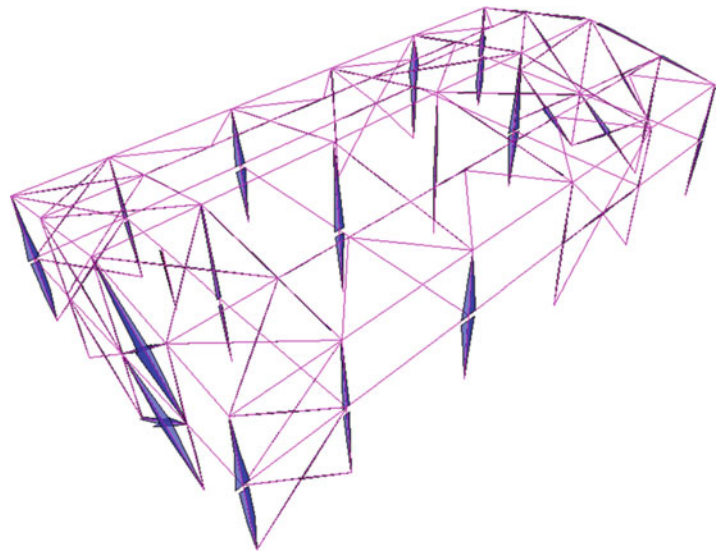
noted that most seismic design codes require capacity design checks, leading, for example, to the necessity of larger bracing sections to satisfy minimum slenderness requirements or larger column sections, to adhere to the weak beam-strong column design approach.

Connection checks (e.g., European Committee for Standardisation 2003c) include strength checks, to ensure safe transfer of internal actions between members, as well as stiffness checks, so that the behavior of actual connections (Fig. 40) is in accordance with the hinged/semirigid/rigid assumptions adopted in

Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 37 Envelope of shear V_2 diagrams of prototype building for ULS combinations
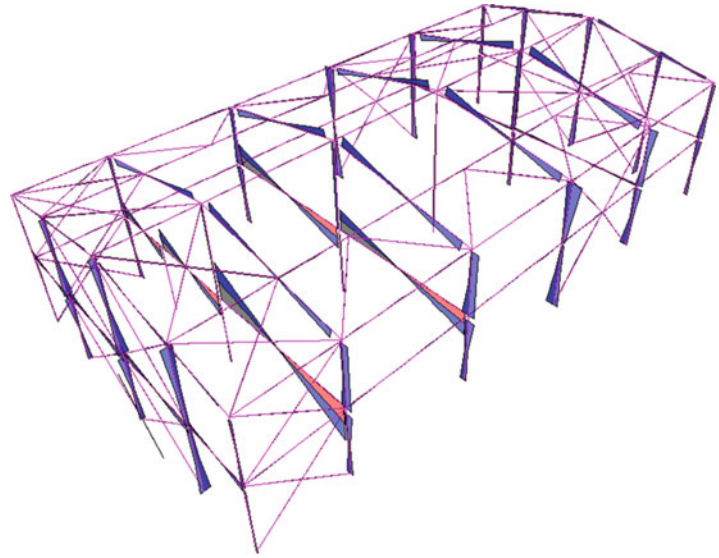
Seismic Analysis of Steel Buildings: Numerical Modeling,
Fig. 38 Envelope of axial force N diagrams of prototype building for ULS combinations
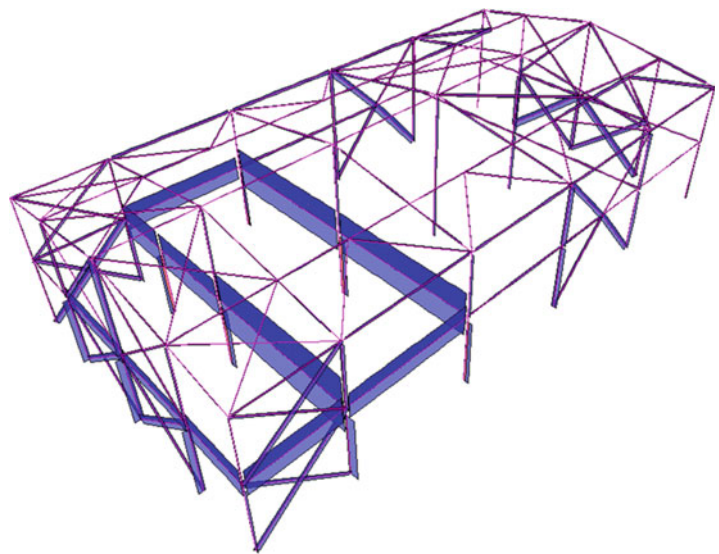
the numerical simulation. As there is so far limited experience with the behavior of semi-rigid connections under cyclic loading, it is recommended to avoid their use in seismic regions. Instead, it is proposed to stiffen the connection so that it can be classified as rigid. Strength calculation and stiffness classification of steel connections are commonly performed by means of dedicated software.

Foundation checks depend on the type of foundation (mat foundation, foundation beams, spread footings, pile foundation, etc.) and consist of general stability checks (overturning, sliding), soil bearing capacity checks, comparison of absolute and differential settlements to allowable values, and calculation of reinforcement for the reinforced concrete elements.
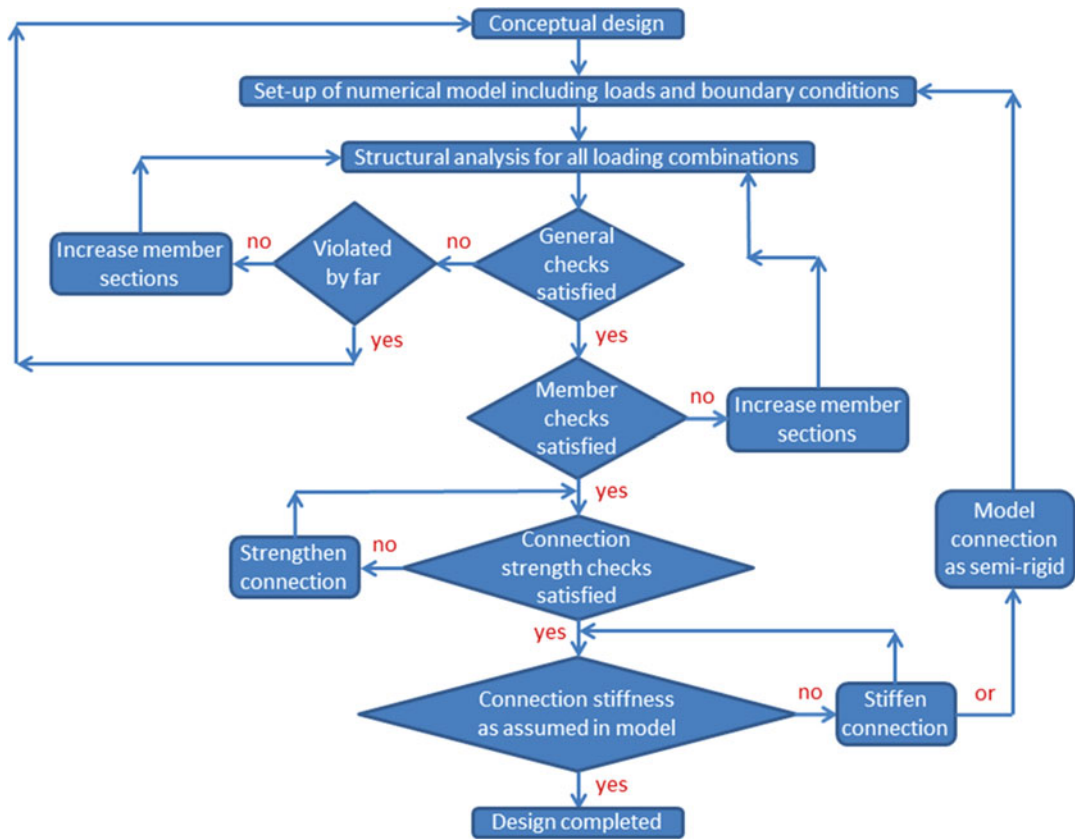
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 39 Flowchart of structural design process for steel buildings
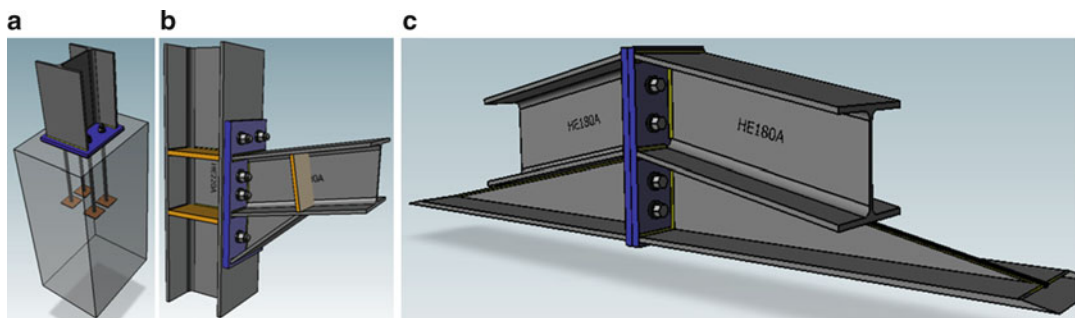
Seismic Analysis of Steel Buildings: Numerical Modeling, Fig. 40 (**a**) Column base, (**b**) girder to column, (**c**) frame ridge connections of prototype building

Summary

The state of the art in numerical modeling of steel buildings has been presented, from the point of view of the practicing structural engineer designing such structures in seismic regions. Conceptual design, numerical modeling, structural analysis, and design checks have been discussed, but emphasis has been directed toward modeling, as the other design phases are covered in detail in other parts of this encyclopedia.

Cross-References

References

European Committee for Standardisation (2001a) Eurocode – basis of structural design

European Committee for Standardisation (2001b) Eurocode 1: actions on structures – part 1–1: general actions – densities, self-weight, imposed loads for buildings

European Committee for Standardisation (2003a) Eurocode 1: actions on structures – part 1–3: general actions – snow loads

European Committee for Standardisation (2003b) Eurocode 1: actions on structures – part 1–5: general actions – thermal actions

European Committee for Standardisation (2003c) Eurocode 3: design of steel structures, part 1–8: design of joints

European Committee for Standardisation (2003d) Eurocode 8: design of structures for earthquake resistance – part 1: general rules, seismic actions and rules for buildings

European Committee for Standardisation (2004a) Eurocode 1: actions on structures – part 1–4: wind actions

European Committee for Standardisation (2004b) Eurocode 3: design of steel structures, part 1–1: general rules and rules for buildings

Gantes CJ, Fragkopoulos KA (2010) Strategy for numerical verification of steel structures at the ultimate limit state. Struct Infrastruct Eng 6(1–2):225–255

Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling

Konstantinos Daniel Tsavdaridis
School of Civil Engineering, University of Leeds, Leeds, UK

Synonyms

Concrete-encased steel sections; Connections; Constitutive models; Cyclic behavior; Damping; Frames; Hysteretic rules; Nonlinear; Numerical modeling; Panel zones; Seismic analysis; Steel–concrete composite; T-stub components

Introduction

Steel–Concrete Composite (SCC) Systems

Composite construction includes a wide range of structural systems, e.g., framed structures employing all steel–concrete composite (SCC) members and components (e.g., composite beam-to-columns and connections) and sub-assemblages of steel and/or reinforced concrete (RC) elements. Such components and elements are employed to optimize the resistance and deformation capacity (Uchida and Tohki 1997). SCC structures have been used

extensively in recent years because of benefits in combining the two construction materials. SCC structures are also known for their excellent earthquake performance owing to their high strength, high ductility, and large energy absorption. Their good structural damping properties arising from the friction between the steel–concrete interfaces make them an even more attractive alternative for seismic resistance.

Consequent effects of combining the two materials are the enhanced lateral strength and stiffness of the frame, with apparent effects of the alteration of the structural natural period of vibration and the complex local behavior of beam-to-column connections. Furthermore, SCC beams subjected to lateral loading show complex behavior due to several factors, including the slip between the concrete slab and the steel beam, the variation of longitudinal stress across the width of the slab, and the overall configuration of the numerous different types of models, while the steel and concrete parts can be subjected to different actions in every case. For the above reasons, the calculation of the seismic response of composite structures is not a straightforward task due to the interaction of local and global effects and hence the unexpected failure modes that might incur. Consequently, it is very important for the analysis of such structures to account for the local interactions (e.g., interface behavior between steel and concrete) as well as the local behavior of structural systems (e.g., beam-to-column and base-to-column response). All these factors make the analysis of SCC structures and their individual components an intriguing but challenging task.

Although experimental procedures can be performed in order to enhance the understanding of the behavior of SCC structures under earthquake loading, they are typically expensive and time-consuming and do not cover a broad range of SCC structures and elements. As a result, numerical modeling procedures have been developed and tested in order to facilitate the analysis of such structures.

Most finite element (FE) packages (e.g., ANSYS, ABAQUS, ADINA, DIANA, LS-DYNA, MIDAS FEA, etc.) rely on the use of constitutive models which emphasize on the description of post-peak material characteristics such as strain hardening and softening, tension stiffening, shear-retention ability, etc. (Cotsovos and Kotsovos 2011). The derivation of such constitutive models has been based on a variety of theories and their combination. However, the application of FE packages in practical structural analysis has shown that the constitutive relationships are case-study dependent, since the solutions obtained are realistic only for particular problems. Therefore, the applicability of packages to a different set of problems requires modifications of the constitutive relationships. This is entirely dependent on the interpretation of the observed material behavior as well as the use of the experimental data to validate the constitutive relationships.

To this end, the aim of the present chapter is to provide an indication of the concepts which are widely used for modeling the steel–concrete composite behavior and to develop numerical guidelines for the nonlinear analysis of such structures and their components, considering the seismic actions under earthquake events. The numerical analyses presented herein model the behavior of SCC structures/components using macro-models (i.e., the use of line elements and spring connections) rather than micro-models (continuum FE models) due to their simplicity and accuracy in nonlinear analysis. Different aspects of modeling including the geometry, material nonlinearity (through the constitutive laws adopted), hysteretic behavior, and geometrical nonlinearity as well as other parameters important for seismic analysis are also presented in this chapter.

Chapter Synthesis

The modeling of SCC elements and frames is based on three approaches. The first one is the simplified modeling approach presented in this chapter, engaging the use of springs and line elements for the elementary simulation of the behavior of each component to the entire frame assembly. The scope is to initiate numerical guidelines based on the simplified approach and to present modeling examples of SCC beam cross sections, flooring systems, fully composite
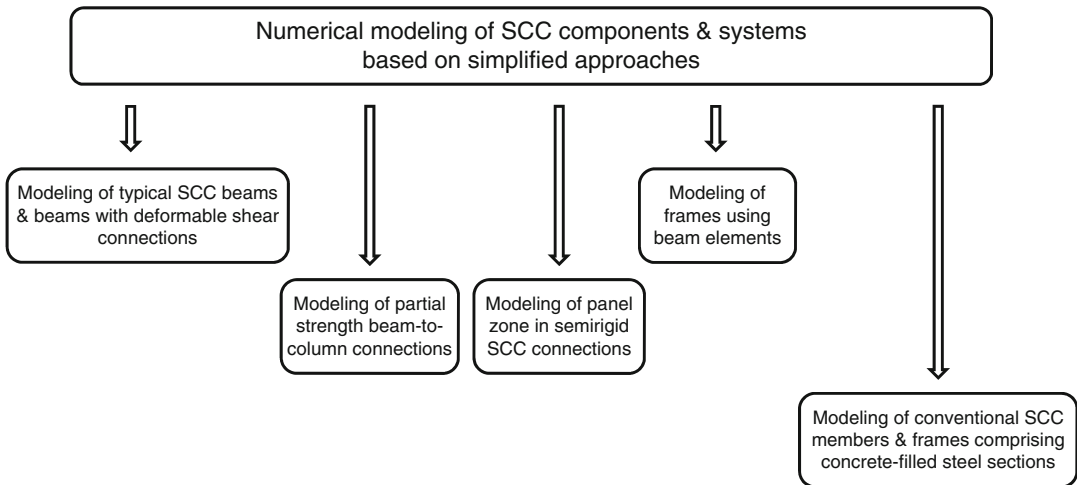
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Chart 1 Structure of the Chapter

members, beam-to-column connections, as well as holistic approaches modeling a frame. The breakdown of this chapter is given in Chart 1.

Based on the bottom-up approach (i.e., combining different structural components starting from the most fundamental of a system and giving rise to grander systems), the engineer will gain a decent understanding on the parameters to be considered during the computational modeling procedure. Modeling of these structural components will enable the computation of their response to different load histories and moreover will enable the engineer to carry out the state determination of a member from a frame assembly.

Requirements for Collapse Analysis of Composite Structures

Numerical modeling procedures should aim to address a number of issues regarding the local, intermediate, and global level of SCC structural design. On the *local* level, aspects such as the cyclic behavior of the steel and the concrete members (including the softening and hardening of the material), the local buckling of steel flanges, the load carrying capacity, the curvature ductility of the components, as well as the effects of

confinement should be carefully studied. On the *intermediate* level, the ductility of the member in terms of rotation/displacement should be established. Additionally, second-order effects (P-Δ) on forces and deformations should be taken into account through the provisions for large displacement analysis. Modeling the beam-to-column connection is also essential when the fully rigid assumption is not suitable. On the *global* level, the overall ductility and strength of the structure should be established through force–displacement relations. The progressive yielding and the hinge formation at the structural frame should also be established through moment–rotation relationships. The complete list of requirements for the collapse analysis and the modeling of SCC structures subjected to earthquake actions is presented below:

- Stress–strain relationships for the steel material including strain hardening and softening
- Provision for the effects of local buckling in the steel section
- Stress–strain relationships for the concrete material including cycling loading regimes and the effect of the confinement on the peak stress and corresponding strain

- Explicit representation of the slip boundary conditions of the shear connection both at local and global levels
- Provision for second-order effects on forces and deformations
- Effective beam-to-column connection models, including panel distortion
- Iterative and advanced dynamic analysis techniques for analyzing the structural response near collapse state

Modeling of Steel–Concrete Composite (SCC) Beams

A variety of different models have been developed by researchers in order to capture the behavior of SCC beams, based on either concentrated or distributed plasticity. In concentrated plasticity models, all the inelasticity is concentrated at the ends of the member; therefore, it deals with material nonlinearity in an approximate but efficient manner. On the contrary, distributed plasticity models simulate the inelastic behavior along the length of the member. This approach is more accurate but at the same time is more computationally demanding. Most of the formulations for both approaches are rather complex and not amenable to generic and routine application in structural engineering design.

The present subchapter presents a simplified (new) modeling approach based on the work of Zhao et al. (2012) for the nonlinear analysis of SCC beams and composite frames with deformable shear connections (based on the distributed plasticity approach) using line elements to simulate the structural beam and column members, layered fiber section to simulate the reinforced concrete slabs, and nonlinear spring elements for the simulation of the interface between the structural steel beams and the reinforced concrete slab. Vertical interactions between the slab and steel beams are not expected to be significant, therefore are not accounted into the analysis. The geometry of the model, along with a simple set of details, is outlined below. The assembled model is shown in Fig. 1.

Model Geometry of a Typical SCC Beam

To model the geometry of the macromodel for a typical SCC beam, the following assemblies should be utilized:

1. Four-node-layered shell elements representing the concrete slab. Reinforcement layers comprising steel material properties should be used to simulate the steel reinforcement located at the top and bottom of the concrete slab.
2. Two-node fiber beam-to-column elements for modeling the steel beam. The reference surface of the slab will be located at the centroid of the steel beam cross section.
3. Dummy nodes at the same locations as the beam-to-column element nodes simulating the connection between the nodes of the steel beam and the shell elements.
4. Rigid beam elements connecting the dummy nodes and the corresponding ones of the shell elements, located on the same x- and z-coordinates.
5. Discrete spring elements with only translation in the z-direction connecting both the dummy nodes and the beam-to-column element nodes in order to control the interface shear–slip surface along the length of the beam.

Model Geometry of Beam with Deformable Shear Connection

Modeling of two-dimensional beams with deformable shear connection is based on the Newmark et al. (1951) model, in which (i) the Euler–Bernoulli beam theory applies to both components of the SCC beam and (ii) the deformable shear connection is represented by an interface model with distributed bond allowing interlayer slip as well as enforcing contact between the steel and concrete components.

A local coordinate system should be established to enhance the understanding of kinematics of Newmark's model. With reference to Fig. 2, Z axis is parallel to the beam axis and the vertical plane YZ is the plane of geometrical and

S
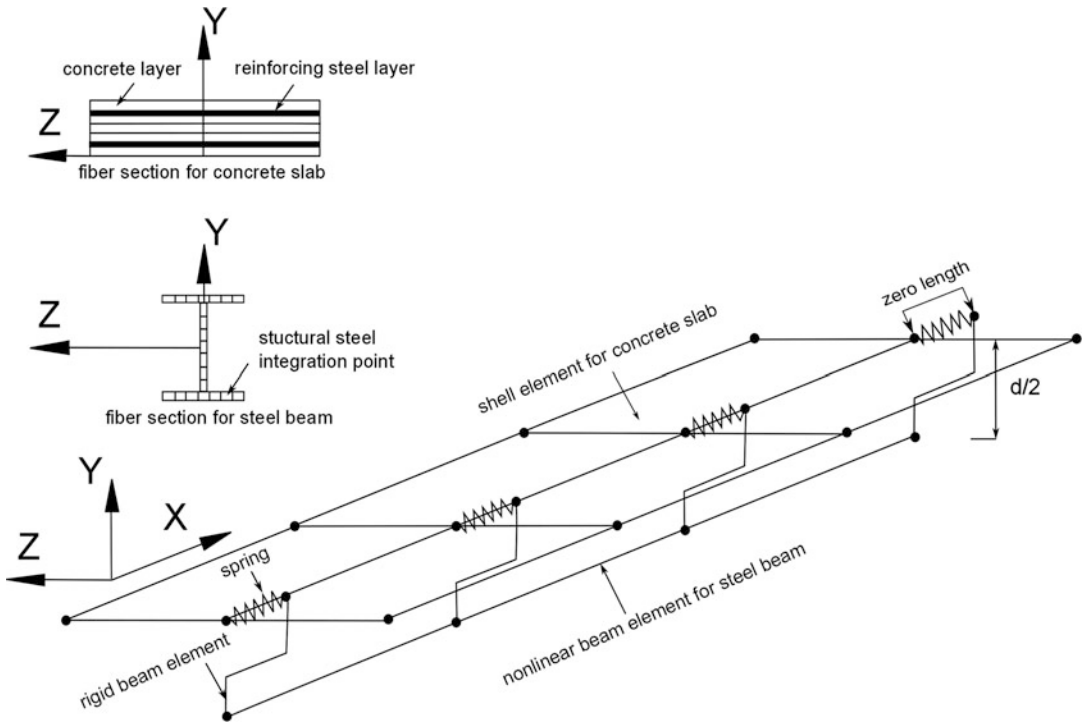
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 1 Assembled macromodel representing a typical steel-concrete composite beam

material symmetry of the cross section. Loads are also assumed symmetric with respect to the YZ plane. The displacement field u of a material point of the beam is given by:

$$u(y,z) = v(z)j$$
$$+ \left[w_a + (y_a - y)v'(z)\right]k \quad \text{on } A_a(a = 1, 2)$$
$$(1)$$

where w_a is the axial displacement of the reference point of domain A_a, the ordinate of which is $y_a(a = 1 : \text{concrete slab}, a = 2 : \text{steel beam})$; v is the vertical displacement of the cross section; and j and k denote the unit vectors along the Y and Z axes, respectively. It is observed that the transverse displacements and rotations of the slab and of the steel beam are equal due to the enforced contact between the two components. The only nonzero strain components are the axial strain ε_{za} and the interface slip δ:

$$\varepsilon_{za}(y, z) = w'_a(z)$$
$$+ (y_a - y)v''(z) \quad \text{on } A_a(a = 1, 2)$$
$$(2)$$

$$\delta(z) = w_2(z) - w_1(z) + hv'(z) \qquad (3)$$

where $h = y_2 - y_1$ is the distance between the reference points (G_1 and G_2 in Fig. 2) of the two components. At the locations of the longitudinal reinforcement, Eq. 2 also provides the strain in the reinforcement, due to the assumption of perfect bond between the steel and the concrete.

FE Formulations

A simple two-dimensional 10 degree-of-freedom (DOF) SCC frame element with deformable shear connection is presented herein, similarly to Zona et al. (2008). With reference to Fig. 3, 8 of the 10DOFs are external (4 DOFs per end node) allowing for the axial displacement, the transverse displacement, and the rotation of the steel beam and 1DOF for the axial displacement
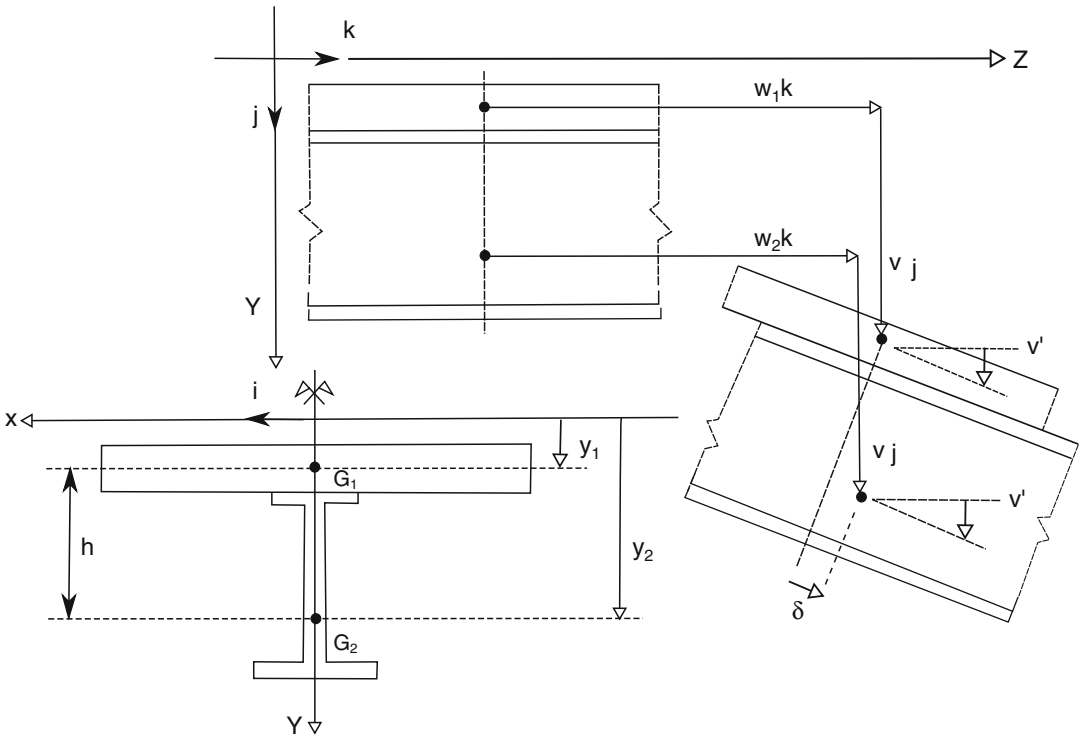
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 2 Kinematics of two-dimensional composite beam model and reference system

Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 3 10DOF SCC beam element used (**a**) references defined at the beam and slab centroids and (**b**) references defined at the slab–beam interface
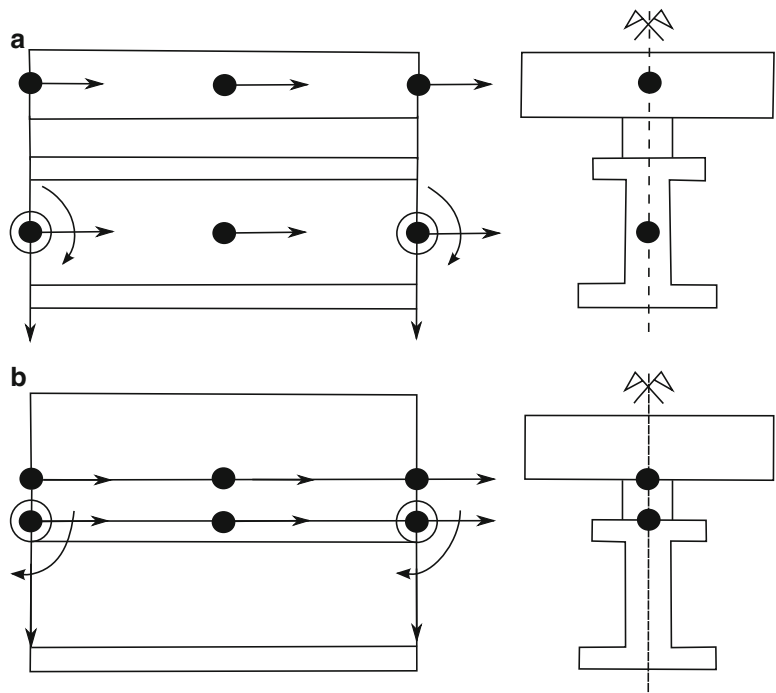
of the concrete slab. The remaining internal 2DOFs allow for axial displacement of the steel beam and the concrete slab (Fig. 3b).

Modeling of Inertia and Damping Properties

Modeling the inertia properties of the frame elements can be achieved using lumped masses at the DOFs of the external nodes. Consequently, the inertia properties of the FE model are independent of the type of finite elements employed (i.e., the structure's mass matrix can be obtained using force, displacement, or mixed-based formulation frame elements).

Even though the friction between steel beams and concrete slabs in SCC frames may be a strong source of structural damping, quantitative information about this energy dissipating mechanism usually referred to as structural damping is limited owing to the partial availability of experimental dynamic data. Consequently, the well-known and widely used Rayleigh damping model can be used by the practicing engineer. In this model, the damping matrix can be obtained using the classical Rayleigh damping relationship (Eq. 4), where the damping matrix is proportional to the mass matrix and the initial stiffness matrix:

$$[\mathrm{C}] = \mu[\mathrm{M}] + \lambda[\mathrm{K}] \qquad (4)$$

where

μ = mass proportional Rayleigh damping coefficient
λ = stiffness proportional Rayleigh damping coefficient
M = system structural mass matrix
K = system structural stiffness matrix

Note: The proposed model presented in the above sections considers only rigid beam-to-column connections. Nevertheless, semirigid connections can be considered in the same numerical procedure by introducing special joint elements with prescribed constitutive behavior.

Constitutive Stress–Strain Relationships

For modeling purposes, the material properties (such as the Young's modulus, Poisson's ratio, elastic and plastic strength, and strain hardening) can be obtained from the uniaxial stress–strain curves derived from coupon tests and then applied to the corresponding fibers across the composite cross section. In order to accurately simulate the behavior of SCC beams under earthquake conditions, robust material models capable of simulating the material nonlinearity as well as other damaging effects under dynamic or cyclic loading (i.e., softening/hardening) need to be employed. Several models have been developed to achieve the aforementioned scope, some of which are presented in the following sections.

Constitutive Law for Concrete Parts (Based on the Kent–Park Model)

The proposed constitutive law modeling the concrete in monotonic compression for the cases of confined and unconfined concrete is the Kent–Park model as described in Park and Paulay (1975). As it is shown in Fig. 4, the material follows a parabolic stress–strain curve up to a maximum stress equal to the cylinder's strength, after which it decays linearly with strain until the residual strength is reached. In tension, the model assumes a linear stress–strain behavior up until the tensile limit of the material is reached, and then the stiffness and strength decays with increasing strain (Fig. 5).

The cyclic behavior of the concrete can be described by the Blakeley–Park model also presented in Park and Paulay (1975). The stress–strain response lies within the Kent–Park envelope; however, the effect of concrete confinement is not taken into account. The model assumes that unloading and reloading takes place along a line without energy dissipation or stiffness deterioration for strains smaller or equal to the strain corresponding to peak stress ($\varepsilon < \varepsilon_c$). Beyond this point, the stiffness deterioration is taken into account through the introduction of reduction factors, given by Blakeley and Park. Along the first unloading branch, the stress is reduced approximately 50 % without any reduction in strain. The reloading branch with slope equal to $f_c E$ extends back to the envelope (Fig. 5).
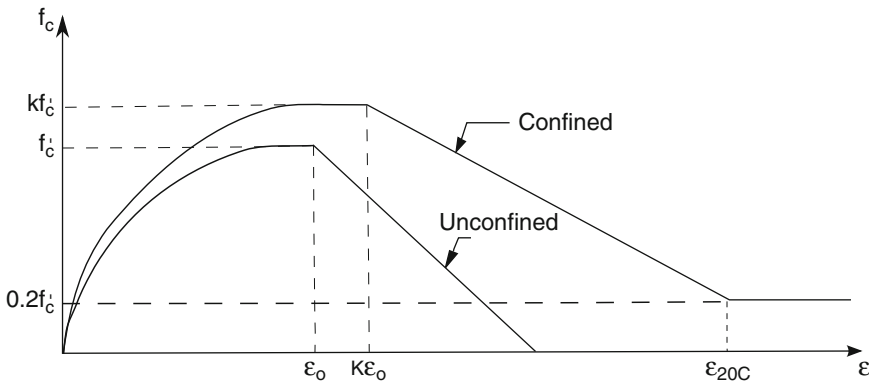
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 4 Monotonic stress–strain Kent and Park model for unconfined and confined concrete

Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 5 Cyclic stress–strain Blakeley–Park model for concrete
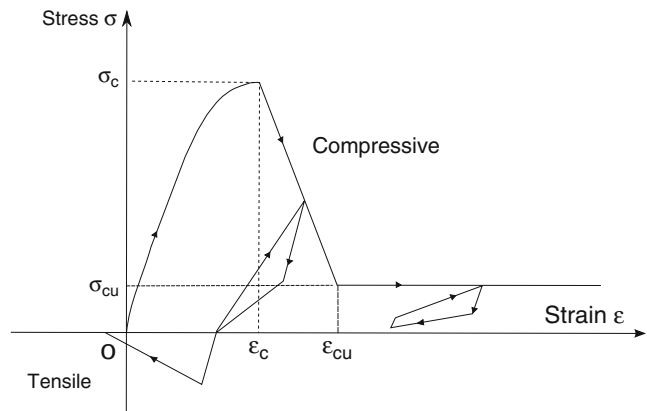
Constitutive Law for Concrete (Based on the Popovics–Saenz Law)

The constitutive law for concrete is a uniaxial cyclic law with monotonic envelope defined by the Popovics–Saenz law (Balan et al. 1997). Linear unloading–reloading branches with progressively degrading stiffness characterize the cyclic behavior of the material. The response of concrete under cyclic loading is shown in Fig. 6. According to the same figure, after each unloading–reloading, the monotonic envelope is reached again when the absolute value of the largest compressive strain attained so far is surpassed. The tensile behavior of concrete is characterized by the same loading–unloading–reloading rules with the same initial stiffness and appropriate values for the other parameters.

Constitutive Law for Steel

Figure 7a, b describes the elastoplastic response of the steel under monotonic and cyclic loading respectively. For monotonic loading, the characteristic yield plateau in the stress–strain model followed by a region of increased strength owing to strain hardening of the material. The unloading from the yielded condition is elastic; thereafter, the Bauschinger effect can be represented by a Ramberg–Osgood relationship (Eq. 5) until the yield stress is reached. This model uses a single nonlinear equation to characterize the observed curvilinear response of steel subjected to monotonic loading:

$$\varepsilon = \frac{\sigma}{E} + K\left(\frac{\sigma}{E}\right)^{\eta} \qquad (5)$$

S

Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 6 Hysteretic concrete material model under compression
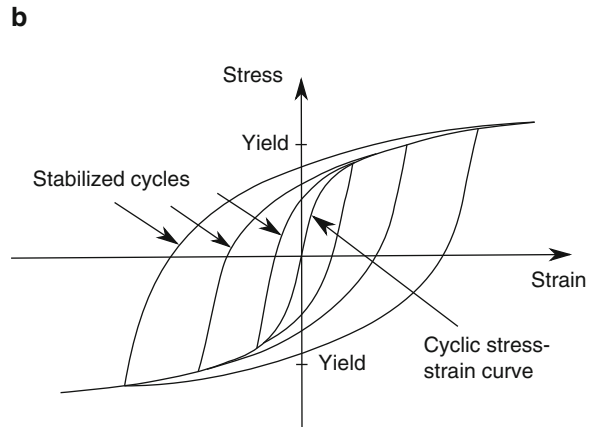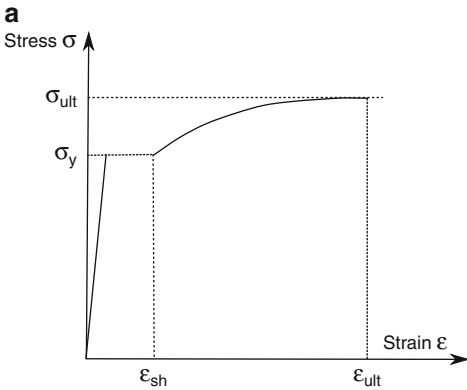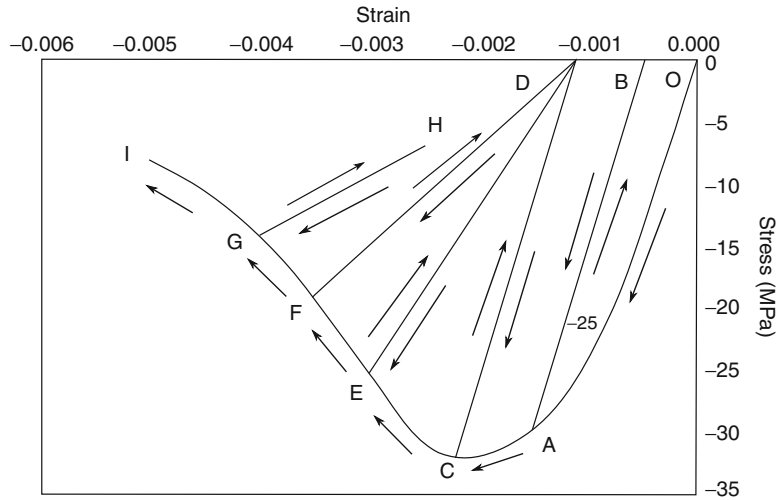
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 7 (**a**) Stress–strain model under monotonic loading. (**b**) Steady-state cyclic response of mild steel

where

$\frac{\sigma}{E}$ = equal to the elastic part of the strain

$K\left(\frac{\sigma}{E}\right)^{\eta}$ = accounts for the plastic part of the strain

K and η = parameters that describe the hardening behavior of the material

When the steel material is subjected to constant strain amplitude under cyclic loading, it exhibits a response that converges to a stabilized saturation loop which depends only on the cycling amplitude (Fig. 7b). As it is shown in Fig. 8, the response of the steel material under constant strain amplitude cycles is described by strain hardening for large amplitudes and strain softening for small amplitudes. For the accurate simulation of the steel material response under an arbitrary load, the constitutive model needs to account for all the monotonic response, the steady-state cyclic behavior, as well as the transient behavior involving softening and hardening. This can be achieved using an efficient simplistic computationally bilinear model.

Bilinear Stress–Strain Steel Model In this bilinear model, the elastic range remains

Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 8 Cyclic response of steel under constant strain amplitude cycles

Softening under small strain amplitudes

Hardening under large strain amplitudes
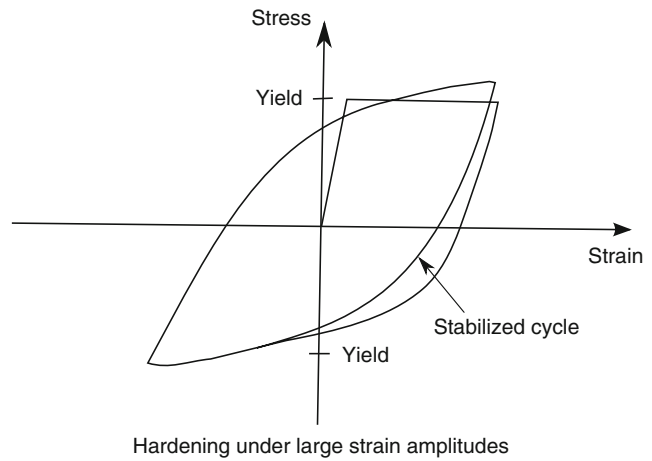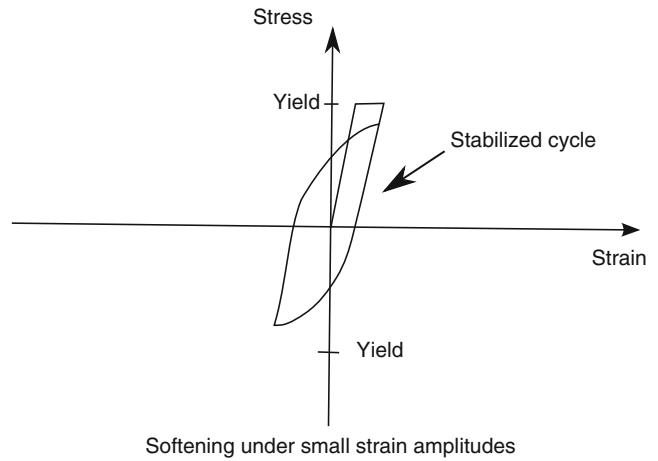
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 9 Loading and unloading paths of bilinear kinematic model
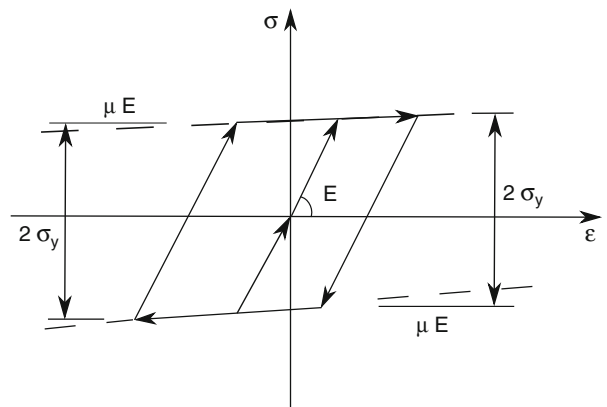
constant throughout the various loading stages. The kinematic hardening rule for the yield surface is assumed to be a linear function of the increment of plastic strain (Fig. 9). The calculation of the current stress state is expressed mathematically using Eqs. 6, 7, and 8, and it is presented graphically in Fig. 10:
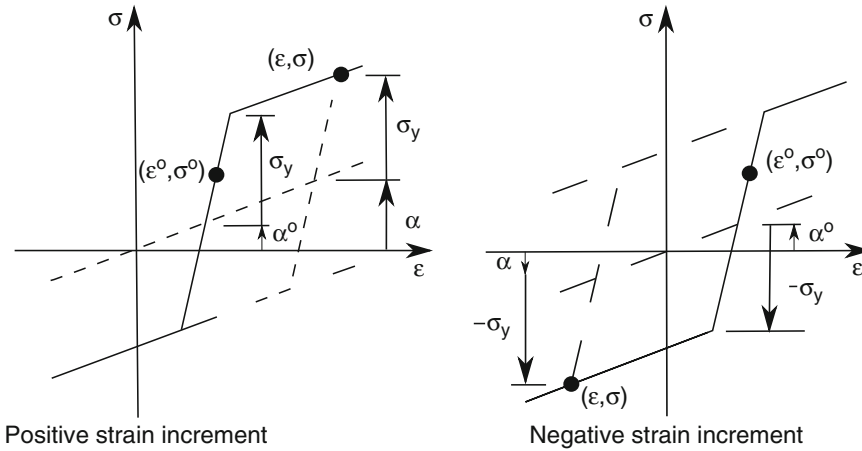
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 10 Stress determination with the bilinear kinematic model

$$\left.\begin{array}{l} \alpha = \alpha^{o} \\ \sigma = \sigma^{o} + E(\varepsilon - \varepsilon^{o}) \end{array}\right\}$$

$$\text{If}\left(\frac{\alpha - \sigma_{y} - \sigma^{o}}{E}\right) \le (\varepsilon - \varepsilon^{o}) \le \left(\frac{\alpha^{o} + \sigma_{y} - \sigma^{o}}{E}\right) \quad (6)$$

$$\left.\begin{array}{l} \alpha = \alpha^{o} + \mu E\left(\varepsilon - \varepsilon^{o} - \dfrac{\alpha^{o} + \sigma_{y} - \sigma^{o}}{E}\right) \\ \sigma = \alpha + \sigma_{y} \end{array}\right\}$$

$$\text{If} \quad (\varepsilon - \varepsilon^{o}) > \left(\frac{\alpha^{o} + \sigma_{y} - \sigma^{o}}{E}\right)$$

$$(7)$$

$$\left.\begin{array}{l} \alpha = \alpha^{o} + \mu E\left(\varepsilon - \varepsilon^{o} - \dfrac{\alpha^{o} - \sigma_{y} - \sigma^{o}}{E}\right) \\ \sigma = \alpha - \sigma_{y} \end{array}\right\}$$

$$\text{If} \quad (\varepsilon - \varepsilon^{o}) < \left(\frac{\alpha^{o} - \sigma_{y} - \sigma^{o}}{E}\right)$$

$$(8)$$

where according to Figs. 9 and 10:

E = Young's modulus
μ = strain hardening parameter
σ_{y} = initial yield surface
ε = current strain
σ = current stress
α = current center of elastic range
Subscript "0" denotes values at the start of an increment

Constitutive Law for Steel (Based on the Menegotto–Pinto Model)

The constitutive law describing the behavior of the steel material is the uniaxial Menegotto–Pinto model (1973). This computationally efficient nonlinear law is capable to model both kinematic and isotropic hardening as well as the Bauschinger effect, allowing for accurate simulation and reproduction of experimental results. The response of the steel material is defined by the following nonlinear equation:

$$\sigma = b\,\varepsilon + \frac{(1 - b)\varepsilon}{(1 + \varepsilon^{R})^{\frac{1}{R}}} \quad (9)$$

where the effective strain and stress (ε, σ) are a function of the unload–reload interval, b is the ratio of the initial to final tangent stiffness, and R defines the shape of the unloading–reloading curves. Figure 11 presents a typical stress–strain response based on the Menegotto–Pinto model. The model assumes a symmetric response for loading in compression and tension.

Interaction of Material Surfaces: Evaluation of Spring Properties

The degree of composite action and interaction between the steel beam and the concrete slab is a fundamental mechanism that needs to be

Fig. 11 Menegotto–Pinto material constitutive model for structural steel; typical cyclic stress–strain response
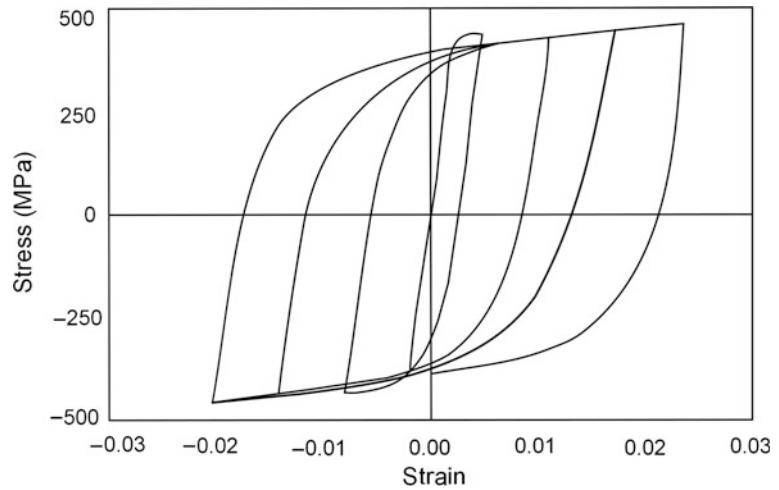
considered by the engineer during the modeling procedure of SCC beams and structural systems owing to the implications on the serviceability and ultimate limit states, the energy dissipation under cyclic loading, and the local stress distribution.

Two different modeling approaches can be used for the description of partial bond in SCC structural systems. The concentrated bond approach is based on the use of concentrated springs for the modeling of the connection. The springs are attached at the location of each connector, modeling either the action of the shear stud connectors between the steel and the concrete slab or the friction in concrete-filled hollow sections and partially encased steel sections. The second approach is based on the distributed bond model, which assumes a continuous bond stress and bond slip along the contact surface. For both approaches, the uplift is typically neglected; therefore, it is considered that the concrete slab and the steel beam have the same vertical displacement and curvature.

Shear–slip relationships are widely available providing information regarding the behavior of the connectors. Figure 12a presents a simplified bilinear shear–slip relationship based on a widely used shear–slip model proposed by Ollgaard et al. (1971). The Ollgaard model is described by the following exponential function (Eq. 10) representing an experimentally observed large reduction of stiffness with increasing slip:

$$\frac{N_v}{N_{vu}} = (1 - e^{-ns})^m \qquad (10)$$

where

N_{vu} = connection (ultimate) strength

N_v = shear load

s = slip between two components of the composite beam (interface slip)

n, m = empirical parameters defining the shape of the curve calibrated from experimental data

In Fig. 12b, the monotonic envelope is presented by the definition of an ultimate slip, S_{ult}. When ultimate slip, the shear force–slip behavior follows zero stiffness with constant shear force $N_v = \pm \tau_{fr}$, where τ_{fr} is the residual shear force.

Modeling of Steel–Concrete Composite Beam-To-Column Partial-Strength Semirigid Connections

Compared to traditional bare steel structures, SCC frames can achieve more effective beam-to-column connections through the contribution of the concrete slab in resisting bending moments under gravitational and lateral loads. Additionally, these structures comprising partial-strength partially restrained beam-to-column joints designed in such a way to exhibit ductile seismic

**Seismic Analysis of
Steel–Concrete
Composite Buildings:
Numerical Modeling,
Fig. 12** Shear–slip
relationships

Equivalent bi-linear model for shear-slip of each connector

Cyclic response with slip
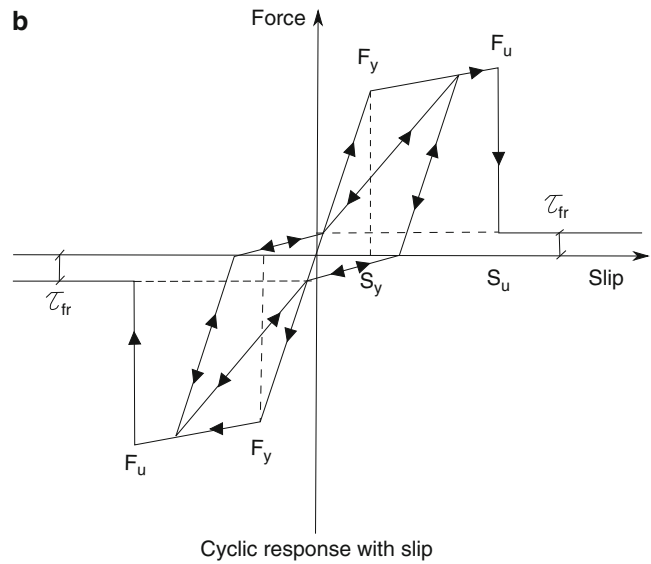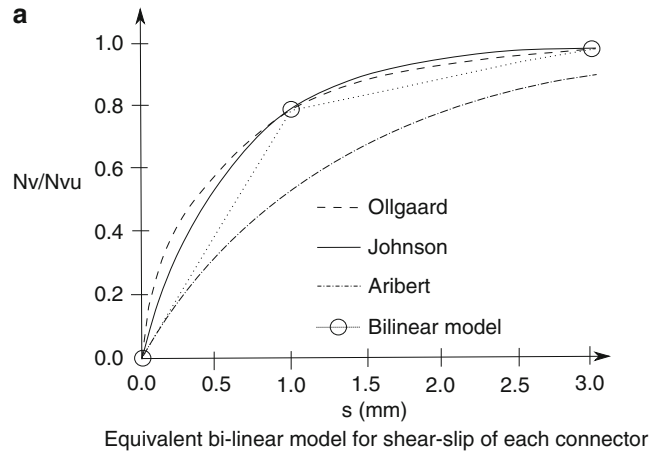
response through the plastic deformation of their components can achieve the formation of a desirable beam hinging global frame mechanism, with large hysteretic energy dissipation capacity and reduced force demand on the columns.

This section of the chapter presents a simplified approach based on the work done by Braconi et al. (2007) for the nonlinear analysis of a partial-strength beam-to-column connection using a component model. The behavior of partial-strength beam-to-column connections under the application of seismic load is described using

nonlinear spring elements as shown in Fig. 13. With reference to the same figure, the elements comprising the model should account for the response of the:

1. Concrete in compression
2. Column web panel in shear
3. Upper T-stub in compression (+Ve moment)
4. Lower T-stub in tension (+Ve moment)
5. Concrete slab in tension
6. The shear studs (+Ve moment)
7. Upper T-stub in tension (−Ve moment)
8. Upper T-stub in compression (−Ve moment)
9. Shear studs (−Ve moment)
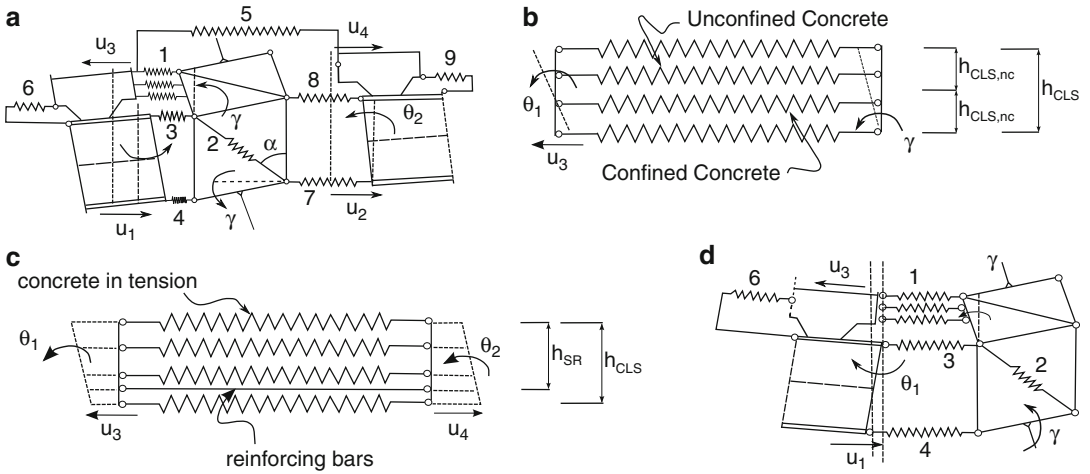
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 13 Component model: (**a**) overall joint model, (**b**) kinematics of the concrete slab in compression (component 1), (**c**) kinematics of the concrete slab in tension (component 5), and (**d**) overall exterior joint model

Kinematics

Equilibrium must be maintained between the force acting in the nine components and the internal and external forces. The response of the assembled model is therefore defined by a set of eight equations related to translational equilibrium between components in the same beam-to-column connection, equilibrium between the shear studs and the beam steel profile, rotational equilibrium between the internal forces in the beam-to-column connection and the bending moment of the beam framing in it, as well as rotational and translation equilibrium acting on the column web panel. The set of eight equations can be then solved using a numerical procedure (i.e., Newton–Raphson) considering the storey deformation (drift, δ) as such an external action in the format of imposed deformations.

On the basis of the small displacement theory, the local kinematics could be described using a total of 7DOFs: the horizontal displacement of the bottom surface of the slab on both sides of the column, u_1 and u_2; the horizontal displacement of the top surface of the slab on both sides of the column, u_3 and u_4; the relative rotation of the two beams with respect to the column faces, θ_1 and θ_2; and the column panel zone shear distortion, γ. The deformations, δ, of the nine components comprising the model are linked to the seven degrees of freedom through the following equations:

$$\delta_{1,i}^{nc} = -u_3 - \frac{\Delta_{nc}}{2} * (2i - 1) \tan(\theta_1)$$
$$+ \left[h_{cs} - \frac{\Delta_{nc}}{2}(2i - 1) \right] \tan(\gamma) \quad (11)$$

$$\delta_{1,i}^{c} = -u_3 - \left[\frac{\Delta_c}{2}(2i - 1) + h_{cls,nc} \right] \tan(\theta_1)$$
$$+ \left[h_{cls} - \frac{\Delta_c}{2}(2i - 1) \right] \tan(\gamma) \quad (12)$$

$$\delta_3 = -u_1 - \left(h_{cs} + \frac{t_{bf}}{2} \right) \tan(\theta_1) \quad (13)$$

$$\delta_4 = -u_1 - \left(h_{cs} + h_b - \frac{t_{bf}}{2} \right) \tan(\theta_1) \quad (14)$$

$$\delta_{5,i} = u_4 - \left[h_{cs} - \frac{\Delta_{ct}}{2}(2i - 1) \right] \tan(\theta_2)$$
$$- u_1 - \left[\frac{\Delta_{ct}}{2}(2i - 1) \right] \tan(\theta_1) \quad (15)$$

$$\delta_{5,\,\mathrm{SR}} = u_4 - \left(h_{cs} + h_b - \frac{t_{bf}}{2} - h_{SR}\right)\, \tan(\theta_2)$$
$$- u_1 - h_{SR}\, \tan(\theta_1)$$
$$(16)$$

$$\delta_6 = u_1 - u_3 \qquad (17)$$

$$\delta_7 = u_2 \qquad (18)$$

$$\delta_8 = u_2 - (h_b - t_{bf})\, \tan(\theta_2) \qquad (19)$$

$$\delta_9 = u_2 - u_4 \qquad (20)$$

In the above equations, h_b is the beam depth and h_{cs} is the overall slab thickness including the depth of the steel deck, but the model assumes that the interaction between the slab and the column only occurs over the slab thickness above the steel deck, h_{cls}, to represent actual test conditions. The parameter, t_{bf}, is the beam flange thickness. For components 1 and 5, the deformation varies among the series of n parallel springs used over the slab thickness. For the concrete in compression (Component 1), a distinction is also made between unconfined ($\delta_{1,i}^{nc}$) and confined ($\delta_{1,i}^{c}$) concrete fibers. Unconfined concrete condition is assumed above the slab reinforcing steel (of thickness $h_{cls,nc}$ in Fig. 13b), and therefore, the parameters Δ_c and Δ_{nc} correspond to the

thickness of each confined and unconfined concrete layers, respectively. Similarly, Δct is the thickness of the concrete layers in tension. Figure 14 presents the deformation of the entire sub-assemblage. The storey drift, δ, is obtained through Eq. 21:

$$\delta = \gamma h_{SP} + \varphi h_{tot} + \delta_{el,\,column} \qquad (21)$$

where the first term represents the storey drift owing to the web panel over the height of the joint, h_{SP}, the second term corresponds to the rotation due to flexural deformations of the beam at beam mid-depth, and the final term corresponds to the elastic flexural deformation of the column. Figure 14 shows the global kinematics of the model.

Component Modeling
Necessary work preceding the assembling of a SCC connection model is the verification of each individual component model. In a semirigid SCC connection, the major components to be considered are outlined below.

Concrete Slab
The nonuniform stress distribution over the slab thickness is taken into account in the model with the use of fibers as shown in Fig. 13b, c for the cases of compression and tension, respectively.

Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 15 Component constitutive relationships: (**a**) concrete in compression, (**b**) concrete in tension, (**c**) reinforcing steel, (**d**) shear stud, (**e**) panel zone in shear, and (**f**) lower T-stub in tension

Concrete slab in compression can be modeled using stress–strain relationships proposed in EC2 or any of the stress–strain relationships presented in sections "Constitutive Law for Concrete Parts (Based on the Kent–Park Model)" and "Constitutive Law for Concrete (Based on the Popovics–Saenz Law)" for both unconfined (concrete above reinforcement level) and confined (concrete below reinforcement level) conditions. A typical example of the stress–strain relationship that can be used in analysis is presented in Fig. 15a for both concrete

conditions. In tension (Fig. 15b), the behavior of concrete can be represented using a linear response until cracking, followed by a softening branch where the tensile resistance reduces exponentially as proposed by Stevens et al. (1991). Linear unloading–reloading branches can be adopted intersecting the deformation axis at a residual plastic deformation ε_{ct}^{pl} given by Eq. 22 and also shown in Fig. 15b:

$$\varepsilon_{ct}^{PL} = 146 \left(\varepsilon_{ct}^{max}\right)^2 + 0.523 \left(\varepsilon_{ct}^{max}\right) \qquad (22)$$

S

Steel Reinforcement

Bilinear stress–strain relationships as shown on Fig. 15c can be adopted for representing the behavior of steel reinforcement.

Shear Connectors

The slip between the concrete slab and the beam owing to the flexibility of the shear stud connectors can be modeled using force–deformation relationships similar to those presented earlier in section "Interaction of Material Surfaces: Evaluation of Spring Properties" ("simulation of composite action"). The recommended force–deformation relationship for this particular model is based on the model proposed by Aribert and Lachal (2000) (and is tailored to the guidelines of EC4 and EC8 for the calculation of the ultimate shear stud resistance), and it is presented in Fig. 15d and Eq. 23.

$$F = Q_u \left(1 - e^{C_1(u_3 - u_4)}\right)^{C_2} \qquad (23)$$

where

Q_u = ultimate shear stud resistance calculated according to EC4 and EC8
C_1, C_2 = coefficients suggested by Aribert and Al Bitar (1989) depending on the height of studs and type of steel profile

Panel Zone in Shear

The behavior of the panel zone plays a significant role in determining the overall stiffness and capacity of the frame. In terms of seismic design, the panel zone can have a significant influence on the distribution of plasticity and energy dissipation on the overall performance of the structure. A multi-linear shear force to shear deformation response is retained for the web panel zone. The following equations describe this multi-linear relationship through the gradients of elastic, post-elastic, and strain hardening stiffness:

$$K_{el,wp} = G_S \frac{A_{vc}}{z} \qquad (23)$$

$$K_{t,wp} = 1.04 \, G_S \frac{b_{cf} \, t_{cf}^2}{z} \qquad (24)$$

$$K_{s,wp} = \frac{G_S}{a_H} \frac{(h_c - t_{cf}) t_{cw}}{z} \qquad (25)$$

where

G_S = shear modulus of steel
A_{vc} = shear area of the column section
z = centerline vertical distance between the column stiffeners
b_{cf} = column flange width
t_{cf} = column flange thickness
h_c = column depth
t_{cw} = column web thickness
a_H = hardening coefficient related to the thickness of the column web panel and the column flanges

Using the elastic and post-elastic stiffness, shear forces in the panel zone can be obtained. The elastic stiffness, $K_{el,wp}$, is obtained from EC3 and is applicable until the shear force, V_{wp}, reaches the yield limit, $V_{el,w}$, which is also specified in EC3 (Eq. 26). The shear forces correspond to the post-elastic branches, described by Krawinkler's model (Krawinkler 1978), in which the shear force is obtained through Eqs. 27 and 28. The three different branches of the shear force–deformation relationship are presented in Fig. 15e (dashed line):

$$V_{el,wp} = \frac{0.9 \, f_{y,cw} \, A_{vc}}{\sqrt{3}} \qquad (26)$$

$$V_{t,wp} = \frac{0.9 \, f_{y,cw} \, A_{vc}}{\sqrt{3}} + 1.04 \, G_S \frac{b_{cf} \, t_{cf}^2}{z} \, 3 \, \gamma_y \qquad (27)$$

$$V_{s,wp} = \frac{f_{u,cw} \, A_{vc}}{\sqrt{3}} + \frac{f_{u,cf} \, b_{cf} \, t_{cf}^2}{4z} \qquad (28)$$

where

γ_y = panel distortion at yield = $V_{el,wp}/K_{el,wp}$
$f_{u,cw}$, $f_{u,cf}$ = ultimate tensile stress for column web and flanges
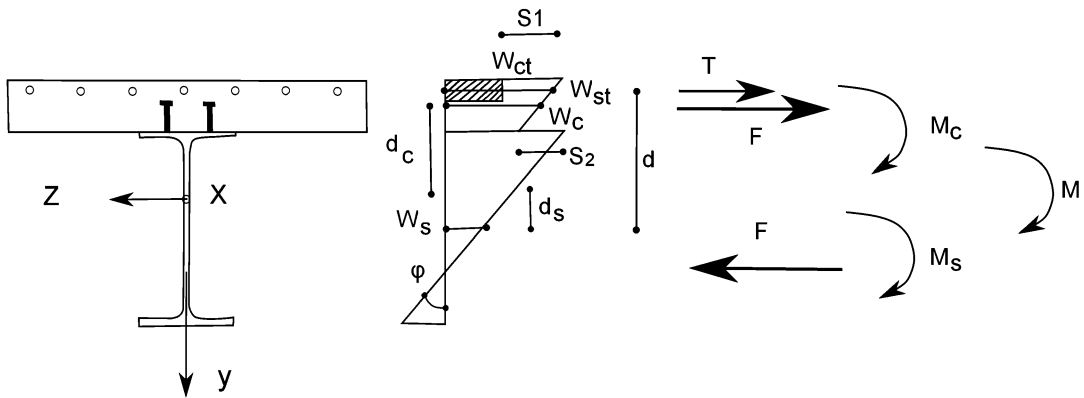
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 16 Cross-sectional model under negative bending

T-Stubs Components

The geometry of the equivalent T-stub components can be determined using the effective width concept presented in EC3 for stiffened columns and end-plates. The force–deformation relationship of T-stub elements required for component modeling can be obtained from the model proposed by Piluso et al. (2001) (dashed line) (Fig. 15f).

Simplifications and Assumptions

1. The difference in column web shear stiffness between the concrete-encased and non-encased segments of the columns should be taken into account in modeling through the assumption of infinite stiffness of the upper encased portion of the column and a flexible diagonal spring in the bare steel column web over the beam depth.
2. Bolt pretension effects need to be considered in modeling through modification of the stiffness of the equivalent T-stub springs in the elastic range.
3. There is an interaction between the connections on both column sides due to the continuity of the slab and the slab longitudinal steel reinforcement between the two beams. Hence, including the anchorage steel bars, the negative moment capacity of one side is dependent on the concrete capacity of the opposite side while essentially transferring the tensile forces acting from the reinforcing steel to the

column. The modeling engineer can use more advanced modeling procedures (i.e., as proposed by Fabbrocino et al. 2002) when continuous SCC beams are considered for the assessment of the behavior of the connection capacity, in terms of global quantities such as the rotations and deflections as well as local quantities such as the slip and the curvature of members, the interaction forces, and the rebar strain.

4. Column web buckling and beam flange buckling under compression are not considered in such modeling procedure.

Fabbrocino et al. 2002

1. Modeling the cross section of composite beams is achieved through a modification of the well-known Newmark's kinematic model, as shown in Fig. 16. This approach requires the definition of the slab effective width depending on the type of loading (hogging or sagging) and on the connection detailing at the beam end. A linear strain pattern is then applied to each component of the cross section. Under the assumption that the curvature and the rotation for each of the components are the same (e.g., for both the concrete slab and the steel profile), the uplift is neglected. Using analytical procedures, the tensile stresses developed in the concrete slab; the slip between the different components of the cross section; the interaction force, F; the

global bending moments in the steel profile, M_s; and the concrete slab, M_c, as well as the moment–curvature relationship of the cross section can be obtained.

2. Modeling the continuous composite beam is based on a combination of the main behavioral aspects on the different regions of the beam (i.e., Newmark's model is used for sagging moments, whereas its modified version is used when cracked zones of the beam are considered). The moment–curvature relationship in each section of the beam can be then defined through an iterative process. Once the generalized moment–curvature relationship is established, rotations and displacements can be obtained by integration of the curvature distribution. The numerical procedure for the solution of a simple structural system of a beam characterized by geometrical and mechanical symmetry is based on the compatibility method; therefore, the support bending moment is the main unknown and the beam is statically determined. The reader is referred to Fabbrocino et al. (2002) for a step-by-step guidance for the solution of the composite section and beam.

Modeling of the Panel Zone in Semirigid Steel–Concrete Composite (SCC) Connections

When moment-resisting frames are subjected to horizontal loading such as earthquake excitation, unbalanced moments occur at the beam-to-column connections resulting to shear deformations in the panel zones of the columns. Therefore, the behavior of the panel zone plays a significant role in determining the overall stiffness and capacity of the frame. In addition to this, in terms of seismic design, the panel zone can have a significant influence on the distribution of plasticity and energy dissipation mechanisms and in turn significant effects on the overall performance of the structure.

This section deals with the modeling of the panel zone region within the beam-to-column joints of SCC moment-resisting frames. This

particular model proposed by Castro et al. (2005) is based on a realistic stress distribution at the edges of the panel, aiming to account for the location of the neutral axis. This methodology enables assessment of the shear stress distribution through the panel depth representing the distribution of the plasticity in the vicinity of that region. Both shear and bending deformations are considered in the elastic and post-elastic stages. The additional resistance of the panel zone owing to the contribution of the column flanges is also taken into account by considering the column depth and flange thickness (Fig. 17).

Procedures and Details

With reference to Fig. 18, the physical dimensions of the panel zone, d_c and d_b, are taken into account. The part of the column being in contact with the slab is modeled through an assemblage of links on top of the panel. This modeling approach is essentially a determination of the spring properties of both the "panel zone" and the "top panel" which are derived analytically allowing for implementation in frame analysis software. The procedure establishes an analogy between the analytical (consisting of the actual connection) and the corresponding numerical model for frame analysis, as shown in Fig. 18.

The location of neutral axis is calculated considering a linear stress distribution, based on the assumption that the SCC beam behaves elastically until the panel yields.

An assessment of the effective width of the slab in the vicinity of the connection is also required. For positive moment (sagging), it is assumed that the slab is limited to the contact width of the column flange width, b_c. For negative moment (hogging), the slab is not considered, under the assumption that the reinforcement is not anchored to the column.

Notion of Calculations

The main aspect of the proposed model is the determination of the spring properties for both the panel zone and the top panel. For a given moment carried by the SCC beam, a corresponding equivalent shear is applied to the panel in the numerical model. The stiffness
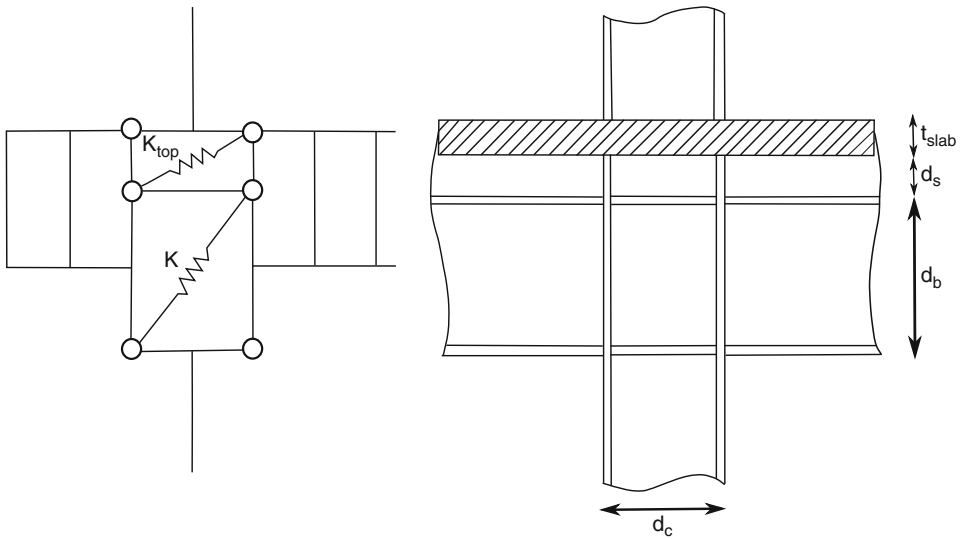
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 17 Numerical and analytical representation of joint models
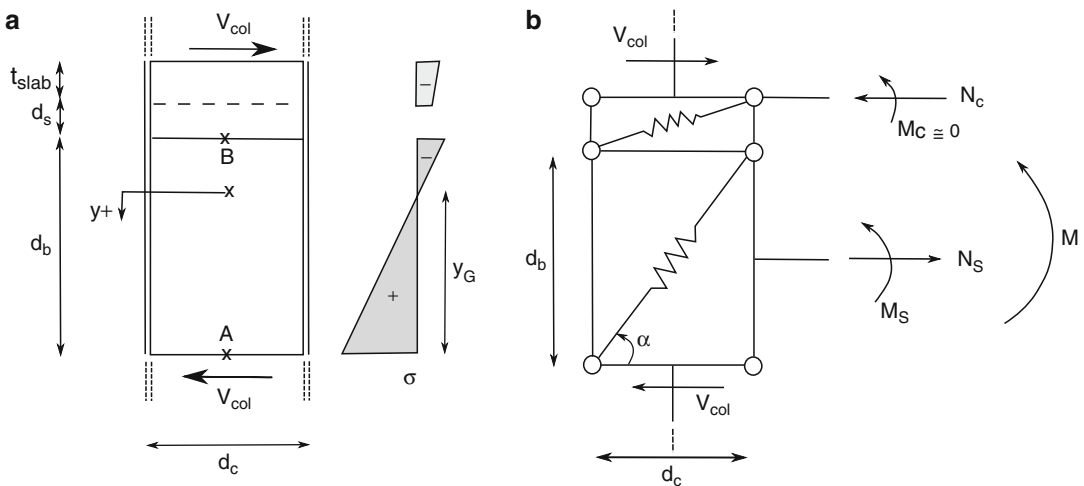
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 18 Analytical and numerical representation of joint models

that should be used in the numerical model can be then determined from the distortion caused to the analytical model owing to the application of moment. Furthermore, the difference between load level corresponding to first and full yielding of the panel can be derived given that the shear stress distribution through the panel is known. The procedure needed to be followed by the engineer for derivation of these important parameters

for both elastic and post-elastic ranges is outlined below.

This section serves as a numerical guideline; the reader is referred to J. M Castro et al. (2005) to exploit the complete set of equations.

Elastic Range
1. The neutral axis location of the steel beam and in turn its second moment of area can be

**Seismic Analysis of
Steel–Concrete
Composite Buildings:
Numerical Modeling,
Fig. 19** The virtual system
used to find panel
deformations
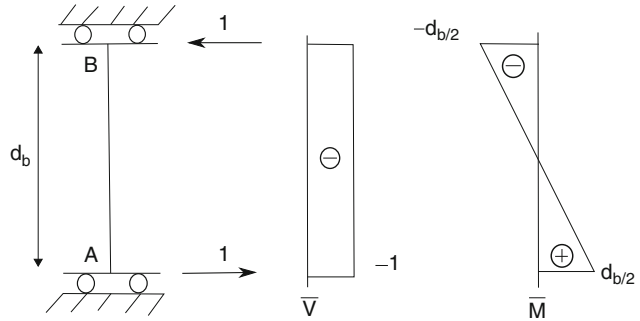
calculated under the assumption that the steel
beam behaves in an elastic manner up to yield-
ing of the panel zone.

2. The ratio of axial force (Ns) and bending
   moment (Ms) carried by the beam can be
   then obtained using analytical procedures.
3. The total moment (M) acting to the connection
   can be then calculated in the numerical model
   using Eq. 29. The moment (Mc) developed in
   the slab is considered to have insignificant
   influence and therefore it is ignored:

$$M = Ms + Ns \left[\frac{d_b}{2} + d_s + \frac{t_{slab}}{2} \right] \qquad (29)$$

4. The equivalent moment carried by the panel
   can be then obtained from Eq. 30:

$$V_{eq} = \frac{Ms}{d_b} + \frac{Ns}{2} - V_{col} \qquad (30)$$

   where $V_{col} = \frac{M}{h_s}$ and h_s is the storey height.

5. From the analytical model, knowing the nor-
   mal stress distribution of the composite sec-
   tion, the shear force and bending moment
   distributions can be obtained.
6. Using the principle of virtual work (Fig. 19)
   and the calculated shear force and bending
   moment distributions, the relative horizontal
   displacement can be obtained. Having the
   application of opposite unit forces in the vir-
   tual system, the internal virtual forces (V_{int},
   M_{int}) can also be obtained.
7. Using the calculated equivalent shear force
   (V_{eq}) and the relative horizontal displace-
   ments ($|\Delta_{shear}|$, $|\Delta_{bending}|$), the elastic stiffness

to be used in the numerical model is
calculated by

$$K_{el} = \frac{V_{eq}}{|\Delta_{shear}| + |\Delta_{bending}|} \qquad (31)$$

8. Using the calculated elastic stiffness, the rela-
   tive drift of the panel ($\Delta_{y,el}$) at the onset of
   yielding can be obtained analytically.
9. Finally, using the calculated elastic stiffness
   and the relative drift of the panel zone, the
   elastic stiffness and the relative drift of
   the spring can be obtained using the Eqs. 32
   and 33:

$$K_{el,\,spring} = \frac{K_{el}}{\cos^2 a} \qquad (32)$$

$$\Delta_{el\text{-}y,\,spring} = \frac{\Delta_{y,\,el}}{\cos a} \qquad (33)$$

where a is the angle of the spring as indicated
in Fig. 18.

Post-Elastic Range

1. Beyond the yielding of the panel, the shear
   stiffness of the column web is assumed to
   drop to the strain hardening stiffness of the
   material. Consequently, the post-elastic stiff-
   ness of the panel is provided by the strain
   hardening of the column web by the flanges
   and a portion of the column web delimiting the
   panel zone, as shown in Fig. 20. Following the
   same assumption for the beam remaining
   largely elastic, and by following similar pro-
   cedures to those described for the elastic

0.9 t$_{cf}$ + 0.05 d$_c$

t$_{cf}$

b$_c$

d$_c$
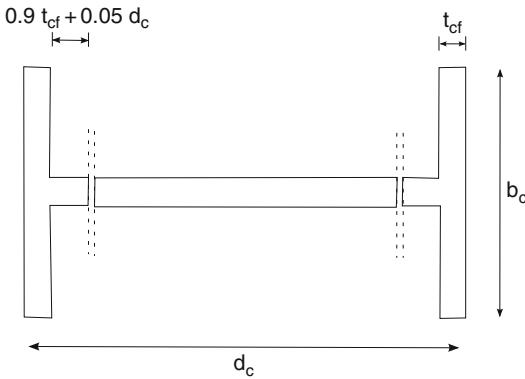
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 20 The cross-sectional definition for the post-elastic range

range, the post-elastic stiffness for the panel can be obtained from Eq. 34:

$$K_{pl} = \left[\mu \ \frac{V_{eq}}{\Delta_{shear}} \right] + \left[\frac{V_{eq}}{\Delta_{bending}} \ \frac{2I_t}{I_{col}} \right] \quad (34)$$

where

μ = strain hardening parameter

I_t = second moment of area of the T-section obtained analytically

I_{col} = second moment of area of the column obtained analytically

2. The relative deformation, $\Delta_{y,pl}$, of the panel zone in the post-elastic range can be obtained using the principle of virtual work and can be readily derived as:

$$\Delta_{y,pl} = \Delta_{y,el} + \frac{\overline{f_y} \ d_b^2}{6E \ d_{CG}} \quad (35)$$

where

d_{CG} = distance from the centroid of the T-section to the external fiber of the column flange

$\overline{f_y}$ = reserve stress in the same fiber after shear yielding of the panel zone

3. Having known the calculated post-elastic stiffness and relative drift of the panel zone, the post-elastic stiffness and relative drift of the spring can be obtained using the Eqs. 36 and 37:

$$K_{pl,\,spring} = \frac{K_{pl}}{\cos^2 a} \quad (36)$$

$$\Delta_{pl-y,\,spring} = \frac{\Delta_{y,\,pl}}{\cos a} \quad (37)$$

4. Finally, the stiffness provided by the panel zone owing to strain hardening in shear (the first term of Eq. 34) is given by:

$$K_{S.Hardening} = \mu \frac{Veq}{\Delta_{shear}} \quad (38)$$

Therefore, the strain hardening stiffness of the diagonal spring can be determined from

$$K_{S.H-Spring} = \frac{K_{S.Hardening}}{\cos^2 a} \quad (39)$$

Moment–Rotation Relationship of Panel Zone

The rotational springs in the panel zone are modeled using trilinear moment–rotation relationships while using the stiffness derived from the above expressions for the elastic and post-elastic range (later in Fig. 22). According to that figure, the panel zone is expected to behave in an asymmetric manner in tension and compression owing to the presence of concrete slab which influences the stiffness, yield moments, and cyclic behavior of the panel zone. Under negative moment, the concrete slab contribution is ignored owing to cracking of the concrete. Yield points $(M_{py}^+, \theta_{py}^+)$ and $(M_{py}^-, \theta_{py}^-)$ are controlled by the steel yielding, and the ultimate points $(M_{pu}^+, \theta_{pu}^+)$ and $(M_{pu}^-, \theta_{pu}^-)$ are controlled by the ultimate strength of the concrete. Under cycling loading, unloading occurs in a straight line with the same slope as the initial stiffness K_e. The reloading is directed towards the previous peak, thereby considering some strength and stiffness degradation.

Modeling of Frames Using Beam Elements

This section presents an analytical approach for the assessment of SCC frames under earthquake excitation using two-dimensional SCC beam

Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 21 Element relative forces and deformations in local coordinate system
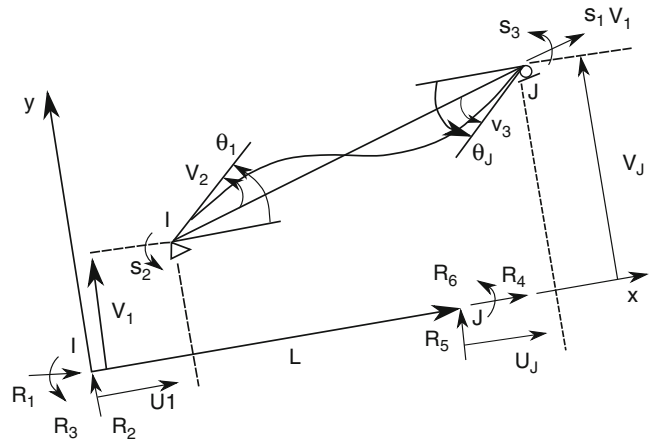
elements (Kim and Engelhardt 2005). In order to model the behavior of SCC beams under earthquake excitation, factors such as the beam-to-column connection details, the local crushing of concrete, the loading pattern on beam, as well as the bond behavior between the reinforcing steel and the concrete need to be considered into the modeling. It is widely accepted that three-dimensional FE models can accurately predict the behavior of SCC beams at the expense of time and computational efficiency. On the other hand, these simpler but reasonably accurate two-dimensional SCC beam models can provide an alternative tool for frame response assessment.

Beam Elements

The beam elements are described as a one-component series hinge-type model combining analytical formulations calibrated against to experimental data and to other data from "sophisticated" described models, while intended to represent the clear span of beams in moment frames (i.e., the length of the beams between column flanges). The two-dimensional SCC beam elements are described by a linear elastic beam with a nonlinear zero-length hinge at each end; the resulting element is referred as "complete element." Each of the hinges is described using nonlinear rigid-plastic moment–rotation relationships in order to simulate the real structural

behavior which was observed from experiments. Each complete element is characterized by two external and two internal nodes. The internal nodes are located between the connection of the linear elastic beam element and the hinges, while the external nodes connect the entire structure. Each of the external nodes has 3DOFs, 2-translations, and 1-rotation in the local coordinate system as presented in Fig. 21.

In the local coordinate system, the element can be considered as a simply supported beam given that the rigid body motions are removed. Based on equilibrium, using the values of relative forces (s_1, s_2, s_3), all the components of local nodal forces (R_1 to R_6) can be calculated. The transformation of forces is defined using the following relationship:

$$R = A\,s \qquad (40)$$

where

A = force transformation matrix (this is well known and can be found in literature, i.e., Przemieniecki (1968))

From the geometry, the transformation from the local displacements, r, to the relative deformations (v_1, v_2, v_3) is performed by:

$$v = A^T r \qquad (41)$$

where

$$r^T = \{U_I, V_I, \theta_I, U_J, V_J, \theta_J\} \qquad (42)$$

Element Stiffness

The initial stiffness of the aforementioned "complete element" is that of the linear elastic beam. As gradual yielding occurs at the hinges owing to increased moments at the element ends, the stiffness of the "complete element" reduces accordingly. In order to obtain the reduced stiffness at any load step after yielding, the instantaneous tangent flexibility of the nonlinear rigid-plastic force–deformation relationship for a hinge is combined with the flexibility of the elastic beam element.

A flexibility matrix, f, is first formed for the elastic element including the effects of elastic shear deformation through the following relationship:

$$dq = f \, ds \qquad (43)$$

where

$dq = (dq_1, \, dq_2, \, dq_3)$ = elastic deformation increment at the internal nodes

ds = action increment in which $ds^T = \{dF, dM^I, dM^J\} = \{ds_1, ds_2, ds_3\}$

For hinges at nodes I and J, the incremental action–deformation relationship is expressed by

$$dw_p = \begin{Bmatrix} 0 \\ d\theta_P^I \\ d\theta_P^J \end{Bmatrix} = \begin{Bmatrix} dv_1 - dq_1 \\ dv_2 - dq_2 \\ dv_3 - dq_3 \end{Bmatrix} = f_p \, ds \quad (44)$$

where

dw_p = vector of plastic hinge deformations at nodes I and J

$d\theta_P^I, d\theta_P^J$ = incremental plastic rotation at nodes I and J

f_p = hinge or plastic flexibility matrix in which nonzero terms are the second and third elements in the diagonal

Using Eqs. 43 and 44, the action–deformation relationship can be obtained for the "complete

element" expressed in terms of degrees of freedom, v, as follows:

$$dv = ds + dw_p = F_t ds \qquad (45)$$

The hinge flexibility coefficients, f_p, can be simply added to the appropriate coefficients of the elastic element flexibility matrix, f, in order to obtain the tangent flexibility matrix, F_t, for the "complete element," as shown in Eq. 46. Once the 3×3 tangent flexibility matrix is obtained, it is then inverted to obtain the 3×3 tangent stiffness matrix, K_t:

$$F_t = \begin{vmatrix} \dfrac{L}{EA} & 0 & 0 \\ 0 & \dfrac{L}{EA^+}F_{ii}+\dfrac{1}{GA_sL}+f_p^i & -\dfrac{L}{EA^+}F_{ij}+\dfrac{1}{GA_sL} \\ 0 & -\dfrac{L}{EA^+}F_{ij}+\dfrac{1}{GA_sL} & \dfrac{L}{EA^+}F_{jj}+\dfrac{1}{GA_sL}+f_p^j \end{vmatrix}$$
$$(46)$$

where

EA^+ = flexural rigidity of the composite beam

EA = axial rigidity of the composite beam

GA_s = effective shear rigidity of the composite beam

$F_{ii} = F_{ij} = 1/3$ and $F_{jj} = 1/6$ for a uniform member

f_p^i, f_p^j = flexibility of hinges at nodes I and J

Hysteretic Rules

Due to the cross-sectional asymmetry of the SCC beam, the response will be different for positive and negative moments. As a result, a hinge must discern the load paths to model the hysteretic behavior of SCC beam for an arbitrary cyclic loading. Apart of the cross-sectional asymmetry, the hysteretic rules employed for the complete element need to take into account factors such as the strength deterioration and the stiffness degradation.

In this part of the chapter, the hysteretic rules are determined from the modification of Lee's model (1987) in order to better fit the curves and the nonlinear behavior of experimental test specimens. This model employs a specified

Fig. 22 Proposed cyclic moment–rotation model of a SCC beam; (**a**) monotonic model and stiffness degradation for negative moment; (**b**) stiffness degradation and pinching for positive moment
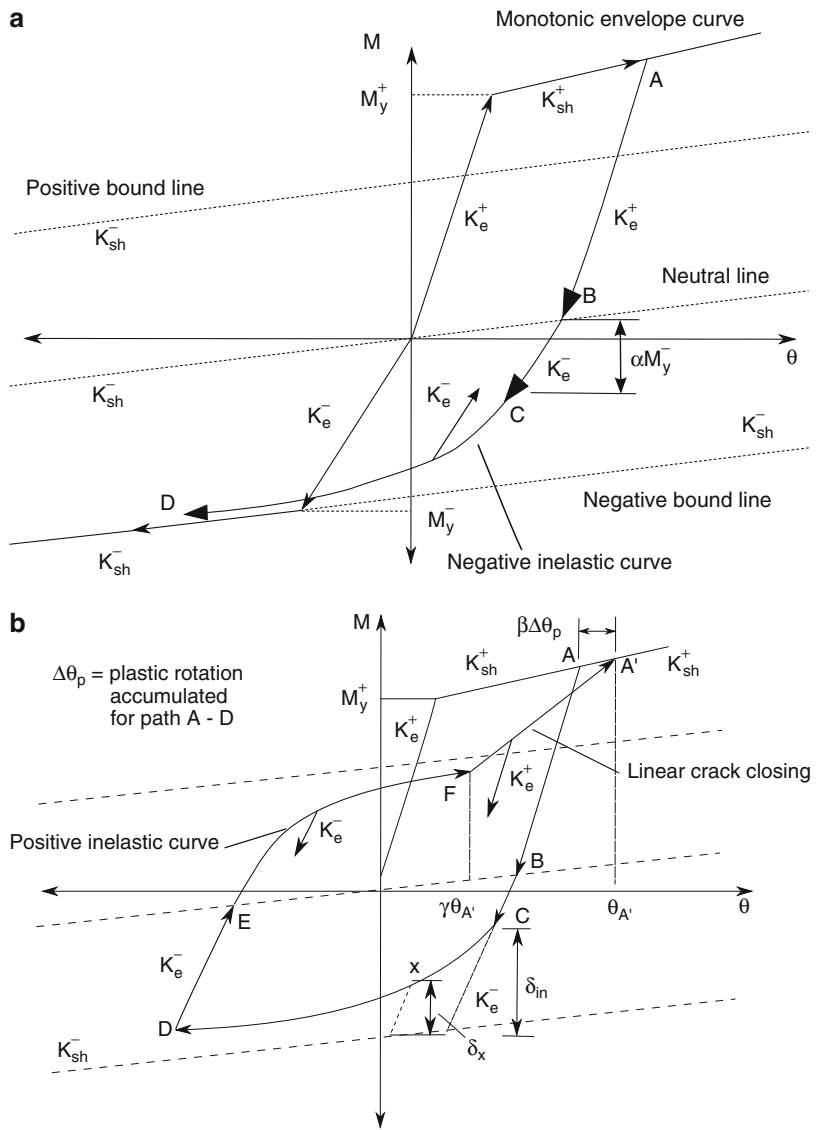
multi-linear moment–rotation relationship based on two bilinear curves.

For the monotonic loading, two bilinear moment–rotation relationships are employed to consider the asymmetrical cross section and the early cracking of the concrete slab under negative moments (Fig. 22a). For the cyclic loading, the moment–rotation relationships of the steel beam are modified in order to account for the effect of the concrete slab (i.e., crack closing and opening) (Fig. 22b).

The basic parameters describing the moment–rotation relation is the moment at the yielding point and the elastic and inelastic stiffness for both positive and negative bending moments. These parameters can be readily obtained analytically. The value of the strain hardening stiffness is expressed as a fraction of the respective elastic stiffness, i.e., the 2.5 % of elastic stiffness in the positive bending and 5 % of elastic stiffness in the negative bending.
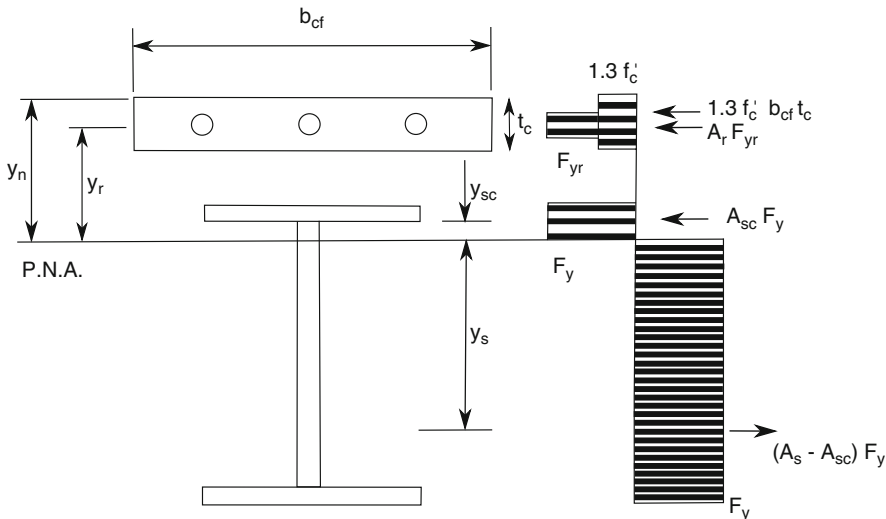
Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 23 Plastic stress distribution for composite beam

The engineer is then required to calculate the effective width of the concrete slab on each side of the beam centerline for computing the positive elastic stiffness. This will enable the calculation of the second moment of inertia of the transformed SCC section. The effects of slip between the concrete and the steel on the positive elastic stiffness are taken into account through the use of a fraction of the second moment of inertia of the transformed SCC section. The calculation of the moment of inertia for the negative elastic stiffness needs to be performed, taking into account the steel beam section and the reinforcing steel bars within the effective slab width. The varying moment of inertia of SCC and cracked sections along the length of the beam is assumed to be equivalent to a uniform moment of inertia of a cracked section with reinforcing bars within the effective width.

Calculation of Moment at the Yielding and Ultimate Point Using a plastic stress distribution for SCC beams (Fig. 23), the ultimate moment (Mmax) at the connection can be obtained. The positive yield moment (My^+) is assumed to be a fraction of the calculated ultimate moment. The negative yield moment (My^-)

is the plastic moment of both the steel beam section and the reinforcing steel bars within the effective width.

The contribution of the concrete slab to the ultimate moment at the connection is determined using the column width and a concrete compressive bearing stress of 1.3f'c, due to concrete confinement near the face of the column.

Modeling of SCC Frames with Concrete-Filled Steel Columns

The advantages of concrete-filled steel (CFS) structures in terms of high strength, high ductility, and large energy absorption led to their extensive use in high-rise structures in earthquake-prone regions.

This section presents a numerical procedure for the nonlinear inelastic analysis of CFS frames based on a fiber beam-to-column element. The nonlinear response of SCC frames is captured through the inelasticity of materials or due to changes in the frame geometry. Global geometric nonlinearities (P-δ effects) are taken into account by the use of stability functions derived from the exact stability solution of a beam-to-column element subjected to axial forces and bending

Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 24 Fiber hinge concept
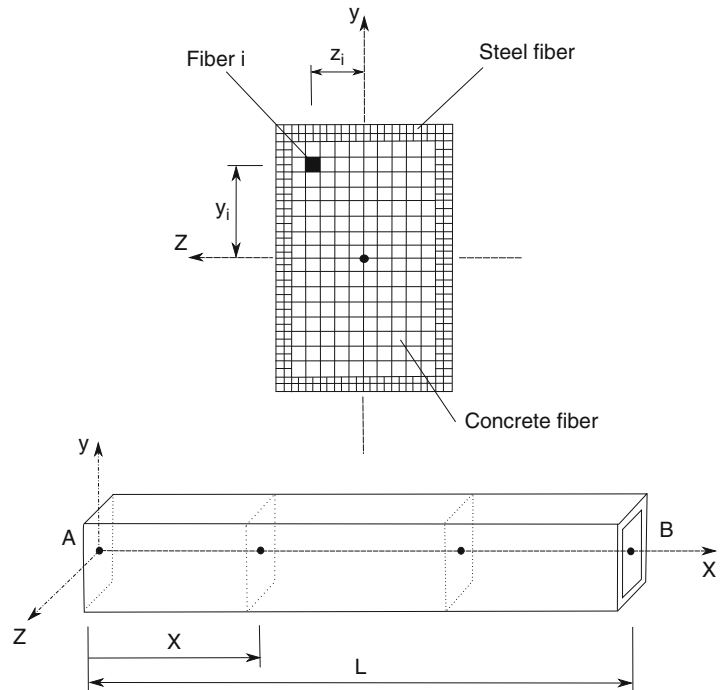
moments. The spread of plasticity over the cross section and along the member length is captured by tracing the uniaxial stress–strain relationships of each fiber on the cross sections located at the selected integration points along the member length. The nonlinear equilibrium equations can be then solved using an incremental iterative scheme, based on the generalized displacement control method.

Fiber Beam-to-Column Element and Material Nonlinearity

The gradual plastification of a composite cross section can be described using the concept of fiber section model, similarly to the modeling of the concrete slabs in section "Introduction." The fiber model is presented in Fig. 24. The concept behind this model is rather simple; the cross-sectional area of the SCC element is subdivided into fibers represented by their area, A_i, and coordinate location (y_i, z_i – with origin the centroid of the section). Different material properties (e.g., concrete confined and unconfined, steel, reinforced steel) can be assigned to each of the fibers. Based on the relevant constitutive material

models, the fiber strains are used to calculate the fiber stresses, which are in turn integrated over the cross-sectional area to obtain stress resultants (i.e., forces and moments).

Using the fiber model, a number of assumptions have been made as follows:

1. Sections remain plane after bending.
2. Due to the latter assumption, cracking is considered to be smeared and normal to the member axis.
3. Torsional and shear effects are ignored.
4. Multi-axial stress states (due to the confinement effects) can be included in the model by increasing the concrete strength and modifying its post-peak response.
5. Local buckling effects or initial stress arising from thermal effects or erection loads are typically not included.

Confinement of Concrete-Encased Steel Sections

In order to utilize the material constitutive models described in the previous sections, the concrete confinement zones need to be identified (Fig. 25).

Seismic Analysis of Steel–Concrete Composite Buildings: Numerical Modeling, Fig. 25 Confinement zones and distribution of monitoring points of SCC sections
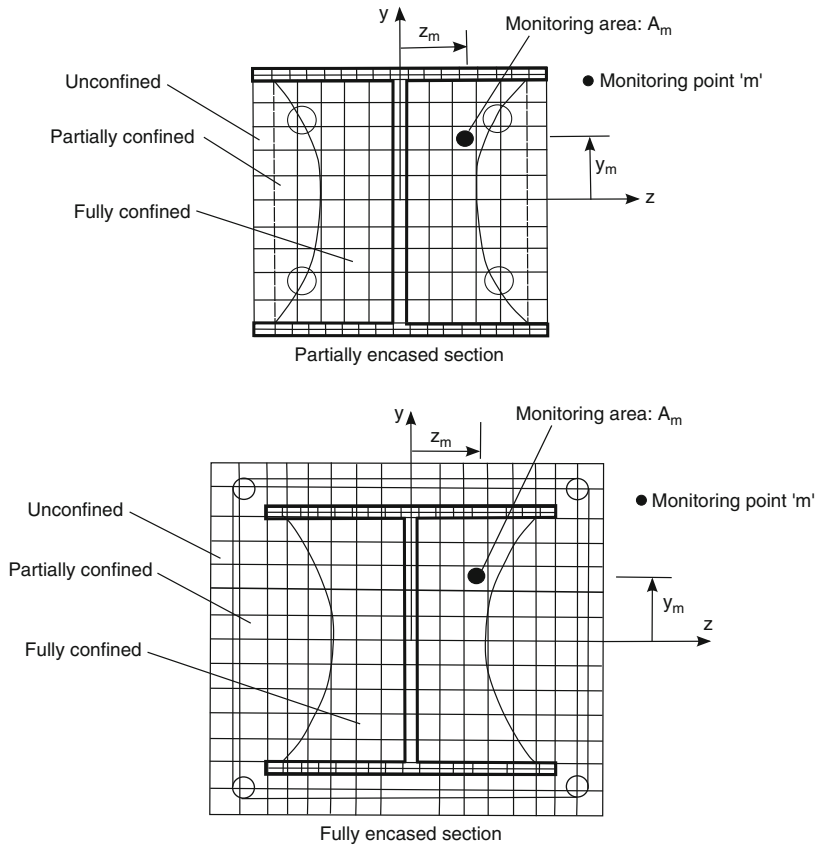
Several methods for the identification of the confined zones have been suggested for partially and fully encased sections, and confinement factors, k, have been developed for use in numerical modeling. According to EC8, the confinement factor, k, is given by:

$$k = 1.0 + 5.0\,\alpha_c\,f_l/f_{co} \qquad \text{for} \quad f_l/f_{co} < 0.05 \tag{47}$$

$$k = 1.125 + 2.5\,\alpha_c\,f_l/f_{co} \qquad \text{for} \quad f_l/f_{co} \geq 0.05 \tag{48}$$

where

α_c = confinement effectiveness coefficient (area of concrete/area of confined concrete)

f_l = lateral confinement pressure from transverse reinforcement

f_{co} = unconfined concrete compressive strength

Local Flange Buckling

The flange buckling is a phenomenon that largely depends on the width-to-thickness ratio, the boundary (i.e., restraint) conditions, and the material properties of the components comprising the section. The effect of the local flange buckling is the reduction of the ultimate strength of the section and/or the diminishing of its rotational capacity in the inelastic range. Additionally, the ductility of encased SCC members is adversely affected by local buckling, and this needs to be considered when estimating the rotational capacity.

To account for the local buckling of bare and encased steel sections, simple methods have been developed and can be readily utilized in a frame analysis software. One of the most popular methods has been developed by Ballio et al. (1987). In this approach, the cross section is divided into a finite number, i, of strips with each strip having an assigned area. If the

compressive strain in any strip exceeds the critical strain, ε_{cr}, the area of the strip reduces to zero for the subsequent load cases. In order to extend the applicability of this method beyond the elastic limit into the inelastic range, the elastic critical stress is divided by the yield strain, as it is represented in the following relationship:

$$\frac{\varepsilon_{cr}}{\varepsilon_y} = \frac{k\,\pi^2\,E}{12\,\sigma_y(1-\nu^2)\left(\frac{\chi_i}{t}\right)^2} \qquad (49)$$

where

χ_i = distance between the centroid of the strip and the plate connection
t = thickness of the plate
k = confinement factor
E = Modulus of elasticity of steel
ν = Poisson's ratio

Geometric Nonlinear P–δ Effect

Geometric nonlinearities can be classified in two categories. The first category is related to the global geometric nonlinearities, usually referred to as P–δ effects. The second category is related to local geometric nonlinearities (i.e., local buckling), which are generally neglected in frame analysis (while they are carefully considered in advanced finite element analyses with discretized models). The global geometric nonlinearities can be incorporated in the models following basic procedures used in nonlinear frame analysis. One of these procedures employs the updated Lagrangian formulation in order to account for geometric nonlinearities such as large displacements and rotations.

In most of the analyses of multi-storey structures subjected to earthquake excitation, the effects of the combination of gravitational forces and lateral displacement are ignored. Such effects are often referred as second-order effects. The reason behind overlooking the second-order effects can be explained by the fact that traditionally, in low-rising reinforced concrete structures

(i.e., structures with low natural period and small lateral displacement response) subjected to earthquake excitation, the second-order effects are insignificant and therefore neglected. As steel structures become taller nowadays, the P–Delta effects are amplified due to the corresponding increase of lateral displacement.

The effect of axial force acting through the relative transverse displacement of the member ends known as P–δ effect can be taken into account in the modeling by using the geometric stiffness matrix, $[K_g]$, as:

$$[K_g] = \begin{bmatrix} [K_s] & -[K_s]^T \\ -[K_s]^T & [K_s] \end{bmatrix} \qquad (50)$$

where

$$[K_s] = \begin{bmatrix} 0 & \alpha & -b & 0 & 0 & 0 \\ \alpha & c & 0 & 0 & 0 & 0 \\ -b & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (51)$$

and

$$\alpha = \frac{M_{zA} + M_{zB}}{L^2} \quad b = \frac{M_{yA} + M_{yB}}{L^2} \quad c = \frac{P}{L} \qquad (52)$$

M_{zA}, M_{zB}, M_{yA}, M_{yB} = end moments with respect to z and y axes, respectively
P = axial force
L = length of the element

The tangent stiffness matrix of a beam-to-column element is then obtained by the following relationship:

$$[K]_{12\times12} = [T]^T_{6\times12} \times [K_e]_{6\times6}[T]_{6\times12} + [K_g]_{12\times12} \qquad (53)$$

where

$$[T]_{6 \times 12} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1/L & 0 & 1 & 0 & 0 & 0 & 1/L & 0 & 0 & 0 \\ -b & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/L & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1/L & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1/L & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix} \qquad (54)$$

and

$[K_e]_{6 \times 6} =$ element stiffness matrix given by $\{\Delta F\}$
$\qquad = [K_e]\{\Delta d\}$

The reader is referred to Thai and Kim (2011) for details on how to derive the element stiffness matrix.

Constitutive Models

Any of the constitutive models presented in the previous sections for steel, concrete, and steel reinforcement can be applied for modeling the material characteristics of the concrete-encased steel sections of the SCC frames.

Summary

The validity of the above case studies has been verified by comparing the numerical predictions with experimental data obtained from a wide range of structural systems subjected to static and hysteretic loadings. Full details of these comparative studies are presented in the literature. However, such constitutive laws and models are usually dependent on parameters which are evaluated through the particular use of the experimental data, and it is in the designers' discretion to chose and interpret these when data are used for specific purposes. In this chapter, it has been attempted to generalize the constitutive models for a number of applications. On the other hand, the lack of generality and objectivity that characterizes most FE packages can only be balanced through the use of material models which are compatible with valid experimental information. In fact, the work presented in

this chapter is considered as a step towards these directions.

The modeling of SCC members in this chapter (i) serves primarily the computation of the response of such members when they are subjected to seismic actions and (ii) acts as a vehicle for carrying out the state determination of the section (or integration point) to a frame element and ultimately to the whole frame assembly. The outcome of the former application is typically the moment–curvature response under a constant axial load. The latter application typically returns section forces that correspond to given section deformations (in uniaxial bending axial strain, and curvature).

In earthquake engineering, the stiffness of the column members is one of the most important parameters of the entire structural systems since it governs the lateral resistance of the frame. The natural period of vibration of the frame decreases with increasing stiffness but also increases with increasing mass. Therefore, the members comprising the frame need to be accurately simulated in order to derive the stiffness and mass matrices to be accounted for the frame analysis. According to the typical response spectra, the acceleration response of a SCC frame reduces with increased natural period. This implies that a composite structure will have to resist lower base shear; therefore, the earthquake effects will be less significant. The displacement response of the structure also increases proportionally with increasing natural period. For the case of increased lateral displacements, second-order effects (P–δ effects) could be developed which will determine the design and amplify the demand on the structure. The engineer is required to reduce the influence of the second-

order effects by controlling the lateral displacement of the frame providing ductility in the beams, columns, and connections. As a result, the accurate analysis of each individual component plays an important role in the aseismic design. In the holistic frame assembly, global geometric nonlinearities can be incorporated in the models following basic procedures used in the nonlinear frame analysis (i.e., modification of the stiffness matrix).

Using nonlinear static analysis, the engineer can obtain information on the global ductility and strength of the structure through force–displacement relations. At each point on the force–displacement curve, the engineer can check the member behavior and see whether the limit states are fulfilled. Weak areas and progressive hinge formation on the structural frame are revealed during the analysis.

Cross-References

References

Aribert JM, Al Bitar A (1989) Optimisation du dimensionnement en connection partielle de poutres de planchers mixtes réalisés avec un bac de tôle mince nervurée (Optimization of the dimensioning in the case of partial connection of beams of composite floors carried out with a thin ribbed sheet vat). Construction Métallique 4 (1989) 12, pp 4–33

Aribert J-M, Lachal A (2000) Cyclic behaviour of the shear connection component in composite joints. In: Behaviour of steel structures in seismic area. Proceedings of STESSA 2000 conference, Montreal, Canada. Balkema, Rotterdam, pp 105–112

Balan TA, Filippou FC, Popov EP (1997) Constitutive model for 3D cyclic analysis of concrete structures. J Eng Mech 123(2):143–153

Ballio G, Calado L, Iori I, Mirabella Roberti G (1987) I problemi delle grandi costruzioni in zona sismica. Associazione Italiana Cemento Armato e Precompresso, Roma, pp 31–44

Braconi A, Salvatore W, Tremblay R, Bursi OS (2007) Behaviour and modeling of partial-strength beam-to-column composite joints for seismic applications. Earthq Eng Struct Dyn 36:142–161

Castro JM, Elghazouli AY, Izzuddin BA (2005) Modelling of the panel zone in steel and composite moment frames. Eng Struct 27(1):129–144

Cotsovos DM, Kotsovos MD (2011) Constitutive modelling of concrete behaviour. In: Papadrakakis M, Fragiadakis M, Lagaros ND (eds) Computational methods in earthquake engineering, Computational methods in applied sciences. Springer, Netherlands, pp 147–175

Fabbrocino G, Manfredi G, Cosenza E (2002) Modelling of continuous steel-concrete composite beams: computational aspects. Comput Struct 80(27–30): 2241–2251

Kim DK, Engelhardt MD (2005) Composite beam element for nonlinear seismic analysis of steel frames. J Struct Eng 131:715–724

Krawinkler H (1978) Shear in beam-column joints in seismic design of steel frames. Eng J 15(3):82–91

Lee SJ (1987) Seismic behavior of steel building structures with composite slabs. PhD thesis, Dept. of Civil Engineering, Lehigh Univ., Bethlehem

Menegotto M, Pinto PE (1973) Method of analysis for cyclically loaded R.C. plane elements including changes in geometry and non-elastic behaviour of elements under combined normal force and bending. In: IABSE, Report of the working commissions, Band 13, symposium on resistance and ultimate deformability of structures acted on by well defined repeated loads, Lisbon

Newmark NM, Siess CP, Viest IM (1951) Tests and analyses of composite beams with incomplete interaction. Proc Soc Exp Stress Anal 9(1):75–92

Ollgaard JG, Slutter RG, Fisher JW (1971) Shear strength of stud connectors in lightweight and normal-weight concrete. AISC Eng J 8(2):55–64

Park R, Paulay T (1975) Reinforced concrete structures. Wiley, New York

Piluso V, Faella C, Rizzano G (2001) Ultimate behavior of bolted T-stubs I: theoretical model. J Struct Eng 127(6):686–693

Przemieniecki JS (1968) Theory of matrix structural analysis. McGraw-Hill, New York

Stevens NJ, Uzumeri SM, Collins MP, Will GT (1991) Constitutive model for reinforced concrete finite element analysis. ACI Struct J 88(1):49–59

Thai HT, Kim SE (2011) Nonlinear inelastic analysis of concrete-filled steel tubular frames. J Construct Steel Res 67:1797–1805

Uchida N, Tohki H (1997) Design of high-rise building using round tubular steel composite columns. In: IABSE international conference on composite construction conventional and innovative, Innsbruck, 16–18 Sept 1997

Zhao H, Yuan Y, Ye Y (2012) Simplified nonlinear simulation of steel-concrete composite beams. J Construct Steel Res 71:83–91

Zona A, Barbato M, Conte JP (2008) Nonlinear seismic response analysis of steel-concrete composite frames. ASCE J Struct Eng 123(6):986

Seismic Analysis of Wind Energy Converters

Giuseppe Failla

Dipartimento di Ingegneria Civile, dell'Energia, dell'Ambiente e dei Materiali, (DICEAM), University of Reggio Calabria, Reggio Calabria, Italy

Synonyms

Aerodynamic damping; Response spectrum; Seismic risk; Soil structure interaction; Wind energy converter

Introduction

Wind power is one of the fastest-growing renewable energy segments on a percentage basis. In 2013, over 35 GW of new wind capacity was installed all over the world, bringing the total wind capacity to 318 GW at the end of 2013 (GWEC 2014). The total installed wind capacity is expected to reach 365 GW by the end of 2014,

Seismic Analysis of Wind Energy Converters, Fig. 1 Horizontal axis wind turbine (HAWT) in Egeln, Germany (photo by Hadhuey)

enough to provide about 4 % of the global electricity demand (GWEC 2014).

While many different design solutions have been considered in the early stages, for commercial use the modern wind industry has now stabilized on horizontal axis wind turbines (HAWT). A typical example is shown in Fig. 1: a land-based tower with a nacelle mounted on the top, containing the generator, a gearbox, and the rotor. Typically, three-bladed upwind rotors are used.

With the continuous increase of wind power production, the search for optimal design is facing new and challenging tasks. The design of land-based HAWTs has been traditionally driven by high wind speed conditions. However, following the introduction of new technologies such as variable pitch and active control in larger, lighter, and cost-effective HAWTs, in some cases the design-driving considerations have been changed, with fatigue and turbulence being considered in addition to high wind speed conditions.

For these lighter HAWTs, especially when installed in seismically active areas, a question has soon arisen as to whether seismic loads shall be considered among design loads. On the other hand, the need to investigate the potential importance of seismic loads has been corroborated by the damage that occurred to land-based HAWTs, following the 1986 North Palm Springs Earthquake, USA, and the 2011 Kashima City Earthquake, Japan. Post-earthquake surveys in the wind farms nearest the epicenter of North Palm Springs Earthquake documented that 48 out of 65 HAWTs were damaged, generally due to buckling in the walls of the supporting tower (photographs are available in the report by Swan and Hadjian (1988)). Earthquake-induced failure may occur also at the foundation level, as for the case of the footing of a HAWT in the Kashima wind farm (photographs are available in the paper by Umar and Ishihara (2012)). In this context, the seismic assessment of HAWTs has drawn an increasing attention in the last years, and as a result, seismic loading has been progressively included in International Standards (ISs) and Certification Guidelines (CGs) (DNV/Risø 2002; GL 2010; IEC 2005; AWEA 2011).

The key points in the seismic assessment of HAWTs can be briefly summarized as:

- Selection of the load combinations
- Use of a specific analysis method
- Definition of the structural model

On these points, sufficient information is generally available in existing ISs and CGs. However, because a certain flexibility is allowed, especially in the definition of the structural model and the selection of an appropriate analysis method, it is important that engineers be aware of the potential options available and how they may affect the reliability of the results. In an attempt to respond to these needs, this entry will provide, first, a preliminary introduction to the relevant issues involved in the seismic assessment of HAWTs. Hence, detailed prescriptions of existing ISs and CGs will be reported and, finally, examples of the possible options for the implementation of the seismic assessment will be presented. Throughout the entry, land-based HAWTs will be referred to.

Seismic Assessment of HAWTs

Load Combinations

The selection of appropriate load combinations for seismic assessment is a relevant issue addressed by ISs and CGs. In general, they are recommended based on the observations that follow.

At sites with a significant seismic hazard, there is a reasonable likelihood that an earthquake occurs while the HAWT is in an operational state, i.e., while the rotor is spinning; in this case, the HAWT is subjected to simultaneous earthquake loads and operational wind loads. It shall be considered, also, the possibility that the earthquake triggers a shutdown and that, as a result, the HAWT is subjected to simultaneous earthquake loads and emergency stop loads. Another possible scenario is that the earthquake strikes when the turbine is parked, i.e., not operating due to wind speeds exceeding the cutoff wind speed of the turbine; specifically, blades may be locked against motion (fixed pitch turbines) or feathered such that no sufficient torque is generated for the rotor to spin (active pitch turbines). In recognition of these observations, the load combinations generally suggested by ISs and CGs for the seismic assessment of HAWTs are:

- Earthquake loads and operational wind loads
- Earthquake loads and emergency stop loads
- Earthquake loads and wind loads in a parked state

Both earthquake loads and wind loads are stochastic processes. The wind process is generally treated as a stationary process. Samples can be generated from well-established power spectral densities (PSDs) in the literature (e.g., Von Karman PSD or Kaimal PSD, see Manwell et al. (2010)), with parameters to be set depending on site conditions. Wind acts on the blades of the rotor and along the tower.

Obviously, wind loading on the blades varies significantly depending on whether the rotor is spinning or not; to generate wind loading on a spinning rotor, concepts of classical aerodynamics are used, for instance, those of Blade Element Momentum (BEM) theory and subsequent modifications (Manwell et al. 2010). The earthquake process is inherently nonstationary. Spectrum-compatible samples may be synthetized from site-dependent response spectra or site-specific historical records may be used, according to the prescriptions of the adopted ISs and CGs.

Structural Analysis Method
The computation of the HAWT response to the different load combinations is a crucial step of the seismic assessment. In general, two approaches can be pursued:

- A fully coupled time-domain simulation computing the response to simultaneously acting wind loading and seismic loading
- A decoupled analysis where the responses to wind loading and seismic loading are computed separately and then superposed

A fully coupled time-domain simulation is the most desirable approach. The reason is that it allows the actual wind loads on the blades to be evaluated correctly, taking into account that the oscillations of the tower top, induced by the earthquake ground motion, affect the rotor aerodynamics (in particular, the relative wind speed at the blades, depending on which lift and drag forces are calculated). However, for the implementation of fully coupled time-domain simulations, dedicated software packages are required, capable of solving the nonlinear motion equations of the structural system under simultaneous wind and seismic excitations.

When performing a decoupled analysis, instead, the responses to wind loading and seismic loading are built separately. This means that wind loads are evaluated as no earthquake ground motion is acting at the tower base. Correspondingly, the response to the earthquake ground motion is computed as no wind loading is acting on the rotor. It is evident that this approach is

approximate since, as explained earlier, the actual wind loads on the blades depend on the oscillations of the tower top, to which contributes also the earthquake ground motion at the tower base. Nevertheless, numerical comparisons with benchmark results obtained by fully coupled time-domain simulations have shown that decoupled analyses can yield accurate results, provided that the separate response to earthquake loading is computed using an appropriate level of damping. In particular, it has been found that:

- A percentage equal to 5 % of critical damping is appropriate when the separate response to earthquake loading is to be combined with the response to operational wind loading, i.e., for the load combination = earthquake loads + operational wind loads (Witcher 2005; Prowell and Veers 2009; Prowell 2011).
- In contrast, a percentage varying between 0.5 % and 2 % of critical damping is appropriate when the separate response to earthquake loading is to be combined with the response to wind loading in a parked state, i.e., for the load combination = earthquake loads + wind loads in a parked state (Prowell and Veers 2009; Prowell 2011; Stamatopoulos 2013).

Such a variability of the damping level, to be considered in a decoupled analysis, can be explained as follows:

- The low damping ratios (0.5–2 %), which are required for the load combination = earthquake loads + wind loads in a parked state, are motivated by the fact that when the turbine is parked, the only damping source is the structural damping of the tower that, as customary in steel structures, is generally low.
- The higher damping ratio (5 %), which is required for the load combination = earthquake loads + operational wind loads, reflects the fact that the earthquake loading significantly affects the aerodynamics of a spinning rotor. In particular, a motion of the tower top due to the earthquake loading, against or in the wind direction, causes, respectively, an increase or a decrease of the instantaneous

S

thrust force, with respect to that computed assuming no earthquake ground motion at the tower base (such increase or decrease of the instantaneous thrust force mirrors an increased or decreased relative wind speed at the blades). Since, in both situations, this alternation of the thrust force is oriented opposite to the tower top motion, it can be understood that its effects are to be modeled by introducing additional damping with respect to the structural damping, when computing the separate response of the HAWT to earthquake loading. The same observation holds true when a decoupled analysis is performed to compute the response of HAWTs to combined wind loading and wave loading (Kuhn 2001).

The difference between damping for the parked state and operational state is generally referred to, in the literature, as *aerodynamic damping*, to mean that its source is essentially the rotor aerodynamics.

Decoupled analyses may be performed in time and frequency domain. Especially frequency-domain formulations have been awarded a considerable attention, because in this case the separate response to earthquake loading can be built by coded response spectra, a concept most engineers are familiar with. However, in light of the earlier observations on the appropriate level of damping, it is evident that particular care shall be taken when following this approach. In fact, the typical 5 % damped response spectra for building structures (ICC 2012) will be suitable only for the load combination = earthquake loads + operational wind loads, while 0.5–2 % damped response spectra shall be used for the load combination = earthquake loads + wind loads in a parked state. It is evident that selecting the correct damping ratio is crucial: if the response to earthquake loading was obtained from 5 % damped response spectra, nonconservative results would be certainly obtained when the turbine is parked. Some ISs and CGs take account of these issues (IEC 2005; AWEA 2011), but no explicit indications on the damping ratio to be adopted are given in the others (DNV/Risø 2002; GL 2010).

Structural Model

A fundamental assumption of existing ISs and CGs, with regard to the structural model, is material linearity. This assumption is essentially justified by the fact that the primary intent is to ensure power production for the design life of the HAWT, usually 20 years, and that nonlinear deformation (damage) to the turbine would interrupt reliable operation. Material linearity means low operational stresses, and this provides some safety margins against failure (Bazeos et al. 2002). Therefore material linearity will be, in general, a prerequisite of ISs and CGs also when assessing the response to seismic excitations.

Starting from the assumption of material linearity, in general, two types of structural modeling are feasible:

- Simplified models which model the tower and consider the rotor-nacelle assembly (RNA) as a lumped mass at the tower top
- Full system models which describe the whole turbine, including the nacelle and rotor with a certain level of detail

Simplified models are appealing since the complexities involved in modeling the rotor are avoided. Full system models include the rotor blades and, in general, turbine components such as power transmission inside the nacelle and pitch and speed control devices, with a different degree of accuracy depending on the specific modeling adopted, for instance, a finite element (FE) or a rigid multi-body modeling.

Simplified or full system models can be used depending on the selected structural analysis method. In particular:

- Fully coupled time-domain simulations involve only full system models as they require modeling the rotor aerodynamics, with the earthquake ground motion simultaneously acting at the tower base.
- Decoupled analyses may be implemented using either a full system model or a simplified model. If a simplified model is adopted, seismic loads are built considering the mass of

RNA lumped at the tower top, while wind loads are obtained by a dedicated software package, capable of modeling the rotor aerodynamics, with no earthquake ground motion at the tower base, since the analysis is decoupled.

International Standards and Certification Guidelines

Guidance for seismic loading on HAWTs can be found in the following ISs and CGs:

- DNV/Risø: Guidelines for design of wind turbines (DNV/Risø 2002). Released by Det Norske Veritas (DNV) and Risø National Laboratory
- GL 2010: Guideline for the certification of wind turbines (GL 2010). Released by Germanischer Lloyd (GL)
- IEC 61400-1: Wind turbine generator systems. Part 1: Safety requirements (IEC 2005). Released by International Electrotechnical Commission (IEC)
- ASCE/AWEA RP2011: Recommended practice for compliance of large land-based wind turbine support structures (ASCE/AWEA 2011). Released by American Society of Civil Engineers (ASCE) and American Wind Energy Association (AWEA)

DNV/Risø Guidelines

DNV/Risø Guidelines are meant to provide a basic introduction to the most relevant subjects in wind turbine engineering (DNV/Risø 2002). Consistently with this general purpose, quite general suggestions are given to deal with seismic loading.

It is prescribed that earthquake effects should be considered for HAWTs located in areas that are considered seismically active based on previous records of earthquake activity (Section 3.2.8). For those areas known to be seismically active but with no sufficient information available for a detailed characterization of seismicity, an evaluation of the regional and local geology is recommended to determine the location of the HAWT relative to the alignment of faults, the epicentral and focal distances, the source mechanism for energy release, and the source-to-site attenuation characteristics. In this case, the evaluation should aim to estimate both the design earthquake and the maximum expectable earthquake, taking into account also the potential influence of local soil conditions on the ground motion.

No specific recommendations are given on the earthquake-wind load combinations to be considered. However, since it is prescribed that in seismically active areas the HAWT should be designed so as to withstand earthquake loads, it is implicit that the three, typical load combinations described earlier (i.e., earthquake loads and operational wind loads, earthquake loads and emergency stop loads, earthquake loads occurring in a parked state) shall be referred to.

As for what concerns the method of analysis, DNV/Risø provides explicit suggestions only for the response spectrum method, as used in a decoupled analysis. In particular, the use of a single degree-of-freedom (SDOF) system with a lumped mass on top of a vertical rod is suggested, with the rod length equal to the tower height and the lumped mass including the mass of the rotor-nacelle assembly (RNA) and ¼ of the mass of the tower. It is prescribed that the fundamental period of the SDOF system is used in conjunction with a design acceleration response spectrum to determine the loads set up by the ground motion, by analogy with the simplified procedures used in building codes. Analyses shall be performed for horizontal and vertical earthquake-induced accelerations. However, no explicit recommendations are given on the criterion to translate the resulting spectral response acceleration into design seismic loads, as well as on the damping ratio to be used. Since in the absence of specific guidance on this matter, a most intuitive choice of engineers could be using the typical procedures of the International Building Code (ICC 2012), it has to be remarked that the 5 % damping ratio, embedded in the standard design response spectrum, is appropriate only for seismic loading acting during an operational state, but overestimates considerably the

S

actual damping in a parked state. This aspect should be well kept in mind when referring to DNV/Risø for seismic assessment of HAWTs.

Regarding the structural model, attention is drawn to the need of including the actual stiffness of the structural component of the foundation and an appropriate model of the supporting or surrounding soil, the latter through a proper soil structure interaction (SSI) modeling (Section 8.4). Although, for this purpose, nonlinear and frequency-dependent models are recommended in principle, appropriate linearized models are allowed, depending on the expected strain level in the soil (typically, it may be up to 10^{-1} for earthquake loading and considerably larger than for other loading conditions). The linearized models consist of translational and rotational springs for circular footings and piles.

GL Guidelines

GL 2010 guidelines aim to set a number of requirements for the certification of wind turbines (GL 2010). For this reason, they are quite prescriptive and provide detailed information on some particular aspects of seismic risk.

In agreement with DNV/Risø, GL 2010 prescribes that seismic loading shall be taken into account in seismically active areas (Section 4.2.4.2.3). Earthquake loading is included in a group of design load cases (Table 4.3.2) classified as load cases accounting for "extended" design situations, including special applications and site conditions. These design load cases are not mandatory for certification purposes, but may be chosen for the verification of the HAWT to complement the applicability in specific design situations. The response to seismic loading is to be assessed both in the operational state and the parked state (Table 4.3.2) under normal wind loading. For the operational state it is also suggested to consider the activation of the emergency shutdown triggered by the earthquake. The safety factor for all the loads to be combined with seismic loading is equal to 1.0 (Section 4.3.5.4). A return period of 475 years is prescribed as the earthquake design level. To model the seismic loading, recommendations of the local building

code should be applied or, in the absence of locally applicable regulations, those of either Eurocode 8 (2004) or American Petroleum Institute (API 2000).

Regarding the method of analysis, GL 2010 specifies that fully coupled or decoupled analyses are possible, with at least three modes in both cases. Time-domain simulations shall be carried out considering at least six simulations per load case. As with DNV/Risø, no guidance is provided on the damping ratio to be adopted when using the design response spectrum in a decoupled analysis. Again, because of the lack of guidance on this matter, it shall be kept in mind that the 5 % damping ratio is appropriate only in the operational state and that lower damping ratios shall be considered in the parked state.

GL 2010 gives no particular prescriptions on the structural model to be adopted. However, because at least three modes have to be included in the vibration response, the use of a multi-degree-of-freedom (MDOF) structural model is implicitly suggested. In general, a linear elastic behavior shall be assumed. A ductile response can be considered only when the support structure has a sufficient static redundancy, such as a lattice tower. However, if ductile behavior is assumed, the structure shall be mandatorily inspected after occurrence of an earthquake.

IEC Standards

IEC 61400-1 Standards aim to specify essential design requirements to ensure structural integrity of wind turbines (IEC 2005). They have the status of national standards in all European countries whose national electrotechnical committees are CENELEC members (CENELEC = European Committee for Electrotechnical Standardization).

IEC 61400-1 recommends that, in seismically active areas, the integrity of the HAWT is demonstrated for the specific site conditions (Section 11.6), while no seismic assessment is required for sites already excluded by the local building code, due to weak seismic actions. The seismic loading shall be combined with other significant, frequently occurring operational

loads. In particular, IEC 61400-1 prescribes that the seismic loading shall be superposed with operational loads, to be selected as the higher of:

(a) Loads during normal power production, by averaging over the lifetime
(b) Loads during emergency shutdown, for a wind speed selected so that the loads prior to the shutdown are equal to those obtained with (a)

No explicit reference is made, however, to the load case of an earthquake loading striking in a parked state.

The safety factor for all load components to be combined with seismic loading shall be set equal to 1.0. The ground acceleration shall be evaluated for a 475-year recurrence period based on ground acceleration and response spectrum requirements as defined in local building codes. If a local building code is not available or does not provide ground acceleration and response spectrum, an appropriate evaluation of these parameters shall be carried out.

Regarding the method of analysis, fully coupled or decoupled analyses are possible (11.6). In time-domain analyses, sufficient simulations shall be undertaken to ensure that the operational load is statistically representative. It is prescribed that the number of tower modes used in either of the above methods shall be selected in accordance with a recognized building code. In the absence of a locally applicable building code, consecutive modes with a total modal mass of 85 % of the total mass shall be used.

IEC 61400-1 gives no particular indications on the structural model for seismic analysis. In agreement with GL 2010, however, it is implicit that the structure shall be modeled as a MDOF system, since the use of consecutive modes with a total modal mass equal to at least 85 % of the total mass is recommended. In general, the response should be linearly elastic, while a ductile response with energy dissipation is allowed only for specific structures, in particular for lattice structures with bolted joints.

Annex C of IEC 61400-1 presents a simplified, conservative method for the calculation of seismic loads. This procedure is meant to be used when the most significant seismic loads can reasonably be predicted on the tower, and shall not be used if it is likely that the earthquake ground motion may cause significant loading on the rotor blades or the structural components of the foundation. The principal simplifications in Annex C are ignoring the modes higher than the first tower bending mode and the assumption that the whole structure is subjected to the same acceleration. Upon evaluating or estimating the site and soil conditions required by the local building code, or adopting conservative assumptions while detailed site data are not available, the simplified method can be applied as follows:

- The acceleration at the first tower bending natural frequency is set using a normalized design response spectrum and a seismic hazard-zoning factor. For this, a 1 % damping ratio is assumed.
- Earthquake-induced shear and bending moments at the tower base are calculated by applying, at the tower top, a force equal to the total mass of the RNA + ½ the mass of the tower times the design acceleration response.
- The corresponding base shear and bending moments are added to the characteristic loads calculated for an emergency stop at rated wind speed, i.e., the speed at which the limit of the generator output is reached.
- The results are compared with those obtained against the design loads or the design resistance for the HAWT. If the tower can sustain the resulting combined loading, no further investigation is needed. Otherwise, a thorough investigation shall be carried out on a MDOF structural model.

With regard to such a simplified method, described in Annex C, it shall be pointed out that ignoring the second tower mode is a significant nonconservative simplification (e.g., see Zhao and Maisser (2006) on the role of the second tower mode in the seismic response of HAWTs).

This is somehow compensated for by incorporating ½ of the tower mass with the tower head mass and prescribing superposition with the characteristic loads calculated for an emergency stop at rated wind speed, which represent quite conservative aerodynamic loads.

ASCE/AWEA Recommended Practice

The general purpose of ASCE/AWEA RP2011 is to clearly identify specific US national recommendations for wind turbine design, which are compatible with IEC 61400-1 but may provide proper recommendations for those cases in which US practice and IEC 61400-1 differ. As for what concerns seismic assessment, ASCE/AWEA RP2011 makes a quite comprehensive effort to harmonize some relevant specific prescriptions of certification agencies with the traditional perspectives of US standards ASCE/SEI 7-05, which sets the minimum design loads for buildings and structures in general (ASCE 2006). For the level of detailed information provided, ASCE/AWEA RP2011 can be considered a very useful and comprehensive reference tool for the seismic assessment of HAWTs.

ASCE/AWEA RP2011 points out that although standard HAWT classes shall be generally designed for normal wind conditions, extreme wind, conditions and other environmental conditions including temperature and air density, specific prescriptions on the criteria for the design of HAWTs subjected to earthquake ground motions are necessary, in recognition of the fact that earthquake events are common in many US jurisdictions. According to ASCE/AWEA RP2011, it is of critical importance to recognize that seismic loads plus operational loads may in some cases govern tower and foundation design. For these reasons, load combinations involving earthquake occurring in an operational state and earthquake triggering emergency stop loads, an earthquake occurring in a parked state should be considered (Section 5.4.4). Seismic ground motion values should be determined based on the acceleration response spectrum or site-specific ground motion procedures as prescribed by ASCE/SEI 7-05 (see Section 11.4 and Chapter 21 in ASCE/SEI 7-05).

Unlike the alternative ISs and CGs, ASCE/AWEA RP2011 provides quite detailed prescriptions on a "best practice" load combination including seismic loads plus operational loads:

$$U = (1.2 + 0.2S_{DS})D + 0.75(\rho Q_E + 1.0M) \tag{1}$$

$$U = (0.9 - 0.2S_{DS})D + 0.75(\rho Q_E + 1.0M) \tag{2}$$

where:

U = factored load effect

D = dead load

M = operational loading equal to the greater of (1) loads during normal power production at the rated wind speed or (2) characteristic loads calculated for an emergency stop at rated wind speed

Q_E = effect of horizontal seismic (earthquake-induced) forces

S_{DS} = design spectral response acceleration parameter at short periods

ρ = 1.0, redundancy factor (for nonbuilding structures not similar to buildings ρ = 1.0, according to Chapter 12.3.4.1 of ASCE/SEI 7-05)

ASCE/AWEA RP2011 suggests Equations 15.4-1 and 15.4-2 of ASCE/SEI 7-05 for the seismic response coefficient Cs (nonbuilding structures), if Equation 12.8-1 of ASCE/SEI 7-05 is used to compute the seismic base shear. Specifically, in Eq. 15.4-2 of ASCE/SEI 7-05 for Cs a response modification factor R = 1.5 is recommended by ASCE/AWEA RP2011. The use of R = 1.5 does not necessarily imply that a ductile response or material overstrength is expected but accounts for a certain conservatism in the seismic response coefficient C_s prescribed for nonbuilding structures.

In Eqs. 1 and 2, the use of a load factor of 0.75 on both seismic loads and operational loads is supported by results of time-domain analyses on HAWTs ranging from 65 to 5 MW, subjected to 100 earthquake ground motion records, for varying orientation of wind and earthquake loads (Prowell 2011). It is observed that when the seismic hazard at a particular site is dominated by

known faults, consideration of site-specific prevailing wind direction and maximum earthquake component direction may be appropriate. In this case no load factor may be applicable, if wind and wave propagation directions are expected to coincide.

ASCE/AWEA RP2011 recommends that for load combinations not including operational loads, the spectral response acceleration parameter should be based on a 1 % damping ratio, which reflects the low inherent damping of typical steel support structures for HAWTs. The multiplicative spectral adjustment factor, B, to adjust spectral response acceleration, S_a, from 5 % (standard IBC value for determining S_a) to 1 % damped values is equal to 1.40 (Table 5-6). For load combinations that include operational loads, the spectral response acceleration parameter should be based on 5 % damped values. ASCE/AWEA RP2011 points out that this increase in damping is based on the aerodynamic damping inherent to an operating HAWT as verified by experimental and numerical results showing that a damping level of 1 % produces overly conservative results (Prowell 2011).

Regarding the method of analysis, according to ASCE/AWEA RP2011 a fully coupled time-domain analysis and decoupled analyses based on equivalent lateral force method or modal response spectrum method are acceptable, as permitted by the local building code. For the specific implementation of each method of analysis, the local building code or ASCE/SEI 7-05 is referred to. In particular, if the equivalent lateral force procedure is used, the vertical distribution of seismic forces should be calculated based on the procedure given in ASCE/SEI 7-05, Chapter 12.8.3, with some modifications: the seismic forces corresponding to the seismic weight of the RNA should be located at the turbine's center of gravity, and those corresponding to the seismic weight of the tower structure (including ladders, platforms, railings, etc.) should be distributed to nodes distributed along the tower height. No further prescriptions are given on a specific structural model to be adopted, when implementing a fully coupled or a decoupled analysis.

As for what concerns decoupled analyses, ASCE/AWEA RP2011 suggests that in those cases when only the peak seismic loads and peak operational loads are available, the proposed combination method for seismic loads and operational loads may be overly conservative, especially in recognition of the fact that the respective peak loads do not occur at the same instant of time and in the same loading direction. Therefore, to reduce potential design conservatism and obtain a more accurate prediction of the response, fully coupled time-domain analyses are suggested, considering earthquake ground acceleration in combination with operational or emergency stop loads. Seismic analysis should comply with the requirements of ASCE/SEI 7-05, Chapter 16, concerning, for instance, the minimum number of simulated earthquake ground motions. It is recommended that time-domain analyses be conducted with analysis software capable of simulating the structural response and global turbine dynamics, including the aerodynamic interaction.

ASCE/AWEA RP2011 gives some interesting points of view on the applicability of typical prescriptions of building codes to HAWTs design. For instance, according to ASCE/AWEA RP2011, enhanced performance objectives may be established to meet specific owner requirements and to improve expected behavior during and after an earthquake. For this purpose the use of a performance factor, similar to an importance factor of 1.5 for essential facilities, is suggested. This performance factor shall be agreed with the wind turbine manufacturer to establish acceleration thresholds for turbine components that will ensure operational performance. Also, according to ASCE/AWEA RP2011, no specific drifts or displacement need to be defined. This is motivated essentially by the fact that thorough analysis and design considerations of the ultimate and fatigue limit states implicitly limit the displacements of the tower. In addition, the wind turbine controls monitor and limit the possible tower top accelerations, to prevent exceeding the design loading.

Another distinctive feature of the ASCE/AWEA RP2011 is the emphasis put on

S

consideration of seismic forces in the foundation design, for areas with historical earthquake activity. Evaluation of earthquake effects should be performed in accordance with the requirements of the local building code or IEC 61400-1. In any case, geotechnical evaluation of earthquake effects should include ground shaking, liquefaction, slope instability, surface fault rupture, seismically induced settlement/cyclic densification, lateral spreading, cyclic mobility, and soil strength loss. In areas susceptible to earthquake effects, appropriate mitigation should be provided for foundations. For projects located near active faults, the characteristics of the fault including type, seismic setting, subsurface conditions, ground motion attenuation, and maximum earthquake magnitude should be considered. At any rate, HAWTs should be located with adequate setbacks from fault zones. Where relatively loose unsaturated cohesionless soils are present at the project site, the effect of ground shaking from a design level earthquake should be taken into account. Also, potential settlement due to cyclic densification of the site soils should be evaluated.

Implementation of Seismic Assessment

Despite the prescriptions given by existing ISs and CGs (DNV/Risø 2002; GL 2010; IEC 2005; ASCE/AWEA 2011), engineers dealing with the seismic assessment of HAWTs may face a few issues, which are only partially addressed by ISs and CGs. This section is meant to provide some insights into these aspects, illustrating the most relevant studies for this purpose. Although investigations are not fully accomplished, it is worth mentioning these studies, as they may provide engineers with very useful data for a correct seismic assessment of HAWTs.

As mentioned earlier, an important step is the construction of the structural model. Simplified models, which avoid the complexities involved in modeling the rotor, are frequently used, especially for a preliminary design. In contrast, for a comprehensive investigation of all factors relevant to seismic risk, full system models are generally considered. They allow prediction of

component loads instead of only tower loads, which cannot be estimated in a simple tower-based model. In this regard, it is remarked that higher modes involving the rotor dynamics may play an important role, as they may fall in the region of maximum spectral response acceleration (Prowell et al. 2010).

Regardless of the structural model adopted, another important issue in the seismic assessment of HAWTs is SSI modeling. Modeling the base as fixed, with no consideration of the SSI, could be justified in the case of overdesigned foundations and stiff soil conditions. However, HAWTs may be installed on relatively soft soils or loose soils containing alluvial deposits, and under these circumstances the SSI modeling, particularly for dynamic loads, could become a major concern in the design of the foundation and, consequently, of the entire support structure of the HAWT (Bazeos et al. 2002; Zhao and Maisser 2006). A proper SSI modeling may play an important role also in consideration of the ground motion amplification effects on soft and loose soils.

In the following, simple and full system models will be described briefly as used in the recent literature, along with relevant information on SSI modeling.

Simple Models

One of the first studies on the seismic response of HAWTs has been carried out by Bazeos and coworkers (Bazeos et al. 2002). They have investigated a 38 m high HAWT resting on a concrete block, located in a site with 0.12 g peak ground acceleration and semi-rock soil conditions. Seismic analyses have been conducted on two different models, a FE model of the tower with shell elements and a simplified FE model with lumped masses along the tower height and 3D beam elements approximately mapping mechanical and geometrical properties of the tower. In both models, a top mass has been added to model the RNA; SSI has been modeled by a set of discrete springs and dashpots and, adding to the mass of the concrete foundation block (modeled as rigid), a virtual soil mass moving in phase (Mulliken and Karabalis 1998). Parked conditions only have been considered, with no aerodynamic loads

along the tower. Seismic analysis has been carried out in the time domain using ground motions compatible with the elastic response spectrum, as prescribed by the Greek Aseismic Code with 0.5 % damping. For the relatively low ground acceleration under consideration (0.12 g), low stress levels have been found due to seismic excitations, with respect to the stress levels due to wind in either operational or emergency states. Results obtained by the time-domain simulations on the two FE models have been validated by a response spectrum analysis on a SDOF system with a mass set equal to the total mass of the system and a stiffness computed from the first natural period of the FE model with shell elements. Interestingly, Bazeos and coworkers (2002) have showed that higher tower modes can be significantly affected by SSI modeling. This result is important in consideration of the fact that the natural frequencies of the higher tower modes may fall within the region of maximum spectral acceleration (Haenler et al. 2006).

Umar and Ishihara (2012) have focused on the construction of a response spectrum for HAWTs under seismic excitations only, i.e., in a parked state. The need for a specific response spectrum is motivated by the observation that the support structures for HAWTs exhibit, unlike buildings, long period, heavy top, and different mass distribution along the height. Besides this general observation, by carrying out numerical simulations using a database of strong earthquake ground motions, they have shown that the very low damping levels in parked conditions determine excessive fluctuations in the response spectrum, and such uncertainty cannot be captured by existing damping correction factors in Eurocode 8 (2004) and Japanese Building Standard Law (BSL 2004). Umar and Ishihara (2012) have modeled the HAWT as a MDOF system with a lumped mass at the top, and a sway-rocking model to take into account SSI effects. They have proposed a modified correction factor for the damping ratio of the BSL response spectrum used in Japan (BSL 2004), depending on the natural period and the targeted reliability. They have shown that the maximum seismic loads, as obtained by a complete quadratic combination of

five modal responses obtained by the specified design spectrum, match very well the corresponding values obtained with time series analyses. Results have been provided for HAWTs with different size.

Stamatopoulos (2013) has addressed the response of HAWTs to near-fault ground motions. He has investigated a 53.95 m tall turbine resting on a circular footing. A FE model of the tower and the circular footing has been used, with a lumped mass at the tower top modeling the RNA; SSI has been modeled by uncoupled nonlinear springs distributed below the footing. The response to near-fault ground motion has been investigated by three methods: a response spectrum method based on the elastic acceleration spectrum provided by the Greek Aseismic Code, suitably increased by 25 % to account for proximity to a seismic fault; a response spectrum method involving an elastic local acceleration spectrum built based on actual records for the project site; and a time history analysis using synthetized ground motions compatible with the elastic local acceleration spectrum. The two response spectrum analyses have been carried out on a FE model with the tower grounded by a linearly elastic rotational spring, with stiffness computed as the ratio of the bottom bending moment to the bottom rotation. An iterative procedure has been implemented, since the bottom bending moment and rotation depend on the seismic loading computed from the spectral acceleration; the latter depends on the first natural period, which depends in turn on the stiffness of the grounded rotational spring. Stamatopoulos (2013) has shown that the acceleration spectrum provided by the Greek Aseismic Code, although appropriately increased to account for proximity of seismic fault, significantly underestimates the shear and bending moment demand at the tower base by 55 %.

Nuta and coworkers (2011) have proposed a methodology to assess the probability of failure of a HAWT under seismic excitations in a parked state. Despite material linearity is generally prescribed for design, in fact, nonlinear behavior has to be taken into account when examining the potential failure mechanisms induced by seismic

excitations. For a HAWT with 1.65 MW rated power and 80 m hub height, they have built lognormally distributed fragility curves to estimate the probability of reaching a defined damage state. Parameters of the lognormal distributions have been obtained from nonlinear incremental dynamic analyses, assuming the magnification factor with respect to the design earthquake as intensity measure, and peak displacement, peak rotation, residual displacement, and peak stress as damage measures. For the specific sites of this study, two in Canada and one in the USA, no significant probability of failure has been found for the 1.65 MW HAWT under consideration (Nuta et al. 2011).

Full Models

In the last few years, a significant effort has been spent on developing advanced tools that may allow a full system modeling of HAWTs. Due to computational complexity involved in a FE modeling of all components of a HAWT, models with a limited number of degrees of freedom have been built, and, in general, a modal approach, a multi-body approach, or a combination of the two has been used. Many high quality full-modeling software packages are now available for the wind industry, such as GH BLADED (Bossanyi 2000) and FAST (Jonkman and Buhl 2005), developed at the United States National Renewable Energy Laboratory (NREL). A few comments on these packages are in order, especially with regard to the options available for seismic analysis.

GH BLADED uses a multi-body dynamics approach in conjunction with a modal representation of the flexible components like tower and blades (Bossanyi 2000). A fully coupled time-domain simulation is feasible, with wind and seismic loadings simultaneously generated. Two methods are available for simulating seismic loading. The first method allows recorded acceleration time histories to be used, while the second method uses an iterative procedure to synthetize acceleration time histories providing an elastic response spectrum that closely matches a specified design response spectrum. SSI can be also modeled. GH BLADED has been validated by GL for calculating operational loads associated

with typical load cases. A combined multi-body dynamics and modal formulation is adopted also by FAST(Jonkman and Buhl 2005), with flexible components modeled based on user-provided mode shapes. A fully coupled time-domain simulation can be implemented, with seismic loading generated as a user-defined loading imposed at the tower base. Like GH BLADED, FAST has been validated by GL.

Using BLADED, Witcher (2005) has compared the results from a response spectrum method and time-domain simulations as applied to a 60 m tall 2 MW turbine subjected to earthquake ground motion in both operational and parked cases. He has found that the elastic design spectrum with 5 % damping ratio yields a maximum bending moment at the tower base in a very good agreement with that computed by time-domain simulations, thus inferring that aerodynamic damping experienced by an operating turbine is quite close to 5 %. However, the results of the response spectrum method and time-domain analyses were very different in the parked case, with the first significantly underestimating the maximum tower base bending moment. This result has confirmed that in the parked case no aerodynamic damping is generated and that using the response spectrum method with a 5 % damping ratio does lead to nonconservative results. Although, in the specific case under examination, the bending moment demand due to earthquake loading in the parked case was lower than that due to earthquake loading in the operational case, Witcher (2005) has drawn the attention to the fact that, in some cases, the driving load can be that corresponding to earthquake loading in the parked case and has recommended further investigations on this issue.

Using FAST, one of the most comprehensive and fruitful studies on the seismic assessment of HAWTs has been carried out by Prowell and coworkers (2013). They have run a multiyear research program including extensive numerical simulations with FAST on HAWTs featuring different sizes and rated power (65 kW; 900 kW; 1,5 MW and 5 MW) and experimental tests on a HAWT (65 kW rated power, 22.6 m hub height, and a 16 m rotor diameter) mounted on the

Fig. 2 HAWT on the outdoor shake table at the University of San Diego, CA. The *arrows* indicate the direction of shaking (From Prowell et al. 2013)

Configuration 1
(FA shaking)

Configuration 2
(SS shaking)
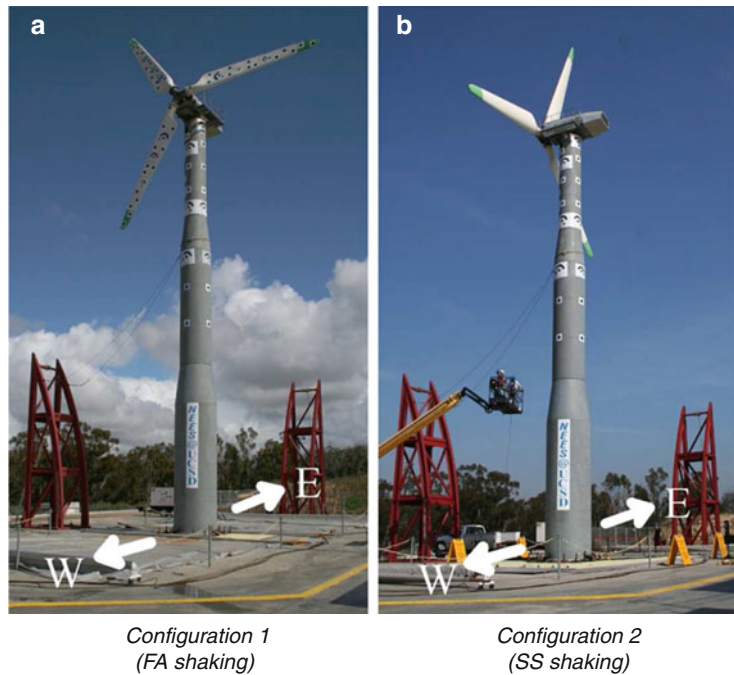
outdoor shake table at the University of California, San Diego (Fig. 2). Prowell and coworkers have collected a considerable amount of data, which has also served as a basis for ASCE/AWEA RP2011 prescriptions. Using a set of 99 ground motions with different magnitude and source-to-recording distance, they have run numerical simulations showing that the considered earthquakes may produce, in the 5 MW HAWT, a bending moment demand at the tower base well above that from extreme wind events. This result has been found for parked, operational, and emergency shutdown simulations and confirmed that seismic loads may be design driving for large turbines in regions of high seismic hazard.

As for what concerns the experimental tests on the 65 kW HAWT, only a few key results are reported here, but interested readers can find detailed information on both numerical and experimental findings in the study by Prowell and Veers (2009), Prowell et al. (2010, 2013), Prowell (2011), and references therein. Experimental tests were carried out in operational and parked states. In each state, shaking has been imparted in two directions, one parallel (FA =

fore-aft) and another perpendicular (SS = side-to-side) to the rotation axis of the rotor. Structural response characteristics have been recorded for motions imparted in both configurations and both operational states (Prowell et al. 2013).

The results have shown that for shaking imparted in the SS direction, no appreciable differences are encountered between operational and parked states, in terms of bending moment envelopes (Fig. 3b). In contrast, for shaking imparted in the FA direction, the bending moment demand at the tower base in the operational state was reduced by approximately 15–33 % from that while parked (Fig. 3a). This reduction of demand confirmed that in the operational state, aerodynamic damping has to be accounted for when performing a decoupled analysis, with separately generated wind loading and seismic loading (Prowell et al. 2013). However, due to the influence of many factors such as wind speed, earthquake magnitude, wind and earthquake relative directions, and SSI modeling, multiple simulations including likely distributions of wind speed and earthquake shaking have been recommended for an accurate quantification of the aerodynamic damping (Prowell et al. 2013).

S

Seismic Analysis of Wind Energy Converters, Fig. 3 Experimental bending moment envelope for three earthquake ground motions (From Prowell et al. 2013)
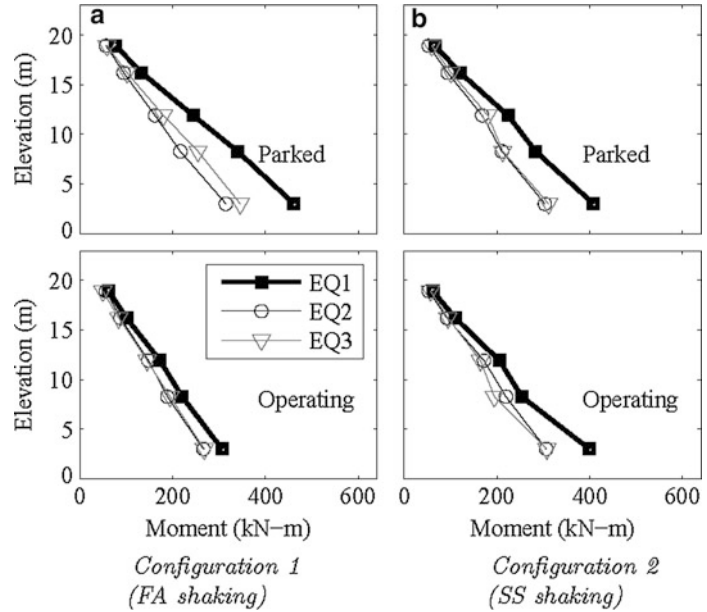
Seismic Analysis of Wind Energy Converters, Fig. 4 Experimental acceleration envelope for three earthquake ground motions (From Prowell et al. 2013)
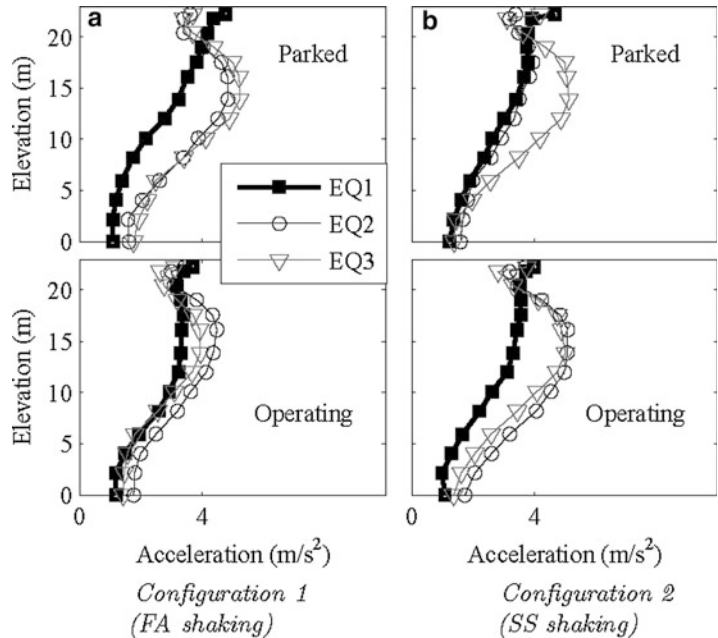
The same simulations should serve for a probabilistic description of the influence of seismic excitations on extreme loads.

A further important result obtained by Prowell and coworkers (2013) concerns the relative contributions of various tower modes. The maximum absolute acceleration envelopes in the FA and SS direction have shown that, in addition to the first mode, the second mode contributes significantly, as indicated by the high acceleration values at two thirds of the tower height (Fig. 4). This result has confirmed what predicted by other

Seismic Analysis of Wind Energy Converters, Fig. 5 SSI modeling for seismic assessment of HAWTs (From Zhao and Maisser 2006)
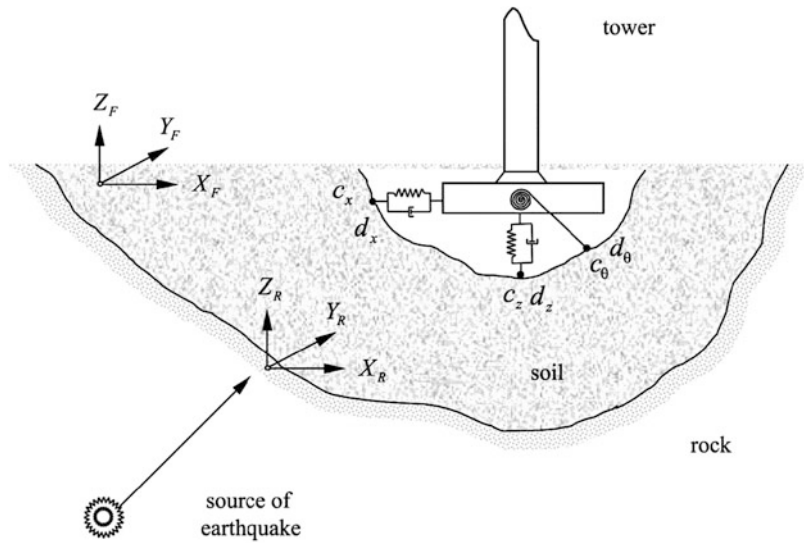
researchers (Haenler et al. 2006) on the importance of the second tower mode for estimation of seismic loads on large turbines. However, on the basis of the response PSDs, it has been found that amplification of energy imparted near the frequency of the second modes was not significantly influenced by operational state, thus implying that aerodynamic damping shall preferably be accounted for only in the first mode response (Prowell et al. 2013).

Besides the software packages used by the wind industry, there exist also a few full models of HAWTs, which have not been translated in software packages for the wind industry yet, but have provided very interesting results on the seismic response of HAWTs.

For instance, Zhao and coworkers (Zhao and Maisser 2006; Zhao et al. 2007) have developed a hybrid multi-body system (MBS) for full modeling of HAWTs. By this approach, the elastic tower is discretized into a series of rigid bodies coupled elastically by constraint joints and springs. The wind rotor is treated as a rigid disk; nacelle and machine carrier are coupled by an in-plane joint and treated as a rigid ensemble connected to the tower top through a revolute joint. The governing equations are derived using Lagrange's equations. This approach, though more mathematically rigorous, does not require external calculation of component mode shapes.

The MBS has been used by Zhao and Maisser (2006) to assess the seismic response of a 65 m high HAWT. SSI has been modeled approximately by a frequency-independent discrete parameter model, as a 3D set of uncoupled spring-damper devices, including translations and rotations (Fig. 5).

Seismic analysis has been carried out in operational conditions with a three-component weak real earthquake record. Numerical results have showed that while the top displacement is dominated by the wind thrust, force and bending moment in the longitudinal direction at the tower base are affected considerably by earthquake loads. In addition, in the lateral direction (where there are no wind loads), the force and bending moment are essentially decided by the earthquake loads and amplify several times.

Another interesting result concerns the lateral reaction force at the main bearing that is found to be significantly increased with respect to the case of no earthquake loading. This increase of the lateral reaction force is attributed to the gyroscopic forces that arise because of the change of the wind rotor direction, due to the earthquake loading. A further result of the study by Zhao and Maisser (2006) is that SSI modeling affects significantly the higher tower modes, especially the second lateral bending mode. In this respect, this result confirms the results obtained by Bazeos

S

et al. (2002) using a simplified model of the HAWT and shall be taken into account in consideration of the fact that especially the higher tower modes may fall within the region of maximum spectral acceleration (Haenler et al. 2006).

Using a multi-body system with flexible parts (tower, blades) described by a variable number of modes and including SSI modeling, Haenler and coworkers (2006) have investigated the seismic response of a HAWT with an 80 m rotor diameter and 60 m hub height, operating under a 13 m/s wind speed and subjected to an earthquake ground motion with a 0.3 g peak ground acceleration (PGA). They have showed that the full system model predicts modes at frequencies in the region of maximum spectral response acceleration for typical design response spectra. An important contribution of this study regards the relative increase in higher mode response. It has been found that, for normal wind loading, 80 % of the tower energy is associated with the first mode, while, during the considered earthquake, the energy in the first mode is reduced to 54 % percent only of the tower energy, thus concluding that higher tower modes are more important for earthquake loading than for typical wind loading.

Another full model of HAWT has been proposed by Ishihara and Sarwar (2008). Starting from the observation that, unlike wind loads, seismic waves may excite a wide range of frequencies including those of higher modes, they have developed a nonlinear FEM code (CAST) for a full FE modeling of the HAWT and its components. Beam elements with a linear material have been used to model tower and blades. Analyses performed by the FE code on HAWTs, in a parked state, have been used in conjunction with the Japanese BSL response spectrum (BSL 2004) to derive design formula for the prediction of seismic loads on two turbines. In particular, following a semi-theoretical codified method provided by the Japan Society of Civil Engineering (JSCE 2007), a profile of seismic loads (shear and moment demand) acting on a HAWT in parked conditions was estimated from a base shear built as the sum of a shear force due the first mode, obtained by using the BSL acceleration response spectrum, and the shear force

contributions due to higher modes, obtained on the basis of the FE analysis. Results have been provided for two HAWTs, one with 400 kW rated power, 36 m hub height, and 31 m rotor diameter and the other one with 2 MW rated power, 67 m hub height, and 80 m rotor diameter. By comparison with the results obtained from time-domain simulations, it has been shown that the BSL 5 % damped response spectrum provides nonconservative seismic load profiles for both the 400 kW HAWT and the 2 MW one, thus confirming that a 5 % damping ratio is not appropriate for HAWTs in parked conditions, which experience much lower damping level. It is worth remarking that the time-domain simulations carried out by Ishihara and Sarwar (2008) highlighted that contribution of higher modes may become significant for large HAWTs (2 MW) under earthquake excitations, confirming similar results obtained by other researchers (Haenler et al. 2006).

Summary

The importance of seismic loading in the design of HAWTs is recognized in existing ISs and CGs, and in seismically active areas attention shall be paid to the possibility that design is driven by load combinations involving seismic excitations.

Existing methods for seismic assessment of HAWTs are fully coupled time-domain simulations, computing the response to simultaneously acting wind loading and seismic loading, and decoupled analyses, where the responses to wind loading and seismic loading are built separately and then superposed. Fully coupled time-domain simulations, although complex and time consuming, remain the indispensable benchmark tool for an accurate seismic assessment of HAWTs. However, decoupled analyses are also important, especially in the early stages of design.

For the implementation of fully coupled time-domain simulations, a full system model must be employed. Simplified models, where the rotor is not modeled, are allowed to build separate seismic loads in decoupled analyses.

Despite the prescriptions of existing ISs and CGs, a few important aspects are actively being investigated, such as the influence of SSI modeling, the importance of higher modes in the seismic response, the potential failure mechanisms under earthquake loading, and the estimation of aerodynamic damping for decoupled analysis. A full understanding of these aspects is desirable for a correct seismic assessment of HAWTs.

Cross-References

▶ Earthquake Response Spectra and Design Spectra
▶ European Structural Design Codes: Seismic Actions
▶ Modal Analysis
▶ Response Spectrum Analysis of Structures Subjected to Seismic Actions
▶ Soil-Structure Interaction
▶ Stochastic Ground Motion Simulation
▶ Time History Seismic Analysis

References

API (2000) Recommended practice for planning, designing and constructing fixed offshore platforms, API RP-2A. American Petroleum Institute, Washington, DC

ASCE (2006) Minimum design loads for buildings and other structures, ASCE/SEI 7-05. American Society of Civil Engineers, Reston

ASCE/AWEA (2011) Recommended practice for compliance of large land-based wind turbine support structures, ASCE/AWEA RP2011. American Society of Civil Engineers/American Wind Energy Association, Reston/Washington, DC

Bazeos N, Hatzigeorgiou GD, Hondros ID, Karamaneas H, Karabalis DL, Beskos DE (2002) Static, seismic and stability analyses of a prototype wind turbine steel tower. Eng Struct 24:1015–1025

Bossanyi EA (2000) Bladed for windows user manual. Garrad Hassan and Partners, Bristol

BSL (2004) The building standard law of Japan. The Building Centre of Japan, Tokyo (in English and Japanese)

DNV/Risø (2002) Guidelines for design of wind turbines. Det Norske Veritas/Wind Energy Department, Risø National Laboratory, Copenhagen/Roskilde

Eurocode 8 (2004) Design of structures for earthquake resistance. Part 1: general rules, seismic actions and rules for building. EN 1998-1. European Committee for Standardization, Brussels

GL (2010) Guideline for the certification of wind turbines. Germanischer Lloyd, Hamburg

GWEC (2014) Global wind report: Annual market update 2013. Global Wind Energy Council, Brussels

Haenler M, Ritschel U, Warnke I (2006) Systematic modelling of wind turbine dynamics and earthquake loads on wind turbines. In: Proceedings of the European Wind Energy Conference & Exhibition (EWEC), 27 Feb–2 Mar 2006. Athens

ICC (2012) International building code. International Code Council, Country Club Hills

IEC (2005) Wind turbine generator systems. Part 1: safety requirements, 3rd edn, IEC 61400-1. International Electrotechnical Commission, Geneva

Ishihara T, Sarwar MW (2008) Numerical and theoretical study on seismic response of wind turbines. In: Proceedings of the European Wind Energy Conference & Exhibition (EWEC), 31 Mar–3 Apr 2008. Brussels

Jonkman JM, Buhl ML (2005) FAST user's guide, NREL/EL-500-38230. NREL, Golden

JSCE (2007) Guidelines for design of wind turbine support structures and foundations, Japan Society of Civil Engineers, Tokyo (in Japanese)

Kuhn M (2001) Dynamics and design optimisation of offshore wind energy conversion systems. Report no. 2001.002, Delft University Wind Energy Research Institute (DUWIND), Delft

Manwell JF, McGowan JG, Rogers AL (2010) Wind energy explained: theory, design and application, 2nd edn. Wiley, Chichester

Mulliken JS, Karabalis DL (1998) Discrete models for through-soil coupling of foundations and structures. Earthq Eng Struct Dyn 27:687–710

Nuta E, Christopoulos C, Parker JA (2011) Methodology for seismic risk assessment of tubular steel wind turbine towers: application to Canadian seismic environment. Can J Civil Eng 38:293–304

Prowell I (2011) An experimental and numerical study of wind turbine seismic behavior. PhD dissertation, University of California, San Diego

Prowell I, Veers P (2009) Assessment of wind turbine seismic risk: existing literature and simple study of tower moment demand, SAND2009-1100 report. Sandia National Laboratories, Albuquerque

Prowell I, Elgamal A, Uang C, Jonkman J (2010) Estimation of seismic load demand for a wind turbine in the time domain. In: Proceedings of the European Wind Energy Conference (EWEC), 20–23 Apr 2010, Warsaw

Prowell I, Elgamal A, Uang C, Luco JE, Romanowitz H, Duggan E (2013) Shake table testing and numerical simulation of a utility-scale wind turbine including operational effects. Wind Energy. doi:10.1002/we.1615

Stamatopoulos GN (2013) Response of a wind turbine subjected to near-fault excitation and comparison

with the Greek Aseismic Code provisions. Soil Dyn Earthq Eng 46:77–84

Swan S, Hadjian AH (1988) The 1986 North Palm Springs earthquake: effects on power facilities, NP-5607 research project 2848. Electric Power Research Institute (EPRI), Palo Alto

Umar AB, Ishihara T (2012) Seismic load evaluation of wind turbine support structures considering low structural damping and soil structure interaction. In: Proceedings of the European Wind Energy Association Conference (EWEA), 16–19 Apr 2012. Copenhagen

Witcher D (2005) Seismic analysis of wind turbines in the time domain. Wind Energy 8(1):81–91

Zhao X, Maisser P (2006) Seismic response analysis of wind turbine towers including soil-structure interaction. Proc Inst Mech Eng K J Multi-Body Dyn 220(1):53–61

Zhao X, Maisser P, Wu J (2007) A new multibody modelling methodology for wind turbine structures using a cardanic joint beam element. Renew Energy 32:532–546

Seismic Anisotropy in Volcanic Regions

Jessica H. Johnson
School of Earth Sciences, University of Bristol Wills Memorial Building, Bristol, UK
School of Environmental Sciences, University of East Anglia, Norwich, UK

Synonyms

Seismic anisotropy/shear wave splitting at (active) volcanoes; Shear wave splitting in volcanic regions; Using seismic anisotropy/shear wave splitting to monitor (active) volcanoes; Using seismic anisotropy/shear wave splitting to track/monitor/detect volcanic/magmatic activity

Introduction

Modern geophysical techniques enable changes to be observed at some volcanoes before magmatic eruptions: detection of seismicity (see "▶ Seismic Monitoring of Volcanoes") from magma pushing through cold country rock is one of the most common and successful monitoring techniques and can lead to short-term

forecasting (see "▶ Volcanic Eruptions, Real-Time Forecasting of"). Another example of geophysical precursors to eruptions is surface deformation from inflation or deflation of a volcano due to magma movement. However, some volcanoes do not display these clues, and there remains a need for techniques that are sensitive to other physical attributes that might change in conjunction with the eruption process. Any overpressured magma storage reservoir, be it a system of dikes, sills, conduits, spherical chamber, or a combination of these, will exert a stress on the surrounding country rock that may or may not be manifest as observable strain. Detecting and understanding this stress may be a key to predicting if and when a volcano will erupt.

Definition

Seismic anisotropy is the variation of seismic wave speed with direction. It is an indicator of geometric ordering in a material, where features smaller than the seismic wavelength (e.g., crystals, cracks, pores, layers, or inclusions) have a dominant alignment. This alignment leads to a directional variation of elastic wave speed. Seismic anisotropy can be measured using many seismological techniques but is most frequently observed using shear wave splitting, which requires clear shear wave energy recorded on three-component seismometers. Measuring the effects of anisotropy in seismic data can provide important information about processes in the Earth such as stress conditions, material flow, and the structure of the subsurface and deep Earth.

Historical Background

Hess (1964) first made a significant observation of large-scale anisotropy when seismic refraction measurements in oceans showed that the P-wave velocity of the upper mantle (Pn) was consistently higher for profiles recorded perpendicular to an oceanic spreading center (i.e., parallel to the direction of spreading or plate movement) than for profiles recorded parallel to the spreading center. These measurements were attributed to the alignment of olivine crystals in the mantle lithosphere because of flow during the formation
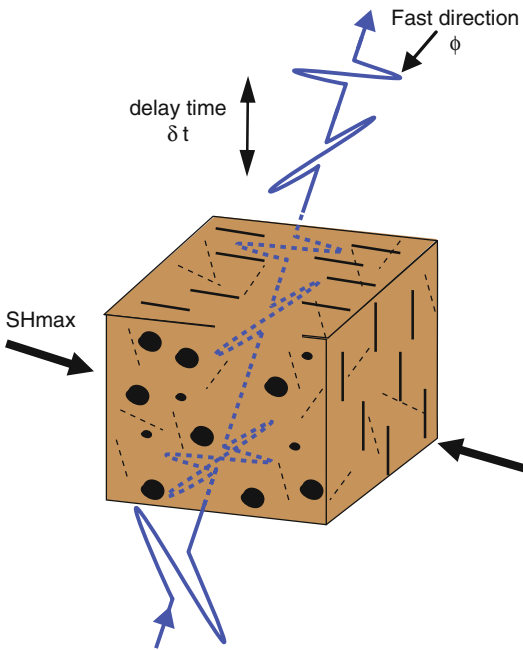
Seismic Anisotropy in Volcanic Regions, Fig. 1 Shear wave splitting in an anisotropic medium. Anisotropy is caused by preferentially aligned cracks due to a maximum horizontal compressive stress (S_{Hmax}). A vertically propagating shear wave that is arbitrarily polarized gets split into a fast wave with polarization (ϕ) parallel to crack alignment and a slow wave, which is polarized at 90° to ϕ. The waves are separated with delay time δt

of the oceanic plate at the ridge. Since the 1970s, improvements in computing power and memory and in seismic field observation have led to a greater understanding of the seismic anisotropy of the Earth at all levels and scales (Savage 1999). Measurements of seismic anisotropy have been used to detect fabric and stress in the Earth's crust, flow in the upper mantle, topography of the core–mantle boundary, and differential rotation of the inner core.

The measurement of seismic anisotropy has been found to be a proxy for determining the direction of maximum horizontal compressive stress (S_{Hmax}) in the crust; applied stress can cause microcracks to preferentially open parallel to the maximum compressive stress, creating an anisotropic medium with the fast direction parallel to S_{Hmax} (Fig. 1). The mechanism of aligned microcracks is thought to be the only one that allows seismic anisotropy to vary on observable

timescales, and temporal changes are traditionally interpreted as stemming from variations in the stress field due to large earthquakes or magmatic intrusions. There is mounting evidence, however, that the dominant mechanism for seismic anisotropy can switch between a static condition, such as aligned fractures in fault zones, and a dynamic process, such as stress causing aligned microcracks to dilate. In areas where there are strong changes in S_{Hmax} direction and magnitude on observable timescales, such as at active volcanoes, seismic anisotropy analysis has proven a valuable tool when combined with ground deformation or other seismological observations for interpretation of volcanic processes such as magma migration (e.g., Gerst and Savage 2004; Bianco and Zaccarelli 2009; Unglert et al. 2011).

Seismic Anisotropy

Shear Wave Splitting

Shear wave splitting occurs when a shear wave travels through a seismically anisotropic medium, i.e., one in which seismic waves travel faster in one direction or with one polarization than another (Fig. 1). In the Earth's upper crust, anisotropy is most likely to be caused either by stress conditions preferentially aligning microcracks parallel to the maximum compressive stress or by pervasive structural features. For a near-vertical propagation direction, the shear wave with the displacement in the plane of the open cracks will travel faster than that crossing the plane of the cracks, and so a fast shear wave with orientation ϕ and a slow shear wave orthogonal to ϕ, separated by a delay time δt, will be observed (Babuška and Cara 1991). Crack-induced anisotropy has in some studies been considered a direct indicator of present-day stress, with ϕ providing information about the orientation of S_{Hmax} and δt giving information about the strength of anisotropy and the amount of time that the wave spent traversing the anisotropic medium. Studies that combine independent stress estimation methods with shear wave splitting results have found that strong geological fabric or aligned structures rather than the maximum

S

stress can govern the observed ϕ in some situations. Both stress-induced and structure-related anisotropies have been observed in the same regions and, in some cases, have been detected at the same station (e.g., Zinke and Zoback 2000; Johnson et al. 2011).

The subject of temporally varying anisotropy is a hotly debated topic (e.g., Crampin and Peacock 2008). Temporal variations in shear wave splitting can take the form of a rotation in the fast direction of anisotropy, an increase or decrease in the delay time, or a combination of both. The changes are thought to stem from perturbations of the elastic properties of the crust due to crack opening either by local concentration of shear stress or by a change in pore-fluid pressure. Monitoring these changes can therefore provide crucial information about the state of stress and pore content around active volcanoes. However, spatial variations of anisotropy masquerading as temporal variations are one of the main caveats associated with the interpretation (Johnson et al. 2011). This can occur when the location of the source earthquakes changes or migrates throughout the duration of the observation period leading to the seismic waves experiencing different anisotropic conditions due to heterogeneity in the anisotropic structure. Temporal variations in anisotropy associated with major magmatic eruptions, such as the 1995/1996 eruptions of Mount Ruapehu, have been subject to doubts about the possibility of shear wave splitting variations being due to changing source earthquake location. Therefore, the methods must be fully understood before changes in anisotropy with time can truly be used as an eruption forecasting tool. To do this, one must first explore the spatial variations in anisotropy at volcanoes before examining the temporal changes.

Mt. Ruapehu volcano in the North Island of New Zealand has been the subject of some of the most comprehensive shear wave splitting studies of any volcano on Earth. Therefore, to illustrate the spatial versus temporal variation issue, anisotropy studies at Mt. Ruapehu will be explored. Temporal variation of shear wave splitting at Mt. Ruapehu was investigated by Miller and Savage (2001) and Gerst and Savage (2004).

Miller and Savage (2001) measured shear wave splitting from earthquakes in 1994 and 1998 and observed a change in the dominant ϕ spanning the time of the last major magmatic eruption in 1995/1996. That study was extended by Gerst and Savage (2004), who used the same techniques and an additional deployment of three-component seismometers in 2002 to observe further changes in ϕ (Fig. 2). The results of both studies were interpreted as being caused by a dike-shaped magma reservoir, or system of dikes, trending NE–SW. According to this model, the magma reservoir was pressurized before the eruption, producing a local stress field different from the regional stress field. This interpretation is favored above one in which dikes are intruded and solidify, causing a new structural anisotropy, because of the lack of detectable deformation and seismicity associated with dike intrusion. The model suggests that following the eruption the reservoir was less full and correspondingly less pressurized, meaning that the local stress returned to that of the surrounding region. The Gerst and Savage (2004) study suggested that the later changes in ϕ were due to repressurizing of the reservoir in response to an increase of magma in the system because ϕ from deep earthquakes displayed the regional trend, while ϕ from shallow earthquakes was oriented to the pre-eruption direction. The return of anisotropy to the pre-eruptive state also supports the interpretation of stress-controlled anisotropy as the intrusion of dikes would be a permanent change.

Johnson et al. (2011) investigated the spatial variations in anisotropy in more detail using a dense seismometer deployment in 2008, to compare future changes in anisotropy and to identify the regions and causes of past changes in anisotropy with more confidence. Johnson et al. (2011) divided the mapped anisotropy into regions in which the fast polarizations agreed with stress estimates from focal mechanism inversions, suggesting stress-induced anisotropy, and those in which the fast polarizations were aligned with structural features such as faults or metamorphic fabric, suggesting structural anisotropy. Using this benchmark of anisotropy,
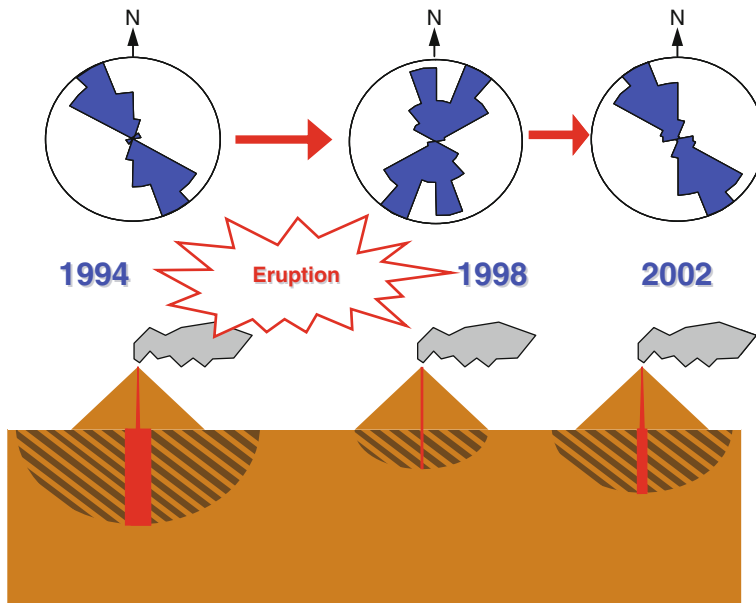
Seismic Anisotropy in Volcanic Regions, Fig. 2 Schematic stress and anisotropy model after Gerst and Savage (2004). In 1994, a pressurized dike system created a local stress field. In 1998, after the eruption, when the dike system was depressurized, stress directions partially returned to the regional trend. In 2002, the dike system refilled, and the stress field in the anomalous region was dominated by the dike again. The alignment of cracks was not as strong as in 1994, so the anisotropy in the anomalous region was not strong enough to affect fast directions from deep events

Johnson and Savage (2012) examined temporal changes in shear wave splitting from 1994 to 2010. They observed a region of strong anisotropy centered on Mt. Ruapehu in 1995, the time of a major magmatic eruption, agreeing with Miller and Savage (2001) and Gerst and Savage (2004), which was interpreted to be due to an increase in fluid-filled fractures during the eruption. They also observed strong anisotropy and a change in fast direction (~80°) at Mt. Tongariro in 2008, which was initially interpreted to be due to a change in the geothermal system but was later the location of a small eruption in 2012.

Measuring Shear Wave Splitting

Many methods have been developed to measure shear wave splitting in the Earth's crust (e.g., Crampin and Gao 2006). The goal is to identify the orientation of the fast split shear wave and the delay time between the fast and slow split shear waves. These parameters can then be used to infer rock properties such as crack distribution and geometry, pore content, or stress regime. Ideally, the procedure will be able to accurately process large quantities of three-component data in an efficient, unbiased, and objective way, without operator intervention, while providing quantitative evaluation of the uncertainties for each measurement of shear wave splitting.

A manual method for analyzing shear wave splitting is that of direct inspection of the 3D particle motion projected onto a horizontal plane (hodograms). This method relies heavily on observer judgment and therefore can produce biased results and is inefficient for large datasets. Other methods employ partially automated programs, where the user must pick a window around the shear wave arrival and/or evaluate the results. There are several approaches commonly used to semiautomatically determine shear wave splitting parameters: the covariance matrix technique searches for nonzero eigenvalues to "unsplit" the shear wave; the cross-correlation technique searches for the rotation

S

and time delay that yields the highest cross-correlation between orthogonal components; the aspect ratio technique searches for the rotation and time delay that yields the most linear particle motion with the maximum aspect ratio; and the vector amplitude technique uses the maximum amplitude in a time window to identify the split shear waves. Any degree of user interaction may introduce some subjectivity and is usually time-consuming, although many practices of auto-mated quality control will result in loss of significant amounts of data. Completely automated shear wave splitting analysis has proven elusive, as the step that is difficult to automate is the picking of the shear wave arrival. Several methods boast full automation other than this problematic step. In general, since phase arrivals are often picked for previous analyses such as hypocenter location, these automated methods are preferable to the ones with a lot of user interaction; more data can be evaluated, reducing the effect of loss of data through the quality control steps.

Interpreting Shear Wave Splitting: Some Assumptions

Most shear wave splitting observations are interpreted under some assumptions, which are rarely completely true in the Earth:

1. **That the medium possesses hexagonal symmetry with a horizontal axis:** The majority of anisotropic rocks in the Earth have, or can be approximated to have, hexagonal symmetry. This is because the most common symmetries have patterns that do not differ significantly for horizontal fast axis alignment and near-vertical incidence angles. The simplest models used to explain variations in two orthogonal directions are hexagonally symmetric models. Therefore, shear wave splitting is usually interpreted in terms of transverse anisotropy with a horizontal symmetry axis.

2. **That the anisotropic medium is in a single, homogeneous layer:** When a shear wave passes through multiple anisotropic layers, the observed splitting parameters depend strongly on the thickness and strength of anisotropy of the layers, on the relative fast directions, and on the wavelength of the wave. If the shear wave has been sufficiently split in the first layer that the fast and slow waves are separated, then when it enters the second layer, which has a fast direction $20-70°$ different, both of the quasi-shear waves will be split again. In this case both waves will now have the fast and slow directions of the second layer. However, when the splitting from the first layer is weak so that the two quasi-shear waves are not more than one wavelength apart, both waves are still resplit, but the result is a complex waveform that is difficult to interpret but can still be meaningful. In general, shear wave splitting fast direction is representative of the last layer that the wave passed through and the delay time can be approximated as accumulating along the path.

3. **That the anisotropic medium is localized beneath the receiver:** As seen in the previous point, in general, the anisotropy parameters are representative of the last anisotropic medium that the wave travelled through. Therefore, the assumption that the medium is local to the station is usually justified. However, in regions with heterogeneous anisotropic structure, data at a single station will not be consistent and the assumption will not be appropriate.

Further discussions of these points can be found in Silver and Savage (1994), Savage (1999), and Johnson et al. (2011).

Other Estimates of Seismic Anisotropy

Shear wave splitting analysis is becoming a very popular method for determining seismic anisotropy in the crust. This is in part due to the abundance of data and methods available and also to the relative insensitivity to the source–receiver geometry (other than deep enough earthquakes for the rays to arrive within the shear wave window) and independence from the need for dense networks of seismometers (see "▶ Seismic Network and Data Quality"). Another benefit of

shear wave splitting analysis is that, even though the results can be averaged over multiple measurements, they represent a snapshot of anisotropy. There are, however, other methods for determining seismic anisotropy. Most of the other methods involve calculating anisotropic velocities using tomographic techniques (see "▶ Seismic Tomography of Volcanoes") with body or surface waves from active or passive sources, or using ambient seismic noise (see "▶ Noise-Based Seismic Imaging and Monitoring of Volcanoes" and "▶ Seismic Noise").

Tomographic methods (see "▶ Seismic Tomography of Volcanoes") require larger amounts of data with relatively even coverage of sources and sensors, which (usually) take longer to acquire, thereby rendering the methods less practical for time-lapse investigations than shear wave splitting investigations. Tomographic methods that are used to calculate seismic anisotropy are different to techniques that conduct tomographic inversions on shear wave splitting data. Anisotropy parameters are derived from the inversions in the former case, while the latter is an inversion of the anisotropy parameters.

Inversions of body wave arrival times for three-dimensional velocity structures are common practice at volcanoes using teleseismic waves, local earthquake sources, and active seismic sources. These inversions can also account for 3D V_p azimuthal anisotropy, which is parameterized with a percent anisotropy and an orientation of the fast axis. Seismic anisotropy can be detected using inversions of Love and Rayleigh surface waves from large earthquakes in the same manner.

Surface waves constructed from cross-correlations of ambient seismic noise can be inverted for 3D seismic velocity structure. Seismic anisotropy from ambient noise tomography can be calculated. These calculations are different from the time-lapse studies that detect temporal variations in isotropic seismic velocities using ambient noise interferometry (see "▶ Tracking Changes in Volcanic Systems with Seismic Interferometry").

Variations in Shear Wave Splitting

Spatial Variation of Shear Wave Splitting

Most shear wave splitting results are plotted as rose diagrams (circular histograms) of fast direction at the station at which the measurements were made (e.g., Fig. 2). This implicitly assumes that the anisotropy is localized beneath the station. In many regions this may be appropriate; however, when there is lateral heterogeneity, the rose diagrams will become scattered or multimodal. Johnson et al. (2011) detected distinct splitting parameters for measurements using earthquakes from several different clusters in the region of Mt. Ruapehu. This backazimuthal dependence implies that the ϕ value obtained from shear wave splitting analysis is highly dependent upon the path that the ray takes, which has also been found in other regions. Furthermore, this suggests that the anisotropy changes over shorter distances in the crust than is often observed for mantle anisotropy and that averaging ϕ over the whole region may not be appropriate. If the causes of the different regions of anisotropy are known, it is easier to identify and map the differences using shear wave splitting analysis. If the crustal stresses or fabrics are more complex or unknown, then it is more difficult to map and interpret heterogeneous anisotropy, although a denser array of seismic stations and a broad range of backazimuths increase the likelihood of identifying the source of heterogeneity.

For time-lapse studies, it is important to mitigate the chance of spatial variations of shear wave splitting being erroneously interpreted as temporal variations. One way to do this would be to map spatial variations in detail at a time when there are no hypothesized temporal changes. This becomes difficult, but essential, when the anisotropy is very heterogeneous and the seismograph array is less dense than the spatial changes. This section outlines some examples of shear wave splitting tomography studies, which attempt to solve the problem of spatial variation.

Shear wave splitting tomography is difficult because of the nonlinear effect of multiple layers
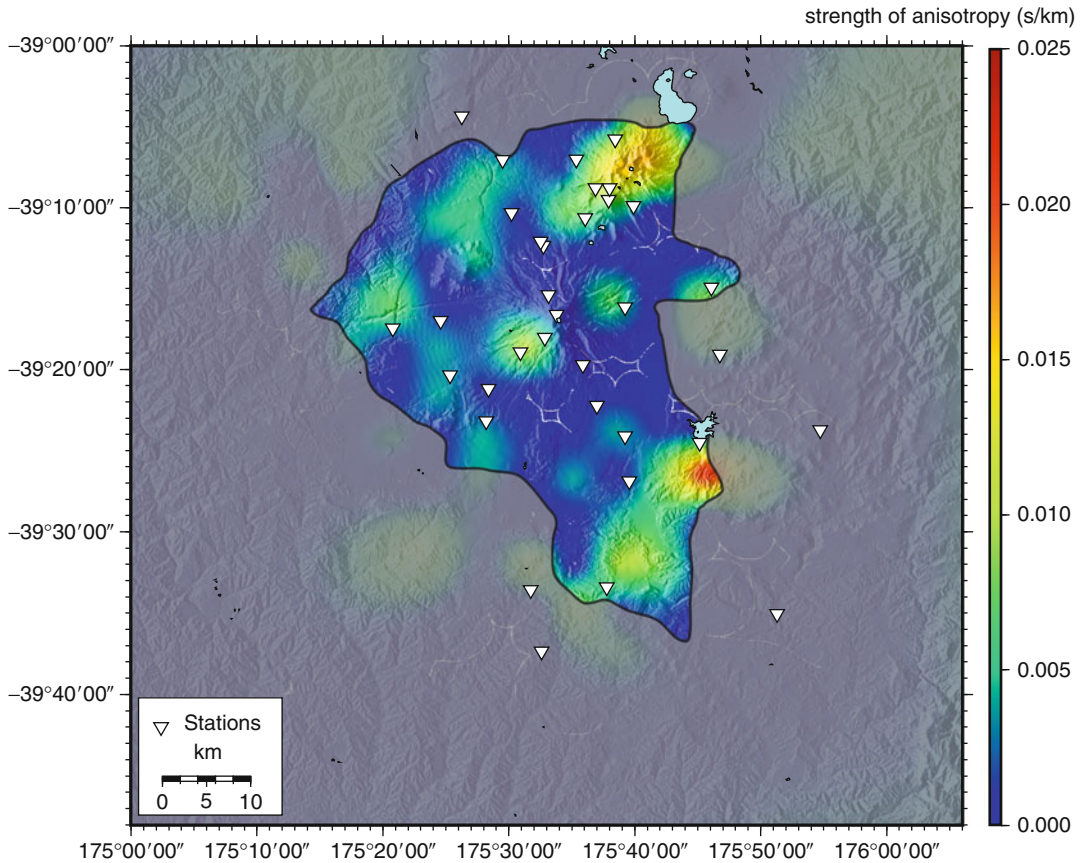
Seismic Anisotropy in Volcanic Regions, Fig. 3 Delay time tomography from the inversion of shear wave splitting data at Mount Ruapehu Volcano, New Zealand. *Warm colors* indicate strong anisotropy, *shaded area* shows the limit of statistical significance calculated from the model variance matrix, *white inverted triangles* are seismic stations (After Johnson et al. (2011))

of anisotropy on waveforms. This is different to travel time tomography, in which the travel time of a wave is often linearly related to the velocity structure of the media it has passed through. Because of this difficulty, many techniques treat ϕ independently from δt. Three-dimensional tomographic inversions can be carried out on the delay time data to characterize fracture density distribution. This tomography uses only δt from local earthquakes to investigate anisotropy strength in the crust. In this way, regions of high anisotropy can be identified, but information regarding fast directions is not accounted for.

Audoine et al. (2004) presented a simple method of 2D spatial averaging to examine heterogeneous anisotropy in the crust. A grid was constructed with nodes regularly spaced between each earthquake and station. This grid was then treated as a new dataset, and ϕ for each node within a polygon, or within a box of a regular lattice, was averaged. This created average ϕ values at regular intervals that could be denser than the station spacing, hereby identifying spatial trends in fast directions, but not accounting for delay times. Johnson et al. (2011) adapted a combination of the above methods to map the heterogeneous seismic anisotropy field around Mt. Ruapehu in New Zealand (Fig. 3). To constrain the locations of high anisotropy, they conducted a two-dimensional tomographic inversion on the delay time estimates. They then used a spatial averaging technique similar to that of

Audoine et al. (2004), but with the fast polarizations weighted according to the strength of anisotropy calculated from the tomography. The method uses a quad-tree gridding system to enable higher resolution where the data permit and couples the two shear wave splitting parameters, even though they are not used in a joint inversion.

Abt and Fischer (2008) carried out full 3D shear wave splitting tomography for the mantle. The method parameterized the mantle as a 3D block model of crystallographic orientations. Nonlinear properties of shear wave splitting were accounted for by applying the inversion iteratively and recalculating partial derivatives after each iteration. Using this method, Abt and Fischer (2008) modeled an idealized subduction zone with uniform stations and sources. When applied to real data, they found that the geometry of stations and observed seismicity in the Nicaragua–Costa Rica subduction zone yielded partial to good resolution. This method also has the potential to be applied to crustal studies such as active volcanoes.

Temporal Variation of Shear Wave Splitting in the Crust

The temporal variation of shear wave splitting and its interpretation is highly controversial. The main point of dispute is whether the accumulation of stress before significant events such as eruptions or large earthquakes allows the time and magnitude of impending events to be stress forecast through shear wave splitting monitoring. The effects have been reported with hindsight before at least 15 earthquakes. Crampin and Peacock (2008) present a summary of observations of temporal variations in shear wave splitting attributed to stress-aligned fluid-saturated microcracks. However, alternative interpretations of these observations of temporal changes suggest there may be observer bias in data selection, unsound statistical analyses, misinterpretation of spatial variation, and lack of correlation with other stress-determining factors/correlation with structural evidence.

Clear evidence has been obtained that crustal shear wave splitting can vary over short distances and can be caused by structural features in the crust, which would not change with changing stress. However, it is also clear that anisotropy due to stress-aligned fluid-saturated microcracks can change with time, as has been demonstrated in industry when small changes associated with injection and removal of fluids from reservoirs were examined. The use of similar earthquakes (i.e., those that have the same source mechanism and location) for shear wave splitting analysis helps to eliminate some of the discrepancies with interpretation. Johnson et al. (2010) used shear wave splitting analysis, a multiplet of 25 similar earthquakes and double-difference relocation to examine temporal variations in seismic properties prior to and accompanying magmatic activity associated with the 2008 eruption of Okmok Volcano. They found a general change in ϕ but could not rule out dependence on backazimuth and no significant change using the multiplet. Using earthquakes originating from the subducted slab in order to reduce the effect from changing paths, several modes of ϕ were identified, relating to the anisotropy of the mantle wedge, regional stress direction, and local stress induced by the pressurization and depressurization of the magma reservoir. These modes were found to have different prominence at different times throughout the eruptive cycle.

Once again returning to the example of Mt. Ruapehu, Miller and Savage (2001) and Gerst and Savage (2004) observed general changes in seismic anisotropy throughout the eruptive cycle. Johnson et al. (2011), however, showed that seismic anisotropy around Mt. Ruapehu is heterogeneous and that averaging the whole region is not appropriate. Therefore, Keats et al. (2011) used seismicity generated from a consistently active area of seismicity about 20 km to the west of Mt. Ruapehu. Shear wave splitting results revealed a decrease in delay time in the 2006–2007 eruption period and a significant variation in the fast shear wave polarization in the same time period (Fig. 4). These changes were attributed to an increase in pore-fluid pressure in the region due to fluid movement, and it was suggested that this fluid movement may be associated with the eruptions in 2006 and 2007.
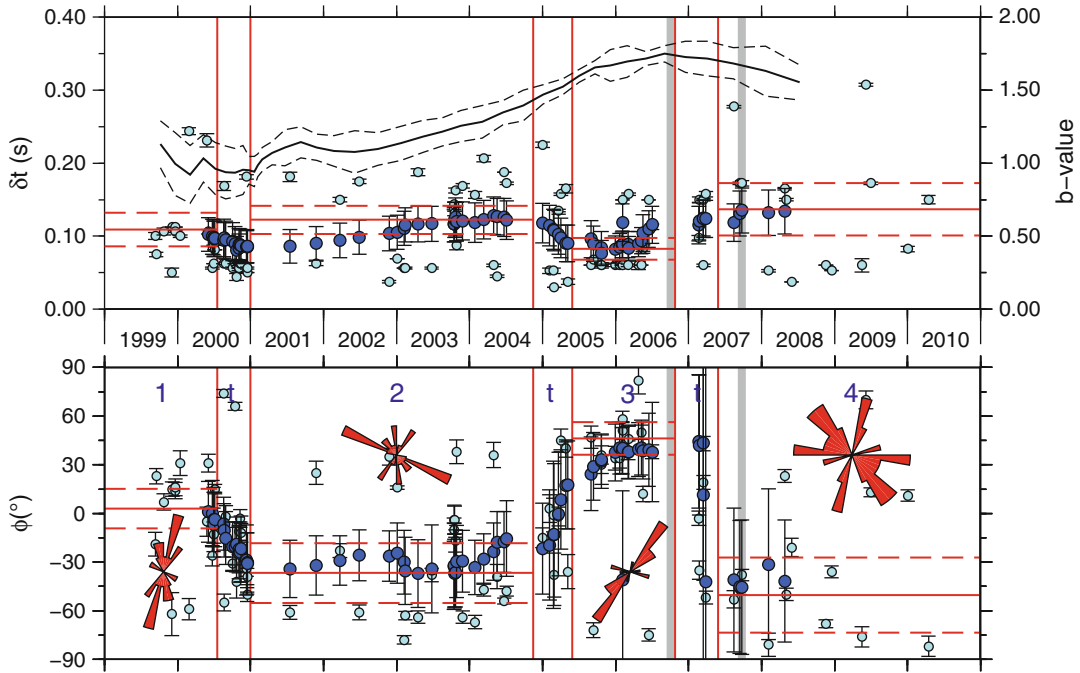
Seismic Anisotropy in Volcanic Regions, Fig. 4 Moving average plot of fast polarization (ϕ) and delay time (δt) using earthquakes within the Erua swarm at station FWVZ at Mount Ruapehu, New Zealand. Individual measurements for ϕ and δt are displayed in *light blue* and 10-point moving averages are displayed in *dark blue*. The error bars indicate 95 % confidence intervals. The four time periods, marked by the numbers 1–4, and three transition zones, marked by a t, are indicated with vertical *red lines* and the mean for each period are shown by the *red horizontal bars* with 90 % confidence interval (*dashed red lines*). The times of the two phreatomagmatic eruptions that occurred are also marked with *grey bars*. Rose diagrams of ϕ are displayed in their respective time periods. The b-value for the Erua swarm catalogue is also plotted against time in *black* at the *top* using a window of 40 events and an eight-event overlap. *Dashed black lines* indicate 95 % confidence interval (After Keats et al. (2011))

Temporal variation of shear wave splitting parameters as a stress indicator has been employed at several other volcanoes. While most studies concentrate on either stress-induced anisotropy or structurally dominated anisotropy, several studies have found that the dominant mechanism can change between the two (Keats et al. 2011; Johnson and Poland 2013) and that more than one mechanism can be dominant over short distances, potentially creating an apparent temporal change that is actually an artifact of changing earthquake location (Zinke and Zoback 2000).

Even though we have seen that shear wave splitting analysis can be used as an indicator of stress and of fluid saturation in the crust, surprisingly few studies have been conducted on shear wave splitting around volcanoes (Fig. 5). This is due, in part, to the generally noisy waveforms and complicated interpretation of such observations when taking into account heterogeneity and complex stress regimes. It should be noted that there is significant literature about shear wave splitting in the mantle beneath active volcanoes because these regions are invariably of interest tectonically, but this entry focuses on crustal studies only.

Relating Shear Wave Splitting to Other Observations

Shear wave splitting analysis has been proven to be a useful indicator of stress and of fluid
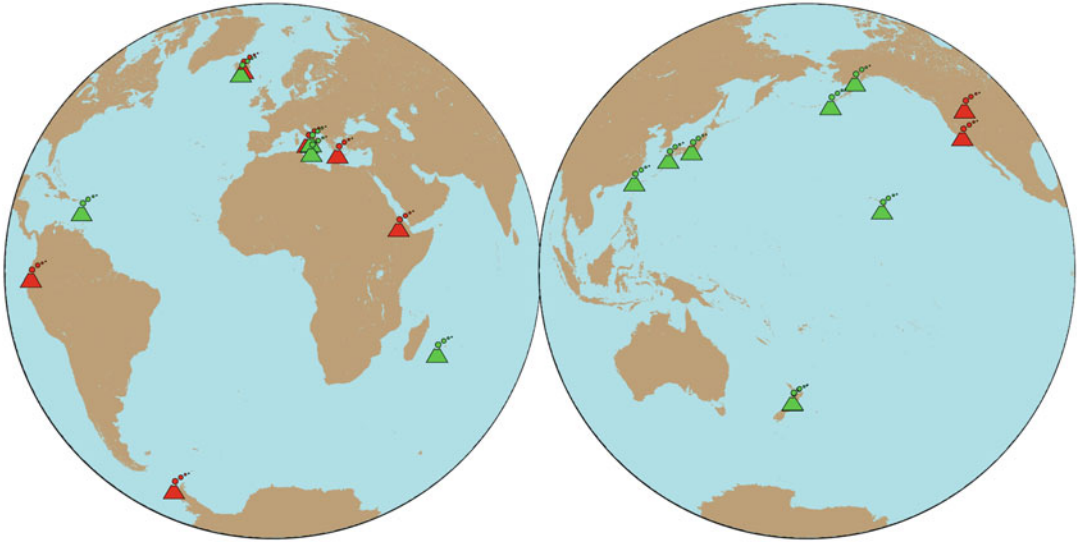
Seismic Anisotropy in Volcanic Regions, Fig. 5 Map showing volcanoes at which shear wave splitting investigations have been carried out. *Green volcano symbol* indicates that temporal variations were observed, *red* indicates that no temporal variations were observed

saturation in the crust but can also be caused by structures such as macroscopic fractures or mineral alignment. At volcanoes, the generally noisy waveforms, complicated heterogeneity, and complex stress regimes can make interpretation of shear wave splitting difficult. For this reason, shear wave splitting investigations are often coupled with other stress or strain indicators and structural geology data to minimize the ambiguity in the interpretation.

Volcanic Seismicity, Fault Plane Solutions, and b-Value

Studies of shear wave splitting in areas of developed industry may use direct indicators of the stress field in the crust, such as borehole breakouts, for comparison with shear wave splitting results. However, few active volcanoes have these clues, so other stress monitoring techniques must be employed.

The occurrence of volcano-tectonic (VT) (see "▶ Volcano-Tectonic Seismicity of Soufriere Hills Volcano, Montserrat") earthquakes is an immediate indicator of differential stress in the crust around a volcano. Changes in the stress field may trigger more VTs, the detection of which is one of the most successful tools of seismic volcano monitoring. For this reason, many shear wave splitting studies have been designed to investigate temporal changes in anisotropy during a volcanic crisis (e.g., Savage et al. 2010; Johnson et al. 2010; Roman and Gardine 2013). Bianco and Zaccarelli (2009) calculated a background seismic anisotropy at Mount Vesuvius of 4 %, but at times of seismic crisis, the average anisotropy was 8 % and the fast directions tended to flip by 90°.

As noted above, however, heterogeneity in the anisotropic media must be carefully considered, especially when there is a migration of seismicity associated with magma movement. These migrating earthquakes are used as sources of shear wave energy, but temporal variations observed in the anisotropy measurements could be due to spatial variations.

VT earthquakes contain additional information about the state of stress in the crust. Fault plane solutions (FPS) provide information about the orientation of the fault plane that slipped, as well as the slip vector for each earthquake. From the FPSs, the pressure and tension axes can be identified, which give information about the stress regime. The FPS P-axis azimuth, which represents the orientation of the principal

S

compressional axis of the moment tensor, may in reality differ significantly from the orientation of maximum compressive stress, depending on the orientation of preexisting planes of weakness (faults). This is sometimes countered by calculating large numbers of FPS P-axis azimuths to highlight the major trends that most likely represent the actual orientations of S_{Hmax} (e.g., Roman and Cashman 2006; Roman and Gardine 2013). Another step can be taken by inverting for the full stress tensor using the FPS results (e.g., Johnson et al. 2011). The true direction of S_{Hmax} can then be compared with shear wave splitting parameters to jointly interpret the results.

The orientations of FPS of VT earthquakes have been observed to display systematic changes related to episodes of magmatic activity at several volcanoes (Roman and Cashman 2006). The pressure axes of FPS have been observed to rotate orthogonal to the dominant regional stress orientation in some cases, indicating that these earthquakes may result from dike inflation in the direction of minimum compressive stress (Roman and Cashman 2006). Roman and Gardine (2013) investigated shear wave splitting and earthquake fault plane solutions at Redoubt volcano around the time of the March 2009 eruption. They found that an approximately ~ 90° change in the polarization of shear wave splitting fast polarization accompanied the earliest signs of seismic unrest in 2008 and continued through the eruption before diminishing in 2009. A similar change in the orientation of fault plane solutions occurred 18–48 h prior to the eruption onset on March 23, 2009, but almost 2 months after a strong increase in the rate of shallow VT earthquakes. The shear wave splitting and FPS results suggested a protracted period of slow magma ascent followed by a short period of rapidly increasing magma pressurization beneath the volcano. Roman and Gardine (2013) found that after the eruption, both shear wave splitting fast polarization and fault plane solution pressure axes had a direction more consistent with the regional stress than before the eruption.

The b-value of a cluster of earthquakes is often used to describe the size distribution. The b-value comes from the Gutenberg–Richter law and is a frequency magnitude (see "▶ Frequency-Magnitude Distribution of Seismicity in Volcanic Regions") relation. For crustal earthquakes, b-values are typically ~ 1 for tectonic earthquakes, though they tend to be higher in volcanic regions (see "▶ Frequency-Magnitude Distribution of Seismicity in Volcanic Regions"). The b-value has been related to physical properties such as stress, material homogeneity, and pore pressure and can therefore be useful in comparison with seismic anisotropy measurements to determine changes in physical properties around volcanoes. Keats et al. (2011) used seismicity from a discrete cluster of earthquakes near Mt. Ruapehu in New Zealand to compare shear wave splitting and b-values over 12 years (Fig. 4). The joint analysis allowed them to attribute temporal changes to an increase in pore-fluid pressure due to fluid movement associated with eruptive activity.

Attenuation, V_p/V_s Ratio, and Noise Cross-Correlations

Perturbations of the elastic properties of the crust around volcanoes are often linked to crack opening either by local concentration of shear stress or by an increase of pore-fluid pressure. Shear wave splitting analysis is one way to monitor these properties, but there are other techniques that can provide additional information about the state of stress or pore content. Earthquake coda (the part of a high-frequency seismogram following the P and S waves) is generated by random scattering processes in the crust. The coda decay parameter (or Q_c^{-1}) is often used to parameterize characteristics of the medium, and temporal changes in Q_c^{-1} have been linked to crack opening (Del Pezzo et al. 2004).

Increased gas content in pores and cracks has the effect of lowering P-wave speed due to the higher fluid compressibility, but not significantly affecting the shear modulus and hence S-wave speed. The ratio between P-wave velocity (V_p) and S-wave velocity (V_s), V_p/V_s, is therefore useful for characterizing pore-fluid content. Gas-enriched pore space has been reported to affect V_p/V_s above magmatic intrusions and has

been combined with shear wave splitting studies to characterize pore content at volcanoes (Unglert et al. 2011; Johnson and Poland 2013).

The perturbations of the elastic properties of the crust around volcanoes caused by changes in stress or pore fluid can be monitored continuously using cross-correlations of ambient seismic noise (see "▶ Noise-Based Seismic Imaging and Monitoring of Volcanoes" and "▶ Seismic Noise"). Monitoring of Q_c^{-1}, V_p/V_s ratio, and isotropic velocity from seismic noise cross-correlations therefore naturally complements shear wave splitting investigations as the mechanisms of change are so similar (Del Pezzo et al. 2004).

Deformation

Geodetic techniques that determine changes in strain, such as GPS and InSAR, can provide some of the stress tensor elements and can therefore be useful to compare with shear wave splitting investigations. Studies such as those by Savage et al. (2010) and Unglert et al. (2011) use strain data to determine that the anisotropy is due to stress-oriented microcracks in the upper crust. At Mount Asama, Savage et al. (2010) correlated GPS baseline length measurements with shear wave splitting measurements to analyze stress changes accompanying the eruption in 2004. They found that the best model from the GPS analysis of a vertical dike and conduit also fit the shear wave splitting measurements, as did the temporal variations. From this, a crack aspect ratio of 2.6×10^{-5} was calculated and a differential horizontal stress of 6 MPa at 3 km depth was inferred. Unglert et al. (2011) performed shear wave splitting analyses on local earthquakes around Aso Volcano between 2001 and 2008 and compared the results to strain from GPS measurements in the area. They observed, using clusters with relatively stable epicenters, that two stations showed a significant change in ϕ in 2004–2005. Models from seismic tomography and receiver functions were found to fit both the anisotropy and strain measurements. Other studies have used the observations of no correlation between geodetic measurements and shear wave splitting results to infer that the mechanism of anisotropy is structurally controlled. An anticorrelation between geodetic measurements and shear wave splitting results was observed at Kīlauea Volcano, implying that some volcanic process was affecting both measurements but that the mechanism was different to the traditional interpretation of a pressurizing magma reservoir. Johnson and Poland (2013) used these observations, combined with V_p/V_s ratio to interpret these changes as being due to increased gas flux.

Geology

Numerous investigations have found that in some regions, the fast direction of anisotropy is oblique to the direction of maximum horizontal compressive stress. In these cases, other clues for the cause of anisotropy are sought and the anisotropy is frequently found to align preferentially with macroscopic structures at the surface. Macroscopic fractures are often aligned because they are caused by faulting in a regional stress field. These macroscopic structures also cause anisotropy and shear wave splitting that often aligns with S_{Hmax}. Exceptions arise when strike-slip faults initially align at $45°$ to the direction of S_{Hmax}, faults are a product of a paleostress, or the rock has been deformed since the faulting. Velocity anisotropy is also strongly dependent on rock fabrics and metamorphic rocks with distinct foliations can have anisotropies of up to 20 % even in the absence of cracks (Babuška and Cara 1991, and references therein). This is mainly due to the preferred alignment of intrinsically anisotropic minerals such as biotite and hornblende. Geological textures can be identified through analysis of in situ rocks or oriented drill cores. Shear wave splitting that is caused by structural features is unlikely to change over time periods of investigation. However, due to changing conditions in stress or pore content, the dominant mechanism of anisotropy can change, rendering a temporal variation in shear wave splitting parameters. Johnson and Poland (2013) investigated shear wave splitting changes at Kīlauea Volcano associated with the onset of the summit eruption in 2008. They found that the orientation of fast shear waves at Kīlauea was usually controlled by structure, but in 2008 showed changes

with increased SO_2 emissions preceding the start of the summit eruption.

Tectonic Versus Magmatic Stress-Controlled Anisotropy

Studies that have recorded rotations in the fast direction of anisotropy due to a localized perturbation in stress often relate the background stress regime to the regional tectonics of the area. Rotations of ϕ by 90° are commonly observed, but several mechanisms for the rotation have been proposed. The 90° rotations should be treated carefully because cycle skipping, when the match of the fast and slow waveforms has a factor of T/2 ambiguity where T is the dominant period, can lead to an error of 90° in the recorded fast direction and a false delay time. When the rotations are real, Crampin et al. (2002) suggest that they are "flips" caused by wave propagation through cracks containing fluids at high pore-fluid pressures. However, a 90° rotation would also be expected due to a dike intrusion (Gerst and Savage 2004). A dike will exert pressure on the surrounding rock, generating a local stress field that is superimposed on the regional stress field. The stresses of such an elongated structure are mainly oriented perpendicular to the strike axis. When the pressure in the dike system is high enough, the generated stress field locally reorients the principal stresses as well as the local crack alignment. The pattern of stress perturbation around an intrusion can be complex, however, and the interaction with the regional, or tectonic, stress has been the subject of several studies (e.g., Roman and Cashman 2006; Vargas-Bracamontes and Neuberg 2012). Vargas-Bracamontes and Neuberg (2012) found that in the presence of a dominant regional stress field, the stress perturbation from an intrusion will be negligible. As the pressure from the intrusion increases, both the regional and the local stress field will coexist. This phenomenon was observed at Okmok Volcano, Alaska, when the shear wave splitting results displayed a mode corresponding with the regional stress direction and one corresponding with the pressurization of the magma reservoir (Johnson et al. 2010). As the pressure is progressively increased,

Vargas-Bracamontes and Neuberg (2012) found that the stress patterns gradually approach those corresponding to the absence of a regional stress field. Therefore, in cases with extremely high magma pressures, such as before large explosions at Soufriere Hills Volcano, Montserrat, the regional stress field may be omitted in numerical models.

The Future of Seismic Anisotropy in Volcanic Regions

Using seismic anisotropy to measure stress at active volcanoes holds enormous potential as a monitoring and eruption forecasting tool. Changes in seismic anisotropy associated with volcanic activity have already been detected; however, interpretation of seismic anisotropy observations is usually qualitative, and researchers struggle to use the results to quantify the magnitude of stress variations and the cause. Eruption forecasting is increasingly evolving from empirical pattern recognition to forecasting based on models of the underlying dynamics. For shear wave splitting monitoring to be beneficial, physical models must be developed that can explain not only changes in general trends but also the scatter in the data and anomalous observations. Central to this are the links between stress changes and fracture or crack compliance in the country rock and the role that fluids in cracks play, particularly the effects of hydrothermal circulation, pore pressure, and gas flux. Another limitation is that the majority of current research seeks to interpret only ϕ or only δt observations, but the two parameters are inherently linked. Models used to explain shear wave splitting observations must predict both parameters.

To quantify the response of seismic anisotropy to pre-, co-, and post-eruption subsurface magma movement, shear wave splitting data should first be used to map the areas affected by changing stress and identify the mechanism of seismic anisotropy in areas that do not have stress-controlled anisotropy. Methods have recently been developed for the inversion of

geomechanical parameters such as crack size, geometry, density, and content (e.g., Wuestefeld et al. 2012). Monitoring of these properties over time will not only indicate the occurrence of changes but will elucidate the nature of the subsurface changes.

Increasingly realistic numerical and experimental models of the fluid dynamics and elastodynamics are becoming possible. Integration of multidisciplinary data within these models will enable greater understanding of the underlying mechanisms and may be used to calculate the shear wave splitting and other important parameters for different volcanic scenarios. Ultimately, it may be possible to use information contained in seismic anisotropy to monitor subsurface magma movement and forecast changes in a volcano's behavior by establishing the characteristic stress field response for a given volcano, or through a deeper understanding of the complex relationships between seismic anisotropy, local crustal stresses, and the physical mechanisms of magma migration.

Summary

Seismic anisotropy is the variation of seismic wave speed with direction. Shear wave splitting occurs when a shear wave travels through a seismically anisotropic medium. Temporal variations in shear wave splitting can take the form of a rotation in the fast direction of anisotropy, an increase or decrease in the delay time, or a combination of both. The measurement of seismic anisotropy has been found to be a proxy for determining the direction of maximum horizontal compressive stress in the crust; applied stress can cause microcracks to preferentially open parallel to the maximum compressive stress, creating an anisotropic medium with the fast direction parallel to the maximum horizontal compressive stress. The mechanism of aligned microcracks is thought to be the only one that allows seismic anisotropy to vary on observable timescales, and temporal changes are traditionally interpreted as stemming from variations in the stress field due to large earthquakes or magmatic intrusions.

The changes are thought to stem from perturbations of the elastic properties of the crust due to crack opening either by local concentration of shear stress or by a change in pore-fluid pressure. There is mounting evidence, however, that the dominant mechanism for seismic anisotropy can switch between a static condition, such as aligned fractures in fault zones, and a dynamic process, such as compressive stress causing aligned microcracks to dilate. In areas where there are strong changes in maximum compressive stress direction and magnitude on observable timescales, such as at active volcanoes, seismic anisotropy analysis has proven a valuable tool when combined with ground deformation or other seismological observations for interpretation of volcanic processes such as magma migration. By inverting shear wave splitting data for geomechanical parameters and integrating the results with numerical and experimental models, it may be possible to monitor subsurface magma movement.

References

Abt DL, Fischer KM (2008) Resolving three-dimensional anisotropic structure with shear wave splitting tomography. Geophys J Int 173(3):859–886. doi:10.1111/j.1365-246X.2008.03757.x

Audoine E, Savage MK, Gledhill K (2004) Anisotropic structure under a back arc spreading region, the Taupo Volcanic Zone, New Zealand. J Geophys Res 109(B11):305. doi:10.1029/2003JB002932

Babuška V, Cara M (1991) Seismic anisotropy in the Earth, vol 10. Klewer Academic, Dordrecht

Bianco F, Zaccarelli L (2009) A reappraisal of shear wave splitting parameters from Italian active volcanic areas through a semiautomatic algorithm. J Seismol 13(2):253–266. doi:10.1007/s10950-008-9125-z

Crampin S, Gao Y (2006) A review of techniques for measuring shear-wave splitting above small earthquakes. Phys Earth Planet In 159(1–2):1–14. doi:10.1016/j.pepi.2006.06.002

Crampin S, Peacock S (2008) A review of the current understanding of seismic shear-wave splitting in the Earth's crust and common fallacies in interpretation. Wave Motion 45(6):675–722. doi:10.1016/j.wavemoti.2008.01.003

Crampin S, Volti T, Chastin S, Gudmundsson A, Stefansson R (2002) Indication of high pore-fluid pressures in a seismically-active fault zone. Geophys J Int 151(2):F1–F5. doi:10.1046/j.1365-246X.2002.01830.x

S

Del Pezzo E, Bianco F, Petrosino S, Saccorotti G (2004) Changes in the coda decay rate and shear-wave splitting parameters associated with seismic swarms at Mt. Vesuvius, Italy. Bull Seismol Soc Am 94(2):439–452. doi:10.1785/0120030141

Gerst A, Savage MK (2004) Seismic anisotropy beneath Ruapehu Volcano: a possible eruption forecasting tool. Science 306(5701):1543–1547. doi:10.1126/science.1103445

Hess HH (1964) Seismic anisotropy of the uppermost mantle under oceans. Nature 203(494):629. doi:10.1038/203629a0

Johnson JH, Poland MP (2013) Seismic detection of increased degassing before Kīlauea's 2008 summit explosion. Nat Commun 4:1668. doi:10.1038/ncomms2703

Johnson JH, Savage MK (2012) Tracking volcanic and geothermal activity in the Tongariro Volcanic Centre, New Zealand, with shear wave splitting tomography. J Volcanol Geotherm Res 223–224:1–10. doi:10.1016/j.jvolgeores.2012.01.017

Johnson JH, Prejean S, Savage MK, Townend J (2010) Anisotropy, repeating earthquakes, and seismicity associated with the 2008 eruption of Okmok volcano, Alaska. J Geophys Res 115, B00B04, doi:10.1029/2009JB006991. (Figure 1)

Johnson JH, Savage MK, Townend J (2011) Distinguishing between stress-controlled and structural shear wave anisotropy at Mount Ruapehu volcano, New Zealand. J Geophys Res-Sol Ea 116:B12. doi:10.1029/2011JB008308

Keats BS, Johnson JH, Savage MK (2011) The Erua earthquake cluster and seismic anisotropy in the Ruapehu region, New Zealand. Geophys Res Lett 38(L16):315. doi:10.1029/2011GL049014

Miller V, Savage M (2001) Changes in seismic anisotropy after volcanic eruptions: evidence from Mount Ruapehu. Science 293(5538):2231–2233. doi:10.1126/science.1063463

Roman DC, Cashman KV (2006) The origin of volcano-tectonic earthquake swarms. Geology 34(6):457. doi:10.1130/G22269.1

Roman DC, Gardine MD (2013) Seismological evidence for long-term and rapidly accelerating magma pressurization preceding the 2009 eruption of Redoubt Volcano, Alaska. Earth Planet Sci Lett 371–372:226–234. doi:10.1016/j.epsl.2013.03.040

Savage MK (1999) Seismic anisotropy and mantle deformation: what have we learned from shear wave splitting? Rev Geophys 37(1):65–106. doi:10.1029/98RG02075

Savage MK, Ohminato T, Aoki Y, Tsuji H, Greve SM (2010) Stress magnitude and its temporal variation at Mt. Asama Volcano, Japan, from seismic anisotropy and GPS. Earth Planet Sci Lett 290(3–4):403–414. doi:10.1016/j.epsl.2009.12.037

Silver PG, Savage MK (1994) The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers. Geophys J Int 119(3):949–963. doi:10.1111/j.1365-246X.1994.tb04027.x

Unglert K, Savage MK, Fournier N, Ohkura T, Abe Y (2011) Shear wave splitting, vP/vS, and GPS during a time of enhanced activity at Aso caldera, Kyushu. J Geophys Res-Sol Ea 116(B11):203. doi:10.1029/2011JB008520

Vargas-Bracamontes D, Neuberg J (2012) Interaction between regional and magma-induced stresses and their impact on volcano-tectonic seismicity. J Volcanol Geotherm Res 243–244:91–96. doi:10.1016/j.jvolgeores.2012.06.025

Wuestefeld A, Verdon JP, Kendall J-M, Rutledge J, Clarke H, Wookey J (2012) Inferring rock fracture evolution during reservoir stimulation from seismic anisotropy. Geophysics 76(6):WC157–WC166. doi:10.1190/GEO2011-0057.1

Zinke JC, Zoback MD (2000) Structure-related and stress-induced shear-wave velocity anisotropy: observations from microearthquakes near the Calaveras Fault in central California. Bull Seismol Soc Am 90(5):1305–1312. doi:10.1785/0119990099

Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring

Vincenzo Gattulli and Francesco Potenza
DICEAA - Dipartimento di Ingegneria Civile, Edile-Architettura, Ambientale, CERFIS - Centro di Ricerca e Formazione in Ingegneria Sismica, University of L'Aquila, L'Aquila, Italy

Synonyms

Masonry failure; Monumental structures; Seismic damage; Seismic vulnerability; Structural analysis; Structural health monitoring

Introduction

The earthquake occurred on April 6, 2009, in Italy at L'Aquila has been a catastrophic event for both the city and the University of L'Aquila (Ceci et al. 2010). Since then, numerous scientific activities have accompanied both the immediate recovery and the long reconstruction program in different fields of earthquake science and engineering. In particular, the development of new

systems and technologies for both understanding and enhancing the structural behavior of significant historical palaces and churches, densely populating the L'Aquila city center, has attracted the attention of specialists and researchers coming from a broadband spectrum of scientific disciplines. This rich multidisciplinary approach has permitted the development of a new paradigm in the preparatory works necessary for planning the retrofitting and the reconstruction of ancient buildings. Therefore, the impact of new technologies in the area of observation, survey, testing, modeling, restoration, retrofitting, and monitoring of historic constructions and monuments, diffusely used at L'Aquila, merits to be discussed and briefly reported, here, as starting point for future novel findings and process optimization.

In this respect, it appears reasonable to mention the geophysical and geotechnical research efforts conducted to characterize various aspects of what has been observed and measured during foreshock, mainshock, and aftershock in the L'Aquila valley. Starting from a historical seismological study regarding the earthquakes that took place in the area of L'Aquila (central Italy) from the ancient Roman period to the late Middle Ages, the persistence and magnitude of earthquakes demonstrated to have a strong bearing on the economy and culture of the communities (Guidoboni et al. 2012). Furthermore, understanding the rupture slip distribution for the 2009 L'Aquila mainshock has allowed for the development of a complex high-resolution 3D FE model incorporating surface topography and rheological heterogeneities, deduced from real tomography (Volpe et al. 2012). The 3D approach provides a more concentrated and localized slip distribution on the rupture plane, evidencing a single area of high-slip release SE of the hypocenter. Moving from the underground level to the ground level, scientific studies have tried to correlate the distribution and the severity of the damage with the geological setting of the area, taking into account the characteristics of the building stock through time (Tertulliani et al. 2012). Strong-motion records and ambient noise measurements taken soon after the mainshock and during the entire aftershock

sequence showed variability in ground motion amplification throughout the city. General conclusions highlight that the building stock of the city suffered different levels of damage that can be partially explained by the combination of building vulnerability and surface geology. In particular, the observation of damages in RC buildings highlighted that the few collapse cases (<1 % of the whole RC buildings stock) are mainly the consequence of structural deficiency, inadequate structural layout, and maybe site effects (Decanini et al. 2012). The majority of RC construction survived the earthquake with minor damages or with damages to nonstructural elements. The extensive observed damage to infill walls that can be associated to a significant amount of energy dissipation prevented the development of plastic hinges in RC columns, thus allowing the maintenance of structural integrity. Concerning the seismic demand, the study highlighted the noticeable difference between the elastic and inelastic demand, indicating that the displacement demand is moderate, associable to a non-extraordinary event, even if, in several cases, it overcomes the threshold of serviceability limit, producing economic losses and potential injuries and deaths (Ceci et al. 2013).

Much more complex than the damage scenario affecting the RC buildings is the case of technical observations made on which occurred specifically in historic constructions and monuments at L'Aquila (D'Ayala and Paganoni 2011; Da Porto et al. 2012; Indirli et al. 2013) and in the small historical centers of the valley (D'Ayala and Paganoni 2011; Carocci 2012; Indirli et al. 2013). These studies show that within the historic city center of L'Aquila, the number of total collapses observed was minor, while the proportions of partial collapses of the upper stories were perhaps greater than would have been expected given the level of shaking (D'Ayala and Paganoni 2011). In a minority of cases, total collapse of the facade was observed, and these were usually associated to substantial alteration of the original structure. Although the number of undamaged masonry buildings is very modest, the majority suffered either minor damage or repairable structural damage.

This situation allowed on one hand to identify and recognize collapse mechanisms and on the other hand to better correlate ultimate capacity to strong-motion characteristics and hence lends itself to a more detailed in situ survey and analytical assessment. In the specific case of churches, a statistical analysis has defined the percentages of possible mechanisms related to local construction practices and architectural features of the region. For each possible mechanism, the relative percentage of activation and calculated average damage allowed some conclusions to be drawn on the seismic vulnerability of the various structural and nonstructural elements making up a church. These conclusions may be regarded as widely valid and, if further corroborated by other post-earthquake surveys, may be used to define hierarchies of interventions, also on non-damaged structures (Da Porto et al. 2012). The study showed that it would be useful to introduce weights to damage mechanisms when calculating damage indices evaluated through survey form compilation, which is still a powerful tool in the management of emergency and post-emergency phases. Furthermore, it has been pointed out that the recorded high-peak ground accelerations reached high values, both in the horizontal and vertical directions. The simultaneous combination of these acceleration components, has produced a reduction of the masonry shear capacity, especially in the upper floors. Consequently, the widespread vulnerability of unreinforced masonry buildings resulted in extensive damage, collapse, and victims in a medium-size city. The underestimation of the earthquake actions is a considerable drawback that can be overcome only by integrating the study of earthquake scenarios with more accurate methodologies. Moreover, with regard to the structural vulnerability, most of the unreinforced masonry structures were built (and structurally modified with subsequent interventions) before the adoption of the 2002 updated requirements; thus, the percentage of unreinforced masonry buildings heavily damaged or collapsed in the epicenter area is relatively high (approximately 45 %). The first cause of widespread overall disruption originates from the very poor quality of the disaggregated masonry, made by arbitrary materials, unable to resist horizontal forces. Several damage/collapse types have been encountered, classified, and associated with the simplified mechanisms foreseen by the Italian MEDEA procedure (Indirli et al. 2013).

The needs of simplified performance-based assessment procedures in order to support the interpretation of observed damage is evidenced in all recent earthquakes, such as those occurred in L'Aquila (2009), Christchurch (2010–2011), and Emilia Romagna (2012), which have caused not only a significant death toll and huge economic losses but also heavy damage to the worldwide cultural heritage (Parisi and Augenti 2013). Earthquake damage to monumental constructions has been discussed and critical issues affecting the seismic response of historic masonry structures are identified such as masonry quality, connections among structural elements, diaphragm flexibility, out-of-plane resistance of masonry walls, structural irregularities, wrong retrofit interventions, and earthquake ground motion characteristics.

Moreover, since 2009, the reconstruction idea of the city of L'Aquila and its surrounding was a main goal, primarily of the citizens and, with them, of the Italian Government. However, during every catastrophic earthquake, Italian citizens and institutions seem to face a series of typical problems due to bureaucratic burdens, corruption, excess legislation, political clientelism, poor performance, and weak public control, and they are the combined effects of cultural, institutional, and political factors at the governance level (Özerdem and Rufini 2013).

Notwithstanding several difficulties, numerous applied research activities have been conducted in the area of new and sophisticated methods for the assessment of masonry complex structures, which need retrofitting interventions. Most of the works conducted at L'Aquila in this area are characterized by a series of correlated activities ranging from historic search, survey, and analysis of the structure with regard to proposals for repair and seismic strengthening. In particular, several problems are generally encountered in the various levels of survey and

damage interpretation and those of modeling and analysis of the structural behavior of complex and clustered masonry buildings (Da Porto et al. 2013). The in-depth comparative studies between the surveyed damages and the results of structural assessment through different approaches have produced a great enhancement in the understanding of the complex behavior of masonry structures, evidencing that use of the "Natural Laboratory of Earth" is the best approach to evaluate the seismic behavior and the performance of structural systems as well as the failures occurring in reality and is a non-substitutable step for the advancement of knowledge in the field of seismic engineering (Brandonisio et al. 2013). For example, the observation made on large historical palace, such as the case of Palazzo Centi, has shown that, during the 2009 earthquake, differently from other historical masonry buildings, it responded "reasonably" well in the main structural parts also due to recent retrofit interventions carried out in 2003. These mechanical interventions proved to be the reason of the good seismic behavior of the building and that "saved" it from more serious damages demonstrating the effectiveness of the traditional techniques in designing and retrofitting masonry buildings (Lucibello et al. 2013).

The structural assessment and retrofitting of important churches in L'Aquila, some of them having a monumental value, has been elaborately studied. Results concerning the seismic behavior of masonry churches have shown that the dynamic excitation due to the seismic ground motion activates many vibration modes of the building structure, though all of them are characterized by small participation factors. Therefore, churches are not behaving through a superposition of global modes but more as dynamic interaction of localized modes. Consequently, in many examined case studies, the ratio between the total base shear and the church total weight ranged between 20 % and 30 %, evidencing a reduction with respect to the plateau value of the spectral acceleration provided by the Italian code. Therefore, appropriate choices of the force reduction factor should be adopted for these monumental buildings different from the case of

traditional residential buildings characterized by shear-type behavior. Further, the activation of many local modes also calls for retrofit interventions, which should "tie up" the building, thus avoiding the local failure modes that are often observed (Brandonisio et al. 2013).

Detailed 3D nonlinear numerical analyses have been performed on several palaces and churches and also on the macroelements composing them such as towers and facades. Linear and nonlinear structural analyses conducted by means of a full 3D finite element model have been used to partially explain the collapse of the transept area of the Basilica of S. Maria di Collemaggio. The transept structural system formed by arches, pendetives, and barrel vaults transferring the vertical loads onto the multilobed pillars has been demonstrated to be vulnerable with respect to both transversal and longitudinal actions. In particular, even in the case of longitudinal actions, the nonlinear static analysis evidences the presence of high tensile stresses in the arches connecting the pillars to the apse walls. This mechanism, observed in the numerical simulations and amplified by the strong vertical acceleration component registered close to the site, together with the minimal resistance of the material inside the core of the collapsed pillars, is the most probable explanation of the implosion of the transept structures in the church (Gattulli et al. 2013).

The bell tower of the San Pietro di Coppito Church in L'Aquila suddenly collapsed during the devastating earthquake, which occurred in April 2009. A series of analyses were performed to provide an ex post evaluation of the causes at the base of the collapse and give operative design information for the reconstruction (Milani et al. 2012). For this aim nonlinear static and limit analysis has been performed evidencing the role played by the actual geometry and by the masonry mechanical characteristics of the tower through reconstructing the failure mechanisms by both modeling approaches. More recently, in a similar modeling approach, an attempt has been made to take into account the effects of fiber reinforcement polymers positioned on a large facade of Palazzo Camponeschi,

S

as example of engineering analyses on the effectiveness of specific retrofitting intervention with new material on ancient masonry structure in order to augment their seismic capacity (Gattulli et al. 2014).

Finally, the dynamic behavior of masonry structures under three-dimensional shaking is not completely understood, and it needs specific investigations. In this sense, the direct observations and measuring of such behavior through the use of novel structural permanent monitoring systems may help to reach such a goal.

Before the 2009 L'Aquila earthquake, very few structures were equipped by permanent structural monitoring systems managed by the Department of Civil Protection (DPC). However, the response of the Pizzoli Town Hall during the mainshock has been recorded and analyzed by DPC, giving special insights on the potentiality of these systems for immediate evaluation of the damages occurring during an earthquake. Successively, the large amount of installed, temporary or permanent, different types of devices (accelerometers, smart wireless devices, displacement and velocity transducers, inclinometers, etc.) reach a number of around three hundred (300) or more, evidencing a large impact of this technology in the post-earthquake emergency phase, especially during the earthquake swarms (Gattulli 2013). In particular several monitoring systems have been installed in the emergency phase, to understand the occurred behavior in damaged building (Cimellaro et al. 2012; Foti et al. 2013), or during the construction of temporary scaffolding, in order to verify the efficacy of the added structural system especially in the case of monumental building (Russo 2013). Because of this scope, in many cases, the permanent monitoring has worked ranging from a limited number of hours up to several months. In other cases, the monitoring system is permanently installed on the structure for years; it can be used also to determine the change that will occur in the structural behavior during the reconstruction phase (Federici et al. 2012; Gattulli 2013).

In several cases, the structural monitoring system uses only accelerometers, starting from very few measurements to a larger number of devices with different characteristics and sensitivities. Instead, more complex monitoring systems are used in monumental churches and buildings where accelerometers are accompanied by crackmeters, inclinometers, temperature measurement devices, etc. (Casarin et al. 2012; Russo 2013; Gattulli et al. 2014).

Within this broad range of activities, in the present chapter are summarized specific findings of the researchers of the University of L'Aquila constantly involved to define sustainable reconstruction interventions, in the framework of broad multidisciplinary collaboration, open to the use of innovative technologies and accompanied by valuable training of young scholars.

Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring

The historic center of L'Aquila, which was widely extended before the earthquake, was inhabited by about 7,000 residents and 8,000 students, constituting a large and widespread university campus. In this situation, dense of monuments of great value, characterized by a particular urban fabric, having an orthogonal system due to the Angioina division into lots even if modified by an accelerated stratification, made by previous earthquake, are placed more than 800 businesses, several professional offices, and headquarters of the administrations of various public institutions. This complicated system is developed for about 10 km, including historical centers, consolidated tissues, parts of the city under construction, and the natural park of Gran Sasso and Sirente-Velino. The earthquake has seriously compromised the system and this area corresponds, quite closely, to the so-called seismic crater in which the most important damage occurred. In the revitalization of this large zone affected by the earthquake, a central role is played by the recovery of the L'Aquila historical center accompanied by the similar ones in the neighboring small villages. Hence, the characterization of this system, using a large scale, starting from an urban point of view up to geomorphological and geotechnical vision, is an element of strong relevance also to understand what happens

to the smaller scale of the individual monument, integrating the specific observations with an overview of the problem. Indeed, the numerous campaigns for the characterization of the seismic behavior of the L'Aquila's valley made by a microzoning mesh of the historical center indicated a number of site effects which are very important for better understanding the L'Aquila earthquake and its effects. Recent studies, characterizing the underground of the historical center, point out a sudden transition between the L'Aquila breccia and silty clay loam through a highly detailed reconstruction.

The problem of the reconstruction of historical buildings and monuments, especially within the walls of the historic center, brings to investigate the main open issues concerning the interventions that ensure a real improvement. The structural behavior of the monuments, comprising a masonry structure, still today open to a full understatement, depends on factors such as the mechanical properties of the different materials composing the masonry walls, the features of the horizontal structure and their connection with the vertical ones, and the used construction methods. Therefore, it is important to realize an optimal use of the wealth in the technical and scientific knowledge, stored over the years, to evaluate properly the physical consistency and safety of historical buildings. The answer, of course, cannot be unique but should result in a critical and competent attitude. Some ideas may be received from the reading of a book dealing with the technical problem of preservation. For example, in a classical book on this issue is reported "...In this type of intervention is of particular importance the respect of the typological, structural and functional quality of the system, avoiding those transformations that alter the character." The matter is still very complicated to be solved. Indeed, the properties of fragility, heterogeneity, and anisotropy of masonry and their different typological varieties make it difficult to well describe unitarily their mechanical behavior. Recent technical codes take into account these uncertainties, allowing to use simplified numerical model, implemented "ad hoc," to evaluate the consistency of masonry structures, avoiding the use of

numerical checks that not always ensure a complete reliability. Moreover, the physical nature of masonry structures, product of improvements obtained following successive attempts, and empirical rules lead to the need to examine each case as a special case, making prohibitive to have a general characterization also due to the historic analyses which produces a series of unique cases. However, the entire reconstruction process must also take into account other problems such as energy efficiency, architectural restoration, and refunctionalization. All these actions have to make considering permanent monitoring of the monument so as to better schedule maintenance and following the target of long-term operation of the object.

A series of research groups of the University of L'Aquila have conducted numerous activities related to earthquake engineering immediately after the earthquake to support the city and the entire community strongly hit by the catastrophic event. Here, we attempt to report this huge effort motivated by the willingness to recover as soon as possible to the original state. This entails a path that goes from collapse observation to permanent monitoring, thus characterizing a broadband research activity and picking, according to the authors' opinion, the most significant results. Therefore, the presentation is subdivided in several specific fields.

Collapse and Damage Observation and Survey

A series of partial or complete collapses have occurred in historical buildings at the center of L'Aquila, some of them having monumental value. Among them, one of the most studied is the collapse of the transept of the Basilica di S. Maria di Collemaggio (Gattulli et al. 2013).

The major damage suffered by the Basilica in 2009, indeed, was probably induced by a mechanism of implosion of all the structures that compose the transept: the great multilobed pillars, the triumphal arch and the wall above, the barrel vaults, the dome, and the roofing structures (Fig. 1). In particular, the failure occurred almost simultaneously in the two large pillars and the arches connecting the nave walls with the chancel, closed at the end by the apse, and it was

Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 1 Basilica of S. Maria di Collemaggio transept collapse: (**a**) aerial view and (**b**) internal view
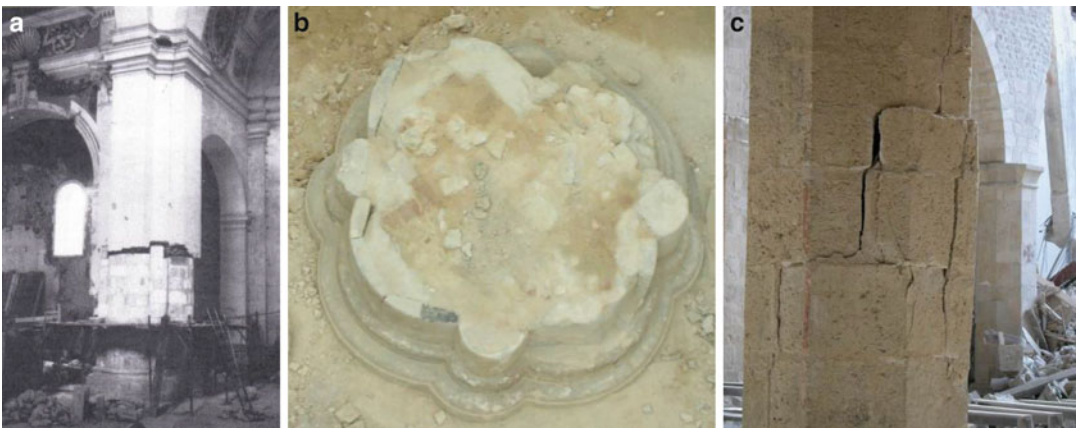
Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 2 Pillars and columns of S. Maria di Collemaggio: (**a**) shaping of the pillars in a multilobed medieval form, (**b**) material pillar core evidenced by the collapse, and (**c**) column damages

followed by the overtopping dome collapse. It should be pointed out that the transept is realized through a structural system in which the arch and vault thrusts, induced by permanent vertical loads, play an important stabilizing role. During the simultaneous vertical and horizontal actions of the 2009 earthquake, the equilibrium state in the masonry element was significantly altered. In particular, due to the proximity to the epicenter, the vertical and horizontal acceleration components had similar amplitudes.

Moretti's restoration intervention in the early 1970s was intended to remove the Baroque decorative features added to the Basilica and to unearth the "powerful original structure" of the thirteenth-century plan, but only the works on the nave had been completed, leaving the transept with its original Baroque decoration. At the same time, he shaped the pillars in a multilobed form, removing material along the height starting from the original multilobed base (Fig. 2). In performing the restoration, the surface of the pillars was realized using blocks of freestone placed around the masonry core of the Baroque pillars having only an aesthetic function with no resistance increase. Small stones placed randomly, and probably coming from previous earthquake debris, were used for the core material.

Moreover, a subhorizontal fracture positioned at the height of around 2 m is present in the north

external wall, hosting the Holy Door, apparently created at the interface between the lower masonry part subjected to grout injection performed during the restoration works conducted in 1999 and the upper masonry part preserving the original workmanship.

Besides that, several of the threaded rods of the bracing system under the central nave roof were broken, almost all in the vicinity of the large crack on the north side wall of the church. In the north aisle, slippage of bolts designed to anchor the wooden trusses was noticed. Both of these cases of damage show the presence of transverse deformations of the longitudinal walls mainly located in the north, which may occur even after the transept collapse.

The damage pattern which affects the columns of the nave walls was developed in a complex manner; there were deep cracks due to heavy compression, initially appearing only in the central columns of the nave (Fig. 2c) and then progressively emerging, with the aftershocks, in almost all the others as well. The progression of damage was very similar to that found during the intervention in the last restoration of the 1970s, which consisted of inserting some missing fragments and grout with cement mortar in the opened joints. The only column that initially exhibited minor damage was the first one towards the transept, which had been completely rebuilt, although simply with brick facing of limited thickness and with an inner core done in small stones immersed in cement mortar.

In all the other columns, large cracks were visible. These columns, in distinction from the first ones, were made using stones shaped like truncated pyramids; these stones were juxtaposed, in an almost dry manner, in order to form the octagonal section. In most cases, cracks occurred in the stone blocks, despite the apparently high resistance of the material. The area of the apse was also characterized by severe damage. This area was affected by clear cracking in the mortar joints of the hewn stone blocks of the main apse, where there were also expelled blocks forming a wedge shape above the mullioned window. Observation of the interior reveals cracks in the apse vault, which is one of the most loaded

elements, a partial detachment, and the presence of permanent relative displacements, which occurred in one of the ribs. The lateral chapels suffered more serious damage. However, the study of the mechanical behavior of the monument cannot avoid to consider the low quality of the masonry, even present in peculiar elements as arches and vaults, exposed clearly in evidence by the collapse (Fig. 3a) and the peculiar realization of the masonry walls at L'Aquila that along the time due to the occurrence of several earthquakes is the product of different craftsmanship (Fig. 3b).

Laser Scanning Measurements

Besides the evident collapses easily visible, the relief of the damage, caused by the earthquake on April 6, 2009, has produced continuous and stimulating activities in damage observation and survey, producing a positive and critical approach about different aspects of the problem. Indeed, in particular regarding this specific area, the catastrophic event was an opportunity to apply the most advanced available technologies in the field. For the first time, the laser scanner technique was extensively used for the surveying of a great portion of L'Aquila territory and in particular for the historical buildings and monuments positioned within the old L'Aquila walls. The survey of the portion of the city hit by an earthquake has been considered, in several cases, a useful support for other structural analysis and conservation and restoration studies. Laser scanning has been used just after the earthquake in L'Aquila and during several summer schools that took place every year since the earthquake, to survey the facades of numerous buildings. The point cloud has been treated and filtered, in order to quantify and verify exactly the deformations due to the earthquake, compared with the original shape and geometry of the building. Laser scanning surveys allow the acquisition of the shape complexity of the building, with its irregularity, not only in some essential lines, as it is in topography. The whole information package is more complex than what is obtained with traditional techniques (topography and photogrammetry), but it is extremely useful for survey and measurements

S

Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 3 Masonry features at L'Aquila: (**a**) masonry quality evidenced by the collapse and (**b**) masonry walls realized with material belonging to different periods and with different craftsmanship

of structural deformations and cracking. Moreover, the data acquisition system permits to acquire information on difficult geometry from a certain distance, avoiding difficulties related to traditional methods of measurement. Moreover, the University of L'Aquila placed important resources in the field, through several labs (CERFIS, CSE, etc.), renewing the technical equipment and using specialist teams, which operate for the geometric and damage relief. Data processing, made by Leica instrumentation (total station HDS2600 and software Leica Cyclone 7.2) and by a scanner velocity max of 1,016,727 point/s and flow rate in the range of ambiguity of 79 m, permits a return of three-dimensional model (see Fig. 4) from which plants, prospects, and sections can be obtained. The opportunity of having this great volume of data has stimulated to develop interactions between different disciplines, and so they constitute a new wealth of information. In particular the detailed reconstruction of crack path in the main facade of historical buildings (like Camponeschi Palace) allowed a verification of the numerical models to reproduce the observed behavior under the earthquake and to predict the retrofitting improvements (Gattulli et al. 2014).

Applications of Infrared Thermography for the Investigation of Historic Structures

Infrared thermography is used as a nondestructive technique to identify the most significant inhomogeneity of materials and properties in the surface and structures that are difficult to appreciate by simple visual inspection but are instead detectable by the relief of anomalies associated to thermal phenomena. A portable instrumentation, which consists in thermal imagers, enables the observation of the materials constituting the main and secondary walls (e.g., infill, detachments, fills, and so on) and naturally the discontinuities. The infrared thermography is a noninvasive technique that uses infrared energy emitted spontaneously by any body or object: electromagnetic radiation whose wavelength depends on the temperature of the same body (Wien's law). Because the temperature of the bodies in ambient condition have a wavelength that belongs to the range of infrared spectrum ($\lambda = 0.3 \div 0.78$ μm), the technique is called
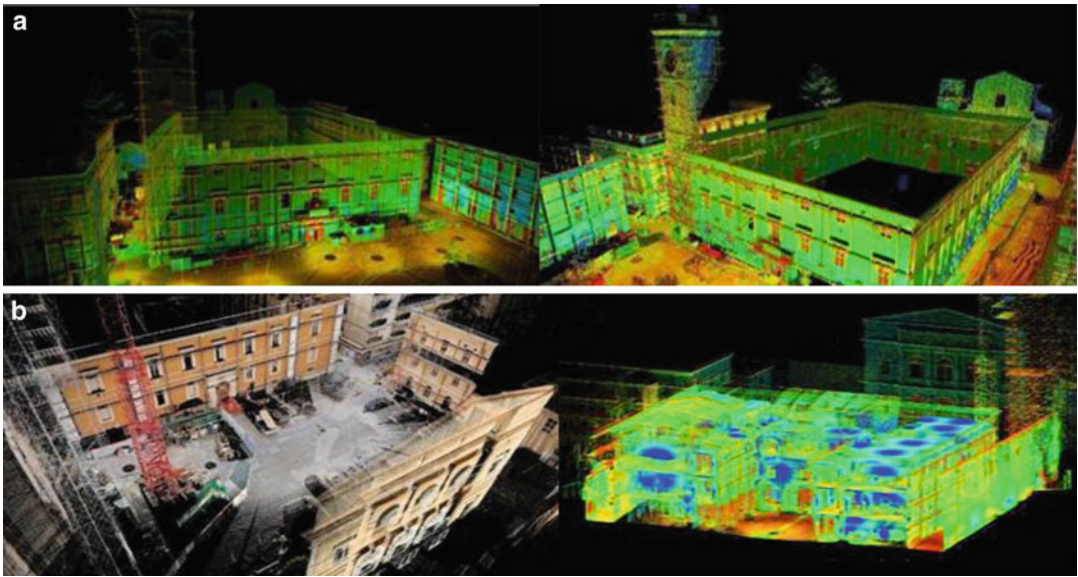
Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 4 Reconstruction of tridimensional geometrical models of the structures damaged by the earthquake using laser scanner techniques: (**a**) Margherita Palace, headquarter of L'Aquila Municipality, and (**b**) De Amicis school building

infrared thermography. The value of radiant energy issued change when the objective of the camera passes from a point at a higher temperature to one with a lower temperature, and so also the wavelength of the radiation assumes two different values (Boltzmann's law). Where there is a thermal gradient, the scanner captures the energy variation that is linked with the body temperature through Stefan's law, and so it can determinate the value of emissivity's coefficient. Using a dedicated scanner optical system that investigates point by point all the recorded fields, the camera reproduces an electrical signal representative of the image's thermal information and proportional to its intensity, which will be shown in a kinescope. In particular for each range of temperature variation, there will be a corresponding range of radiant intensity variation to which the instrument assigns a particular shade of color. The main characteristics of the instrumentation used are the following: minimum focusing distance, 0.3 m; temperature range, from -20 °C to $+120$ °C; field of view (FOV), $24° \times 18°$; thermal sensitivity (NETD mK), <40 mK @ $+30$ °C ($+86$ °F); accuracy, ± 1 °C o ± 1 % of readings;

resolution IR, 640×480 pixel; and spectral field, from 7.5 to 13 m. An extensive campaign of infrared thermography has been conducted to characterize the masonry of Camponeschi Palace, where the Letter and Philosophy Faculty is located. Figure 5 shows a series of findings. Moreover, this technique can be used without any interruption of the current activities and works. The fast data processing allows for the recognition of peculiar masonry zones with doors or windows under the plaster that are partially or completely closed, as well as with the presence of chimneys and underground utilities, masonry wall textures, warping of the floor, nonhomogeneity of the materials, structural damages, delamination of plaster as well as malfunctions of electrical and plumbing systems, infiltration of moisture, leaks, and thermal bridges.

Ultrasonic Testing

In ultrasonic testing (UT), very short ultrasonic pulse waves with center frequencies ranging from 0.1 to 15 MHz and occasionally up to 50 MHz are transmitted into materials to detect internal flaws
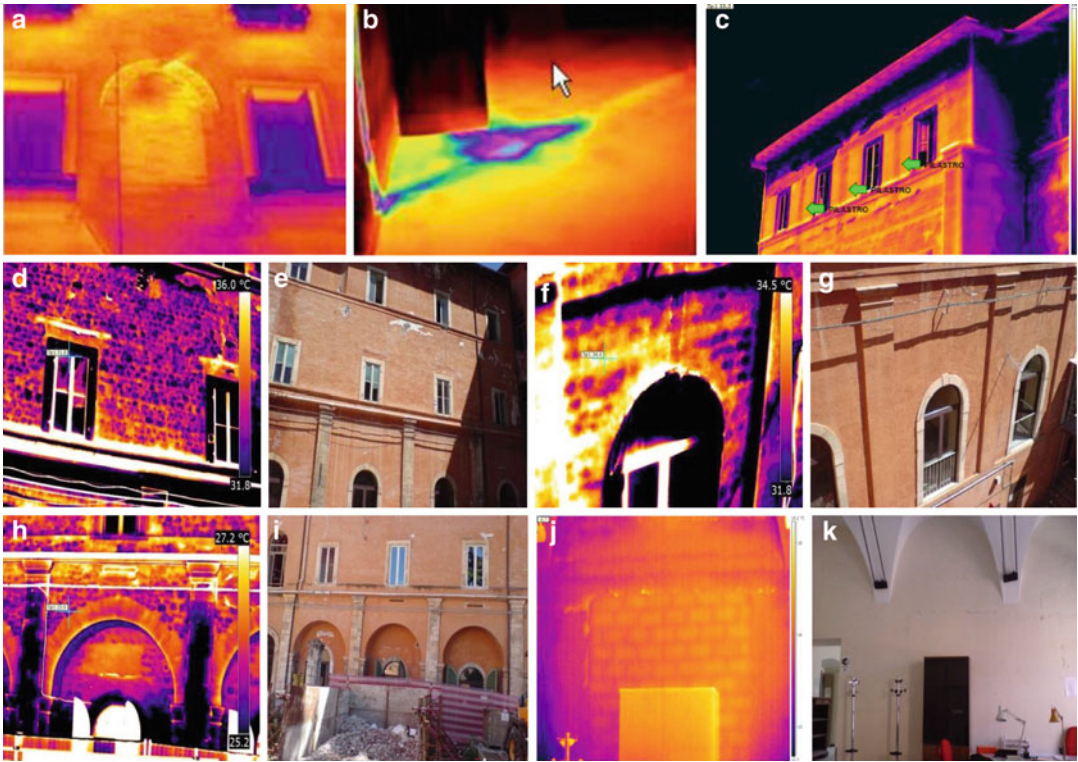
Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 5 Infrared thermography results obtained in Camponeschi Palace at L'Aquila: (**a**) evidence of opening closed with arched lintel, (**b**) failure of pipes, (**c**) identification of concrete elements, (**d–g**) reading of textures walls under plaster, (**h, i**) reading of weaving arched lintels, (**l, m**) presence of openings in the internal masonry walls transversal to the perimeter walls

or to characterize materials. An ultrasound transducer connected to a diagnostic machine is passed over the object being inspected.

There are two methods of receiving the ultrasound waveform: reflection and attenuation. In reflection (or pulse-echo) mode, the transducer produces and receives pulsed waves when the "sound" is reflected back to the device. Reflected ultrasound comes from an interface, such as the back wall of the object or from an imperfection within the object. The diagnostic machine displays these results in the form of a signal with amplitudes representing the intensity of the reflection and the distance and registering the arrival time of the reflection. In attenuation (or through-transmission) mode, a transmitter sends ultrasound through one surface, and a separate receiver detects the transmitted signal on another surface after traveling through the medium. Imperfections or other conditions in the space between the transmitter and receiver reduce the amount of sound transmitted, thus revealing their presence.

Ultrasonic testing, used on masonry, permits to obtain information about the homogeneity and composition of the masonry wall of the analyzed building. This technique has the benefit to be a nondestructive test and can be done in a very rapid and simple way in wide region of structural parts, but on the contrary, this large volume of data needs a specific interpretation. For this reason, it is important to combine the sonic tests with destructive ones, performed in a limited and significant structural zone, to have a well-specific and useful reference which defines the right relation between the speed of sound and the mechanical characteristics of the investigated masonry. Therefore, it is useful to highlight the main

technical properties of the instrumentation used in a lot of campaign tests at L'Aquila by the University:

- Acquisition, measuring range from 100 mv to 20 v
- Time bases from 20 ns to 81.9 s
- Sample resolution 8 bit
- Samples for event 8,192 for contact measurements, 640 for log, bandwidth 50 mhz, filter for ultrasounds: central frequency 50 kHz, measurement channel 1
- Probes through contact and sonic test with hammer: resonant frequency 53 kHz, diameter

48 mm, maximum frequency for pulse emission: 1 per second, minimum measurement pitch: 10 mm

In Fig. 3 images obtained from sonic tests are reported. From the comparison of the graphs reported in Fig. 3a, b, the distribution of pulse wave speed in the same portion of masonry before and after fluid mortar injection grouting is highlighted. In the first case the maximum speed is 250 m/s, while in the second case it grows up to 1,709 m/s. The different velocity transmission indicates clearly the increase of stiffness due to the grouting (Fig. 6).
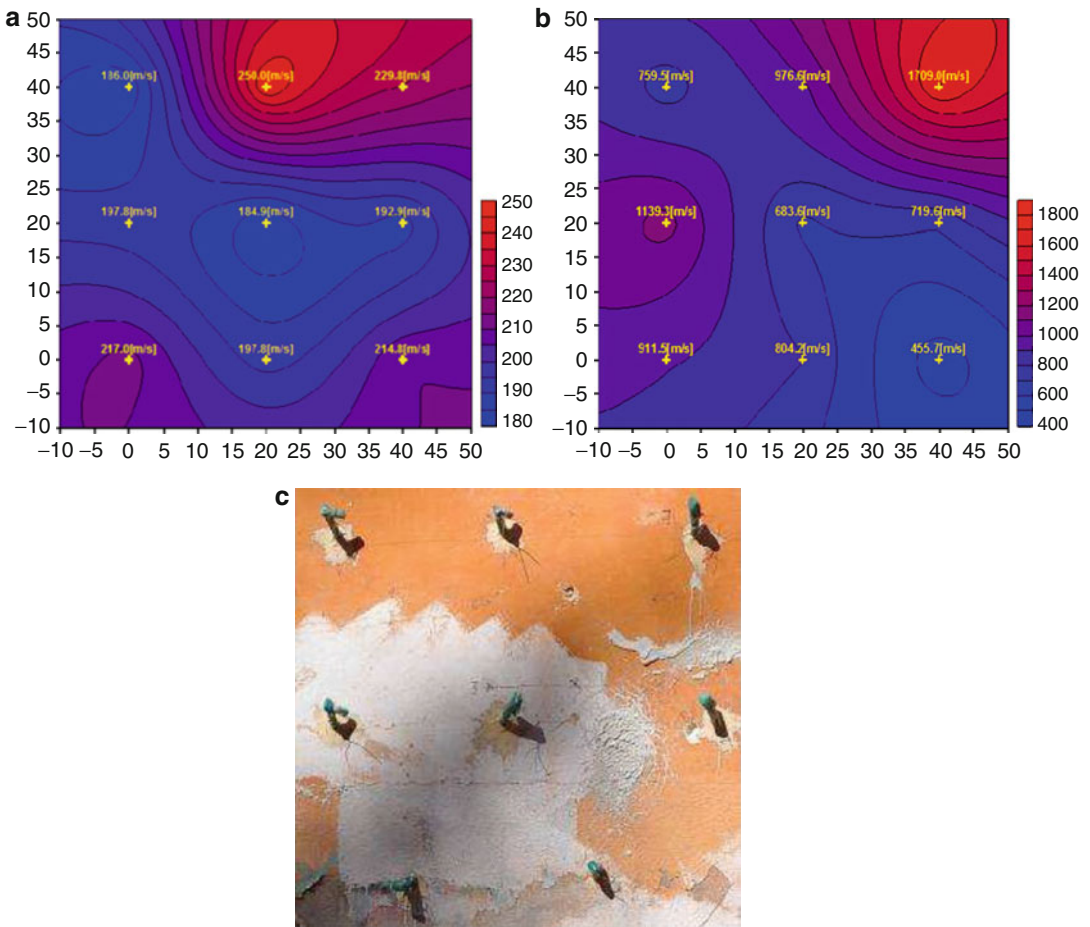
Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 6 Sonic tests on a wall of Camponeschi Palace: distribution of the velocity of propagation (**a**) before and (**b**) after the fluid mortar injection grouting and (**c**) locations of injection pipes

Georadar (Ground Penetrating Radar: GPR) Tests

The GPR methodology that employs electromagnetic waves permits the boundary definition of mechanical characteristics of the investigated body by the identification of the interface between layers that have different resistivity and dielectric constant. Structural degradation is classified (four classes of QRI "Quality Radar Index") by a qualitative analysis of the power radar signal behavior with the depth by both radar anomalies in the section (presence/absence, point 0 and 1) and energy recovery, intense and localized (also called "pick") in the graphs (point 1 for the pick up to 6 dB, 2 point up to 12 dB, and 3 point over 13 dB). The total QRI is the sum of the two scores described above. The method is calibrated using correlation between radar data and a direct destructive test made on hundred scans acquired on masonry tester. This method permits a direct or relative classification depending on whether there are direct data on each structure or not. The relative classification is divided in four classes: QRI 4, class of low quality; QRI 3, class of medium-low quality; QRI 2, class of medium quality; and QRI 1, class of medium-high quality. In Table 1 the main results of the georadar tests are reported and the obtained classification, in particular, is highlighted in the last column.

In Fig. 7a a grid in the scanned surface of a portion of a masonry wall of the Camponeschi Palace is presented while the obtained results through the georadar test are described in Fig. 7b in terms of QRI distribution, evidencing a relevant variability of QRI ranging from 1 to 3.25.

Removal of Plaster and Endoscopic Tests

Among minimal invasive tests, the removal of plaster rendered a zone of inhomogeneity, eventually pointed out by infrared thermography, visibly permitting to check the true nature of the walls and of the composing elements. Successively, endoscopic tests have been usually made in several historical masonry walls at L'Aquila. Endoscopy is simply an extension of the essential visual survey into areas inaccessible to the naked eye. The equipment ranges from relatively simple borescopes, consisting of a light source, a small-diameter rigid tube with built-in optics, and an eyepiece, to complex controllable systems with numerous specialized attachments. By drilling a hole (normally less than 12 mm) and inserting the tube, it is possible to inspect voids under floors or behind paneling, for example. Any hidden problems such as fungal growth can, in theory, be identified. The more sophisticated and expensive equipment is fully flexible and can be steered by wires built into the casing. Systems are available down to 6 mm in diameter, and more specialized systems may reach less than 2 mm. It is possible to attach still or video cameras to the eyepiece to record the findings. The theory is fairly simple, but in practice it can be very difficult to retain a sense of scale of the observed image and keep track of the location and orientation of the tip. The focal range, depth of field, and strength of

Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Table 1 Main information deduced by GPR tests

| Scansions | | GPR graphics | | | GPR sections | QRI index |
| | | Max | | | Presence/absence | |
Codex	Length	db	Max height of the picks equalized	Point	Localized anomalies	QRI index
LEA10000	0.6	5	5.0	1	0	1
LEA10001	0.6	2	2.0	1	0	1
LEA10002	0.6	5	5.0	1	1	2
LEA10003	0.6	7	7.0	2	1	3
TEA10000	0.6	9	9.0	2	0	2
TEA10001	0.6	8	8.0	2	1	3
TEA10002	0.6	5	5.0	1	0	1
TEA10003	0.6	7	7.0	2	0	2
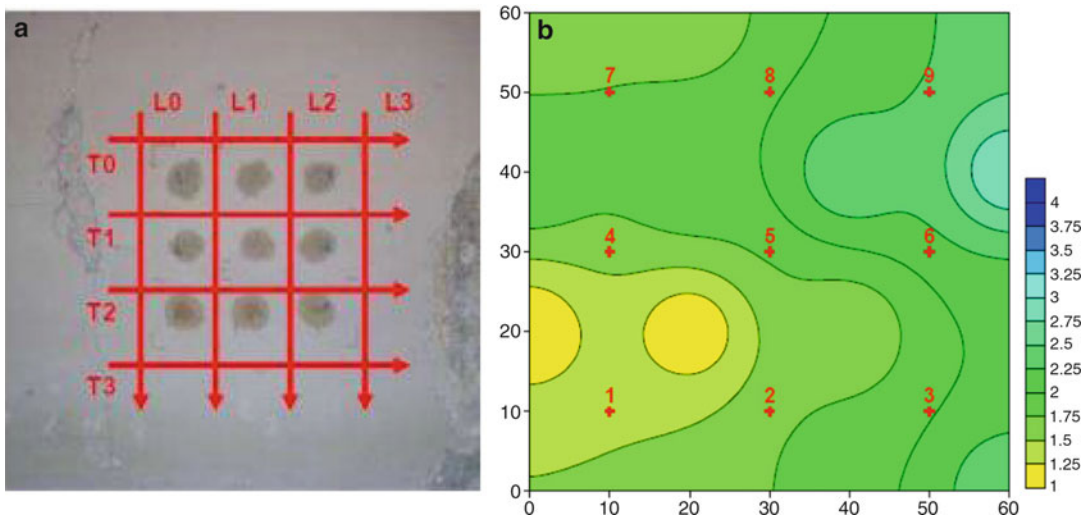
Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 7 Georadar tests on masonry of Camponeschi Palace: (**a**) grid scans and (**b**) QRI distribution

light are greatly reduced in the smaller diameter systems. It is not unknown for insulation lagging to be misidentified as dry-rot. However, reliable reconstruction of the stratification in the main masonry walls has been obtained, to know the construction technique and to assess qualitatively the physical conditions. The employed instrumentation has permitted to register short movies of what is visible inside the cavities, producing a large quantity of data in digital form to be processed.

Flat-Jack Tests

Several tests of single and double flat jacks were carried in different historic buildings in the center of L'Aquila including Camponeschi Palace, De Amicis school, the headquarter of the Carispaq Bank and Margherita Palace, and others. Moreover, diagonal tests have also been performed to characterize the shear strength in selected masonry panel individuated in existing walls in order to characterize the mechanical properties of the main typology of historical masonry walls at L'Aquila. A probabilistic characterization of the available data on the mechanical features of the historical masonry at L'Aquila is going on in order to have the possibility to investigate the structural assessment of historical buildings,

taking into account explicitly the inherent uncertainties of the problem. A more reliable description of the seismic behavior of the masonry mechanical characteristic will realize a well-balanced design of retrofitting interventions.

Figure 8 summarizes the results obtained through double flat-jack tests realized at different locations of the Camponeschi Palace. The three performed tests evidence the variability in both elastic modulus and measured strength through the adopted testing procedure. Indeed, in the analyzed case, the average strength value is 2.47 N/mm^2, while the variance is 0.736 N/mm^2, evidencing the large scatter in the material properties belonging to the same historical building.

Computational Mechanics and Masonry Structural Modeling

Structural analysis and numerical modeling of the seismic behavior of masonry structures is still an open research topic, especially in Italy and in particular after the L'Aquila earthquake. Relevant issues pertain to the inherent uncertainties related to the evaluation of the seismic adequacy of complex masonry buildings through reliable numerical models. Starting from the limit analysis in which the simplified assumption of describing the incipient collapse behavior through chains
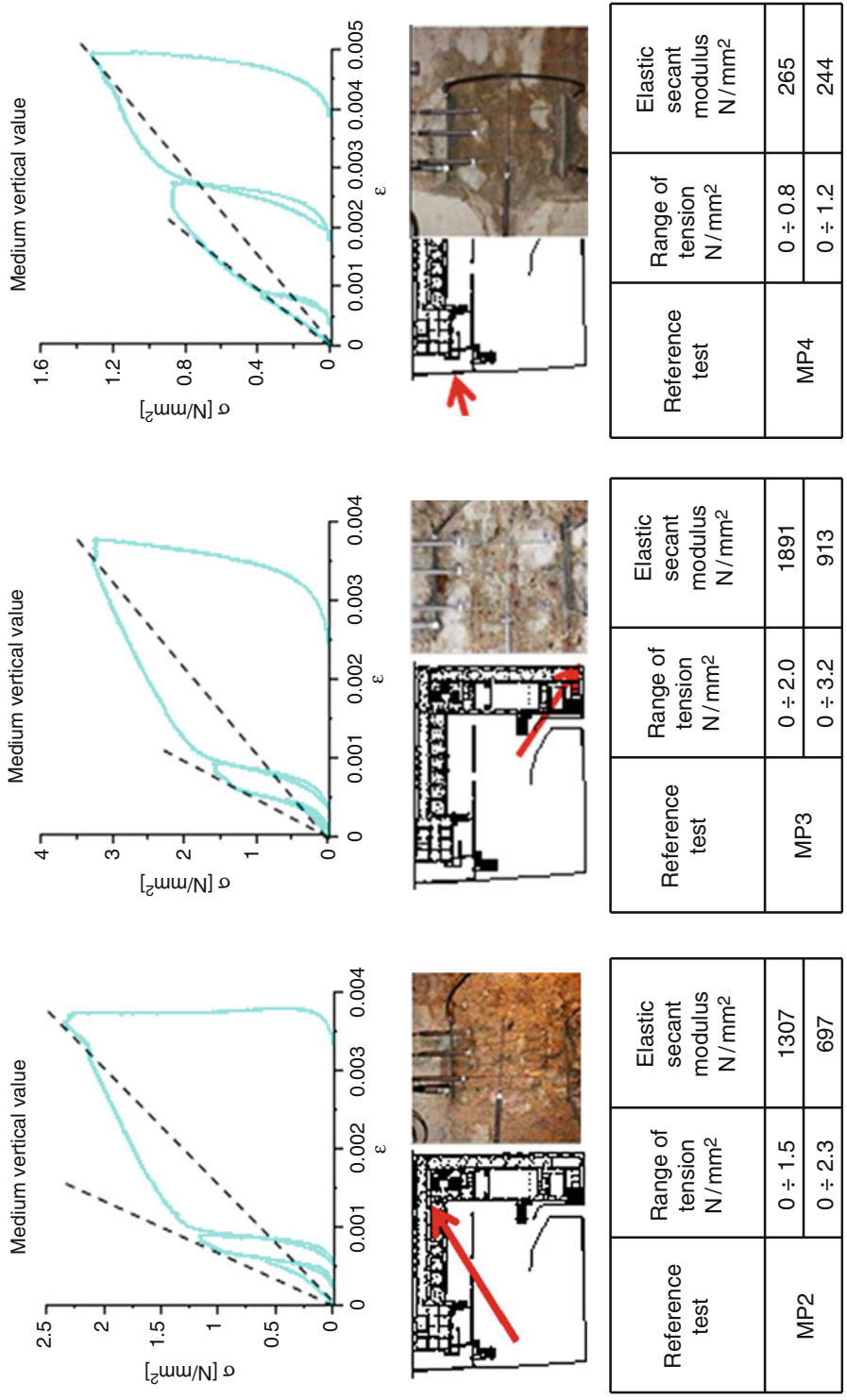
Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 8 Results of double flat-jack tests on Camponeschi Palace
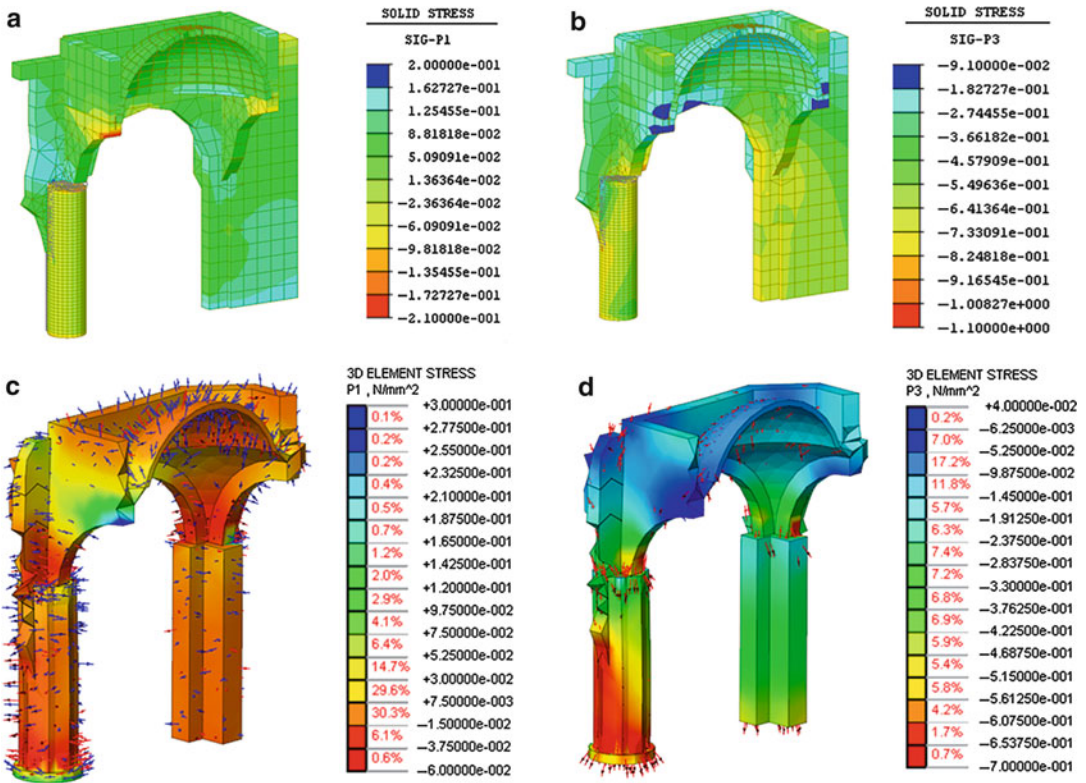
Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 9 Principal stresses due to static loads in the transept: model A, (**a**) maximum P1 and (**b**) minimum P3; model B, (**c**) maximum P1 and (**d**) minimum P3

of rigid masonry blocks, a series of different modeling approaches have been experimented within the activities of the computational Mechanics Lab of the University of L'Aquila directed by the first author. On this respect, the experience gained at L'Aquila benefits from the large amount of data available for possible comparison with the numerical simulations. Therefore, potentiality and drawbacks of several modeling assumptions have been evidenced along the path, which have shown to be particularly useful also to enrich the tools for retrofitting design. The structural modeling of the Basilica of S. Maria di Collemaggio constitutes a relevant case study in this field. Figure 9 reports, synthetically, the results of two stress analyses induced by live loads conducted by means of two different finite element models of the entire Basilica: in the first case (model A), four-node quadrilateral isoparametric plate-shell elements have been

used to model the main walls in which a prevalent bidimensional behavior has been recognized (Fig. 9a, b) (see also Gattulli et al. 2013); in a second case (model B), a more refined model has been considered in which all the masonry elements have been modeled using eight-node isoparametric solid elements (Fig. 9c, d). It is evident as the stress flux is more clearly defined in the arches and in the pillars by the more refined model, even if the level of stress amplitude is comparable in the two cases. At the same time, the analysis evidences that in the pillars is present an important flexural deformation induced by the thrust differences in the longitudinal direction between the last arch belonging to the nave walls and the arch of the transept on which is laying the dome and the vault. Even if the results are obtained under the hypothesis of modeling the masonry material as an elastic and homogeneous solid, they evidence the role played by the

three-dimensional geometry and the consequent internal action distribution. Figure 9c evidences as the maximum tension flux in the column is prevalently radial, while the minimum one in Fig. 9d is vertical.

A different activity has been carried out on the Camponeschi Palace which represents a good example of how damage observation and experimental survey, performed using the techniques described above, can be a valuable aid in supporting the modeling choices for both damage interpretation and simulating the effect of retrofitting interventions.

The Camponeschi Palace is an important historical building located in the historical city center of L'Aquila. The palace was built near the end of the sixteenth century by the Jesuit order that established and hosted in the structure an institution of higher education. Several modifications to the original configuration and structural interventions have been made during the years. The palace sets upon a rather sloping site and it is characterized by an L-shaped plan with two arms of equal length delimiting an internal courtyard. The building was hosting the Faculty of Philosophy and Letters of University of L'Aquila at the time of the 2009 earthquake.

Due to the relevant structural damage, which occurred, the Department of Civil Protection classified the palace structurally unsafe. The crack pattern of the facade due to the 2009 L'Aquila mainshock and successive earthquake swarms was detected through automated laser-based surveys. The crack pattern distribution of the main facade is shown in Fig. 10a. The major damage was concentrated on the first story. On the left side of the facade, subvertical and diagonal cracks developed at the end sections of the spandrels above and below the openings.

Single and double flat-jack tests were carried out on site in order to evaluate the compressive strength and the deformability properties of the masonry for the entire building.

A numerical model was built specifically to simulate the nonlinear static behavior of the facade and of a retrofitting system using fiber reinforcement polymers for the masonry panels. Eight-node isoparametric plane stress elements

with a 3 × 3 Gauss integration scheme were used. Figure 10b shows the adopted mesh. The incremental/iterative solution procedure employed is a modified Newton–Raphson approach. A constitutive law, comprising a parabolic hardening rule and a parabolic exponential softening branch, modeled the compression behavior of the masonry; a linear hardening branch followed by a linear softening branch characterized the tension behavior. The FRP reinforcement was modeled using truss elements directly connected to the nodes of the mesh of the panels without using interface elements. Thus, perfect adhesion between the nodes of trusses and the corresponding nodes of the mesh was considered. The strips were considered not able to carry compression loads. The tensile behavior of the FRP was represented with a stress–strain relationship, characterized by a linear elastic behavior in tension, followed by exponential degradation.

Nonlinear incremental static (pushover) analyses were carried out on unreinforced and reinforced facade, separately for the positive (Fig. 10c) and negative x direction (the horizontal one) (Fig. 10d) and for both the unreinforced and FRP-reinforced masonry (Gattulli et al. 2014).

In general, a correct and reliable implementation of the model allows performing in-depth analyses of the expected structural behavior under seismic loads enabling robust design choices.

Structural Monitoring for Historic Structures

Structural health monitoring is an emerging tool for reliable structural assessment. Efficient monitoring programs may help in the characterization of the progressive decay of structural performance, both in the short and in the long term, providing useful information for optimized maintenance and safety. Generally, electronic equipment deployed with the purpose of structural health monitoring is implemented via wired systems. In this case, the development of the system can be limited by a series of factors such as the high cost of measurement equipment, difficulty in installation due to large instruments size, and need of a wired communication and power infrastructure. In recent years, great attention has been
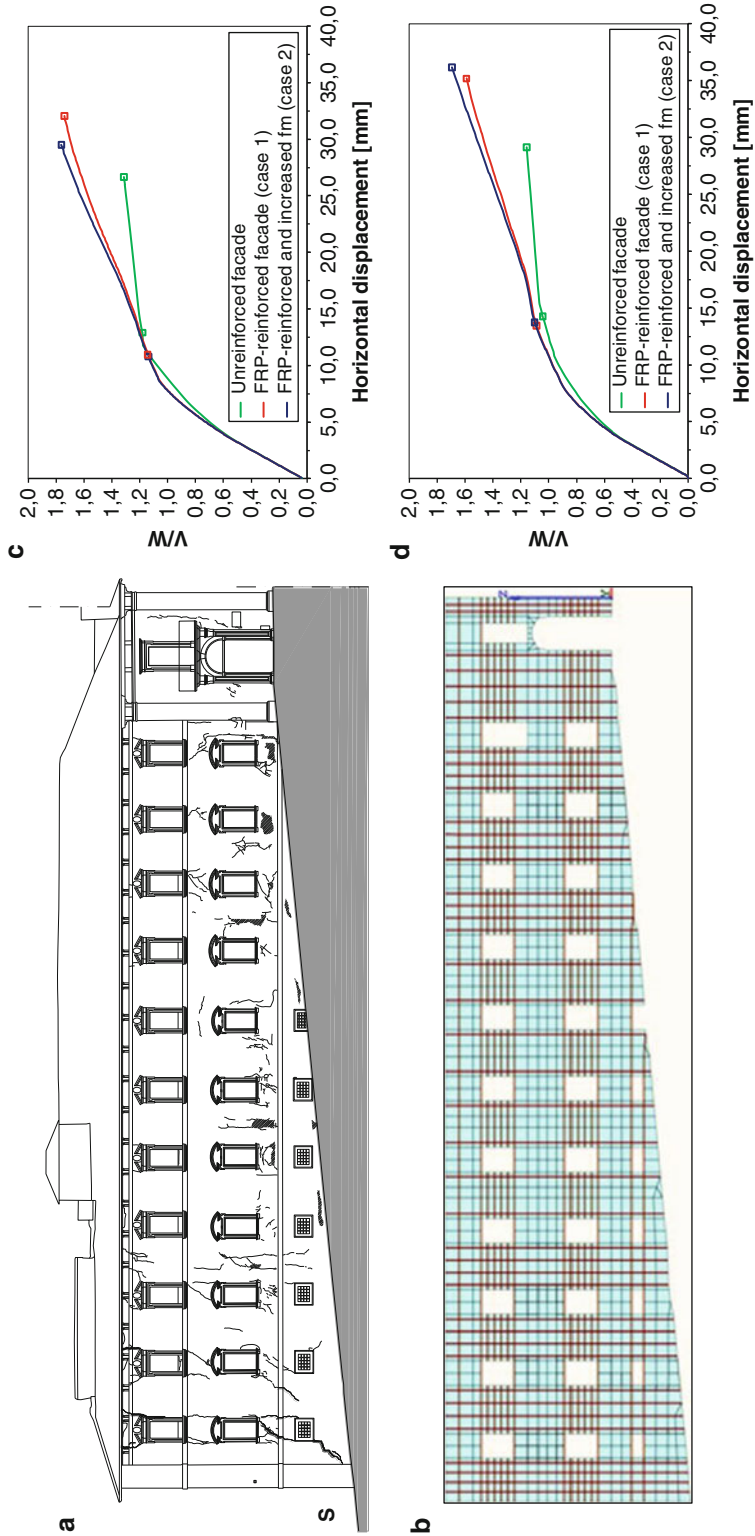
Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 10 Camponeschi Palace: (**a**) damage pattern on the facade (**b**) mesh of the FEM model with the FRP layout; numerical pushover curves for horizontal loads (**c**) +x direction and (**d**) −x direction

S

paid to the development of wireless sensor networks that are able to overcome several problems linked to the development of the entire layout but evidencing other issues related to the new technology. One of which is connected to the collection and transmission of a great quantity of data. In fact relevant information can be obtained by the registrations of the vibration data induced by ambient sources, wind, earthquakes with low intensity, operation of machines, and traffic. The sensors are placed in specific points of the structure and appropriately selected and the choice of which is based, for example, on numerical modeling. Using the acquired data, it is possible to identify dynamical models, which predict the observed behavior on the basis of error minimization. Differences in the identified systems produced by data acquired in different periods separated by a long interval of time may permit to identify relevant changes in the behavior of the systems which can be linked to degradation of the structural performance or incurred damages. For example, in largely damped, like masonry, structures, it is convenient to use high sampling times for acquisition as opposed to frame structures. Although the interesting structural modes can be placed within an interval 0–10 Hz, the structural response may be highly damped and attenuated with a reasonable amplitude level contained in few seconds. The acquired time histories under seismic action will be composed of a few samples that make difficult the reconstruction of the frequency spectrum. Some important benefits are that smart sensors include microprocessors that are able to run some basic algorithmic tasks and then store some condensed information to be exchanged between the other sensors (nodes) that comprise the wireless network (WN).

The long-term activity at the University of L'Aquila aims to develop innovative automated techniques, which permit to automatically update the appropriately selected models, which in turn may address questions regarding the seismic safety of the monitored structures. In particular, following a parametric identification approach, using the distributed available measurements from sensors embedded in the reinforcements, the mechanical parameters characterizing the local behavior of the reinforcement either in the reinforcing fibers and in their support may be opportunely monitored. Subsequently, through information data fusion made available by arrays of sensitized reinforcements distributed on a masonry macroelement, mechanical parameters influencing the global behavior may be identified. Possible developments concern the use of measurements at different levels of excitation amplitudes, which are available in seismic areas. They may be used to define a methodology for identifying nonlinear behavior, distinguishing between scenarios where damage occurs in the fiber, in the support, or in the masonry structure.

The implementation of a full SHM system in historical buildings is still an open problem especially in aspects relating to the correct fusion of vibration measurements, integrated with other type of measurements. For example, in concrete structures it has proven useful to follow the trend of humidity and temperature in order to evaluate deterioration and updated material performances. In this manner early assessment and warning of incipient problems may be used to drive quick interventions or the planning and scheduling of the maintenance programs. To the contrary, in monumental structures often numerous cracks are present which can be the objective of permanent static monitoring. Indeed, following the slow opening and closing of the crack may be used as an important correlation measurement for validating information provided for the identification of other parameters. The opening and closing may depend on the cyclic trend of the average temperature corresponding to a certain time interval (day, month, season). Similarly to this trend, it may possible to observe a reduction or increase of stiffness or a change of configuration that may induce also a cyclical trend in the identification of natural frequencies.

Of course, this set of actions permits both a rationalization of the cost and an improvement of the service-life predictions, potentially also for monumental structures, especially if they have been recently restored and retrofitted with a consequent important investment.

In general there are no particular rules for the design of a fully operating SHM system, but each case is unique and it requires particular care especially with respect to the goals to be achieved. The observation and the investigation on the damage reported by the structure after a seismic event can lead to a customized system able to provide specific information for continuously assessing structural condition. Correlations among obtained results should be sought for and only after analyzing all available information; the responsible team, formed by professionals with different skills (structural, geotechnical, architectural, electronic, and informatics), on the basis of their experience and sensibility, can formulate a final diagnosis and define the actions to be performed.

Within such a context, in the first phase of the SHM system deployment, at the Basilica di Santa Maria di Collemaggio, different actions have been performed including design, implementation, installation, and test of a network monitoring accelerations. The main goal was an accurate measurement of the basic structural response, both to environmental actions and to low-amplitude seismic events, which occasionally continue to occur at the site. A wireless sensor network was used as a platform combining a MEMS tri-axial accelerometer and also sensors to measure humidity, temperature, and luminosity (Gattulli et al. 2014). The proper positioning of the sensors has been designed through the study of different numerical models accounting for the interaction between the main wall element and the temporary scaffolding (Fig. 11a). In particular, the dynamic behavior can be observed looking at the principal modal characteristics (frequency, modal shape, and percentage of participating mass) of the linear dynamic response (Fig. 11b, c). The analyses suggest, for example, that, by placing the sensors along the internal longitudinal facade, it is possible to survey the transversal displacements especially during a seismic event. Until today, after almost 3 years of permanent monitoring, the system is still able to capture relevant data during the occurrence of several seismic events coming from far and near fields. Generally, the registered maximum response amplitude in the acquired time history (one of this shown in Fig. 11f) determines the possibility to get information regarding the main modal features.

Nowadays, in order to derive the processed information, from experimental data, different modal identification procedures can be used which are generally classified based on whether they operate in the time domain or in the frequency domain. However, the amount of contained information is relatively independent of the analysis domain.

Among several procedures probably, the two most popular algorithms are the Enhanced Frequency Domain Decomposition (EFDD) and the Stochastic Subspace Identification (SSI) (Peeters and De Roeck 2001), and the first operates in the frequency domain and the second in the time domain.

The two procedures assume that the system input (loads) is white noise that can be thought as a representative model of environmental noise. However, even if the adoption of the white noise model in the case of a seismic event may be questionable, the straightforward applications of the abovementioned procedures to response data acquired during small earthquake still allow to derive valuable information especially if special care is taken, such as in the case of the studied system at the Basilica (Gattulli et al. 2015a, b). Indeed, here, using the SSI procedure, for example, the evaluation of the identified stable poles may include dominant components of the input, producing possible errors or misinterpretation of the dynamic behavior. However, the direct comparison with results obtained by numerical models driven by engineering judgment permits to extract realistic structural modal data. Moreover, the robustness of this data can be pursued through the use of several different registrations generated by earthquakes with different characteristics dependent on the source location. Following the reliable identification of the modal characteristics, the second step is the updating of the numerical model parameters in order to minimize the error between measured and numerical modal features. In the error index formulation, natural frequencies and/or modal shapes
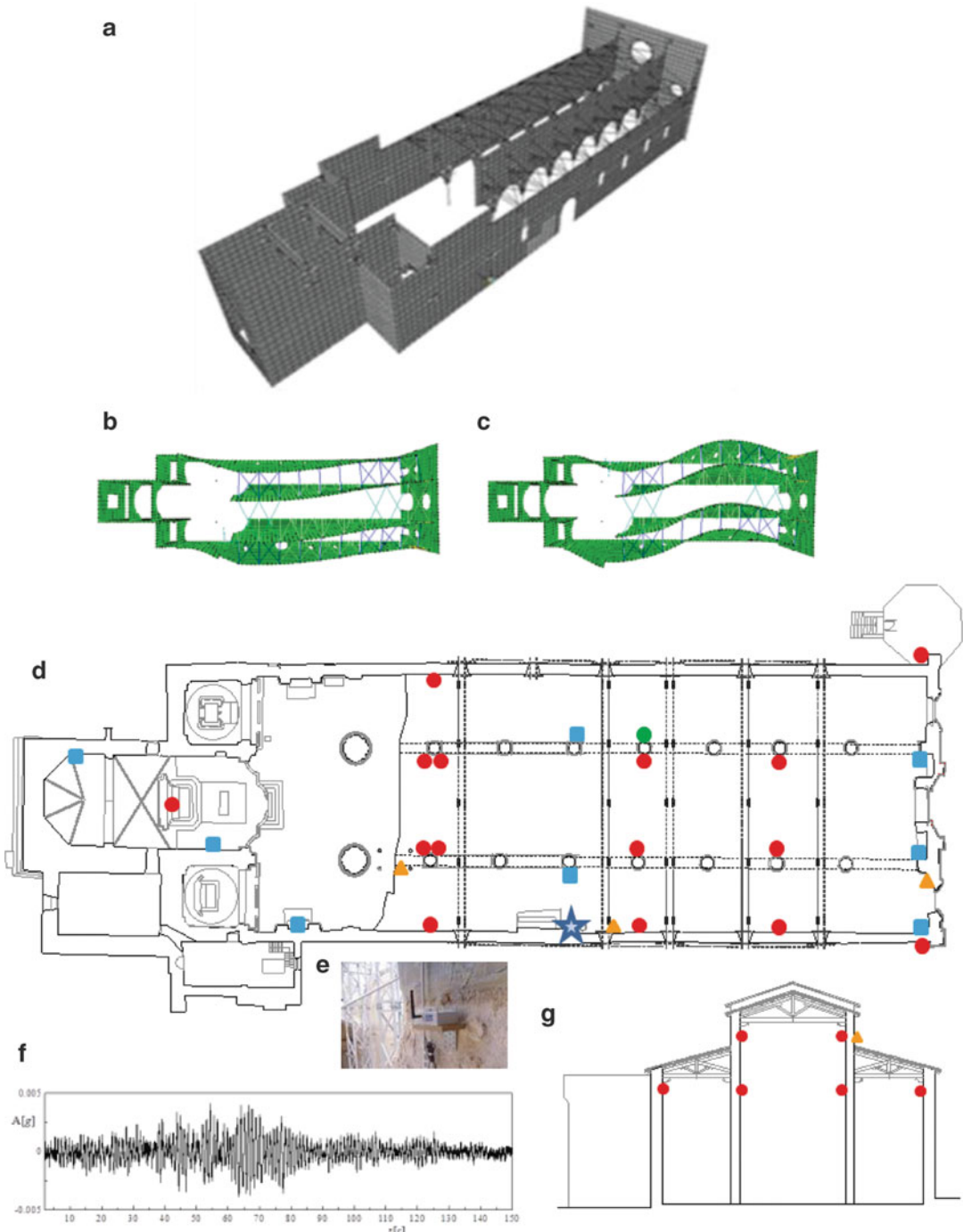
Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 11 Structural monitoring at the Basilica of S. Maria di Collemaggio: (**a**) finite element model accounting for the interaction between the main wall element and the temporary scaffolding; (**b**) first (1.09 Hz) and (**c**) second (1.97 Hz) mode shape; layout of the multifunction wireless sensor network: (**d**) plant, (**e**) sensor, (**g**) section, and (**f**) acquired registration during a seismic event
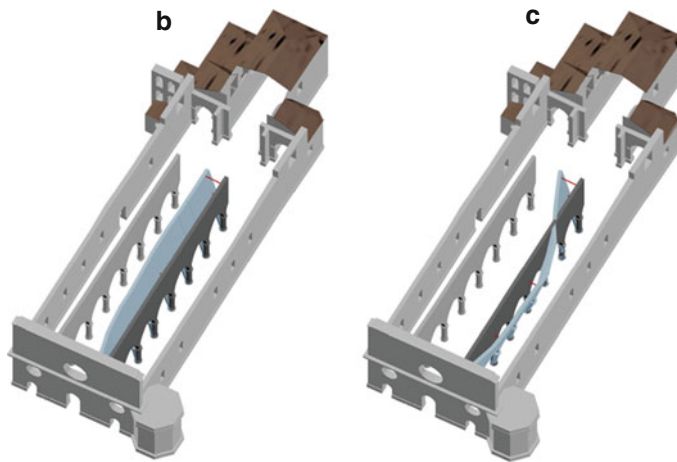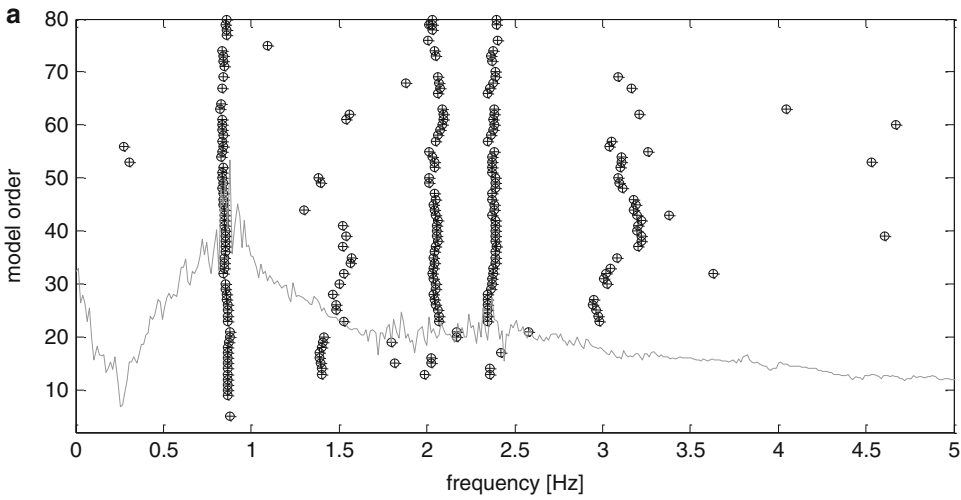
Seismic Behavior of Ancient Monuments: From Collapse Observation to Permanent Monitoring, Fig. 12 Identified modal model by structural monitoring:

(**a**) stabilization diagram of the SSI procedure; (**b**) first (0.90 Hz) and (**c**) second (1.45 Hz) identified mode shape

may be considered. In the first case, a simple difference in percentage is generally considered, while the Modal Assurance Criteria (MAC) is used to compare measured and numerical modal shapes. The MAC criterion indicator permits to evaluate the differences between similar modes but also the cross-orthogonality between different numerical and experimental modal shapes.

As an example, Fig. 12a illustrates the stabilization diagram obtained from the time histories acquired by the wireless sensor network using SSI during a seismic event. The ordinates include the model order, which simply refer to the size of

the Hankel matrix considered in the identification process, while the abscissa is reporting the frequency; the circles with a cross inside indicate stable modes, while the others, colored in gray, are unstable. At least one stable mode is present at different frequencies, but in the case of two specific frequencies, stable modes seem to be present at almost all different model orders considered in the analysis. These stable poles are present at 0.90 Hz and in the range between 1.45 and 2 Hz. In the latter case, a comparison with numerical models permits possible interpretation of the modal identification. The first two

identified modes are illustrated in Fig. 12b, c illustrating the deformed shape of the portion of the Basilica in which the response used in the identification procedure has been registered. The second phase of the monitoring system development involves the deployment of crackmeter and inclinometer sensors at the Basilica. Figure 11d summarizes the employed instrumentation, where the red circles indicate the wireless sensor network relative to the accelerometer; the blue squares and the orange triangles correspond to the crackmeters and the inclinometers, respectively. Moreover, the green circles refer to an accelerometer placed at the base of a pillar necessary for acquiring the acceleration of the seismic input, while the star indicates the positioning of the node gateway that collects the data of each sensor and sends them to the server. The data fusion of the two sets of information is currently under investigation.

Summary

The earthquake, which occurred on April 6, 2009, has been a catastrophic event for both the city and the University of L'Aquila. Nevertheless, the disaster has transformed itself into a tremendous opportunity to revitalize the area, providing the potential for the national and international scientific community to test the effectiveness of new systems and technologies and consequently to establish main directives for the monitoring of monumental structures. The applicability of new technologies in the areas of observation, surveying, testing, modeling, restoration, retrofitting, and monitoring of ancient monuments, of which the city center is densely populated, is extensively documented herein, through an organized outline and a series of case studies.

Cross-References

- ▶ Ambient Vibration Testing of Cultural Heritage Structures
- ▶ Damage to Ancient Buildings from Earthquakes
- ▶ Masonry Modeling
- ▶ Retrofitting and Strengthening Masonries of Heritage Structures: Materials Used
- ▶ Seismic Vulnerability Assessment: Masonry Structures
- ▶ Stochastic Structural Identification from Vibrational and Environmental Data

References

Brandonisio G, Lucibello G, Mele E, Luca AD (2013) Damage and performance evaluation of masonry churches in the 2009 L'Aquila earthquake. Eng Fail Anal 34:693–714

Carocci CF (2012) Small centres damaged by 2009 L'Aquila earthquake: on site analyses of historical masonry aggregates. Bull Earthq Eng 10(1):45–71

Casarin F, Benetta MD, Modena C, da Porto F, Valluzzi MR, Cantini L, Tedeschi C, Condoleo P (2012) Structural monitoring and investigation campaign in the church of St. Giuseppe dei minimi in L'Aquila, after the 6th of April 2009 earthquake. RILEM book series, Springer, Berlin, vol 6. pp 1187–1193

Ceci AM, Contento A, Fanale L, Galeota D, Gattulli V, Lepidi M, Potenza F (2010) Structural performance of the historic and modern buildings of the University of L'Aquila during the seismic events of April 2009. Eng Struct 32(7):1899–1924

Ceci AM, Gattulli V, Potenza F (2013) Serviceability and damage scenario in RC irregular structures: post-earthquake observations and modelling predictions. J Perform Constr Facil 27(1):98–115

Cimellaro GP, Piantà S, De Stefano A (2012) Output-only modal identification of ancient L'Aquila city hall and civic tower. J Struct Eng 138(4):481–491

D'Ayala DF, Paganoni S (2011) Assessment and analysis of damage in L'Aquila historic city centre after 6th April 2009. Bull Earthq Eng 9(1):81–104

Da Porto F, Silva B, Costa C, Modena C (2012) Macroscale analysis of damage to churches after earthquake in Abruzzo (Italy) on April 6, 2009. J Earthq Eng 16(6):739–758

Da Porto F, Munari M, Prota A, Modena C (2013) Analysis and repair of clustered buildings: case study of a block in the historic city centre of L'Aquila (Central Italy). Construct Build Mater 38:1221–1237

Decanini LD, Liberatore L, Mollaioli F (2012) Damage potential of the 2009 L'Aquila, Italy, earthquake. J Earthq Tsunami 6(3):art. no 1250032, 32 pp

Federici F, Graziosi F, Faccio M, Gattulli V, Lepidi M, Potenza F (2012) An integrated approach to the design of wireless sensor networks for structural health monitoring. Int J Distrib Sens Netw Art. No. 594842, 16 pp

Foti D, Gattulli V, Potenza F (2013) Output-only identification and model updating by dynamic testing in

unfavorable conditions of a seismically damaged building. Comput Aided Civ Infrastruct Eng 29 (9):659–675

Gattulli V (2013) Advanced applications in the field of structural control and health monitoring after the 2009 L'Aquila earthquake. In: Sebastiano D'Amico (ed) Engineering seismology, geotechnical and structural earthquake engineering. InTech, Croatia, pp 208–235. ISBN: 978-953-51-1038-5, doi:10.5772/55438

Gattulli V, Antonacci E, Vestroni F (2013) Field observations and failure analysis of the Basilica S. Maria di Collemaggio after the 2009 L'Aquila earthquake. Eng Fail Anal 34:715–734

Gattulli V, Lampis G, Marcari G, Paolone A (2014) Simple modeling of FRP reinforcement: from masonry panels to a historic palace facade. Eng Struct 75:604–618

Gattulli V, Potenza F, Federici F, Graziosi F, Colarieti A, Faccio M (2015a) Distributed structural monitoring for a smart city in a seismic area. Key Eng Mater 628:123–135

Gattulli V, Potenza F, Lepidi M, Federici F, Graziosi F, Colarieti A (2015b) Seismic structural monitoring by wireless sensor networks of the Basilica S. Maria di Collemaggio after the transept collapse. Earthq Eng Struct Dyn (in preparation)

Guidoboni E, Comastri A, Mariotti D, Ciuccarelli C, Bianchi MG (2012) Ancient and medieval earthquakes in the area of L'Aquila (northwestern Abruzzo, central Italy), A.D. 1–1500: a critical revision of the historical and archaeological data. Bull Seismol Soc Am 102(4):1600–1617

Indirli M, Kouris LAS, Formisano A, Borg RP, Mazzolani FM (2013) Seismic damage assessment of unreinforced masonry structures after the Abruzzo 2009 earthquake: the case study of the historical centers of L'Aquila and Castelvecchio Subequo. Int J Archit Herit 7(5):536–578

Lucibello G, Brandonisio G, Mele E, Luca AD (2013) Seismic damage and performance of Palazzo Centi after L'Aquila earthquake: a paradigmatic case study of effectiveness of mechanical steel ties. Eng Fail Anal 34:407–430

Milani G, Russo S, Pizzolato M, Tralli A (2012) Seismic behavior of the San Pietro di Coppito Church bell tower in L'Aquila, Italy. Open Civ Eng J 6(1):131–147

Özerdem A, Rufini G (2013) L'Aquila's reconstruction challenges: has Italy learned from its previous earthquake disasters? Disasters 37(1):119–143

Parisi F, Augenti N (2013) Earthquake damages to cultural heritage constructions and simplified assessment of artworks. Eng Fail Anal 34:735–760

Peeters B, De Roeck G (2001) Stochastic system identification for operational modal analysis: a review. J Dyn Sys Meas Control 123(4):659–667

Russo S (2013) On the monitoring of historic Anime Sante church damaged by earthquake in L'Aquila. Struct Control Health Monit 20(9):1226–1239

Tertulliani A, Leschiutta I, Bordoni P, Milana G (2012) Damage distribution in L'Aquila city (central Italy) during the 6 April 2009 earthquake. Bull Seismol Soc Am 102(4):1543–1553

Volpe M, Piersanti A, Melini D (2012) Complex 3-D finite element modelling of the 2009 April 6 L'Aquila earthquake by inverse analysis of static deformation. Geophys J Int 188(3):1339–1358

Seismic Collapse Assessment

Christoph Adam[1] and Luis F. Ibarra[2]
[1]Department of Engineering Science, Unit of Applied Mechanics, University of Innsbruck, Innsbruck, Austria
[2]Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, USA

Synonyms

Collapse criterion; Collapse variability; Component deterioration; Nonlinear dynamic analysis; P-delta effect; Structural modeling strategies

Introduction

In earthquake engineering, structural collapse is defined as the local or global failure of a system that occurs due to the loss of vertical load-carrying capacity in the presence of seismic events. The two primary modes of global collapse are sidesway and vertical collapse. Sidesway collapse is the global failure of the system caused by a reduction of the lateral load-bearing capacity due to large horizontal displacements, whereas vertical collapse is caused by a direct loss of the gravity load-bearing capacity in one or several structural components (Krawinkler et al. 2009). Vertical collapse is a type of progressive collapse, which is the total or disproportionate failure of the system triggered by an initial local failure that spreads out from element to element.

Commonly, in ductile frame structures, sidesway collapse is the predominant mode of

collapse during catastrophic earthquake events. These systems are affected by component material deterioration and/or P-delta effects. Material deterioration reduces the component's strength under large inelastic displacements. Moreover, the strength and stiffness of the structural components deteriorates in each load cycle, reducing the load-bearing capacity. The importance of material deterioration decreases for long-period structures, where the P-delta effects (also referred to as second-order effects) may lead to a negative post-yield stiffness that in combination with large inelastic displacements accelerates the onset of collapse.

In most non-ductile frames, however, columns lose the capacity to sustain gravity loads due to shear and subsequent axial-load failures, and as a result, the collapse mechanism is controlled by the loss in vertical load-bearing capacity. As outlined in Baradaran Shoraka (2013), this type of collapse of older concrete buildings may be precipitated by axial-load failure of columns, punching shear failure of slab-column connections, failure of slab-diaphragm connections, or axial-load failure beam/column joints.

Prediction of earthquake-induced collapse requires the identification of all possible modes of collapse, appropriate structural modeling of building components under cyclic loading, and suitable representation of earthquake excitation including the consideration of epistemic and aleatory uncertainties. Krawinkler et al. (2009), Haselton et al. (2009), and Baradaran Shoraka (2013) provide insights into the phenomenon of dynamic collapse of earthquake-excited structures. For basic concepts in prediction of collapse and its implementation to performance-based earthquake engineering, the reader is referred to Zareian et al. (2010). Villaverde (2007) presents a comprehensive literature that provides a path through which seismic collapse assessment has evolved in the last decades.

The required accuracy and the available resources determine the selection of a seismic collapse assessment procedure. In a first step, the governing seismic modes of collapse such as vertical collapse or sideway collapse are identified. Then a structural model is created that
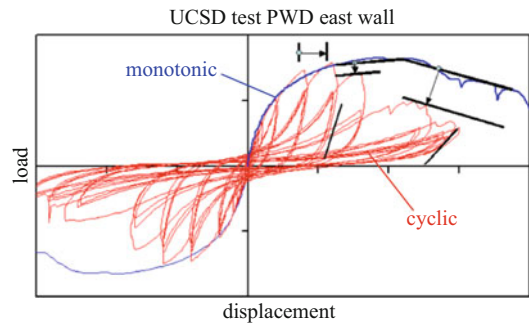
Seismic Collapse Assessment, Fig. 1 Force–displacement relations from a monotonic and a cyclic test (Modified from Gatto and Uang 2002; With permission from Prof. Uang, UC San Diego)

captures these failure modes, and modeling uncertainties are quantified. Evaluation of seismic collapse should be based on structural analyses that incorporate strength and stiffness material deterioration, as well as geometric nonlinearities (P-delta effects). Collapse of single-degree-of-freedom (SDOF) systems was first studied by including only P-delta effects in seismic response, which may cause a negative inelastic tangent stiffness that eventually leads to system collapse (Jennings and Husid 1968; Bernal 1987; MacRae 1994). The development of hysteretic models that include strength and stiffness deterioration (Kunnath et al. 1990; Sivaselvan and Reinhorn 2000; Song and Pincheira 2000; Ibarra and Krawinkler 2005) improved the assessment of collapse capacity. As observed in Fig. 1, collapse assessment requires models that can predict strength reduction under large displacements, as well as strength and stiffness deterioration due to cyclic loading.

This article focuses on sideway collapse of frames, which has been traditionally estimated by means of non-degrading systems, in which judgmental limits are assigned to response quantities such as roof drift, story drift, ductility, and hysteretic energy. These parameters are usually referred to as engineering demand parameters (EDPs) (Moehle and Deierlein 2004). The development of degrading models permits tracking the collapse limit state, but EDPs become very sensitive when the system is close to collapse, and

Seismic Collapse Assessment, Fig. 2 IDA curve for a single ground motion up to collapse: elastic, inelastic, and EDP-sensitive domain; collapse (With permission from Adam 2014)
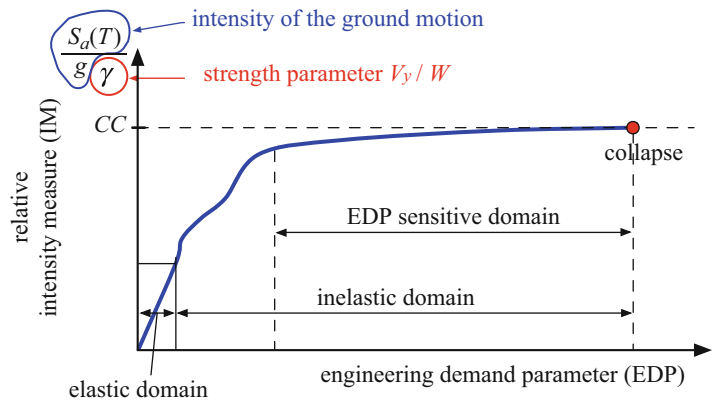
small perturbations in input parameters lead to large variations in the response. For this reason, global collapse methodologies are commonly based on an intensity measure (IM) instead of an EDP. A relative IM is the ratio of ground motion intensity (demand) to a structure strength parameter (capacity). For a given structure and ground motion, collapse evaluation consists of a series of response history analysis starting with a relative intensity that produces an elastic response of the system. Then the relative intensity is increased to cover the full range of interest or until collapse takes place. This series of response history analysis is called incremental dynamic analysis (IDA) (Vamvatsikos and Cornell 2002). As an example, Fig. 2 shows an IDA curve that results from increasing the IM up to collapse. Collapse is imminent when a very small increment in relative intensity causes a very large increment in the EDP of the system, a condition that indicates dynamic instability. In Fig. 2 dynamic instability occurs when the slope of the IDA curve approaches zero.

The global seismic collapse capacity of a system is defined as the largest earthquake intensity at which the structure still maintains dynamic stability (Krawinkler et al. 2009), denoted as CC in Fig. 2. This definition is associated with a specific structure under a given ground motion. Dynamic analysis requires the specification of system properties and input ground motions, but variations in these quantities may produce large dispersion in the response. Therefore, collapse

needs to be evaluated in a probabilistic framework that includes uncertainties in the frequency content of the ground motions (e.g., record-to-record (RTR) variability) and in the input parameters of the system (e.g., epistemic uncertainty). Then, nonlinear response history analyses usually include a set of ground motion records that represents the RTR variability at the site of the structure. The effect of uncertainty in nonlinear deterioration parameters on the variance of collapse capacity can be quantified by means of approximate methods (e.g., FOSM method) or carrying out Monte Carlo simulations (Ibarra and Krawinkler 2011; Ugurhan et al. 2014).

The variance in collapse capacity can be used to generate collapse fragility functions that provide the probability of collapse given that an event with a certain IM occurred. Shome and Cornell (1999) summarize the arguments used to consider the collapse fragility as a log-normally distributed function. For simplified procedures, only the median, 16th, and 84th percentiles of the individual collapse capacities are required to find an approximation of the collapse fragility. A final product of collapse assessment is the mean annual frequency of collapse, which is obtained by combining fragility curves with the hazard information at the site.

In this contribution some of the issues involved in seismic sidesway collapse assessment of moment-resisting frames are addressed in more detail, including nonlinear dynamic analysis, structural modeling, strength and stiffness

material deterioration, and P-delta effect. The probabilistic framework used to evaluate collapse and the effect of epistemic uncertainties and aleatory variability on the collapse capacity variance are discussed, although ground motion selection is not part of this article. The last sections deal with the development of simplified methods to evaluate collapse capacity.

Nonlinear Dynamic Analysis

In the last two decades, significant progress has been achieved in the nonlinear dynamic analysis of structures. It is now established as the most reliable tool to predict realistically the seismic collapse capacity of a building, and several provisions such as FEMA P-695 (2009), ATC-58 (2012), and ATC-78 (2013) propose collapse assessment methodologies based on nonlinear dynamic analysis.

Incremental Dynamic Analysis

If sideway collapse is the governing failure mechanism, the IDA procedure can be used to determine the largest earthquake intensity at which the structure still maintains stability. In this approach nonlinear dynamic response history analyses are performed repeatedly, increasing in each subsequent run the ground motion intensity. The IDA curve of the structure for the considered earthquake record is obtained by plotting an appropriate record intensity measure against its corresponding EDP (Vamvatsikos and Cornell 2002). The analysis can be stopped when the EDP satisfies a certain failure criterion that may correspond to structural sideway collapse. In this case, the corresponding intensity of the ground motion is referred to as collapse capacity of the building for this specific ground motion record.

IDA is commonly performed for a representative set of ground motions to account for RTR variability. Examples of such sets are the LMSR ground motions (Medina and Krawinkler 2003) and FEMA P-695 sets for far-field and near-field ground motions (FEMA P-695 2009). In this sense, the ground motions are used as standardized excitation protocols, and the effort of

selecting ground motions is avoided (Baker 2013). These records need to be scaled to the same IM, and since there is not a unique definition, several IMs have been proposed: (a) elastic ground motion-based scalar IMs such as peak ground acceleration (PGA), peak ground velocity (PGV), and peak ground displacement (PGD); (b) elastic and inelastic spectral-based IMs such as spectral acceleration and spectral displacement at the fundamental period of the structure, as well as spectral values related to higher mode effects or period elongation; and (c) vector valued IMs (e.g., Baker and Cornell 2005). Currently, the most widely used IM is the 5 % damped spectral pseudo-acceleration $S_a(T_1, \zeta = 0.05)$ at the fundamental period T_1 of the structure. For example, if $S_a(T_1, \zeta = 0.05)$ is normalized by the product of the acceleration of gravity g and the base shear coefficient γ, the relative seismic collapse capacity of the building subjected to the ground motion "i" reads as

$$CC_i = \left. \frac{S_{a,i}}{g\gamma} \right|_{\text{collapse}} \qquad (1)$$

The base shear coefficient γ is defined as the base shear V_y at onset of yielding divided by the total weight W.

The collapse fragility function is obtained by counting the fraction of collapse capacities CC_i ($i = 1, 2, \ldots, n$) that correspond to the n ground motions of the considered set. When all ground motions lead to collapse, the collapse capacity distribution is usually log-normal distributed (Ibarra and Krawinkler 2005). As an example, Fig. 3 shows the IDA curves of a structure subjected to the 44 ground motions of the FEMA P-695 far-field ground motion suite and the corresponding collapse fragility function.

Multiple Stripe Analysis

The IDA approach cannot consider certain site-specific ground motion conditions. As outlined by Baker (2013), some of the ground motion properties relevant for collapse assessment are the spectral shape of the elastic response spectrum, the duration, and the effects of near-fault ground motions and velocity pulses. The
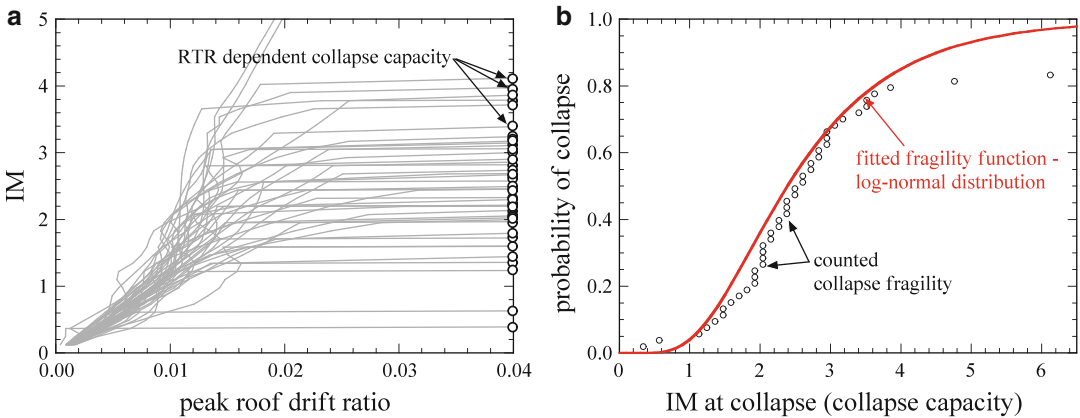
Seismic Collapse Assessment, Fig. 3 (**a**) IDA curves of a structure for all records of a ground motion set. (**b**) Counted collapse fragility and corresponding fitted fragility function according to a log-normal distribution (With permission from Adam 2014)

expected spectral shape of a ground motion does not only vary with location but it also depends strongly on the ground motion amplitude (Baker and Cornell 2006). For instance, at different seismic intensity levels, the spectral shape of the expected earthquake is dissimilar. These differences in the spectral shapes can be ascribed to the soil condition, magnitude, distance, and parameter ε, which quantifies the difference between the selected ground motion's spectral acceleration at a specific period and the median of the ground motion spectral acceleration obtained from attenuation relationships (Baker 2013). Because the same set of records are often used for different locations, however, IDA yields the same risk of sidesway collapse for two equal buildings located in different locations with similar seismic hazard, although the anticipated ground motion spectral shapes are different.

To address some of the above issues, Jalayer and Cornell (2009) developed a multiple stripe analysis (MSA) that accounts for specific site conditions and for ground motion properties that change with increasing magnitude. In this approach, the ground motions are reselected at each IM level according to the changing spectral shape. For each record the peak response of the EDP derived from nonlinear response history analysis and the number of events that leads to sidesway collapse is plotted as a function of the corresponding IM, thus leading to "stripes."

The counted fraction of observed individual collapses at each IM level leads then to the collapse fragility curve.

Collapse Criteria

The numerical sidesway collapse prediction depends on the model capabilities. Historically, the first nonlinear numerical models did not include deteriorating characteristics, and damage indexes were developed to predict collapse based on maximum displacements and/or dissipated energy. The most widespread cumulative damage model was the Park–Ang model (Park and Ang 1985), which was developed specifically for reinforced concrete components. The damage measure (*DM*) consisted of a linear combination of displacement and energy demands:

$$DM = \frac{\delta_{\max}}{\delta_{\mathrm{ult}}} + \frac{\beta}{F_y \delta_{\mathrm{ult}}} \int \mathrm{d}HE \qquad (2)$$

where $\delta_{\max}$ is the maximum displacement of the system, δ_{ult} denotes the monotonic displacement capacity of the system, β represents the structural performance parameter, and *HE* is the hysteretic energy. In concept, the parameters δ_{ult} and β of this cumulative damage model are equivalent to the deteriorating parameters of recent hysteretic models that include cumulative deterioration. However, damage index models were based on non-deteriorating models, leading to unrealistic

S
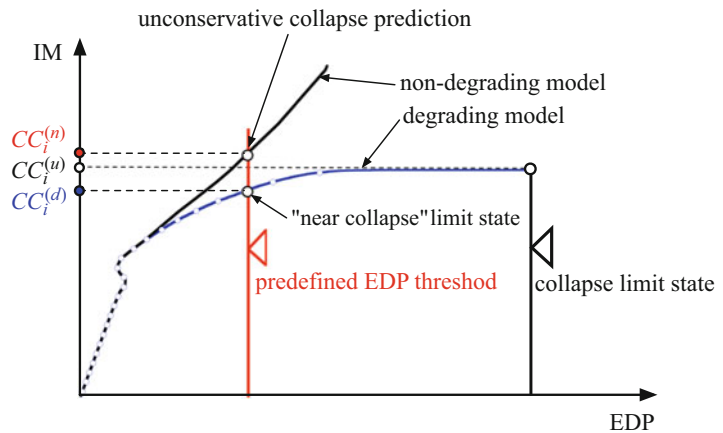
displacement–force relationships for large inelastic excursions.

In recent practical applications, when non-degrading models are used to predict collapse, this limit state is associated with the deformation of a single structural component that exceeds a predefined EDP threshold, often represented by a plastic rotation demand (Haselton et al. 2009). Eurocode 8 (CEN 2005) provisions, for instance, are based on capacity thresholds for the deformation of RC members at yielding and failure developed by Fardis and collaborators (e.g., Panagiotakos and Fardis 2001; Fardis and Biskinis 2003). As a further example, ASCE/SEI 41-13 (2014) procedure compares the component demands of an existing building with component acceptance criteria that define collapse prevention. In such an approach, the component that fails first governs the state of the entire structure. That is to say, component capacity thresholds do not capture the global behavior and the capability of the structure to redistribute forces after failure of one component.

A structure may exhibit several modes of seismic-induced collapse such as sidesway collapse or vertical collapse, depending on the history of spread of inelastic deformations, load distribution, etc. If the corresponding degrading structural model captures only a specific type of collapse such as sidesway collapse, the other collapse mechanisms that might occur are referred to non-simulated modes of collapse. That is, they only can be predicted by an artificial criterion such as the exceedance of a predefined EDP threshold. In particular non-ductile collapse modes such as column shear failure and subsequent vertical collapse are frequently considered as non-simulated collapse modes (e.g., Liel and Deierlein 2013).

The IDA curves of Fig. 4 illustrate sidesway collapse prediction for degrading and non-degrading structural models. The intensity $CC_i^{(u)}$ is the ultimate collapse capacity associated with infinite EDP deformations for the degrading model, and it corresponds to the "exact" collapse capacity. However, if a predefined deformation threshold is used to define collapse, the use of the degrading model always results in a conservative near-collapse prediction. That is, the intensity at this EDP threshold, $CC_i^{(d)}$, is smaller than $CC_i^{(u)}$. On the other hand, the use of a non-degrading model may underestimate or overestimate the actual collapse capacity, depending on the preestablished EDP threshold. In Fig. 4 the corresponding intensity denoted as $CC_i^{(n)}$ results in a nonconservative prediction because the predefined EDP threshold is rather large.

Baradaran Shoraka (2013) developed a robust collapse assessment framework to predict vertical collapse, such as column shear failure and subsequent axial failure in non-ductile reinforced concrete frames. In this approach, progression of damage is tracked throughout the numerical analysis by comparing stepwise floor level gravity demands with corresponding capacities. Explicit modeling of progressive collapse involves

element (commonly column) removal until the structure attains a state that cannot resist the vertical load demand. Baradaran Shoraka et al. (2014) also compared the results of explicit gravity load collapse modeling and those obtained from nonlinear acceptance collapse prevention criteria specified in ASCE/SEI 41-13 (2014) for retrofit of non-ductile frames.

Structural Modeling Strategies for Assessing Sidesway Collapse

Structural Modeling on the System Level

The structural model may be two-dimensional or three-dimensional, but when the building exhibits a coupled bending-torsional response, a three-dimensional system better predicts the building seismic behavior. Such a system requires sophisticated modeling of the nonlinear constitutive behavior at the component level that is not well understood. Thus, three-dimensional models are avoided for collapse analysis as long as it is feasible, and a more generic two-dimensional structural model is utilized.

As outlined by Krawinkler et al. (2009), there are two practical modeling strategies for frame structure evaluation: fiber models and lumped plasticity models. In a fiber model, an axial stress–strain relation is separately assigned to each fiber of the cross section. For example, different constitutive laws can be used in reinforced concrete sections to represent confined and unconfined concrete and the longitudinal steel bars. In this manner, the spread of inelastic deformations within the structural component is captured. In many situations this approach can accurately predict earlier limit states such as yielding and crack initiation, among others. However, several problems have been observed in fiber models. Since these models usually are based on axial stress–strain relations, shear deformations cannot be modeled appropriately. According to Krawinkler et al. (2009), further phenomena that are difficult and often impossible to capture are, for example, bond slip, rebar fraction and rebar buckling, fracture of concrete, and local and lateral torsional buckling in steel. Since

these phenomena may govern the deterioration process, strain "artificial" limits need to be specified in fiber models (Krawinkler et al. 2009).

In a lumped plasticity model, locations within the structure that might undergo inelastic deformations in a severe seismic event are identified up front. Inelastic nonlinear zero-length springs are assigned to these locations, while the rest of the model components remain elastic. These simple phenomenological component models are calibrated with experimental outcomes to represent appropriate cyclic response behavior.

For instance, in a moment-resisting frame, rotational springs are placed at the end of beams and columns. Frames are designed according to the strong column-weak beam philosophy, in which beams and bottom end of base columns are intended to exhibit inelastic deformations at their ends when subjected to strong ground motions. However, many columns may undergo inelastic deformations at the story level when approaching the collapse limit state. Thus, nonlinear spring elements need to be placed at these locations to accurately predict the seismic collapse capacity.

A concentrated plasticity model includes beam/column elements of different degree of sophistication. In the simplest case, all elastic, inelastic, and deteriorating properties of the element are assigned to rotational springs at both ends of the element, whereas the element itself is considered to be rigid; see Fig. 5a. This model may capture the development of plastic hinges. As a disadvantage the spring stiffness must be defined a priori, which is a constant regardless of the potential changes in the moment gradient of the element. To avoid this drawback, Ibarra and Krawinkler (2005) suggest to keep the beam/column element elastic and to add rotational springs at the ends, whose rotational stiffness is several times stiffer than the rotational stiffness of the elastic beam/column element. In this model, an elastic beam/column element is connected to two springs at the ends, and thus, the properties of the parameters can be calibrated according to the component "real" behavior (Fig. 5b). Jäger and Adam (2012) showed that the latter element can describe column and beam behavior sufficiently

| beam/column element model | constitutive behavior of the element between the springs | constitutive spring behavior |
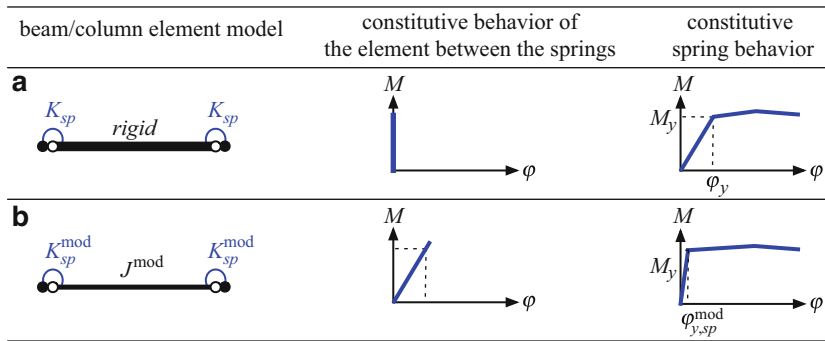
Seismic Collapse Assessment, Fig. 5 Different beam/ column elements and their constitutive behavior to monotonic loading. (**a**) Rigid element with two nonlinear rotational springs. (**b**) Modified elastic element with two nonlinear rotational springs (Modified from Jäger and Adam 2012)

accurate for collapse assessment. The rigid element with two nonlinear springs at its ends should be used for beams only because when it is used in columns, the elastic axial-bending loading interaction cannot be taken into account. A shortcoming of lumped plasticity models is that the nonlinear column spring strength remains constant during the response history analysis, independently of the moment - axial force interaction.

Figure 6a shows a concentrated plasticity model for a multistory three-bay steel moment-resisting frame, whose plan configuration is shown in Fig. 6b. At both ends of each beam and column, a nonlinear rotational spring is arranged with an offset from the node to account for the member's cross section. In contrast, generic frame models used for research purposes are often "centerline models," i.e., the lumped plasticity elements are located at the nodes without any offset from the nodal point.

The structural model of Fig. 6a includes a leaning column with zero flexural stiffness to account for the P-delta effect resulting from gravity loads that are not directly applied to the frame (Haselton et al. 2009). This leaning column is loaded with a vertical load at each story level representing half of the total system gravity load that is not directly tributary to the columns of the frame (NIST GCR 10-917-8 2010), as shown in Fig. 6b.

Another issue that needs to be properly addressed is modeling of viscous damping for

nonlinear dynamic analysis. Several studies have revealed that inappropriate consideration of viscous damping leads to spurious damping forces that distort the nonlinear dynamic response prediction. Particularly affected are lumped plasticity models with elements that exhibit large initial stiffness (Charney 2008). According to the ATC-72-1 report (ATC-72-1 2010), modal damping is the preferred choice for modeling of damping in inelastic structures, but in many software packages, this is not possible. In such cases a combination of mass and stiffness proportional damping (e.g., Rayleigh damping) should be used. Modeling guidelines for damping in inelastic structures are provided in ATC-72-1 (2010).

Structural Modeling of Cyclic Component Behavior

Sidesway collapse assessment based on lumped plasticity structural models requires hysteretic models that include all the important modes of deterioration observed in experimental tests. As illustrated in Fig. 7, monotonic tests on structural components show that strength of the backbone curve is "capped" and is followed by a negative tangent stiffness. In addition, the cyclic response indicates that strength and stiffness deteriorate with the number and amplitude of cycles. As a result, the strength predicted by the monotonic backbone curve is usually not reached during cyclic loading. Several hysteretic models have been developed to represent this behavior (e.g., Song and Pincheira 2000). However, few models
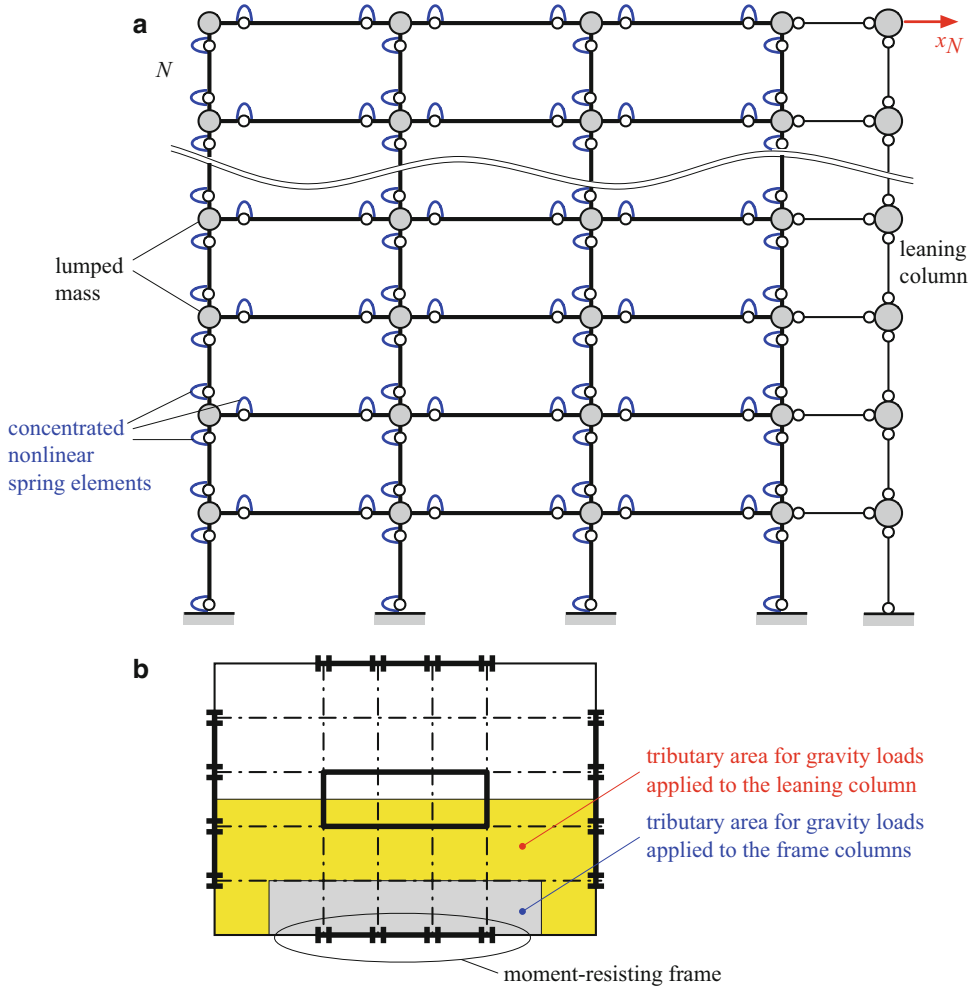
Seismic Collapse Assessment, Fig. 6 Example of structural modeling of a frame building. (**a**) Overall lumped plasticity structural model of the moment-

resisting steel frame. (**b**) Rectangular plane configuration of the frame building (Modified from NIST GCR 10-917-8 2010)

integrate strength deterioration in the backbone curve, as well as cyclic deterioration of strength and stiffness. This section describes exemplarily the deteriorating models developed by Ibarra et al. (2005) for bilinear, peak-oriented, and pinched hysteretic systems.

Figure 8 presents bilinear, peak-oriented, and pinching deteriorating models based on a back-bone curve that includes a post-capping negative slope, and sometimes a residual strength, as shown in Fig. 7. The deteriorating models pre-serve most of the rules of the original non-deteriorating models. For instance, the

bilinear model (Fig. 8a) is based on the standard hysteretic bilinear rules with kinematic strain hardening, but a "strength limit" is introduced when the backbone curve includes a branch with negative slope. The limit for strength is the smallest strength reported on the post-capping branch in previous excursions. If this condition were not established, the strength in the loading path could increase in later stages of deteriora-tion. Note that the original peak-oriented and pinching models exhibit stiffness deterioration in the reloading path. However, in this article the term "deterioration" is used for hysteretic
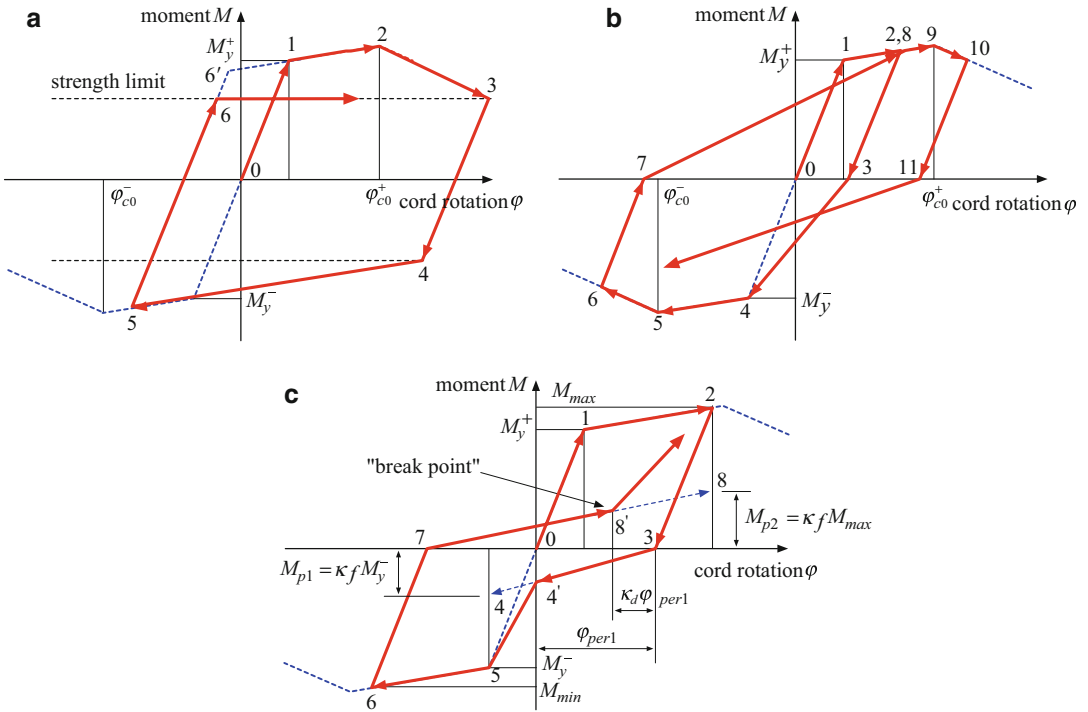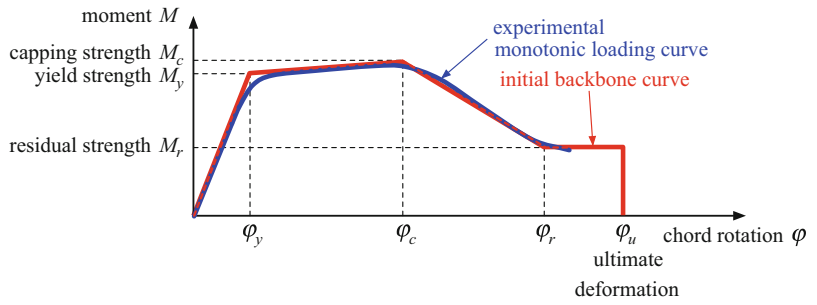
Seismic Collapse Assessment, Fig. 8 Basic deteriorating hysteretic models: (**a**) bilinear, (**b**) peak oriented, and (**c**) pinching (Modified from Ibarra and Krawinkler 2005)

models that possess a post-capping stiffness branch in the backbone curve. These models may also be subjected to cyclic deterioration.

The peak-oriented model (Fig. 8b) keeps the basic hysteretic rules proposed by Clough and Johnston (1966), but the backbone curve is modified to include strength capping and residual strength. The negative post-capping stiffness does not modify any basic rules of the model. The pinching model is similar to the peak-

oriented one, but the reloading consists of two parts. First the reloading path is directed toward a "break point," which is a function of the maximum permanent deformation and the maximum load experienced in the direction of loading. The break point is defined by the parameter κ_f that modifies the maximum "pinched" strength (points 4 and 8 of Fig. 8c) and the parameter κ_d that defines the displacement of the break point (points 4′ and 8′).

Component Deterioration Based on Hysteretic Energy Dissipation

Consideration of cyclic component deterioration is important for collapse prediction, especially for short-period highly ductile systems. The Ibarra–Krawinkler hysteretic model includes four cyclic deterioration modes based on energy dissipation that are activated once the yielding point is surpassed. As observed in Fig. 9, basic strength and post-capping strength deterioration effects translate the strain-hardening and post-capping branch toward the origin, unloading stiffness deterioration decreases the unloading stiffness, and reloading (accelerated) stiffness deterioration increases the target maximum displacement.

The amount of deterioration depends on the parameter $\beta_{m,i}$, which is based on the hysteretic energy dissipated when the component is subjected to cyclic loading (Ibarra and Krawinkler 2005),

$$\beta_{m,i} = \left(\frac{E_i}{E_{t,m} - \sum_{j=1}^{i} E_j} \right)^{c_m} \quad (3)$$

where E_i is the hysteretic energy dissipated in excursion i, $\sum E_j$ represents the hysteretic energy dissipated in all previous positive and negative excursions, and $E_{t,m}$ denotes the hysteretic energy dissipation capacity for the cyclic deterioration mode "m." The original hysteretic rules defined $E_{t,m}$ as a function of twice the elastic strain energy, $E_{t,m} = \gamma_m M_y \varphi_y$ (Ibarra and Krawinkler 2005). The modified Ibarra–Krawinkler model defines $E_{t,m}$ as a function of the plastic chord rotation φ_p, $E_{t,m} = \lambda_m M_y \varphi_p$, or $E_{t,m} = \Lambda_m \varphi_p$ (Lignos and Krawinkler 2012). The parameters γ_m, λ_m, and Λ_m are calibrated from experimental tests, and reasonable results are obtained if all cyclic deterioration parameters have the same value. Exponent c_m is the rate of deterioration parameter. $c_m = 1$ implies a constant rate of deterioration.

The parameter $\beta_{m,i}$ modifies the strength and stiffness properties of the previous cycle. For basic strength deterioration, for instance, the strain-hardening branch moves toward the origin by an amount equivalent to reducing the yield strength as follows (for this mode index $m = s$):

$$M_{y,i}^+ = \left(1 - \beta_{s,i}\right) M_{y,i-1}^+ \quad (4)$$

where $M_{y,i}^+$ is the deteriorated yield strength after excursion i and $M_{y,i-1}^+$ is the deteriorated yield strength before excursion i. Except for the unloading stiffness mode, a positive and a negative value is defined for each deterioration parameter because the algorithm deteriorates the strength independently in both directions.

Seismic Collapse Assessment, Fig. 9 Four modes of component deterioration in a pinching component (Modified from Lignos and Krawinkler 2013; With permission from American Society of Civil Engineers)
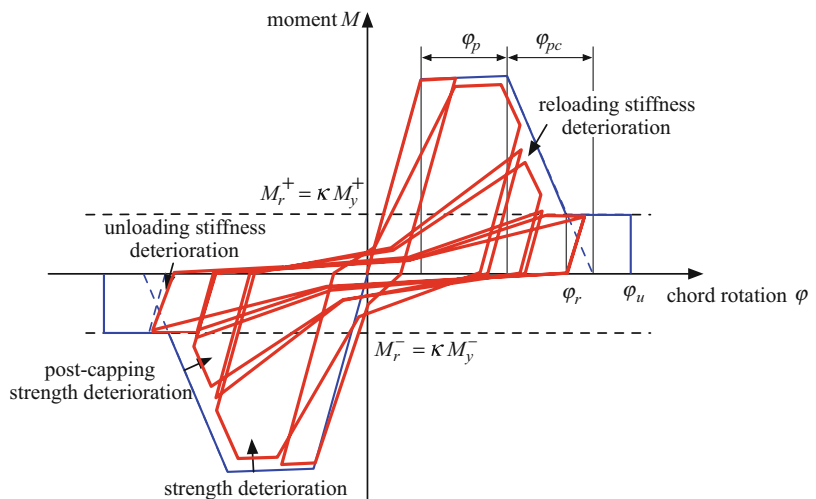
Seismic Collapse Assessment, Fig. 10 Inelastic structural response history with ratcheting effect due to P-delta (With permission from Adam 2014)
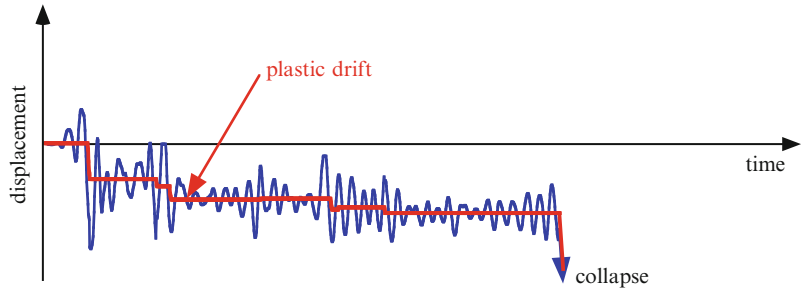
That is to say, $M_{y,i}^-$ is updated after every positive inelastic excursion, and $M_{y,i}^+$ is updated after every negative inelastic excursion. For example, $\beta_{s,i}$ is calculated each time the inelastic path crosses the horizontal axis. These parameters are calibrated from experiments. For instance, Lignos and Krawinkler (2012, 2013) conducted deterioration modeling of steel components based on an experimental database.

Global Seismic Sidesway Collapse and the P-Delta Effect

In a vibrating structure, gravity loads acting through large horizontal deflections induce additional bending moments, which in turn increase the lateral deflections. In structural analysis, this destabilizing effect of gravity loads (i.e., global P-delta effect) is usually interpreted as a reduction of the lateral structural stiffness by the so-called geometric stiffness. Generally, for a real building within its elastic range of deformation, this stiffness decrease is of minor importance because its magnitude is small compared to the first-order elastic stiffness. During severe seismic excitations, however, inelastic deformations combined with gravity may cause a structure to approach a state of dynamic instability if the post-yield tangent stiffness becomes negative. In such a condition, the displacement response tends to amplify in a single direction due to the ratcheting effect, as shown in Fig. 10. Very flexible structures are especially affected by seismic-induced sidesway collapse in the P-delta mode, and in many of these cases, cyclic component material deterioration can be disregarded.

However, this consideration not only depends on the structure characteristics but also on the ground motion applied to the structure. For example, if a strong-motion earthquake causes structural collapse after a few load cycles, component degradation is of minor significance compared to a situation where a long-duration earthquake exhibits many cycles before the structure fails.

In an inelastic SDOF system, the gravity load generates a uniform shear deformation of its hysteretic force–displacement relationship, as shown exemplarily in Fig. 11, for the backbone curve of an inelastic SDOF system. Characteristic displacements (such as the yield displacement) of this relationship remain unchanged, whereas the characteristic forces (such as the strength) are reduced. As a result, the slope of the elastic and post-elastic branch rotates. The magnitude of this reduction can be expressed by means of the stability coefficient θ (MacRae 1994). The parameter θ is a function of the gravity load, geometry, and stiffness. For instance, for the bilinear SDOF system shown in Fig. 12a, the stability coefficient reads as

$$\theta = \frac{Ph}{k_r} \qquad (5)$$

where P is the gravity load, h the length of the rod, and k_r the initial stiffness of the elastoplastic rotational spring at the base. In the example of Fig. 12, the post-yield stiffness is negative, because the stability coefficient θ is larger than the hardening ratio α. In a nonmaterial degrading SDOF system, a negative post-tangential stiffness ratio, i.e., $\theta - \alpha > 0$, is a necessary condition for potential structural collapse under severe earthquake excitations.

P-delta effect on
the backbone curve of an
inelastic SDOF system
(Modified from Ibarra and
Krawinkler 2005)
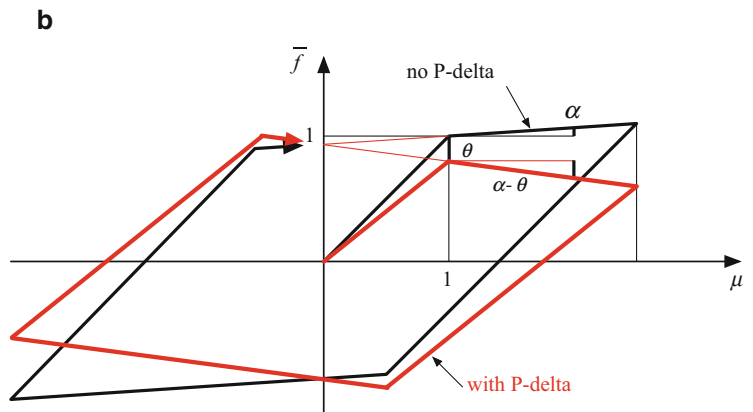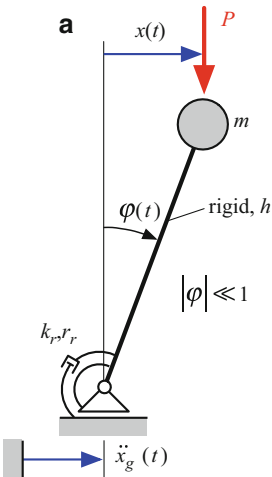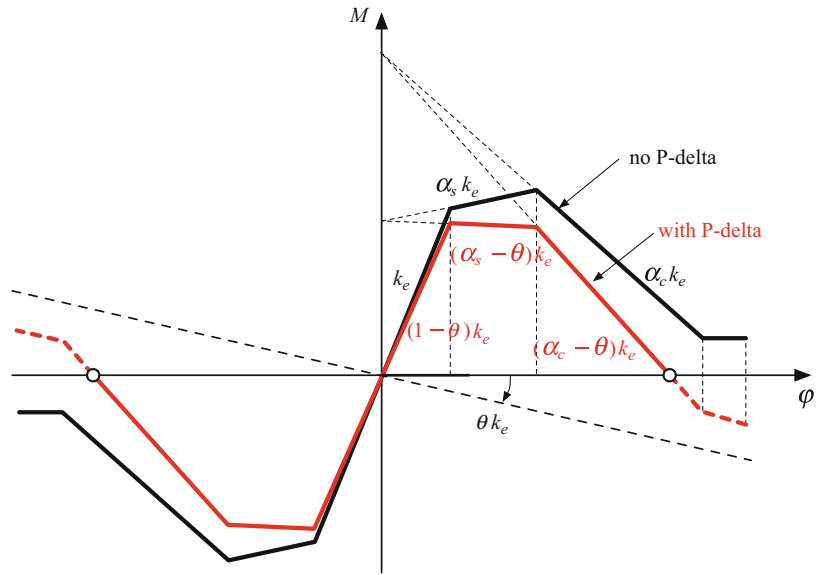
Seismic Collapse Assessment, Fig. 12 (**a**) Model of an SDOF system vulnerable to P-delta. (**b**) Normalized bilinear cyclic structural behavior of an SDOF system with and without P-delta (With permission from Adam and Jäger 2012a)

Adam and Jäger (2012a) showed that the collapse capacity of a P-delta vulnerable inelastic SDOF system with non-deteriorating hysteretic behavior and bilinear backbone curve is a function of the negative slope of the post-tangential stiffness ratio $\theta - \alpha$, the elastic structural period of vibration T, the viscous damping coefficient ζ, and the shape of the hysteretic loop such as bilinear, peak oriented, or pinching. Based on this observation, Adam and Jäger (2012a) and Jäger

and Adam (2013) conducted a rigorous parametric IDA study to reveal the impact of characteristic parameters (i.e., period T, negative post-yield stiffness ratio $\theta - \alpha$, damping ζ, and hysteretic loop) on the collapse of those systems, presenting the outcomes in the form of "collapse capacity spectra."

In a collapse capacity spectrum, the P-delta vulnerable SDOF system collapse capacity is represented as a function of the initial period of
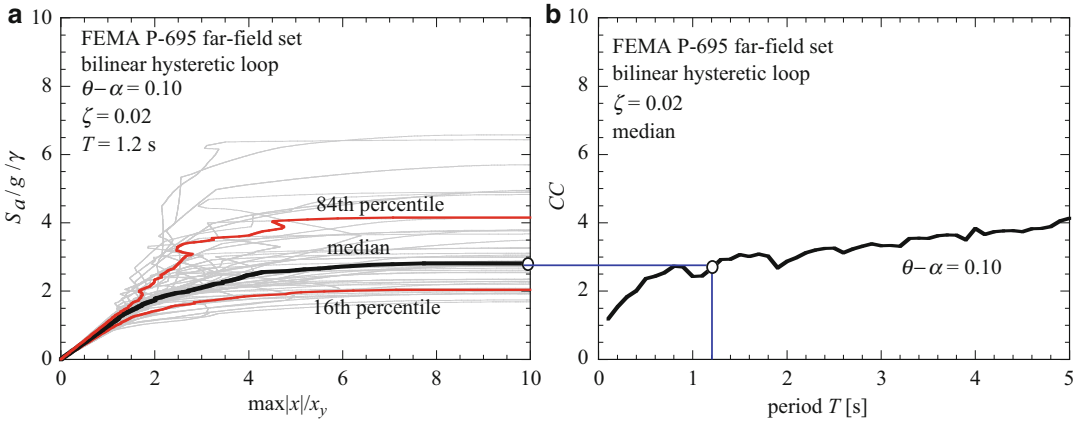
Seismic Collapse Assessment, Fig. 13 (**a**) IDA curves of an SDOF system with parameters defined in the figure for all records of the FEMA P-695 far-field ground motion set and statistically evaluated results. (**b**) Corresponding median "instability collapse" capacity spectrum (With permission from Adam 2014)

Seismic Collapse Assessment, Fig. 14 Median collapse capacity spectra (*blue lines*) for a set of negative post-yield stiffness ratios and corresponding smooth collapse capacity spectra (With permission from Adam 2014)
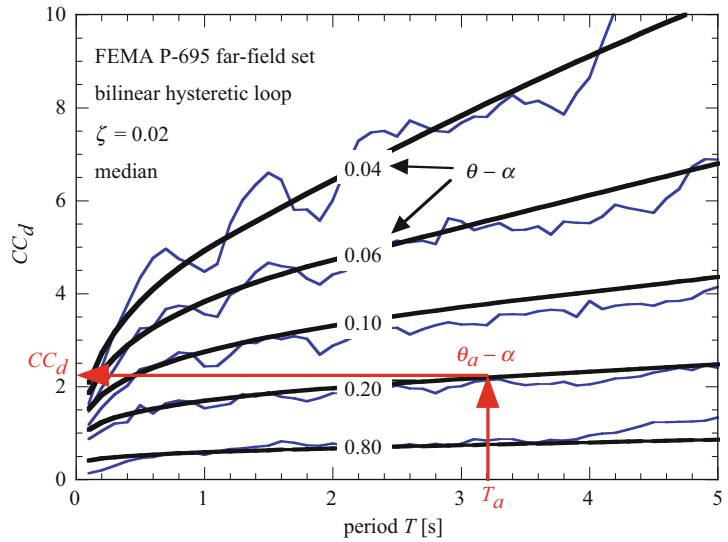
vibration T for a fixed set of characteristic parameters. For example, Fig. 13a shows the individual IDA curves of an SDOF system for the 44 ground motions of FEMA P-695 far-field set (in Adam and Jäger (2012a) referred to as ATC63-FF set) up to "instability collapse," as well as median, 16th, and 84th percentile curves. Collapse capacity spectra are obtained by repeating this procedure for different vibrational periods within a certain interval. As before, collapse capacity spectra can be computed for a specific suite of earthquake records, leading to median, 16th, and

84th percentile collapse capacity spectra (Fig. 13b). Based on a collapse capacity spectrum database, Adam and Jäger (2012a) and Jäger and Adam (2013) formulated analytical expressions of median, 16th, and 84th percentile collapse capacity spectra with a "smooth" shape via nonlinear regression analyses, as shown in Fig. 14.

In multistory frames, gravity loads may substantially impair the complete structure or only a subset of stories (Gupta and Krawinkler 2000). The local P-delta effect may induce collapse
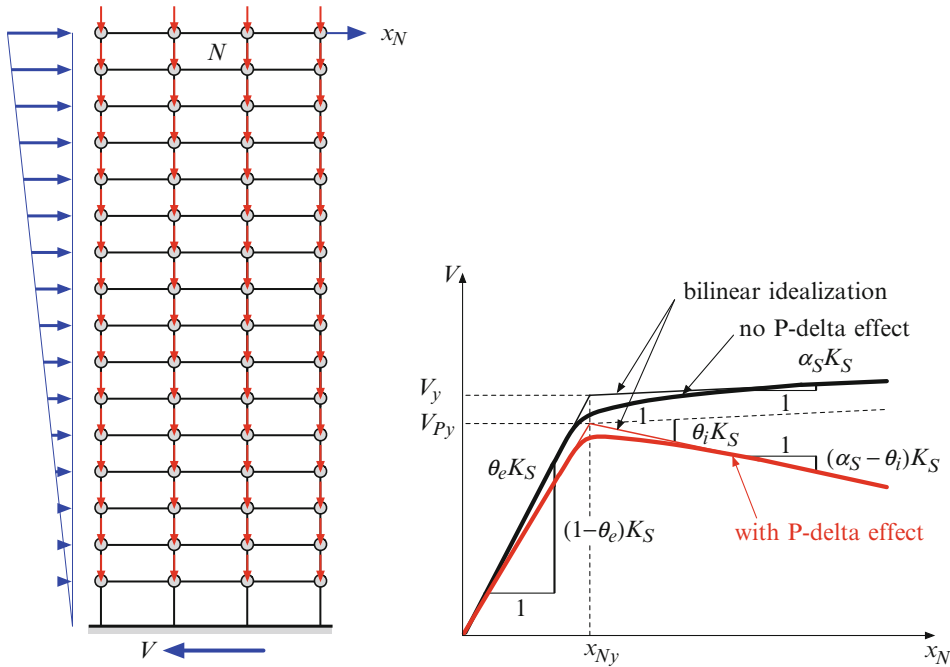
Seismic Collapse Assessment, Fig. 15 Global pushover curves disregarding P-delta (*black graphs*) and considering P-delta (*red graphs*) (With permission from Adam 2014)

of a structural element, without necessarily affecting the stability of the complete structure. An indicator of the severity of the local P-delta effect is the story stability coefficient proposed in the Eurocode 8 (CEN 2005). However, a consistent relationship between the local P-delta effect and the global P-delta effect, which characterizes the overall impact of gravity loads on the structure, cannot be established due to dynamic interaction between adjacent stories in a multistory frame structure (Gupta and Krawinkler 2000).

Strong evidence for the vulnerability of a regular frame structure to P-delta-induced global seismic collapse is obtained from a first-mode pushover analysis (Adam and Jäger 2012b). For instance, the red curve of Fig. 15 illustrates the global pushover curve of a very flexible structure that exhibits a negative post-yield stiffness due to effect of gravity loads. When severe seismic excitation drives the structure into its inelastic branch, a state of dynamic instability may be approached, and the global collapse capacity is attained at a rapid rate. The pushover curve plotted in black

disregards gravity loads in the nonlinear static analysis.

Global pushover curves of multi-degree-of-freedom (MDOF) structures do not exhibit a uniform stability coefficient in the entire range of elastic and inelastic deformations, unlike SDOF systems. As observed in Fig. 15, two stability coefficients can be identified in a bilinear approximation of pushover curves with and without P-delta effect, that is, a stability coefficient in the elastic range of deformation (θ_e) and a stability coefficient in the post-yield range of deformation (θ_i). According to Medina and Krawinkler (2003), θ_i can be much larger than θ_e: $\theta_i > (>)\theta_e$.

Variance of Collapse

Earlier studies on quantification of epistemic uncertainties on the seismic response concluded that the contribution of epistemic uncertainties to the variance of the system's seismic performance is in general much smaller than that from aleatory uncertainties, and thus it can be omitted

(e.g., Esteva and Ruiz 1989). In contrast, recent studies indicate that the epistemic uncertainty caused by degrading nonlinear system parameters is very relevant when assessing the total variance of the collapse limit state (Ibarra and Krawinkler 2005; Zareian et al. 2010). For example, Haselton et al. (2011) evaluated the total variance of collapse capacity of a four-story reinforced concrete special moment frame, including RTR variability and modeling uncertainties. He estimated the logarithmic standard deviation associated with modeling uncertainties as 0.45, a value comparable to the standard deviation of the logarithm of RTR variability, which varies in general between 0.25 and 0.45 (Liel et al. 2009; Adam and Jäger 2012a). As an example, Fig. 16 shows the collapse fragility of a P-delta vulnerable SDOF structure as shown in Fig. 12a, considering the RTR uncertainty only, and the total variability considering RTR and epistemic uncertainties.

The effect of RTR variability on the variance of collapse is directly obtained by performing IDAs for a set of ground motions on a deterministic system. The variance of collapse capacity of nonmaterial deteriorating systems is also affected by uncertainty in parameters, such as the yield moment and post-yield hardening ratio, but their effect on the variance of collapse capacity is small compared to that originated by RTR variability.

For deteriorating models, however, epistemic uncertainty can be relevant for collapse capacity evaluation because of the large uncertainty on degrading nonlinear parameters, such as plastic rotation and post-capping stiffness (Fig. 17) (Ibarra and Krawinkler 2011; Ugurhan et al. 2014).

The first-order second-moment (FOSM) method (e.g., Melchers 1999) can be used to compute the additional variance of collapse capacity due to uncertainty in the system parameters. The total variance of the collapse capacity, $\sigma^2_{\ln CC(\text{TOT})}$, based on FOSM is (Ibarra and Krawinkler 2011):

$$\sigma^2_{\ln CC(\text{TOT})} \cong \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\partial g(\overline{X})}{\partial x_i} \frac{\partial g(\overline{X})}{\partial x_j} \right)_{\overline{X} = \overline{\mu}_x}$$
$$\rho_{x_i, x_j} \sigma_{x_i} \sigma_{x_j} + \sigma^2_{\ln CC(\text{RTR})}$$

(6)

The first term on the right hand side of Eq. 6 represents the contribution to the variance of collapse capacity due to uncertainty of the system parameters, whereas $\sigma^2_{\ln CC(\text{RTR})}$ is the contribution to the variance of collapse capacity due to RTR variability, and $\overline{X}$ represents the set of system random parameters. The function $g(\overline{X})$ is the collapse capacity as a function of the variation on system parameters, and the variance is computed from the gradient of $g(\overline{X})$, which is

Fig. 17 Uncertain parameters of the backbone curve (Modified from Ibarra and Krawinkler 2005)
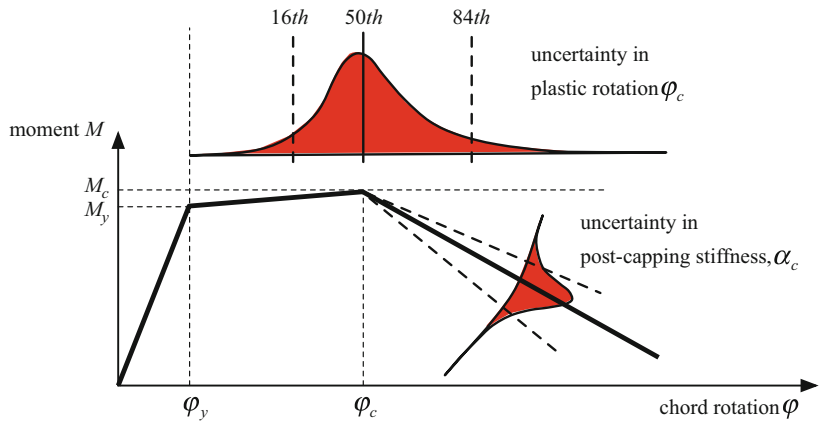
linearized using a Taylor series expansion about the mean $\overline{X}$. The derivative $\partial g(\overline{X})/\partial x_i$ can be approximated by evaluating the performance function at two system parameter values. The derivative is estimated as the slope of the straight line that connects the pairs $[x_i, g(x_i)]$. The standard deviation of the system parameters, σ_{x_i}, and the correlation coefficients, ρ_{x_i, x_j}, need to be estimated in advance.

Ibarra and Krawinkler (2005, 2011) used the FOSM method to evaluate the effect of uncertainty in system parameters on the collapse capacity of SDOF systems. They found that uncertainty in the displacement at the peak (cap) strength and the post-capping stiffness contributes significantly to the variance of collapse capacity. The uncertainty in deterioration parameters on the variance of collapse capacity may be comparable to that caused by RTR variability.

Although the FOSM method is relatively simple to implement, it cannot predict the shift in the median caused by system uncertainties. Vamvatsikos and Fragiadakis (2010) concluded that this shortcoming should not be considered important for practical applications because for more cases the shift in the median is less than 10 %. Liel et al. (2009) found shifts in the median closer to 20 % and proposed a simplified method, termed ASOSM (approximate second order second moment), that uses FOSM to predict the increase in fragility's logarithmic standard deviation and the shift in the median of the limit state fragility. The total variance can also be approximated using the point estimate method, in which

the first moments of a function are calculated in terms of the first moments of the random variables (Rosenblueth 1981). The total variance in collapse capacity can be more accurately predicted using Monte Carlo simulation, but this method is computationally expensive. An alternative is to use the Latin hypercube sampling (LHS) method to reduce Monte Carlo realizations. LHS is a stratified sampling that allows efficient estimation of the quantity of interest by reducing the variance of Monte Carlo simulations (Vamvatsikos and Fragiadakis 2010). In the example of Fig. 16, the FOSM method and the LHS method deliver almost the same prediction of the collapse fragility function.

Simplified Methods

IDA and MSA procedures require the numerical solution of the governing equations of motions of the structural model together with constitutive equations in each time step for each response history analysis. As a consequence, it is time-consuming and computationally expensive. Thus, there is a need for simple to apply, but yet sufficient accurate methods to predict the global collapse of MDOF structures under seismic excitations (Lignos and Krawinkler 2012).

Methods Based on Nonlinear Static Analysis
One of the simplest approaches is to conduct a nonlinear static analysis to estimate the seismic collapse capacity. During this so-called pushover
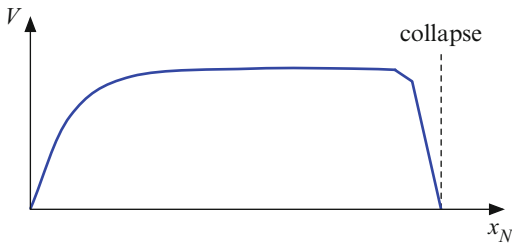
Seismic Collapse Assessment, Fig. 18 Global push-over curve up to global collapse (With permission from Adam 2014)

analysis, gravity loads are applied to the structural model that considers explicitly the inelastic constitutive behavior of its elements. Subsequently the model is subjected to lateral forces with a predefined invariant load pattern that is amplified incrementally in a displacement-controlled procedure. A first-mode load pattern may be used. Alternatively, the load pattern recommended in ASCE 7-10 (2010) can be utilized. As a result the global pushover curve of the structure is obtained, at which the base shear is plotted against the displacement of the control node (usually the roof displacement). The global pushover curve represents the global capacity of the building against horizontal loads, and it is assumed that it reflects the global or the local mechanism involved when the structure approaches dynamic instability. In the perspective of a "near-collapse" limit state, collapse is assumed if one of the selected EDPs, such as story drifts, plastic hinge rotations, etc., exceeds a certain threshold. When the base shear drops to zero (Fig. 18), global "instability collapse" is assumed to be attained.

Pushover analysis was originally developed for first-mode dominated structures and later refined to account also for higher mode effects (e.g., Chopra and Goel 2002). The basic assumption of these procedures is that the nonlinear static response can be related to the nonlinear dynamic response. This is, however, not the case for many structures because the dynamic response is strongly path dependent, and effects, such as cyclic deterioration, damping, and duration of an earthquake, among others, cannot be captured. Nonetheless, nonlinear static analysis

may be reasonable applied as an ingredient for the collapse assessment of regularly shaped low-rise buildings whose seismic response is dominated by the first mode. Even if the pushover procedure is not appropriate for a full collapse assessment, it can be used to identify weaknesses in the structural design, to debug a structural model used for collapse assessment, and to obtain a better understanding of the strength and deformation demands of the structure (Deierlein et al. 2010). In contrast, Villaverde (2007) states that it is doubtful that nonlinear static methods can be used reliably to predict the seismic collapse capacity of structures and to estimate their margin of safety against a global collapse. These methods, however, are commonly utilized in practical procedures. For example, Shafei et al. (2011) proposed a simplified methodology for collapse assessment of degrading moment-resisting frames and shear walls utilizing closed-form equations, given that the global pushover curve is provided. These equations correlate median and dispersion of collapse fragility curves and were generated through multivariate regression analysis from a comprehensive database of collapse fragilities and pushover curves.

Methods Based on Equivalent Single-Degree-of-Freedom Systems

Several studies use equivalent single-degree-of-freedom (ESDOF) systems to predict the global seismic collapse capacity (e.g., Bernal 1998; Fajfar 1999; Adam et al. 2004). Application of ESDOF systems implies that the first mode dominates the dynamic response and the collapse mode. A global collapse capacity assessment of low- to medium-rise buildings by means of an ESDOF system is straightforward when the story drifts remain rather uniformly distributed over the height, regardless of the extent of inelastic deformation. Adam and Jäger (2012b) showed that the application of these simplified systems yields reasonable collapse capacity predictions for P-delta-sensitive regular high-rise buildings, because P-delta-induced collapse is primarily controlled by the first mode. However, if a partial mechanism develops, the global collapse capacity is greatly affected by the change of the

deflected shape, and it will be amplified in those stories in which the drift becomes large (Bernal 1998).

Capacity Spectrum Method (N2 Method)

In the last decade, the capacity spectrum methodology became a popular tool for assessing the seismic performance of regular first-mode dominated structures. It represents a compromise between nonlinear dynamic analyses with complex modeling strategies and simplified linear static analysis methods. In Eurocode 8 (CEN 2005), a version of this method, i.e., the N2 method (Fajfar 1999), is implemented.

In this approach a global lateral load–displacement relation, which is the outcome of a pushover analysis, represents the seismic capacity of the structure and is referred to as capacity curve. The base shear V and the roof displacement x_N of a bilinear idealization of this curve are transformed into the domain of the corresponding ESDOF system according to the following Eurocode 8 equations (CEN 2005) (see Fig. 19):

$$f^* = V/\Gamma, \qquad D = x_N/\Gamma,$$
$$\Gamma = L^*/m^*, \quad L^* = \sum_{i=1}^{N} m_i \phi_i, \quad m^* = \sum_{i=1}^{N} m_i \phi_i^2 \tag{7}$$

Here, f^* is the equivalent spring force, and D the displacement of the ESDOF system. The equation includes the story masses m_i, $i = 1,\ldots,$ N, of the ductile N-story structure, and the N components ϕ_i of the shape vector ϕ (with $\phi_N = 1$), which prescribes the vertical distribution of the displacements of the structure in its ESDOF approximation (Fajfar 2002). Note that ϕ should be affine to the horizontal load pattern of the pushover analysis.

Dividing f^* by the equivalent mass L^* renders its ordinate in the "acceleration" dimension; see Fig. 19. This transformation permits the comparison of the capacity curve with the seismic demand represented by the response spectrum in the ADSR format of the actual site. In this format for an inelastic SDOF system with a target ductility μ, the spectral acceleration S_{av} at yield is plotted against the corresponding peak spectral displacement S_d (Fajfar 1999), as shown in Fig. 20a. Note that the ductility defines the ratio of the maximum imposed (inelastic) deformation to the deformation at onset of yield.

In the subsequent step, the intersection point between a bilinear approximation of the capacity curve and the response spectrum is obtained. If this point is in the inelastic branch of deformation, the elastic spectrum must be reduced such that the ductility of the capacity curve at this point and of the corresponding inelastic spectrum coincide, Fig. 20b. For periods larger than the corner period T_c, this intersection point, which is referred to as performance point, is found by application of the equal displacement rule.

According to Eurocode 8 (CEN 2005), structural stability under seismic excitation cannot be verified if the displacement demand at the performance point is larger than two-thirds of the ultimate deformation capacity of the structure or if no performance point can be found. Consequently, collapse is indicated, as presented in Fig. 20b.

Seismic Collapse Assessment, Fig. 19 Capacity curve With permission from Adam (2014; Modified from CEN 2005)

a S_{ay} response spectrum in ADRS format

b S_{ay}
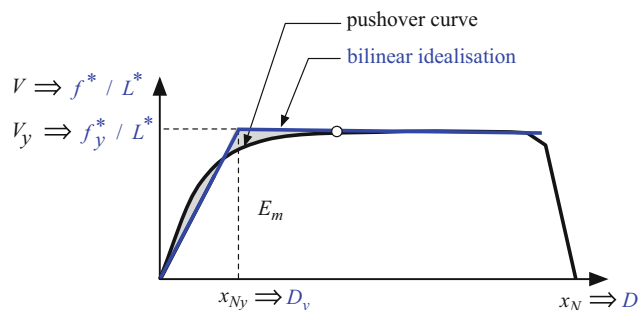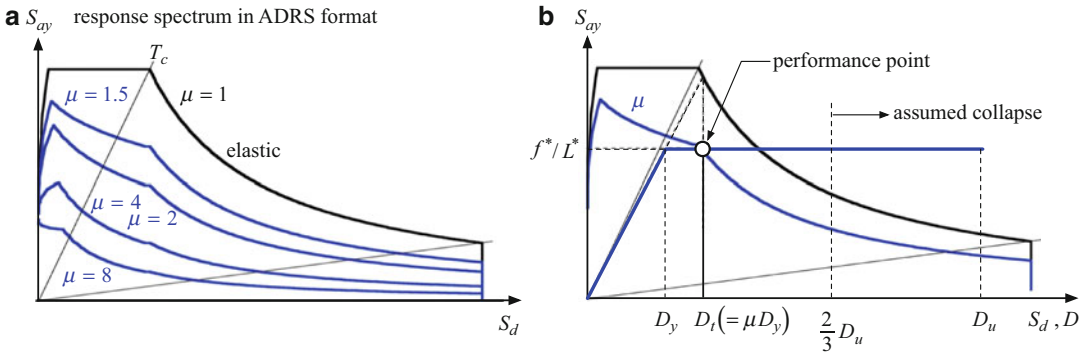
Seismic Collapse Assessment, Fig. 20 (**a**) Elastic and corresponding constant ductility response spectra in the ADRS format. (**b**) Performance point With permission from Adam (2014; Modified from Fajfar 1999)

Collapse Capacity Spectrum Method

The collapse capacity spectrum method is another example of a simplified ESDOF system-based method. This method predicts the global collapse capacity of regular P-delta-sensitive moment-resisting frame structures, where cyclic component deterioration can be omitted (Adam and Jäger 2012b). In a first step, the vulnerability of the structure to P-delta-induced collapse is evaluated based on global first-mode pushover curve of the structure with applied gravity loads. When this curve exhibits a negative post-yield stiffness, as shown in Fig. 15, the building may collapse under severe earthquake excitation in the P-delta mode. In this situation the structure is transformed into an ESDOF system.

Collapse capacity spectra (Adam and Jäger 2012a) are the further ingredient of this method. However, application of collapse capacity spectra is not straightforward for an MDOF structure, because the backbone curve of the ESDOF system is derived from the global first-mode pushover curves with and without considering P-delta. As discussed before, bilinear idealization of these curves does not exhibit a uniform stability coefficient as in a real SDOF system, but an elastic θ_e and an inelastic stability coefficient θ_i that is always larger, in some cases even much larger than θ_e: $\theta_i > (>)\theta_e$. Assigning a specific auxiliary backbone curve (Ibarra and Krawinkler 2005) to the ESDOF system solves this problem because its rotation by the uniform auxiliary stability coefficient θ_a,

$$\theta_a = \frac{\theta_i - \theta_e \alpha_s}{1 - \theta_e + \theta_i - \alpha_s} \quad (8)$$

yields the backbone curve considering P-delta in analogy to a real SDOF system; see Fig. 21. The parameter α_S is the strain-hardening coefficient of the global pushover curve without P-delta. The discussed backbone curves of the ESDOF system are depicted in Fig. 21: The bilinear curve with the largest strength represents the auxiliary backbone curve, and the graph with the smallest strength is the backbone curve with P-delta. Subsequently the negative post-yield stiffness ratio $\theta_a - \alpha_S$ of the auxiliary ESDOF system is evaluated. Since the initial stiffness of the auxiliary backbone curve is larger than the original one, the period of the ESDOF system with assigned auxiliary backbone curve is (Adam et al. 2004)

$$T_a = 2\pi \sqrt{\frac{1 - \alpha_S}{1 - \theta_e + \theta_i - \alpha_S}} \sqrt{\frac{x_{N_y}}{V_y}} \sqrt{\sum_{i=1}^{N} \phi_i m_i} \quad (9)$$

In Eq. 9 x_{N_y} and V_y denote the roof displacement and the base shear, respectively, at the onset of yield of the global pushover curve without P-delta; see Fig. 15.

Now the appropriate collapse capacity spectrum with respect to the underlying ground motion set, viscous damping ζ, the hysteretic loop, and the negative post-yield stiffness $\theta_a - \alpha_S$ is consulted to obtain the median (CC_d), 16th

Seismic Collapse Assessment, Fig. 21 Backbone curves with and without P-delta and auxiliary backbone curve (Modified from Adam and Jäger 2012b)
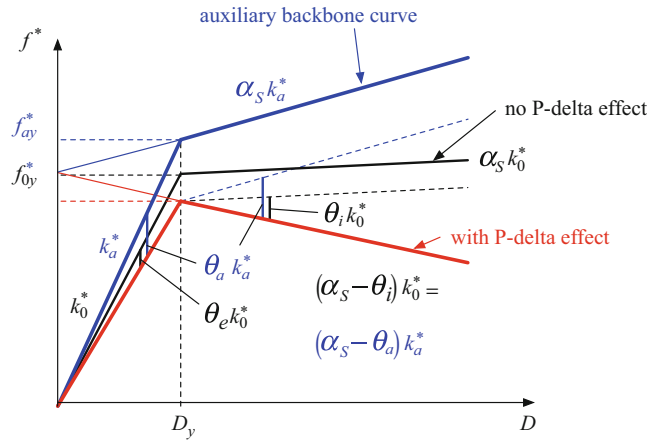
(CC_d^{p16}), and 84th percentile (CC_d^{p84}) collapse capacities of a SDOF system at the period T_a of the auxiliary ESDOF system. For instance, in Fig. 14 CC_d is read at a period $T_a = 3.2$ s for $\theta_a - \alpha_S = 0.20$. These collapse capacity quantities are transformed into the domain of the ESDOF system:

$$CC_{\text{ESDOF}} = \frac{CC_d}{\lambda_{\text{MDOF}}},$$

$$\lambda_{\text{MDOF}} = \left(\sum_{i=1}^{N} \phi_i m_i\right)^2 \Big/ \left(\sum_{i=1}^{N} m_i \sum_{i=1}^{N} \phi_i^2 m_i\right) \tag{10}$$

λ_{MDOF} is the transformation coefficient that relates the actual structure with the corresponding SDOF system. For details it is referred to Fajfar (2002). CC_{ESDOF} is an estimate of the median collapse capacity of the MDOF system, $CC_{\text{MDOF}} \approx CC_{\text{ESDOF}}$. The same applies to the percentiles: $CC_{\text{MDOF}}^{p16} \approx CC_{\text{ESDOF}}^{p16}$ and $CC_{\text{MDOF}}^{p84} \approx CC_{\text{ESDOF}}^{p84}$. Since the collapse capacities follow generally a log-normal distribution, an approximation of the collapse fragility function is derived from

$$\ln N(m, \sigma^2), \quad m = \ln(CC_{\text{ESDOF}}),$$
$$\sigma = \ln\sqrt{CC_{\text{ESDOF}}^{p16} CC_{\text{ESDOF}}^{p84}} \tag{11}$$

Application of this method for assessing the seismic collapse fragility of a series of generic and real moment-resisting frame structures showed that the collapse capacity can be estimated quick, but yet accurate, without conducting time-consuming dynamic analyses (Adam and Jäger 2012b). The collapse capacity spectrum method is particularly useful in engineering practice because the structure can be evaluated with respect to its seismic collapse capacity in the initial design process without a detailed dynamic analysis.

Summary

This article summarizes procedures currently used for assessing global collapse of structures induced by strong-motion earthquakes. Through intensive research in the last two decades, substantial progress was achieved in this field, and the outcomes of these studies are compiled in numerous publications. Thus, in the present contribution, only some of the issues involved in seismic collapse assessment are elaborated in detail. The selection of the presented material is naturally biased by research and experience of the authors.

Seismic collapse assessment is a branch of earthquake engineering that combines multidisciplinary fields of research such as seismology, structural dynamics, materials science, applied mathematics, and computational mechanics. It includes adequate prediction of the seismic hazard, ground motion selection, identification of

possible modes of collapse, modeling of cyclic component deterioration, appropriate consideration of hysteretic and viscous damping, quantification of modeling and parameter uncertainties, and nonlinear dynamic analyses based on stable numerical algorithms.

For accurate assessment of seismic structural collapse, however, there are still challenges that need to be addressed. For example, understanding of nonlinear cyclic component behavior in a condition of oblique bending requires experimental tests and advances in constitutive modeling. Another issue is the quantification of the contribution of nonstructural elements on the seismic collapse capacity. On the other hand, from engineering practice, there is a strong demand on simplified assessment strategies that are simple to apply and at the same time reasonably accurate.

Cross-References

▶ Analytic Fragility and Limit States [P(EDP|IM)]: Nonlinear Dynamic Procedures
▶ Analytic Fragility and Limit States [P(EDP|IM)]: Nonlinear Static Procedures
▶ Conditional Spectra
▶ Performance-Based Design Procedure for Structures with Magneto-Rheological Dampers
▶ Seismic Risk Assessment, Cascading Effects
▶ Site Response for Seismic Hazard Assessment

References

Adam C (2014) Structural dynamics and earthquake engineering (in German). Internal report. University of Innsbruck

Adam C, Jäger C (2012a) Seismic collapse capacity of basic inelastic structures vulnerable to the P-delta effect. Earthq Eng Struct Dyn 41:775–793

Adam C, Jäger C (2012b) Simplified collapse capacity assessment of earthquake excited regular frame structures vulnerable to P-delta. Eng Struct 44:159–173

Adam C, Ibarra LF, Krawinkler H (2004) Evaluation of P-delta effects in non-deteriorating MDOF structures

from equivalent SDOF systems. In: Proceedings of the 13th world conference on earthquake engineering (13 WCEE), Vancouver, 1–6 Aug 2004, digital paper, 15 pp

ASCE/SEI 7-10 (2010) Minimum design loads for buildings and other structures. American Society of Civil Engineers, Reston

ASCE/SEI 41-13 (2014) Seismic evaluation and retrofit of existing buildings. American Society of Civil Engineers, Reston

ATC-58 (2012) Seismic performance assessments of buildings, 100% draft. Applied Technology Council, Redwood City

ATC-72-1 (2010) Modeling and acceptance criteria for seismic design and analysis of tall buildings. Applied Technology Council, Redwood City

ATC-78 (2013) Identification and mitigation of seismically hazardous older concrete buildings: interim methodology evaluation. Applied Technology Council, Redwood City

Baker JW (2013) Trade-offs in ground motion selection techniques for collapse assessment of structures. In: Adam C, Heuer R, Lenhardt W, Schranz C (eds) Proceedings of the Vienna Congress on recent advances in earthquake engineering and structural dynamics 2013 (VEESD 2013), Vienna, 28–30 Aug 2013, digital paper, paper no. 123, 10 pp

Baker JW, Cornell CA (2005) A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon. Earthq Eng Struct Dyn 34:1193–1217

Baker JW, Cornell CA (2006) Spectral shape, epsilon and record selection. Earthq Eng Struct Dyn 35:1077–1095

Baradaran Shoraka MB (2013) Collapse assessment of concrete buildings: an application to non-ductile reinforced concrete moment frames. PhD thesis, The University of British Columbia, Vancouver

Baradaran Shoraka MB, Elwood KJ, Yang TY, Liel AB (2014) Collapse assessment of non-ductile, retrofitted and ductile reinforced concrete frames. In: Elwood KJ, Kim InSung, Dragovich J (eds) Seismic assessment of existing reinforced concrete buildings: new developments, ACI special publication ACI SP 297, SP297-8, 19 pp., American Concrete Institute, Farmington Hills, MI

Bernal D (1987) Amplification factors for inelastic dynamic P-Δ effects in earthquake analysis. Earthq Eng Struct Dyn 15:635–651

Bernal D (1998) Instability of buildings during seismic response. Eng Struct 20:496–502

CEN (2005) Eurocode 8: design of structures for earthquake resistance. Comité Européen de Normalisation (CEN), Brussels

Charney FA (2008) Unintended consequences of modeling damping in structures. J Struct Eng 134:581–592

Chopra AK, Goel RK (2002) A modal pushover analysis procedure for estimating seismic demands for buildings. Earthq Eng Struct Dyn 31:561–582

Clough RW, Johnston SB (1966) Effects of stiffness degradation on earthquake ductility requirements. In: Proceedings of the 2nd Japan earthquake engineering symposium, Tokyo, October 1966, pp 227–232

Deierlein GG, Reinhorn AM, Willford MR (2010) Nonlinear structural analysis for seismic design. A guide for practicing engineers. NEHRP Seismic Design technical brief no. 4, produced by the NEHRP Consultants Joint Venture, a partnership of the Applied Technology Council and the Consortium of Universities for Research in Earthquake Engineering, for the National Institute of Standards and Technology, Gaithersburg, NIST GCR 10-917-5

Esteva L, Ruiz S (1989) Seismic failure rates of multistory frames. J Struct Eng 115:268–284

Fajfar P (1999) Capacity spectrum method based on inelastic demand spectra. Earthq Eng Struct Dyn 28:979–993

Fajfar P (2002) Structural analysis in earthquake engineering – a breakthrough of simplified non-linear methods. In: Proceedings of the 12th European conference on earthquake engineering, London, 9-13 September 2002, CD-ROM paper, Paper ref. 843. Elsevier, p 20

Fardis MN, Biskinis DE (2003) Deformation capacity of RC members, as controlled by flexure or shear. In: Toshimi K, Hitoshi S (eds) Proceedings of the international symposium on performance-based engineering for earthquake resistant reinforced structures, a volume honoring Prof. Shunsuke Otani, Tokyo, 8-9 September 2003, pp 511–530

FEMA P-695 (2009) Quantification of building seismic performance factors. Federal Emergency Management Agency, Washington, DC

Gatto KS, Uang CM (2002) Effects of loading protocol and rate of loading on woodframe shearwall response. In: Seventh U.S. national conference on earthquake engineering (7 NCEE), Boston, MA, 21-25 July 2002

Gupta A, Krawinkler H (2000) Dynamic P-delta effects for flexible inelastic steel structures. J Struct Eng 126:145–154

Haselton CB, Liel AB, Deierlein GG (2009) Simulating structural collapse due to earthquakes: model idealization, model calibration, and numerical simulation algorithms. In: Papadrakakis M, Lagaros ND, Fragiadakis M (eds) Proceedings of the 2nd international conference on computational methods in structural dynamics and earthquake engineering (COMPDYN 2009), Rhodes, 22–24 June 2009, CD-ROM paper, paper no. CD497

Haselton CB, Liel AB, Deierlein GG, Dean BS, Chou JH (2011) Seismic collapse safety of reinforced concrete buildings. I: assessment of ductile moment frames. J Struct Eng 137:481–491

Ibarra LF, Krawinkler H (2005) Global collapse of frame structures under seismic excitations. Report No. PEER 2005/06, Pacific Earthquake Engineering Research Center, University of California, Berkeley

Ibarra LF, Krawinkler H (2011) Variance of collapse capacity of SDOF systems under earthquake excitations. Earthq Eng Struct Dyn 40:1299–1314

Ibarra LF, Medina RA, Krawinkler H (2005) Hysteretic models that incorporate strength and stiffness deterioration. Earthq Eng Struct Dyn 34:1489–1511

Jäger C, Adam C (2012) Seismic collapse assessment of non-deteriorating frames with irregular structural properties vulnerable to P-delta. In: Proceedings of the 15th world conference on earthquake engineering (15 WCEE), Lisbon, 24–28 Sept 2012, digital paper, paper no. 0245, 10 pp

Jäger C, Adam C (2013) Influence of collapse definition and near-field effects on collapse capacity spectra. J Earthq Eng 17:859–878

Jalayer F, Cornell CA (2009) Alternative non-linear demand estimation methods for probability-based seismic assessments. Earthq Eng Struct Dyn 38:951–972

Jennings PC, Husid R (1968) Collapse of yielding structures during earthquakes. J Eng Mech Div Proc ASCE 94:1045–1065

Krawinkler H, Zareian F, Lignos DG, Ibarra LF (2009) Prediction of collapse of structures under earthquake excitations. In: Papadrakakis M, Lagaros ND, Fragiadakis M (eds) Proceedings of the 2nd international conference on computational methods in structural dynamics and earthquake engineering (COMPDYN 2009), Rhodes, 22–24 June 2009, CD-ROM paper, paper no. CD449

Kunnath SK, Reinhorn AM, Park YJ (1990) Analytical modeling of inelastic response of R/C structures. J Struct Eng ASCE 116:996–1027

Liel AB, Deierlein GG (2013) Cost-benefit evaluation of seismic mitigation alternatives for older reinforced concrete frame buildings. Earthq Spectra 29:1392–1411

Liel AB, Haselton CB, Deierlein GG, Baker JW (2009) Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. Struct Saf 31:197–211

Lignos DG, Krawinkler H (2012) Sideway collapse of deteriorating structural systems under seismic excitations. Report no. TR 177, The John A. Blume Earthquake Engineering Research Center, Department of Civil and Environmental Engineering, Stanford University, Stanford

Lignos DG, Krawinkler H (2013) Development and utilization of structural component databases for performance-based earthquake engineering. J Struct Eng ASCE 139:1382–1394

MacRae GA (1994) P-Δ effects on single-degree-of-freedom structures in earthquakes. Earthq Spectra 10:539–568

S

Medina RA, Krawinkler H (2003) Seismic demands for nondeteriorating frame structures and their dependence on ground motions. Report no. 144. The John A. Blume Earthquake Engineering Research Center, Department of Civil and Environmental Engineering, Stanford University, Stanford

Melchers RE (1999) Structural reliability analysis and prediction, 2nd edn. Wiley, Chichester

Moehle J, Deierlein GG (2004) A framework methodology for performance-based earthquake engineering. In: Proceedings of the 13th world conference on earthquake engineering (13 WCEE), Vancouver, 1–6 Aug 2004, digital paper

NIST GCR 10-917-8 (2010) (ATC-76-1). Evaluation of the FEMA P-695 methodology for quantification of building seismic performance factors. National Institute of Standards and Technology, Gaithersburg

Panagiotakos TB, Fardis MN (2001) Deformations of reinforced concrete members at yielding and ultimate. ACI Struct J 98:135–148

Park Y, Ang A (1985) Mechanistic seismic damage model for reinforced concrete. J Struct Eng ASCE 111:722–739

Rosenblueth E (1981) Point estimates for probability. Appl Math Model 5:329–335

Shafei B, Zareian F, Lignos DG (2011) A simplified method for collapse capacity assessment of moment-resisting frame and shear wall structural systems. Eng Struct 33:1107–1116

Shome N, Cornell CA (1999) Probabilistic seismic demand analysis of nonlinear structures, reliability of marine structures program report no. RMS-35, Department of Civil and Environmental Engineering, Stanford University

Sivaselvan MV, Reinhorn AM (2000) Hysteretic models for deteriorating inelastic structures. J Eng Mech 126:633–640

Song J, Pincheira J (2000) Spectral displacement demands of stiffness and strength degrading systems. Earthq Spectra 16:817–851

Ugurhan B, Baker JW, Deierlein GG (2014) Uncertainty estimation in seismic collapse assessment of modern reinforced concrete moment frame buildings. In: Proceedings of the tenth U.S. national conference on earthquake engineering (10 NCEE), Anchorage

Vamvatsikos D, Cornell CA (2002) Incremental dynamic analysis. Earthq Eng Struct Dyn 31:491–514

Vamvatsikos D, Fragiadakis M (2010) Incremental dynamic analysis for estimating seismic performance sensitivity and uncertainty. Earthq Eng Struct Dyn 39:141–163

Villaverde R (2007) Methods to assess the seismic collapse capacity of building structures: state of the art. J Struct Eng 133:57–66

Zareian F, Krawinkler H, Ibarra L, Lignos D (2010) Basic concepts and performance measures in prediction of collapse of buildings under earthquake ground motions. Struct Des Tall Spec Build 19:167–181

Seismic Design of Dams

Guoxi Wu
Wutec Geotechnical International, New Westminster, Metro Vancouver, BC, Canada

Synonyms

CSR; Design seismic parameters; Earthfill dams; Factor of safety; Granular and plastic soils; Limit equilibrium; Newmark sliding block; Nonlinear finite element analysis; Performance-based design; PWP model; Risk analysis; Soil liquefaction; SPT; Time history

Introduction

Design of new dams or safety evaluation of existing dams for seismic loads is standard practice and routinely required. In a broad term, dams can be classified into three types – concrete dam, rock-fill dam, and earthfill dam – and seismic design can be at a dam safety level and/or at a serviceability level.

This entry provides contents and discusses methods for design and analysis of earthfill dams at a dam safety level where the ultimate limit state is applied for the highest level of design earthquake loads. For dam safety, dam failure is the primary concern; for serviceability consideration, the dam should remain functional and any damages should be easily repairable under this level of earthquake loads.

Dam Performance in Past Earthquakes

Damages to dams under earthquakes can result from ground shaking, soil liquefaction, ground cracking, ground displacements (lateral spreading and settlement), and in extreme cases surface rupture along an earthquake-active fault. Examples of recent big and devastating earthquakes include the 2008 Sichuan earthquake (a crustal

earthquake with a magnitude of M8.0 occurred in China) and the 2010 Chile earthquake (a subduction earthquake with M8.8 occurred off the coast of central Chile).

Historical data show that dams around the world have performed well and satisfactory and that the probability of a dam failure under strong ground motion shaking is low. Nearly all well-built and well-compacted embankment dams can withstand moderate earthquake shaking with peak ground accelerations (PGA) greater than 0.2 g. Dams constructed of clay soils on clay, or on rock, or on overburden foundations resistant to liquefaction have withstood (with no apparent damages) extremely strong shaking with PGAs from 0.35 to 0.8 g.

Soil liquefaction, either in the dam fills or in the foundation, is the most damaging factor that affects the performance of dams under earthquakes. Dams built of sandy soils, especially hydraulic or semi-hydraulic fills, or built on foundations of loose (low density) sandy soils are highly susceptible to earthquake damage or vulnerable to failures due to the potential for soil liquefaction. A famous case history of such is the near failure of the Lower San Fernando Dam in California, USA, in a 1971 earthquake (Seed 1979).

Design Seismic Loads Based on a Risk Analysis

Public tolerance to seismic risk for the consequence of dam failure ultimately determines the adequacy of an existing dam or the criteria for the design of a new dam. For instance, at a specific location with the same tectonic setting or geological condition and thus the same earthquake hazard condition, a nuclear reactor facility may be designed using a PGA of 1.0 g for a very low probability earthquake event, a hydroelectric dam using a PGA of 0.7 g for a low probability earthquake event, a residential building using a PGA of 0.5 g for a median probability of happening, and a temporary bridge for construction traffic using a PGA of 0.3 g for a relatively high probability of

occurring. This is mainly because the societal risk tolerance on a relative scale increases in order from nuclear radiation leak to flood from a dam breach, to collapse of a residential building, and to loss of a temporary structure.

The first factor contributing to seismic risk, such as from an earthquake event with a PGA of 0.7 g, is the occurrence rate of such earthquake which is often measured by annual exceedance probability (AEP). The other factor is the consequence as a result of an earthquake, which is ultimately measured in terms of loss of life and sometimes economic loss. At a conceptual level, seismic risk can be expressed as the product of seismic hazard probability and consequence, and it represents the probabilistic expectation of the consequence.

Ideally, seismic design of a dam should be based on a risk analysis, including calculation of the actual probability of a dam failure and its consequence. The adequacy of the dam would then be judged from the seismic risk (such as annual probability of single or multiple fatalities) that is acceptable to the society for the loss of life involved in a dam failure.

The risk-based approach to dam safety evaluation should balance the public risk and the limited societal resources available to manage the particular risk. As shown in Fig. 1, it is considered generally acceptable that the maximum level of societal risk to fatality is less than 10^{-3} per annum for loss of one life and the risk is less than 10^{-5} per annum when more than 100 lives would be lost in the event of a dam failure. The principle that the risks should be as low as reasonably practicable (ALARP) is generally followed in practice, and it is thus reasonable to use an annual probability of 10^{-5} and 10^{-6} for 100 and more fatalities.

In mathematical terms, seismic risk to loss of life (fatality) is calculated using the following two equations:

$$P_{Fatality} = P_{Failure} \times P_{Fatality/Failure}$$
$$\text{and}$$
$$P_{Failure} = P_{Earthquake} \times P_{Failure/Earthquake}$$

**Seismic Design of Dams,
Fig. 1** An example
relationship to manage
societal risk for dam safety
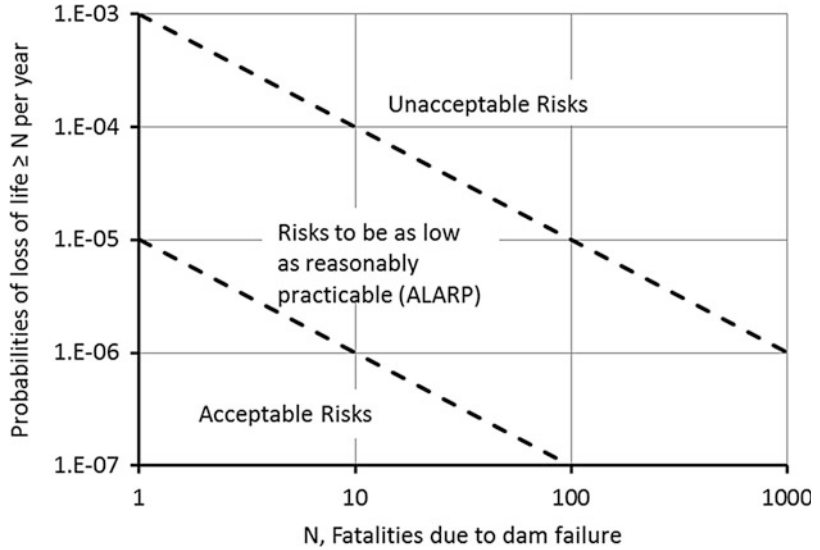
where:

$P_{Fatality}$ = unconditional probability of fatality
from earthquake, i.e., seismic risk.

$P_{Earthquake}$ = unconditional probability for an
earthquake to occur, such as the one with a
PGA of 0.7 g.

$P_{Failure/Earthquake}$ = conditional probability of a
dam failure in event of the earthquake.

$P_{Fatality/Failure}$ = conditional probability of loss of
life in the event of a dam failure.

The probabilities caused by all dam failure
initiating events (failure modes) need to be aggre-
gated or compounded in order to obtain the total
probability to a dam failure. Failure modes (such
as soil liquefaction) for earthquake loading should
be identified in advance in order to perform a risk
calculation. The risk calculation for a particular
failure mode ($P_{Failure}$) is carried out using an event
tree approach where significant events are
sequenced in levels. Construction of an event
tree for a dam requires special knowledge in geo-
technical earthquake engineering and a compre-
hensive understanding of the dam.

Using "soil liquefaction" as an example fail-
ure mode, an event tree for dam failure could
consist of earthquake acceleration (Level 1), con-
tributing earthquake magnitude (Level 2), lique-
faction analysis method (Level 3), soil

liquefaction capacity (Level 4), dam crest settle-
ment magnitude (Level 5), reservoir water level
(Level 6), and dam damage level (Level 7). For
illustration purpose, at Level 1 for earthquake
acceleration, it may have four scenarios with a
PGA of 1.0, 0.7, 0.5, and 0.3 g, respectively, and
each PGA scenario has its probability of occur-
rence (AEP). At Level 3 assuming two models for
liquefaction analysis, the weighting factors (e.g.,
0.55 and 0.45) would be assigned to each model
to make a total weighting of 1.

In many cases where a dam consequence class
is available after completion of a life safety
model and analysis, it can be conservatively
assumed that the conditional probability
($P_{Fatality/Failure}$) equals to 1.

Although in recent years the risk-based seis-
mic dam safety evaluation is increasingly used in
dam safety management, it is not widely used in
engineering design or seismic safety evaluation
due to the limited ability to perform such a com-
plex risk-based analysis. As described below, the
standard-based or traditional approach is more
commonly used in seismic design.

Standard-Based Method and Design Earthquake AEP

The standard-based method is a semi-
probabilistic method that defines the seismic

hazard using the probabilistic approach, but does not explicitly calculate the probability of a dam failure and its consequence. Without any quantifiable risk calculation, this is an empirical or experience-based approach to risk evaluation and management. The consequence, i.e., conditional fatalities in the event of a dam failure, is indirectly evaluated by using a dam consequence class scheme (FEMA 1998; CDA 2007). In some countries the highest consequence class is "Extreme" when potential (expected) fatalities are 100 and more.

With this approach, a seismic hazard level (AEP) is selected according to the consequence class of a dam so that the selected AEP conforms to the societal acceptable risk level. For a dam under an "Extreme" class, it would be reasonable to target a dam failure probability of 10^{-5} to 10^{-6} per annum, and an earthquake with an AEP of 10^{-5} or higher could be adequate, taking into account the satisfactory post-earthquake performance of dams around the world, i.e., the low conditional probability of a dam failure under strong shaking. For background information, the nuclear industry uses the 84th percentile spectra at AEP of 10^{-4} or approximately median-mean spectra at a probability of 10^{-5} per annum. With less stringent safety criteria than the nuclear industry, dam safety design can adopt a design earthquake with an AEP of 1/10,000 (i.e., 10^{-4}) based on the mean spectra. The mean is the expected value given the epistemic uncertainties, and the mean hazard value typically such as in Canada varies from 65th to 75th percentiles in the hazard distribution.

The design earthquake AEP for an "Extreme" class dam can vary from one region to another or from one country to another depending on the risk tolerance ability, and it is normally jointly selected by the dam owner and the government regulatory agency. In some developing countries with lower societal reliability than the developed countries, an earthquake AEP of 1/1,000 is used for dam safety design. On the other hand, even in developed countries such as Canada, an earthquake AEP of 1/2,475 is used for building safety design as a result to balance public risk and economic cost.

Once a seismic hazard level is selected, conventional analyses such as limit equilibrium or finite element analyses are conducted using seismic loads corresponding to the selected AEP. The results (such as stresses, ground displacements, or stability in a dam or its foundation) and consequences are evaluated deterministically using standards, specifications, and design codes. The potential for a dam failure is evaluated by comparing deterministically the resulting stresses and displacements with ultimate stability and established failure criteria. The evaluations of results are primarily done by empirical evidence, past experience, and engineering judgment.

Seismic Hazard Evaluations

A deterministic seismic hazard analysis (DSHA) is normally conducted for each known earthquake source to determine site ground motion parameters such as PGA, response spectra at 50th and 84th percentile, magnitude of earthquake, and duration of strong shaking. This would include assessment of potential seismic hazards from earthquake activities along local or regional known faults for the life of a dam.

A probabilistic seismic hazard analysis (PSHA) is usually conducted to determine design seismic parameters at various AEP levels, in additional to DSHA where known active faults or subduction zones are identified. A PSHA normally consists of **identification** of earthquake sources by developing a source model and the earthquake occurrence rate for each source zone with the understanding of local geology and regional past earthquake history, **application** of ground motion prediction equations (Earthquake Spectra 2008) that are appropriate to the region (e.g., hardness and shear wave velocity of the bedrock) and the types of seismic sources, and **determination** of earthquake response spectra at various levels of probability by the integration of hazard contributions over all earthquake magnitudes and distances for all seismic source zones.

The response spectra, either from DSHA or from PSHA, are normally defined for a series of discrete natural periods or frequencies (such as

Seismic Design of Dams, Fig. 2 Example results of seismic hazard analyses (**a**) probabilistic method and (**b**) deterministic approach
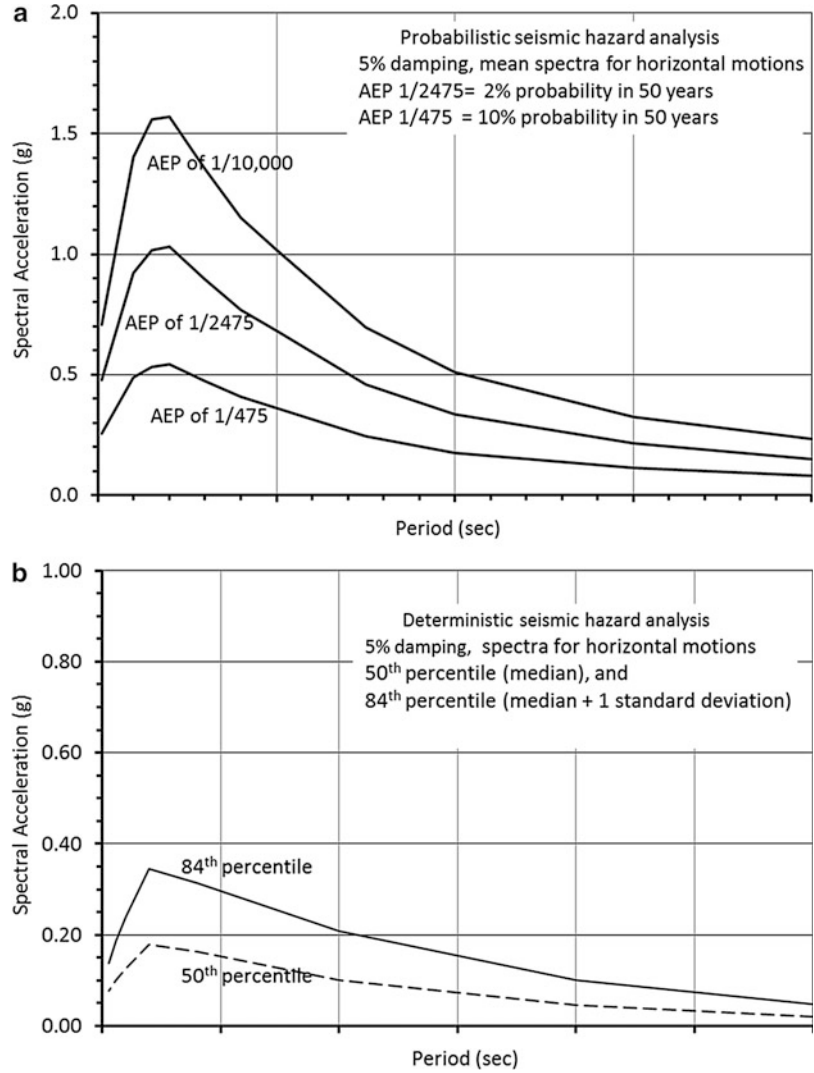
0.1, 0.2, 0.5, 1, 2, 5, 10, and 33 Hz; the frequency for PGAs can range from 33 to 100 Hz) that are used in the hazard calculation, and they are always associated with an uncertainty level (e.g., median, mean, or 84th percentile).

A seismic hazard analysis determines the potential intensity of seismic loads that could hit a damsite; and it always proceeds to, but does not always relate to, the design of a new dam or performance evaluation of an old dam for earthquake loading. However, seismic design of dams uses results of a seismic hazard analysis. Figure 2 shows

typical response spectra from DSHA and PSHA at the most commonly used damping ratio of 5 %.

Design Seismic Parameters

Seismic parameters for design of dams consist mainly of peak ground accelerations (PGAs), site-specific response spectra for horizontal and vertical accelerations, magnitude and site-source distance for the design earthquake, duration of strong shaking, ground motion time histories as

required for dynamic analyses, and fault displacements in rare situations when a dam is (or to be) on an active fault.

Design Response Spectra and PGAs

Response spectra and PGAs from DSHA can be used directly in seismic design, and they are conventionally computed at an 84th percentile value in the hazard distribution, i.e., one standard deviation (using $\varepsilon = 1$ in ground motion prediction equations) above the predicted median value. In some cases for seismic faults with low rate of activities, the 50th percentile values (the median) are used in design.

When seismic hazard is evaluated using a probabilistic approach, response spectra and PGAs from PSHA are used for seismic safety design of dams, using the results for the selected AEP (e.g., mean spectra with AEP of 1/10,000).

In current practice, the vertical ground motions do not seem to have much effect on the performance of earthfill dams and thus are not normally included in design analysis.

Design Earthquake Magnitude and Site-Source Distance

The PSHA results represent at a specific AEP level a composite of hazard contributions from earthquakes of all magnitudes and distances. The response spectra from a PSHA are also called the uniform (or equal) hazard response spectra (UHRS).

Deaggregation of the composite seismic hazard is performed to identify relative contributions of individual earthquakes or scenario earthquakes with various magnitudes and distances. For seismic design of a dam, representative earthquake magnitudes (M) and site-source distances (D) are obtained by deaggregation of the uniform hazard at natural periods that are significant and critical to the dam seismic response.

The representative M and D for the design earthquake are then used in seismic analysis where magnitude and duration of the earthquake are needed such as in the soil liquefaction analysis, or for selection of ground motion records needed in a time-history analysis.

Input Ground Motion Time Histories

The current trend in seismic design of dams is to conduct linear or nonlinear time-history analysis to obtain dynamic response of dam to earthquake loads. Time-history analysis of dam requires input ground motion time histories (acceleration, velocity, and displacement).

For dynamic analysis of earthfill dams, multiple ground motion records (five to eight commonly adopted) from past earthquake are selected, and each record is then linearly scaled to fit the design response spectra (such as a UHRS) over the range of natural periods that are appropriate for the dam. Using this method, an individual ground motion record is multiplied by a single scale factor to increase or decrease the magnitude of the motion and its spectrum, without modifying the shape of its spectrum or its frequency contents.

The selected ground motion records should be consistent with the seismic parameters and representative to the design earthquake in terms of magnitude, site-source distance, duration of strong shaking, tectonic setting and source mechanism, and consistency of site conditions between the recording station and the dam.

In some cases when uniform scaling of recorded ground motions is unable to meet the requirements, the design earthquake ground motions would be obtained by modifying a recorded ground motion in the time or frequency domain.

General Approach to Seismic Design and Analysis

For dams under static conditions, limit equilibrium analyses for slope stability are normally conducted for long-term operating reservoir conditions (steady-state seepage) and for short-term reservoir drawdown conditions. In general, for satisfactory stability factor of safety (FS) would be minimum 1.5 for long-term conditions (FS > 1.5) and 1.3 for short-term conditions (FS > 1.3).
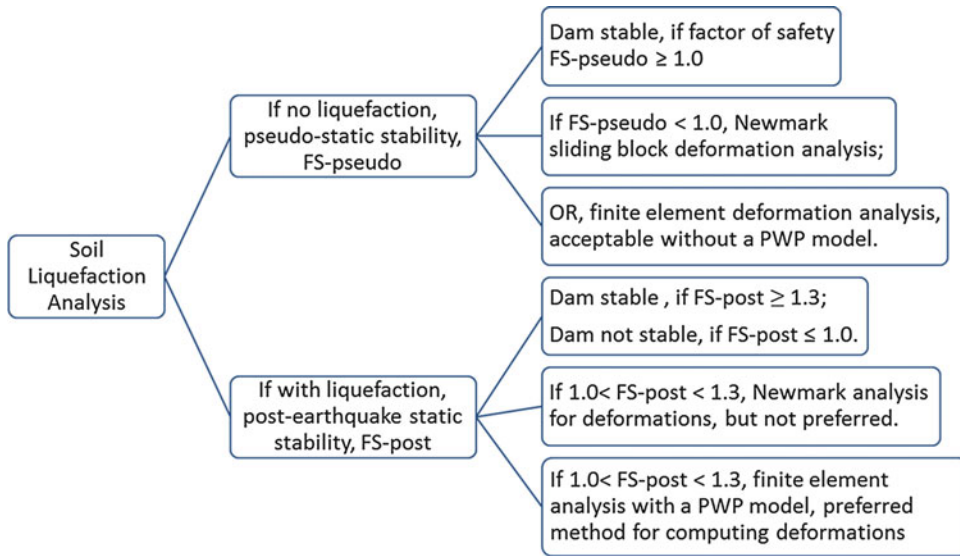
Seismic Design of Dams, Fig. 3 Types of seismic analyses based on liquefaction susceptibility (PWP = excess pore water pressures from earthquake loads)

For dams under seismic loading, two primary modes of dam failure must be addressed, the overtopping failure caused by excessive settlement of dam crest and the internal erosion and piping failure caused by cracks in dams, damage of filter layer upstream, or drains downstream of the core zones.

Some of the common design measures to mitigate these failures include the following: remove or improve problematic foundation soils by ground treatment and adequately compact the dam fills, use wide core zones of plastic soils that are resistant to erosion, use well-graded wide filter zones upstream of the core, construct chimney drains downstream of the core to lessen soil saturation and reduce downstream seepage, and use dam crest and downstream slope details to provide protection of the dam in the event of an overtopping.

Seismic analyses appropriate to a dam and its foundation are conducted using design seismic parameters to provide adequate information on expected stresses and ground displacements for evaluation of expected performance of the dam. Seismic analyses (see Fig. 3) can include a soil liquefaction evaluation, a pseudo-static stability analysis, a Newmark sliding block deformation analysis, a post-earthquake static stability analysis, and a finite element dynamic analysis for computing permanent ground deformations.

Evaluation of Soil Liquefaction

Liquefaction of Sands Under Cyclic Loads

Sand liquefaction is a fluidization process of saturated sand mass subject to cycles of shear stress. Under shaking, it can be easily observed that loose sands in a dry container will experience volume contraction and settle to a more compact state. In a saturated condition, immediate volume change of the sand would not occur because water in the pore does not drain quickly enough under the rapid earthquake loading. Instead, the potential for volume change translates into a quick increase in excess pore water pressure (PWP) in the sand mass. Liquefaction occurs when the PWP exceeds a threshold value that the pore water effectively suspends sand particles. Sand boiling to ground surface is a surficial expression of liquefaction of sands in the ground.

Liquefaction strength, or capacity, or resistance (more commonly used) of sands can be measured by laboratory testing using

Seismic Design of Dams, Fig. 4 Liquefaction resistance trend curves for clean sands (<5 % fines)
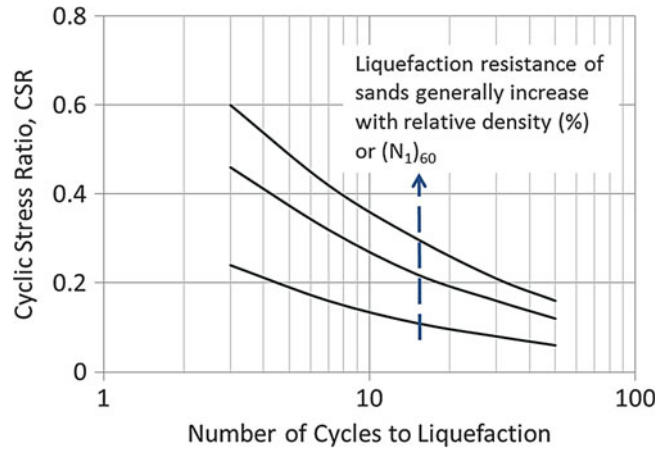
reconstituted samples prepared to a target relative density. It is found (see Fig. 4) that liquefaction resistance generally increases with relative density as expected, and under the same relative density, it decreases with more cycles of shear stress used in the testing.

Because liquefaction is triggered by multiple cycles of shear stresses, liquefaction resistance of soils is represented using both the cyclic shear stress level (τ_{cyc}) and the number of cycles for such uniform (constant amplitude) stresses to trigger a liquefaction failure. Cyclic stress ratio (CSR) refers to the ratio of the cyclic shear stress to the initial vertical normal stress (σ'_{v0}) at which the sample is under cyclic shearing, i.e., CSR = τ_{cyc}/σ'_{v0}. For earthquake loading, the number of cycles is related to earthquake magnitude and duration. In engineering analysis, nonuniform cycles of shear stress from a M7.5 earthquake are artificially set equal to, in terms of net effect on soil, a shear stress level with 15 uniform cycles. Thus, liquefaction resistance of soil is typically characterized by CSR at 15 cycles or CSR_{15} as shown in Fig. 5.

Liquefaction Resistance of Sandy Soils

Except with ground-frozen sampling technique that is practically not applicable, liquefaction resistance of in situ sands is considered not measurable from laboratory testing.

The most widely adopted test for field evaluation of relative density for sands is the standard penetration test (SPT) with split-barrel soil sampling. A key result of SPT is SPT N value that is the number of blow counts for a penetration depth of 305 mm (1 ft.), and $(N_1)_{60}$ is defined as the N value corrected for the soil under an in situ effective vertical normal stress of about 100 kPa (1 atmospheric pressure) and to the standard SPT hammer energy level (i.e., at 60 % of the theoretic maximum hammer energy). The general understanding is that the measured SPT N value is corrected downward (i.e., reduced) when the test is performed for soil at a depth with a confining stress higher than 100 kPa (e.g., at 300 kPa).

Current technology or methodology for evaluating liquefaction resistance of sandy soils was developed empirically by indirect studies of performance case histories of dams in historical earthquakes, where in some cases soil liquefaction was assessed to have occurred and in other cases liquefaction was believed not triggered.

Field soil liquefaction resistance of sands (particle size from 0.075 to 4.75 mm) has been derived from field performance of soils under various levels of historical earthquakes, and CSR_{15} from case history studies is plotted with in situ measured $(N_1)_{60}$ (Seed and Harder 1990) as an equivalence to the relative density of sand. Similar to the curve shown in Fig. 5, sands with $(N_1)_{60} > 30$ are considered to be dense and generally not liquefiable; sands with $(N_1)_{60} < 10$ are generally loose and highly susceptible to liquefaction.

Seismic Design of Dams,
Fig. 5 An example
liquefaction resistance
curve for clean sands
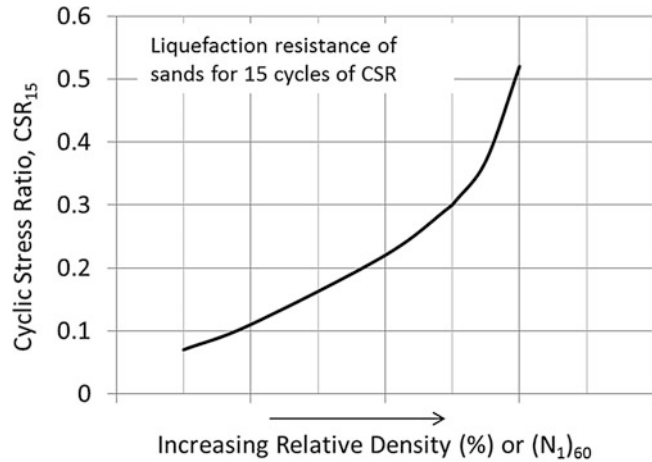
Unless completely confined with no drainage path, gravels (particle size from 4.75 to 75 mm) are not so much vulnerable to soil liquefaction due to its high permeability that allows fast dissipation of excess pore water pressures. However, sandy soils containing some amount of gravels are liquefiable, and their liquefaction resistance can be tested in the field using a large penetration hammer such as the Becker Hammer Penetration Test (Youd et al. 2001).

At a given relative density or $(N_1)_{60}$, sandy soils containing significant amount of fines (particle size <0.075 mm) are known to have higher liquefaction resistance than clean sands with no fines (Youd et al. 2001). Nonplastic silts, containing 100 % fines but with very low plasticity (such as plasticity index PI < 5 %), are considered to behave under cyclic loading in a similar manner as sands. The liquefaction resistance of nonplastic silts is evaluated as sands with upward correction on fines content.

Liquefaction Resistance of Plastic Soils

A more descriptive terminology for liquefaction failure of plastic soils (silts or clayey silts), containing 100 % fines and with relatively high PI (such as PI > 10 %), is "cyclic strain softening" which emphasizes more on the structural breakdown of the material under cyclic loading than on the buildup of excess pore water pressure. Evaluation of liquefaction resistance for plastic soils is an area that requires more research. In the

1980s and 1990s of the last century, soil index parameters (water content, liquid limit, PI, and fines content) were used as basis for liquefaction assessment, but the current trend method is to use results of laboratory cyclic tests (Finn and Wu 2013), especially for evaluation of more critical structures such as for dam safety.

Liquefaction resistance of in situ plastic soils can be measured directly by laboratory testing on in situ Shelby samples. Soil samples in a cyclic direct simple shear test are normally first consolidated to its pre-consolidation pressure (the highest historic pressure experienced by the soil or σ'_p) and then sheared by cycles of τ_{cyc} under σ'_{v0} and initial shear stress (τ_{st}) similar to the in situ stresses of the sample.

As compared to sands where the relative density or $(N_1)_{60}$ is the single parameter, there are more parameters and factors affecting and contributing to the liquefaction resistance for plastic soils (see Fig. 6). The curves for plastic soils are generally flatter than curves for sands; that is, the resistance decreases less with increasing number of cycles. The two key factors are the overconsolidation ratio (OCR = σ'_{v0}/σ'_p) and PI. Liquefaction resistance of plastic soils generally increases with increasing OCR and PI but decreases with increasing τ_{st}.

As a general observation, plastic soils having CSR_{15} > 0.4 (for 15 cycles) are considered to be hard, insensitive to earthquake loading, and unlikely to experience much strength loss. Plastic

Seismic Design of Dams,
Fig. 6 Liquefaction
resistance trend curves for
plastic soils, generally
flatter than sands
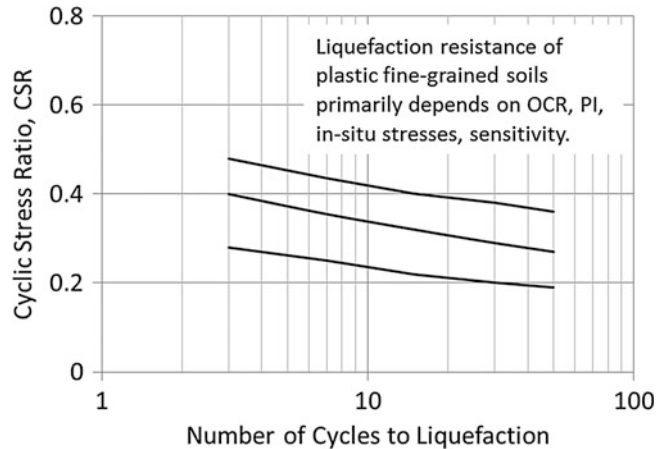
soils having CSR_{15} near or less than 0.15 may be sensitive to shaking and can experience significant loss of strengths in a strong earthquake.

Earthquake-Induced CSR

For liquefaction analysis, seismic load in a soil element is represented by time history of shear stresses (τ_{dyn}) imposed by earthquake shaking, and they can be calculated from a site response analysis using the input ground motions at the damsite and dynamic properties of dam fills and foundation soils.

Soil dynamic properties for each of the soil zones in the dam and its foundation primarily consist of low-strain shear modulus (G_{max}) and the damping characteristics (Seed et al. 1986). The low-strain shear modulus (G_{max}) is commonly computed from shear wave velocity (V_s) that can be measured by seismic downhole or crosshole surveys. Shear modulus would reduce from G_{max} as shear strain increases; reduction of soil shear modulus (stiffness) with increasing strains is more significant for gravels than for sands and more for sands than for clays. Soil damping ratio would be less than 5 % at low strain, but it increases with increasing strain. The maximum damping ratio is about 25 % for gravels and sands and in the order of 20 % for clays and silts.

A seismic site response analysis (▶ Site Response: 1-D Time Domain Analyses) is commonly performed in a total stress analysis, without including PWP effect, using a 1D soil column analysis for a low dam or using a 2D finite element dynamic analysis to model the geometric effect for a high dam. A site response analysis, either 1D or 2D model, can be conducted using the equivalent linear method (Idriss et al. 1974) or a true nonlinear method (Finn et al. 1986) for the simulation of shear modulus degradation (decrease) and damping ratio increase with the increase of shear strain. A true nonlinear approach is considered more appropriate when the ground shaking level is high, such as with a PGA of 0.4 g or higher.

Factor of Safety Against Liquefaction

Upon completion of a site response analysis, cyclic stress ratio ($CSR = \tau_{dyn}/\sigma'_{v0}$) in each soil zone is calculated by converting the nonuniform cycles of shear stresses from the earthquake to equivalent cycles of uniform shear stresses (Wu 2001) and then normalized to the in situ vertical normal stress (σ'_{v0}).

The liquefaction resistance (CSR_{15}) of soils is determined, for sands using $(N_1)_{60}$ from SPT and for plastic soils using liquefaction resistance curves from laboratory cyclic shear tests.

A factor of safety against soil liquefaction for each soil zone is calculated to be the ratio of liquefaction resistance of soil and seismic shear stress from earthquake, i.e., $FS_{LIQUEFACTION} = CSR_{15}/CSR$. A computed $FS_{LIQUEFACTION}$ near or less than 1.0 indicates triggering of soil liquefaction, and $FS_{LIQUEFACTION} > 1.5$ means not liquefiable.

Pseudo-Static Limit Equilibrium Stability Analysis

A pseudo-static analysis is a limit equilibrium method which includes additional seismic inertia forces in a conventional static slope stability analysis. Seismic coefficients (k_h for horizontal and k_v for vertical) are often used in such analysis, and they are normally taken as the peak ground accelerations (PGAs) as a fraction of the gravity acceleration. The seismic inertia forces are represented by $k_h W$ and $k_v W$ in horizontal and vertical direction, respectively, where W is the weight of a sliding block. Due to the nature of seismic shaking, shear strengths of soils for rapid loading conditions are conventionally used for characterization of soil resistance to earthquake loadings.

In a scenario that soils in a dam and its foundation would not develop significant PWP from shaking, a pseudo-static factor of safety greater than 1 (FS-pseudo >1) is a very strong indication that there would be little or no damage to the dam from an earthquake.

However, a pseudo-static factor of safety less than 1 (FS-pseudo <1) does not necessarily represent dam instability or unsatisfactory performance; instead due to the transient nature of earthquake motions, the dam would undergo some deformations for the short time interval when the ground acceleration exceeds the yield acceleration. In such case seismic ground deformations are computed and used for performance assessment.

It is not possible to predict failure of a dam or estimate seismic ground deformations by a pseudo-static analysis, and therefore other types of more comprehensive analysis are required to provide a more reliable basis for seismic design of dams.

Newmark Sliding Block Seismic Deformation Analysis

Newmark (1965) pointed out that for seismic loading permanent ground deformations of a dam, in addition to the pseudo-static factor of safety, should be considered to be tolerable or not for dam performance assessment. The Newmark deformation method assumes that deformation of a dam is modeled by a rigid block sliding on an assumed failure surface under the design ground motions at the damsite.

Sliding deformation is assumed to occur whenever the acceleration of a sliding block exceeds the yield acceleration, which is the horizontal acceleration that results in a factor of safety of 1 in a pseudo-static slope stability analysis, and the sliding stops when the acceleration falls below the yield acceleration. Mathematically, the Newmark sliding block displacements are computed by double integration of acceleration time history on the sliding block for the portion of accelerations that exceeds the yield acceleration (Fig. 7).

Although it has been widely used at a time when there is lack of direct methods for computing seismic ground deformations, the Newmark sliding block model is generally recognized to be not a very good representation of how embankment dams deform under earthquake, especially for dams with low factors of safety. Due to its simplicity, the method may provide a range of expected ground deformations, but it would neither give distribution of ground displacements in the dam nor provide direct calculation of dam crest settlement.

Post-Earthquake Static Stability Analysis

A direct consequence of soil liquefaction is a great reduction of shear strength, in particular for loose and contractive sands (Fig. 8), for which the shear strength may become a small fraction of its static and drained value due to the buildup of excess pore water pressures and fluidization. In design practice, shear strength of liquefied soil is also described as residual strength and sometime as post-liquefaction strength.

As a frictional material and under static loading, loose sand can have a friction angle possibly ranging from 30° to 35°; however, the post-liquefaction strength of the same loose sand

Seismic Design of Dams,
Fig. 7 Newmark sliding
block deformation
calculation diagram
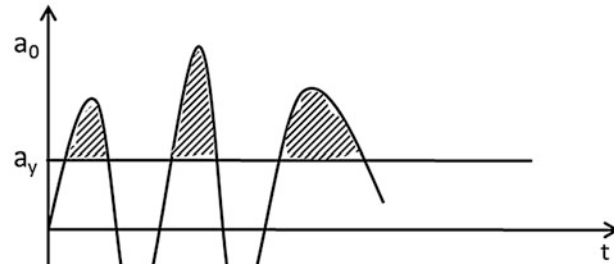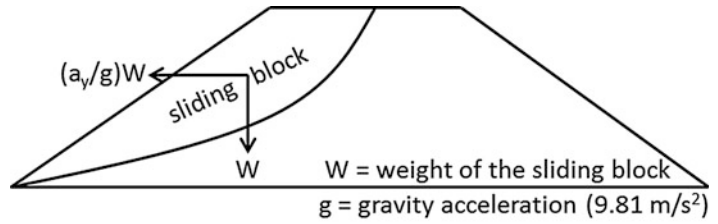
$a_0(t)$ - acceleration time history at center of a sliding block.
a_y = yield acceleration resulting a pseudo-static factor of safety of 1.
Permanent displacement is generated from the net portion (shaded) of $a_0(t)$ with $a_0 \geq a_y$.
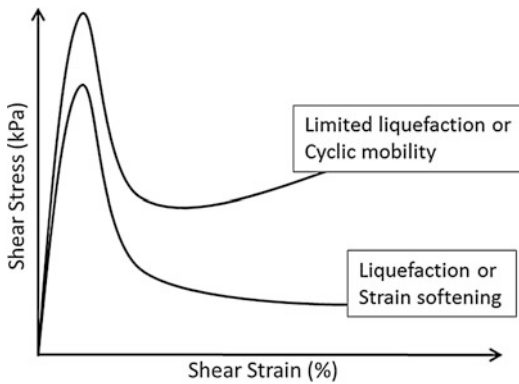
Seismic Design of Dams, Fig. 8 Shear stress – strain response indicating great strength loss after liquefaction and some strength loss with cyclic mobility

may decrease to an equivalent friction angle of $5°$ or perhaps less (Seed and Harder 1990), indicating the degree of shear strength reduction due to liquefaction.

Residual strength of plastic soils after initial strain softening could decrease to 50–80 % of its pre-earthquake static value, and after sufficient shaking disturbance, it can eventually decrease to the remolded strength value. Depending on sensitivity, remolded strength of silts and clayey silts may range from about 25 % of its peak static strength for low-sensitivity soils to only 10 % for high-sensitivity soils.

A static limit equilibrium analysis, without seismic loads but using the post-earthquake soil strength, is conducted in order to assess the post-earthquake static stability of the dam. The analysis would use residual strength for liquefied soils and reduced strength in soils that do not liquefy but with the buildup of excess pore water pressures from the design earthquake shaking.

A factor of safety less than 1 from a post-earthquake static stability analysis (FS-post <1) is a strong indication of dam instability and a potential failure during or immediately after the design earthquake. As such, the design for a new dam would be improved and ground strengthening for an existing dam would be required, in order to increase the post-earthquake dam stability and meet the dam safety requirements.

If FS-post >1 is calculated, ground deformations from the design earthquake are often computed for dam performance evaluation, preferably using a nonlinear finite element dynamic analysis. The Newmark sliding block deformation analysis should be avoided for computing deformations of dams involving soil liquefaction.

Seismic Deformations by Finite Element Dynamic Analysis

A more rigorous and direct method for computing earthquake-induced ground deformations, with and without soil liquefaction, is a nonlinear finite element dynamic analysis (Finn et al. 1994) that overcomes the uncertainties associated with the approximation embedded in the pseudo-static stability calculation and in the Newmark deformation analysis. With PWP calculations built into effective stress soil models, the finite element method has the capability to simulate the coupled effects from dynamic site response, generation of excess pore water pressures, soil liquefaction, and post-liquefaction behavior, and this is a trend method for quantitative evaluation of earthquake-induced ground deformations.

A well-known analysis procedure for the finite element approach is the one developed by Finn et al. (1986) using soil data from centrifuge tests and then verified using soil data and field measured deformations of the Upper San Fernando Dam in California, USA, in a 1971 earthquake (Wu 2001), and recently extended to include a dilative silt model (Finn and Wu 2013). This analysis procedure is highly regarded by practicing engineers and referenced for seismic design of dams by the US Federal Emergency Management Agency (FEMA 2005).

The Finn-Wu finite element procedure is simple and well suitable for practical engineering analysis. This analysis method adopts Mohr-Coulomb criteria for modeling shear strength of soils and uses a nonlinear hyperbolic model to simulate hysteresis response of soils under cyclic loads. Excess pore water pressures (PWP) caused by earthquake loads can also be computed using effective stress models with calculation of factor of safety against liquefaction for each soil element. In addition, the procedure provides two options for applying input ground motions: (1) ground accelerations applied at the model's rigid base and (2) outcropping ground velocities applied at the model's viscous and elastic base. The elastic base option is often used when the model base stiffness is not rigid relatively to the model body that prevents incident seismic waves from effectively reflecting back to the model body.

As an example of Finn-Wu finite element procedure, factors of safety against liquefaction computed from a dynamic analysis using a subduction earthquake input motion are shown in Fig. 9 (Finn and Wu 2013). The results showed that almost the entire saturated zone (below the water table shown in blue) under the downstream slope of the dam would undergo large cyclic strains with $FS_{LIQUEFACTION} < 1$. A computed deformed mesh of the dam, with soil material zones shown in colors, immediately after the earthquake using the same subduction ground motion is shown in Fig. 10, indicating a deep-seated ground deformation pattern and significant lateral movement and settlement of the downstream dam crest. The original dam surface is outlined in blue in the figure.

Seismic Performance Evaluation

In some cases, expected performance of dam can be assessed to either clearly safe or clearly unsafe, based on results of simple limit equilibrium stability analyses. In majority cases, however, dam performance is assessed from estimated ground deformations and strains. Ultimately, the expected ground deformations and the associated ground strains, either from a Newmark analysis or a nonlinear finite element analysis, should not exceed what the dam can safely withstand without catastrophic release of the reservoir water due to overtopping or internal
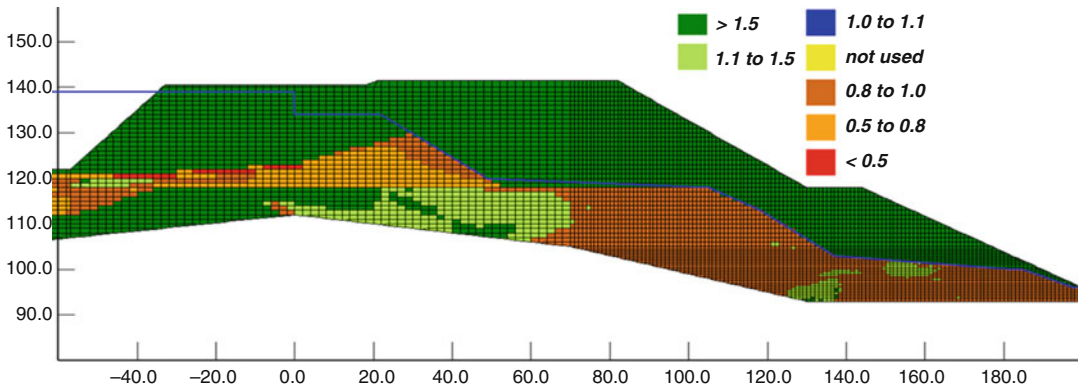
Seismic Design of Dams, Fig. 9 Computed FS$_{\text{LIQUEFACTION}}$ shown in colors for an example finite element model of an earthfill dam (undeformed, partial model, dimension in m)
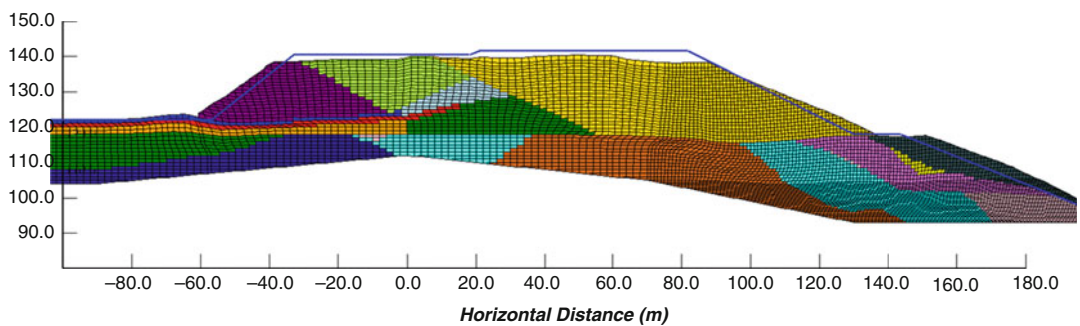
Seismic Design of Dams, Fig. 10 Deformed dam cross section after a subduction earthquake (displacement scale factor 1.0, elevation for vertical axis in m)

erosion. In the process, uncertainties associated with the calculated factor of safety or with the computed ground deformation need to be understood to achieve the confidence on the expected performance. In some cases, the consequences of misjudging uncertainty levels would also be taken into considerations.

Two key factors contributing to uncertainties in predicting the expected performance of dam under earthquake are related to soil liquefaction. The first is "Will liquefaction be triggered?", and the second is "What is the residual strength after soil liquefaction?" Ideally, a risk analysis would be beneficial to quantify the uncertainties. However, state of practice is mostly based on a deterministic approach for dam performance assessment, such as using analysis steps outlined in Fig. 11.

Overtopping of a dam in earthquake can be caused either by slope instability (failure) or by excess ground deformations (settlement) at dam crest. If post-earthquake stability analyses indicate a factor of safety well above 1 such as FS-post > 1.3, historical dam performance experience in earthquakes would indicate that the dam will have limited or small deformations and will perform satisfactorily. When liquefaction is not involved, factor of safety greater than 1 with seismic force in a pseudo-static analysis (FS-pseudo > 1) would also indicate that the dam will perform well in a design earthquake.

On the other hand, confidence in dam safety decreases and probability of a dam failure due to overtopping or internal erosion increases when a post-earthquake factor of safety (FS-post) near or less than 1.0 is calculated using residual strengths
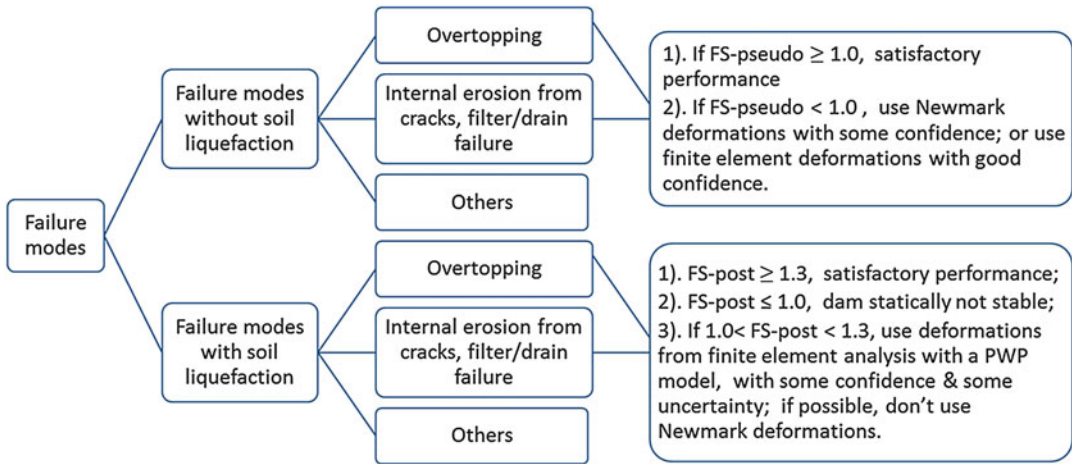
Seismic Design of Dams, Fig. 11 Evaluation of dam performance for critical failure modes under earthquake

for liquefied soils. A FS-post less than 1.0 would indicate slope instability under post-earthquake static conditions, and a dam slope failure can occur immediately after the earthquake.

In general, when a wedge or circular sliding surface has a low post-earthquake factor of safety, the deformations along the slip surface would be large or sometime excessive. In these cases, ground deformations calculated preferably from a finite element dynamic analysis are used to assess the potential of a dam failure by overtopping, or internal erosion, or loss of soils in piping.

Engineering judgment is carefully applied in assessing the level of uncertainties in design parameters and analytical methodology and thus the confidence level in the use of calculated deformations. When soil liquefaction is not an issue, deformations estimated from a Newmark sliding block analysis would prove to be adequate for many cases. For dams with liquefaction issues, seismic design or evaluation would use deformations computed from a finite element dynamic analysis, which has adequate soil models for simulation of soil liquefaction and post-liquefaction behavior and also can compute large-strain ground deformations. Even using a finite element approach, computed deformations for dams without involving soil liquefaction would be more reliable (with less uncertainties) than those involving liquefaction.

Summary

This entry describes basic principles and methodology used in seismic analysis and design for earthfill dams.

The risk-based approach for dam safety evaluation is introduced to indicate that seismic design criteria are governed by societal tolerance of seismic risk which consists of both seismic hazard and its consequence. In engineering practice, a full risk analysis is not commonly performed due to its complexity; instead, consequence of a dam failure in an earthquake is often represented using a dam consequence class scheme. A design level of seismic hazards is then selected from the consequence class for a dam. Seismic hazard of a damsite can be evaluated on a probabilistic approach. Once design seismic parameters are determined using the selected seismic hazard level, seismic analysis and design are conducted using a deterministic approach where seismic demands are compared with the ultimate capacity of soils.

State-of-practice methodology for evaluation of soil liquefaction is described, which include measurement of liquefaction resistance of soils by laboratory tests, assessment of liquefaction resistance from in situ tests in the field, soil parameters that impact the liquefaction resistance, and typical values of liquefaction resistance. Liquefaction potential of a soil element is

evaluated by comparing the seismic shear stress from earthquake with the liquefaction resistance of the soil.

Seismic analyses discussed in this entry include pseudo-static stability analysis, Newmark sliding block deformation analysis, post-earthquake static stability analysis, and finite element dynamic analysis for computing permanent ground deformations.

It is pointed out that soil liquefaction is the key factor contributing to uncertainties in predicting the expected performance of dam under earthquake, which are based on results of limit equilibrium stability analyses and in majority cases from estimated ground deformations. Engineering judgment should be applied in assessing the level of uncertainties in design parameters and analytical methodology and thus the confidence level in the use of calculated factors of safety or ground deformations.

Cross-References

▶ Probabilistic Seismic Hazard Models
▶ Site Response: 1-D Time Domain Analyses

References

Canadian Dam Association (CDA) (2007) Dam safety guidelines. CDA

Earthquake Spectra (2008) Special issue on the next generation attenuation project. Earthq Spectra Spec Issue 24(1):639–644

Federal Emergency Management Agency (FEMA) (1998) Federal guidelines for dam safety: hazard potential classification systems for dams, FEMA 333, Oct 1998. Reprinted April 2004

Federal Emergency Management Agency (FEMA) (2005) Federal guidelines for dam safety: earthquake analyses and design of dams, FEMA 65, May 2005

Finn WDL, Wu G (2013) Dynamic analyses of an earthfill dam on over-consolidated silt with cyclic strain softening, Keynote lecture. In: 7th international conference on case histories in geotechnical engineering, Chicago, Apr 2013

Finn WDL, Yogendrakumar M, Yoshida N, Yoshida H (1986) TARA-3: a computer program to compute the response of 2-D embankments and soil structure interaction systems to seismic loadings, Department of Civil Engineering, University of British Columbia, Vancouver

Finn WDL, Ledbetter RH, Marcuson WF (1994a) Seismic deformations in embankments and slopes. In: Balasubramaniam AS et al (eds) Proceedings symposium on developments in geotechnical engineering from Harvard to New Delhi 1936–1994, Bangkok. Balkema Publication. ISBN 90 5410 522 4, pp 233–264

Finn WDL, Ledbetter RH, Wu G (1994b) Liquefaction in silty soils: design and analysis. Ground failures under seismic conditions. ASCE Geotech Spec Publ 44:51–76

Idriss IM, Seed HB, Serff N (1974) Seismic response by variable damping finite elements. J Geotech Eng ASCE 100(1):1–13

Newmark NM (1965) Effects of earthquakes on dams and embankments. Geotechnique 15(2):139–160

Seed HB (1979) 19th Rankine lecture: considerations in the earthquake resistant design of earth and rockfill dams. Geotechnique 29(3):215–263

Seed RB, Harder LF (1990) SPT-based analysis of cyclic pore pressure generation and undrained residual strength. In: Proceedings of H. Bolton seed memorial symposium, vol 2. BiTech Publishers, Vancouver, pp 351–376

Seed HB, Wong RT, Idriss IM, Tokimatsu K (1986) Moduli and damping factors for dynamic analyses of cohesionless soils. J Geotech Eng ASCE 112(11):1016–1032

Wu G (2001) Earthquake induced deformation analyses of the Upper San Fernando dam under the 1971 San Fernando earthquake. Can Geotech J 38:1–15

Youd TL et al (2001) Liquefaction resistance of soils: summary report from the 1996 NCEER and 1998 NCEER/NSF workshops on evaluation of liquefaction resistance of soils. J Geotech Geo-Environ Eng ASCE 127(10):817–833 (21 authors)

Seismic Design of Earth-Retaining Structures

Luigi Callisto
Department of Structural and Geotechnical Engineering, Sapienza Università di Roma, Rome, Italy

Introduction

Earth-retaining structures provide support to excavations or earth fills: the earth pressure is either transmitted to the soil located below the foundations of the structure, as in gravity

retaining walls, or balanced by a combination of horizontal passive forces in the ground and the reaction forces of additional structural elements, as in the case of embedded retaining structures. Seismic forces have a detrimental effect on these structures, producing an increase in the earth pressure and a decrease in the available soil resistance. Commonly, the seismic design of an earth-retaining structure is carried out using a pseudo-static approach, in which the seismic forces are derived from a uniform acceleration field and are applied statically to the soil-structure system. This approach has the advantage of being relatively simple to implement, but the rationale for the choice of the constant acceleration values needs careful consideration. Therefore, the following section provides a simple conceptual model explaining how the pseudo-static acceleration values can be derived from the estimated maximum acceleration and from the acceptable seismic performance of the structure. The remaining sections provide the essential formulas for the evaluation of the seismic earth pressure and deal with specific issues relative to the seismic design of gravity and embedded retaining structures.

Conceptual Model for the Seismic Behavior of Retaining Structures

During an earthquake, the structure and the retained soil are accelerated and are therefore subjected to the corresponding inertial forces that, combining with the preexisting gravitational forces, produce additional loading on the retaining structure. It is important to note that the seismic inertial forces are different from the gravitational forces, as they vary cyclically in amplitude and direction, and have a transient nature. This has important consequences on the seismic design of a retaining structure.

Displacing Retaining Structures

An effective illustration of the seismic behavior of a retaining structure is provided by the simple scheme of a rigid block with a mass m resting on a horizontal base (Fig. 1a). The base is subjected to a simple dynamic excitation, consisting of a horizontal acceleration a_0 that is kept constant over a time interval t_0, and is then removed (Newmark 1965). If a_b is the block's acceleration, then the corresponding inertial force is ma_b that must be balanced by the resultant force T applied by the

Seismic Design of Earth-Retaining Structures,
Fig. 1 Schematic representation of a rigid block subjected to an acceleration pulse at the base
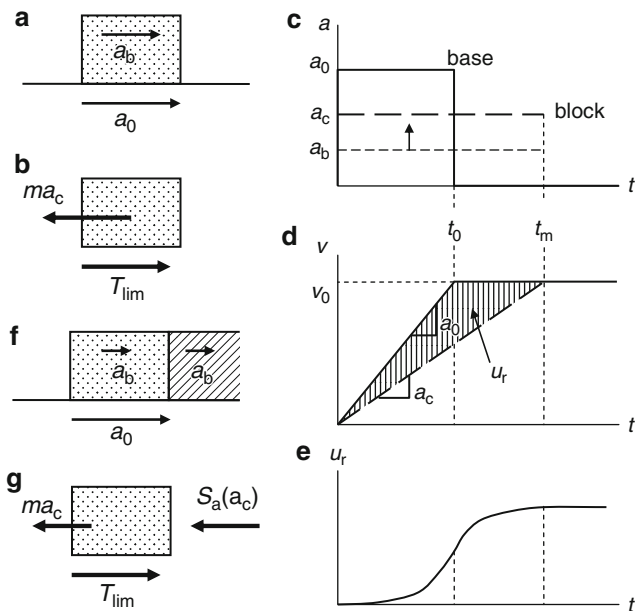
base to the bottom surface of the block (Fig. 1b). If the capacity at the block-base contact is T_{lim}, then the acceleration of the block cannot be larger that its critical value $a_{\text{c}} = T_{\text{lim}}/m$.

Now suppose that a_0 is larger than a_{c} (Fig. 1c): since the acceleration of the block cannot go beyond a_{c}, a relative acceleration occurs between base and block and the corresponding velocity plots tend to diverge, as shown in Fig. 1d. The shaded area between the two velocity plots represents the relative displacement u_{r} between base and block; if the horizontal acceleration at the base of the block were kept constant at a_0 for a long period of time, the system would be loaded permanently by the inertial forces, and the relative displacement would increase rapidly with an upward concavity. This would be by definition a collapse mechanism: the system capacity is attained under constant forces, and the displacements accelerate.

In the present case, since a_0 is maintained only for the time interval t_0, it can be seen that as t gets larger than t_0, the sign of the relative acceleration changes, and the relative velocity decreases. Therefore, the relative displacements increase with a downward concavity and eventually stop when the relative velocity equals zero (Fig. 1e): at the end of this specific event, since the system's capacity has been attained only transiently, the system has not collapsed, but rather suffered a damage in the form of a permanent relative displacement. Whether this damage is acceptable or not depends on the specific function of the system.

This schematic illustration shows that because of the transient nature of the seismic forces, the attainment of the full strength of the system during an earthquake does not imply a collapse, but rather a damage deriving from the temporary activation of a plastic mechanism. In principle, it would be possible to design the structure in such a way that its resistance be never attained during the seismic event ($a_{\text{c}} = T_{\text{lim}}/m > a_0$), but this approach to the design is not particularly desirable, especially for severe, low probability events, because it may be uneconomical and also because the seismic behavior would not benefit

from the dissipation of energy taking place during the activation of the plastic mechanism.

In addition to the inertial forces related to its own mass, a retaining structure is loaded by the earth thrust that in turn increases because the masses of the retained soil are accelerated. Therefore, a scheme more appropriate to a retaining structure is that of Fig. 1f, where the block is loaded additionally by $S_{\text{a}}(a_{\text{b}})$, which is the resultant force of the total earth pressures corresponding to the acceleration a_{b}. (Expressions for $S_{\text{a}}(a_{\text{b}})$ are given in the next section.) Hence in this case a given acceleration produces larger inertial forces. The critical acceleration a_{c} is derived from the equilibrium equation (Fig. 1g):

$$m\, a_{\text{c}} + S_{\text{a}}(a_{\text{c}}) \;= T_{\text{lim}} \qquad (1)$$

and is smaller than that of the previous scheme of Fig. 1a: the conceptual model is essentially the same, but from Fig. 1d and e, it is evident that a smaller critical acceleration corresponds to larger relative velocities and hence to larger displacements.

The schematic model of Fig. 1 evidences an important principle: since the block's acceleration cannot exceed a_{c}, the internal forces acting within the block cannot go beyond those evaluated for $a = a_{\text{c}}$. Since a_{c} is related to the seismic resistance of the system, it follows that stronger systems are called to resist larger internal forces.

With reference to the seismic design of a retaining structure, one can depict two different cases:

1. The critical acceleration of the system a_{c} is larger than the maximum acceleration at the base a_0; in this case, the maximum acceleration of the block is a_0, there are no relative displacements, and the maximum internal forces are evaluated with $a = a_0$.
2. The critical acceleration of the system a_{c} is smaller than the maximum acceleration at the base a_0; in this second case, the maximum acceleration of the block is a_{c}; the

displacements increase as a_c decreases; the internal forces are evaluated with $a = a_c$ and increase with a_c.

This conceptual model can be used to illustrate a performance-based approach to the design of earth-retaining structures. A given limit state corresponds to a certain design seismic action that must be found accounting for the local site response. The required seismic performance of a retaining structure can be expressed by its permanent displacement and can be chosen by the designer for compatibility with the limit state under consideration. The design consists on one hand in endowing the system with a critical acceleration sufficiently large to produce displacements that do not exceed the specified limit and on the other hand in protecting the structural elements from a premature yielding by making sure that their structural capacity is sufficient to carry the maximum internal forces occurring during the development of the displacements.

Evaluation of Seismic Displacements

The conceptual model of Fig. 1 can be used directly to evaluate, for a given acceleration time history at the base, the seismic displacements of a retaining structure. This is accomplished as shown in Fig. 2: whenever the base acceleration $a(t)$ exceeds a_c, the time history of the relative acceleration $a(t) - a_c$ is integrated to provide the relative velocity v_r. The integration is halted when v_r equals zero and is restarted when $a(t)$ becomes again larger than a_c. It is recognized that a retaining wall can move only downhill; therefore, the positive and the negative portions of the acceleration time histories are considered one at a time. The displacements are obtained from the integration of the time history of the relative velocity.

The direct evaluation of the displacements depicted in Fig. 2 is very simple, but can be performed only if the seismic action is described as a set of accelerograms. This implies the evaluation of the outcrop seismic motion that must be inserted into a detailed site response calculation, which in turn requires an accurate geotechnical
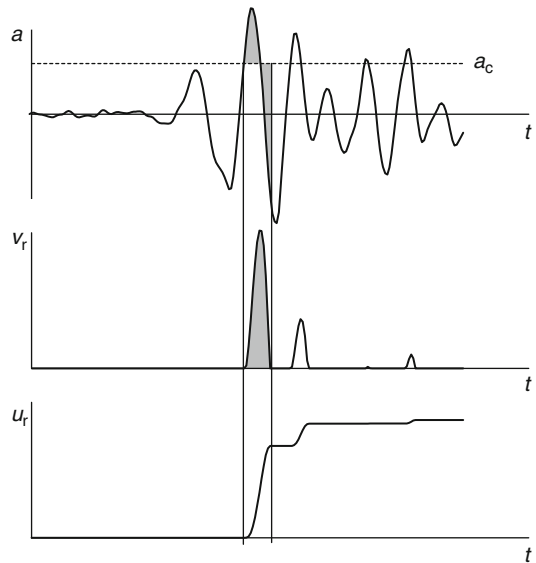
Seismic Design of Earth-Retaining Structures, Fig. 2 Illustration of the Newmark (1965) integration scheme

characterization of the soil deposits, starting from the bedrock.

For simple earth-retaining structures, the outcrop seismic action is more commonly expressed as an elastic response spectrum, and the local site response is considered through simple amplification coefficients, usually based on the classification of the subsoil according to the shear wave velocity of the topmost, say, 30 m. In this case, acceleration time histories are not available: a direct calculation of the displacements is not possible, and it is necessary to use a simplified approach.

The final displacement induced by the acceleration time history of Fig. 1 can be evaluated as

$$u_r = \frac{v_0^2}{2a_c}\left(1 - \frac{a_c}{a_0}\right) \qquad (2)$$

where $v_0 = a_0 t_0$ is the maximum acceleration at the base. Newmark (1965) argued that a given acceleration time history can be approximated as a_{max}/a_c cycles equivalent to the simple impulse of Fig. 1, where a_{max} is the maximum recorded acceleration. This assumption results in the following expression for the final displacement:
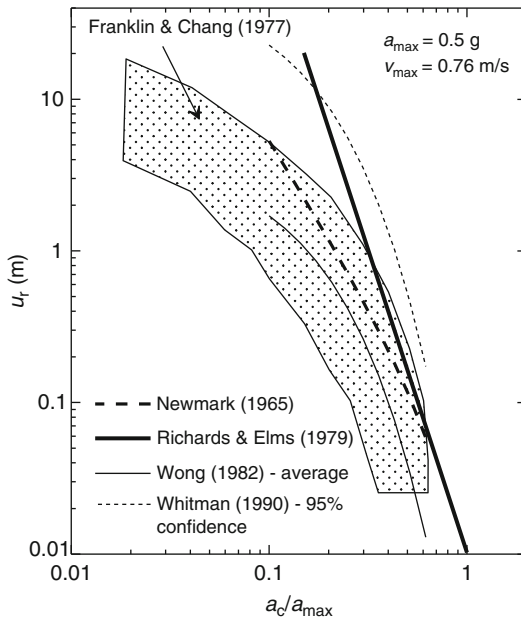
Seismic Design of Earth-Retaining Structures, Fig. 3 Seismic displacements plotted as a function of the critical to maximum acceleration ratio

$$u_r = \frac{v_{max}^2}{2a_c}\left(1 - \frac{a_c}{a_{max}}\right) \cdot \left(\frac{a_{max}}{a_c}\right) \qquad (3)$$

where v_{max} is the maximum recorded acceleration.

Franklin and Chang (1977) selected a number of accelerograms scaled to $a_{max} = 0.5$ g and $v_{max} = 0.76$ m/s; they produced a parametric integration of these accelerograms for different values of the ratio a_c/a_{max}. Richards and Elms (1979), based on the same acceleration time histories, produced the following relationship for the prediction of the seismic displacements:

$$u_r = 0.087\frac{v_{max}^2}{a_{max}} \cdot \left(\frac{a_{max}}{a_c}\right)^4 \qquad (4)$$

The displacements predicted by the above relationships are plotted in Fig. 3 as a function of the ratio a_c/a_{max}, together with the results by Franklin and Chang (1977) and with the curves proposed by Wong (1982) for the average values and by Whitman (1990) for the values associated to a 95 % confidence. In principle, the use of the previous relationships or of the diagram of Fig. 3 is quite immediate; the critical acceleration expresses the resistance of the retaining structure to the seismic action and can be evaluated using the earth pressure formulas given in the following; a_{max} is the maximum expected acceleration for the limit state under consideration; the maximum velocity v_{max} can be estimated, for instance, from the constant-velocity portion of the elastic spectra provided by the construction codes (e.g., EN 1998-5 2003).

However, Fig. 3 shows that for a given ratio a_c/a_{max}, the available relationships result in a very large uncertainty for the displacement. The reason for this substantial scatter is that the maximum recorded acceleration alone is not sufficient to represent the characteristics of an accelerogram. While in principle it is possible to produce more accurate expressions that relate the displacements to additional properties of the seismic motion (frequency content, Arias intensity, duration, etc.), often the designer that wishes to adopt a simplified procedure is provided only with an estimate of a_{max}. An alternative way to obtain a more accurate prediction of the seismic displacement is to use the structure of the above predictive equations for the development of local relationships based on regional seismic records.

Equivalent Seismic Action on Retaining Structures

Consider a relationship between the displacement u_r and the ratio a_c/a_{max} that can be deemed appropriate to the local seismicity (Fig. 4). If u_0 is the tolerable displacement for the limit state under consideration and β is the corresponding value of the ratio a_c/a_{max}, then the requisite that $u \leq u_r$ implies that the critical acceleration a_c must be larger than β a_{max} (with β < 1). This means that a check that the desired performance is met can also be done by controlling that under the conventional seismic forces associated to the acceleration βa_{max}, the resistance of the system is not exceeded.

In other words, the performance of the retaining structure that in principle should be evaluated in terms of displacements can also be checked in terms of equivalent forces. This is
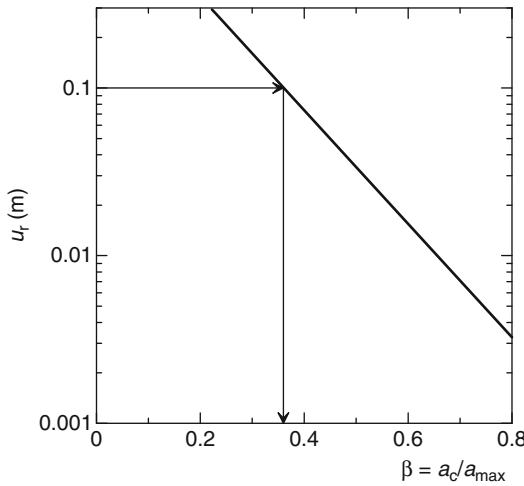
Seismic Design of Earth-Retaining Structures, Fig. 4 Relationship between seismic displacements and the ratio a_c/a_{max} developed by Rampello et al. (2010) for stiff soils based on the Italian seismicity

commonly expressed in terms of seismic coefficients: seismic forces deriving from a horizontal seismic coefficient,

$$k_h = \beta a_{max}/g, \tag{5}$$

generally used in conjunction with a vertical seismic coefficient expressed as a fraction of k_h, must be such that the system capacity is not exceeded.

Note that no global or partial safety coefficient should be used in this check, as this is not an assessment of the safety with respect to a collapse mechanism, but rather an indirect evaluation of the seismic performance of the system.

Non-displacing Retaining Structures

In some circumstances, a retaining structure cannot undergo permanent deformation resulting from the activation of a plastic mechanism. This may be due either to the design requirements for the specific limit state implying a negligible damage to the structure or to the retaining structure being significantly constrained by additional structural elements, like props, rakers, etc.

In the absence of displacements, the coefficient β of Eq. 5 is equal to 1, and therefore the seismic earth pressure can be evaluated using the formulas given in the following section, using a horizontal seismic coefficient corresponding to the maximum acceleration. However, there are cases in which a different approach is necessary, and the soil should be regarded as an elastic material. These are dealt with at the end of the next section.

Evaluation of Earth Pressure

Earth Pressure in Active and Passive Limit States

When a plastic mechanism is mobilized, forces and/or stresses transmitted by the soil in contact with the retaining structure may be evaluated by assuming that the soil is in limit conditions, that is, its shear strength is completely mobilized. If the retaining structure displaces away from the soil, the soil is in an active limit condition; conversely, if the soil is contrasting the displacement of the retaining structure, it is in a passive limit condition.

For a purely frictional strength criterion, the resultant force applied by the soil to the retaining structure may be found using the Mononobe-Okabe approach (Okabe 1924; Mononobe and Matsuo 1929), which is an extension to seismic conditions of the Coulomb's limit equilibrium method. In active limit conditions, the total earth thrust S_{aE} is provided by a volume of soil that slides downward along a planar surface with an inclination α to the horizontal; this mechanism is shown in Fig. 5, where φ is the angle of shearing resistance of the soil and δ is the angle of friction at the soil-structure contact. The earth thrust is found by maximizing with respect to α the force that ensures the translational equilibrium for the soil wedge of Fig. 5, subjected to the bulk force W_E and to the resultant force R applied by the in situ soil onto the sliding plane. Since the shear strength of the soil is fully mobilized along the sliding surface, R is inclined by φ with respect to the normal to the

Seismic Design of Earth-Retaining Structures, Fig. 5 Illustration of the Mononobe-Okabe approach for the evaluation of the active seismic force on a retaining structure
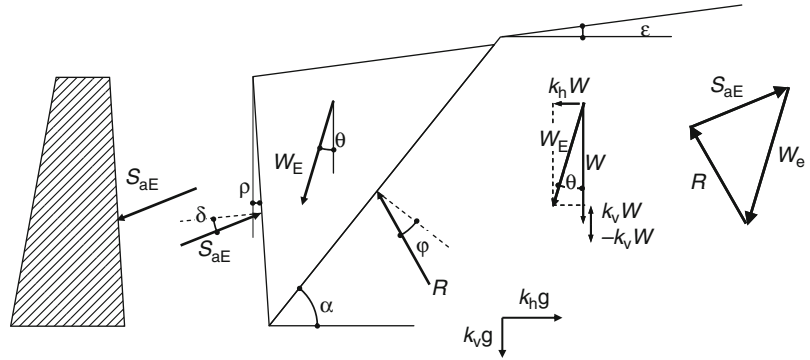
sliding plane and S_{aE} has an inclination δ with respect to the normal to the retaining structure. The seismic forces are included in the analysis by considering the horizontal and vertical seismic coefficients k_h and k_v in the evaluation of W_E.

The value of S_{aE}, maximized with respect to α, is given by the expression

$$S_{aE} = \frac{1}{2}\gamma \cdot (1 - k_v) \cdot H^2 \cdot K_{aE} \qquad (6)$$

where γ is the soil unit weight, H is the height of the retaining structure, and K_{aE} is the coefficient of active thrust, evaluated with the following equation:

$$K_{aE} = \frac{\cos^2(\varphi - \beta - \theta)}{\cos\theta \cos^2\beta \cos(\delta + \beta + \theta)\left[1 + \sqrt{\frac{\sin(\varphi + \delta)\ \sin(\varphi - \varepsilon - \theta)}{\cos(\delta + \beta + \theta)\cos(\varepsilon - \beta)}}\right]^2} \qquad (7)$$

in which the angles β and ε have the meaning depicted in Fig. 5 while the angle θ defines the direction of W_E:

$$\theta = \arctan\frac{k_h}{1 - k_v} \qquad (8)$$

The inclination α_{cr} of the sliding surface that maximizes S_{aE} can be evaluated with the following expression:

$$\alpha_{cr} = (\varphi - \theta) + \arctan\left[\frac{\{\tan a(\tan a + \cot b)[1 + \tan(\delta + \beta + \theta)\cot b]\}^{\frac{1}{2}} - \tan a}{1 + \tan(\delta + \beta + \theta)(\tan a + \cot b)}\right] \qquad (9)$$

with

$$\begin{aligned} a &= \varphi - \varepsilon - \theta \\ b &= \varphi - \beta - \theta \end{aligned} \qquad (10)$$

The active pressure S_{aE} given by Eq. 6 is a function of the horizontal acceleration, through

the dependency of K_{aE} on k_h (Eqs. 7 and 8); therefore, it can be regarded as the function $S_a(a_b)$ introduced in the previous section and used accordingly.

Equations 6, 7, and 8 imply quite obviously that S_{aE} increases with the horizontal seismic coefficient k_h. Conversely, it can be seen from

Eqs. 9 and 10 that the angle α_{cr} decreases with k_h, implying that the dimensions of the sliding wedge increase with the earthquake intensity. In other words, for larger accelerations, the volume of soil that interacts with the retaining structure is larger, and this finding has several implications, for instance, in the selection of the relevant strength parameters or for the design of a drainage system that needs to be effective in controlling the pore pressure within the entire volume of soil interacting with a retaining wall.

Since the Mononobe-Okabe method considers only translational equilibrium, it provides no direct information on the point of application of S_{aE}.

For passive limit conditions, the Mononobe-Okabe approach is not recommended, because

the assumption of a planar sliding surface is not conservative, especially when the soil-wall frictional coefficient δ approaches φ.

A more general approach to evaluate the seismic pressure in active and passive limit conditions is to use the lower-bound theorem of limit analysis. A convenient expression for vertical walls was derived by Lancellotta (2007) for passive limit conditions and extended by Rampello et al. (2011) to active limit conditions. The effective normal stress acting on the retaining structure in active or passive limit conditions $\sigma'_{a,p}$ is expressed as the product of a suitable thrust coefficient $K_{a,p}$ and the notional vertical effective stress σ'_v; the thrust coefficient has the expression

$$K_{a,p} = \left[\frac{\cos\delta}{\cos(\varepsilon - \theta) \pm \sqrt{\sin^2\varphi - \sin^2(\varepsilon - \theta)}} \cdot \left(\cos\delta \mp \sqrt{\sin^2\varphi - \sin^2\delta} \right) \right] \frac{\cos\varepsilon}{\cos\theta} (1 - k_v) \cdot e^{\mp 2\psi \tan\varphi}$$

(11)

with

$$\psi = 0.5 \left\{ \arcsin\left(\frac{\sin\delta}{\sin\varphi} \right) \mp \arcsin\left[\frac{\sin(\varepsilon - \theta)}{\sin\varphi} \right] \right.$$
$$\left. \mp \delta + (\varepsilon - \theta) + 2\theta \right\}$$

(12)

In the above expressions, the upper and lower operators apply to active and passive conditions, respectively. Equations 11 and 12 imply that K_a increases and K_p decreases with θ, that is, with the seismic acceleration.

Providing normal stresses that are proportional to the vertical effective stresses, this lower-bound solution yields the entire distribution of soil-wall contact stresses that can be used to study the equilibrium of the retaining structure for the selected limit state.

Effect of Pore Water Pressure

The presence of pore water pressures produces a change in the seismic contact stresses. This effect

is better evaluated using the above lower-bound solution. The total contact stress $\sigma_{a,p}$ is equal to the sum of the effective contact stress $\sigma'_{a,p}$ and the pore pressure u:

$$\sigma_{a,p} = K_{a,p}\,\sigma'_v + u = K_{a,p}\,(\sigma_v - u) + u \quad (13)$$

The change in total contact stresses $\sigma_{a,p}$ is not caused only by the introduction of u in the above expression but also by the coefficients of earth pressure in seismic condition depending on the pore water pressure.

For most practical purposes, it can be assumed that during the seismic motion, the pore water vibrates in phase with the soil skeleton (Matsuzawa et al. 1984 suggest that this is the case for permeability coefficients smaller than 10^{-5} m/s). The evaluation of the seismic pressure should consider that the inertial forces in the soil are proportional to the total mass of the soil skeleton and the pore water while the shear strength is proportional to the effective stresses. Callisto and Aversa (2008) showed that this

effect can be taken into account by modifying the expression of θ in Eq. 8, which becomes

$$\theta' = \arctan\left(\frac{\sigma_v}{\sigma'_v} \frac{k_h}{1 - k_v}\right) \quad (14)$$

where σ_v is the total vertical stress. This expression leads to larger values of K_a and to smaller values of K_p, showing that pore water pressures have a detrimental effect on the coefficients of earth pressure. Expression (14) implies that K_a and K_p may vary with depth even in a homogeneous soil. For the particular, and quite uncommon, case of a hydrostatic pore water pressure distribution with the total hydraulic head at the ground surface, expression (14) becomes independent of depth:

$$\theta' = \arctan\left(\frac{\gamma}{\gamma'} \frac{k_h}{1 - k_v}\right) \quad (15)$$

where γ is the unit weight and γ' is the submerged weight of the soil. Use of expression (15) is explicitly suggested by Eurocode 8 part 5 (EN 1998-5 2003).

Effect of Cohesion

The relationships listed above were developed for a purely frictional material. For some of these solutions, the possibility exists to account rigorously for the effect of cohesion. However, small values of the cohesion c' can be incorporated in the analysis with sufficient accuracy using the Rankine expression:

$$\sigma'_{a,p} = \mp 2c'\sqrt{K_{a,p}} + K_{a,p} \, \sigma'_v \quad (16)$$

Drainage Conditions

For a saturated soil, it can be assumed that the seismic event occurs in undrained conditions, even for coarse-grained materials. The excess pore water pressures Δu generated during the earthquake can be easily incorporated into Eq. 13, to provide (for a purely friction material)

$$\sigma_{a,p} = K_{a,p} \left[\sigma_v - (u + \Delta u)\right] + (u + \Delta u) \quad (17)$$

If a wall retains an excavation in a fine-grained soil, in the long-term, the removal of soil produces a significant over-consolidation and the volumetric response of the soil becomes dilatant. In this case, the seismic event is likely to produce a decrease of the pore water pressures ($\Delta u < 0$), and this beneficial effect may be neglected. However, for relatively loose coarse-grained soils, the earthquake may produce a significant increase Δu of the pore water pressure. The evaluation of Δu for coarse-grained material is beyond the scope of this entry; the reader may refer to the sections devoted to liquefaction and to Seed and Booker (1977).

An excavation in fine-grained material can be taken to occur in approximately undrained conditions, generating negative excess pore water pressures. Traditionally, the static design of an excavation in these conditions is carried out expressing the shear strength of the soil in terms of total stresses, using the Tresca criterion and a limiting shear stress equal to the undrained shear strength S_u. Yet for the analysis of the seismic conditions, a total stress analysis is seldom appropriate, because the time interval corresponding to a significant dissipation of the excess pore water pressure may be considered to be much smaller that the return period of any severe seismic event. Therefore, the seismic design is controlled by the seismic events that, having a large return period (e.g., 475 years for a 10 % probability over a life span of 50 years), may be deemed to produce effects on a retaining structure that is already in drained conditions.

Evaluation of Earth Pressure for Non-displacing Retaining Structures

Although the case of zero displacement can be addressed simply by inserting $\beta = 1$ into Eq. 5 and therefore evaluating the earth pressure using the maximum predicted acceleration, there are cases in which active limit conditions may not develop at the rear of a retaining structure, because of the very small deformation allowed during the excavation or because the wall has been backfilled after its construction. In these situations, if the deformation of the retaining structure is impeded during the seismic event, a

S

different approach is in order to evaluate the earth pressure, based on the hypothesis that the soil does not attain its strength and can be modeled as a linearly elastic material. The increase in the horizontal stress produced by the earthquake can then be expressed as

$$\Delta\sigma_{hE} = \gamma H \frac{a_{max}}{g} \cdot f \qquad (18)$$

where H is the height of the retaining structure. For an infinitely rigid wall, the coefficient f may be taken equal to 1 (Wood 1973). For a flexible wall, the coefficient f is a function of the soil-wall relative stiffness d_p, expressed as (Younan and Veletsos 2000)

$$d_p = \frac{G_s H^3}{(EI)_w} \qquad (19)$$

In the above relationship, G_s is the shear modulus of the backfill, and the quantity $(EI)_w$ is the bending stiffness of the retaining structure (per out-of-plane length). Figure 6 shows the coefficient f plotted as a function of the relative stiffness d_p as obtained by Younan and Veletsos (2000) for two different restraining conditions. This graph,
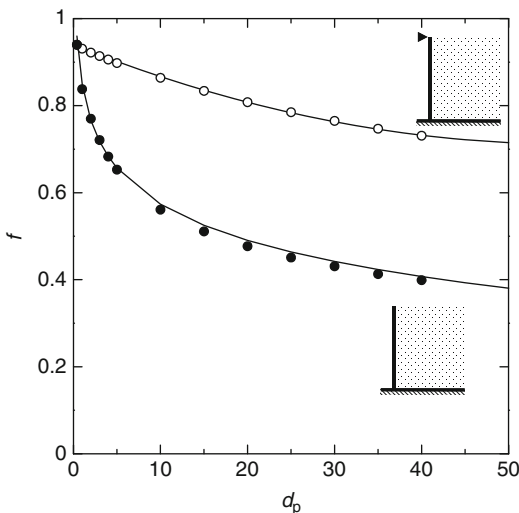
Seismic Design of Earth-Retaining Structures, Fig. 6 Coefficient f of Eq. 18 plotted as a function of the relative stiffness d_p

together with Eqs. 18 and 19, may be used to evaluate the design earth pressures.

Gravity Retaining Walls

In static conditions, gravity retaining walls transfer the sub-horizontal earth thrust S to the foundation soil, thanks to their considerable self-weight. Figure 7 shows that the weight of the wall W combines with S to yield a resultant force R that (i) goes through the base of the wall and (ii) has a small inclination to the vertical. As shown in Fig. 7, different types of gravity retaining walls have a similar global behavior, the only difference being the relative amount of soil contributing to the total weight of the structure.

Since these structures are endowed with a significant mass, seismic accelerations produce large inertial forces that must be taken into account, considering both the horizontal and vertical acceleration components.

Construction codes provide the values of the seismic coefficient that should be used for the design, essentially in the form of the reduction coefficient β of Eq. 5, depending on the allowable displacement at the end of the earthquake. Using the appropriate seismic coefficient, the design of a gravity retaining wall is carried out studying global and local mechanisms. The global mechanisms are essentially the same for the different types of retaining structures depicted in Fig. 7: (a) sliding along the base, (b) bearing capacity of the foundation, (c) overturning, and (d) overall slip-circle mechanism (Fig. 8). Local mechanisms are relative to the attainment of the structural strength or to the mobilization of the reinforcement strength in a reinforced earth structure.

The details of each safety check are given by the construction codes and therefore are not dealt with here. However, a general appreciation of the seismic design of a retaining wall is provided by the example of Fig. 9, relative to a cantilever retaining wall with a height $H = 3.5$ m subjected to a seismic event characterized by a maximum acceleration $a_{max} = 0.25$ g. Restricting the
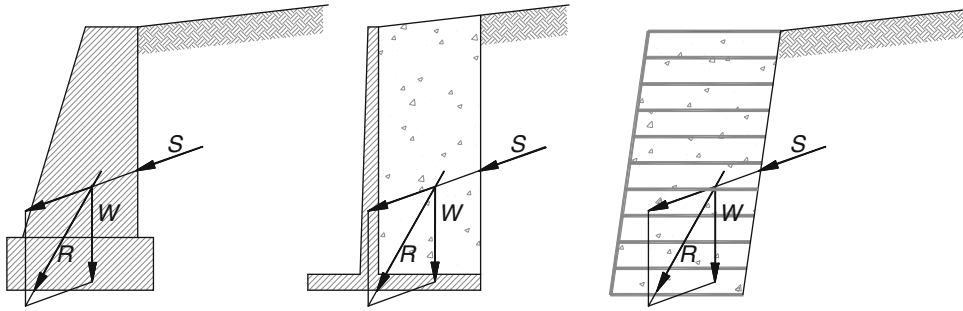
Seismic Design of Earth-Retaining Structures, Fig. 7 Transfer of earth pressure to the foundation soils for gravity retaining walls
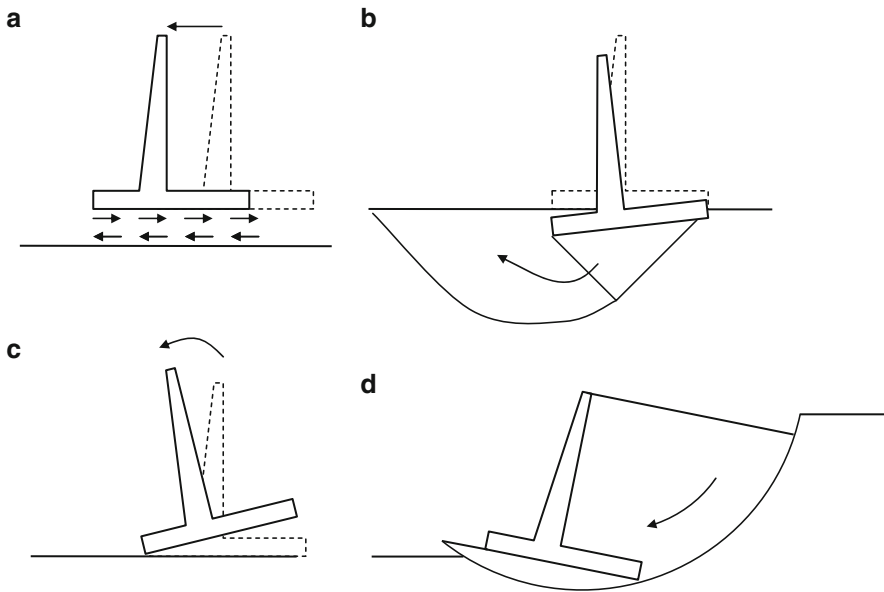
Seismic Design of Earth-Retaining Structures, Fig. 8 Global limit states for a gravity retaining wall

attention to the sliding mechanism only, Fig. 9 shows the values of the critical acceleration a_c plotted as a function of the width B of the wall base and of the corresponding displacements evaluated using the relationship of Fig. 4. It is evident that wider walls have a larger critical acceleration and therefore undergo smaller displacements. The figure also shows the maximum bending moment $M(a_c)$ in the wall stem, evaluated when the wall is critically accelerated. It is evident that since the bending moments increase with a_c, they must show an increasing trend with B. Hence, if a retaining wall is designed

with a large base to have a particularly good seismic performance, it must be endowed with a correspondingly large structural strength to resist the internal forces associated with its large critical acceleration. Following this line of thought, it is evident that a wall with a critical acceleration equal to a_{max} will suffer negligible displacement and will be subjected to internal forces that cannot be larger than those evaluated with an acceleration equal to a_{max}.

It should be noted that any "safe" provision, such as the use of strength or resistance factors, or an underestimation of the soil strength, leads to

S
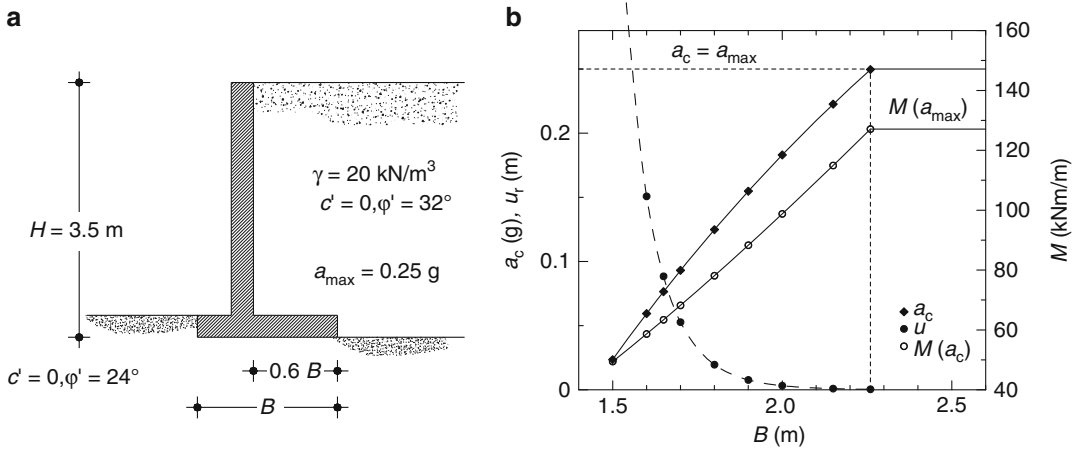
Seismic Design of Earth-Retaining Structures, Fig. 9 (**a**) Example of a retaining wall; (**b**) critical acceleration, displacement, and maximum bending moment plotted as a function of base width

Seismic Design of Earth-Retaining Structures, Fig. 10 Reference schemes of a cantilevered (**a**) and a singly propped (**b**) embedded retaining wall, showing the direction of the inertial forces activating the plastic mechanisms
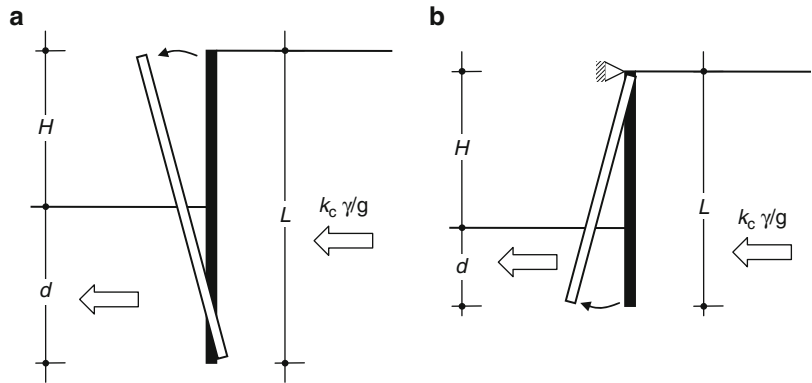
design acceleration smaller than the critical one. Figure 9 shows clearly that a larger critical acceleration implies on one hand smaller displacements and on the other hand larger internal forces. For the structural design of the wall, it is then recommendable to evaluate always the internal forces corresponding to the actual critical acceleration of the wall.

Embedded Retaining Walls

The seismic behavior of embedded retaining walls is different from that of gravity retaining walls, in that the earth thrust is resisted by the passive strength of the soil in front of the wall and by additional constraints (anchors, props); therefore, it is not necessary that the structure

possesses a large mass and the corresponding inertial forces are small; the effect of vertical acceleration can usually be neglected.

On the other hand, cantilevered or singly propped walls (Fig. 10) bear some similarities with gravity retaining walls, because if the strength of the soil interacting with the wall is attained during the seismic event, a plastic mechanism may develop, and the earth pressure relationships may be used in a limit equilibrium computation to evaluate the critical acceleration that activates the mechanisms. Figure 11 shows an example in which a cantilevered retaining wall is subjected to a maximum acceleration $a_{max} = 0.5$ g. For this wall, the critical acceleration can be obtained studying the rotational mechanism of Fig. 10 using, for instance, the Lancellotta (2007) closed form expressions for
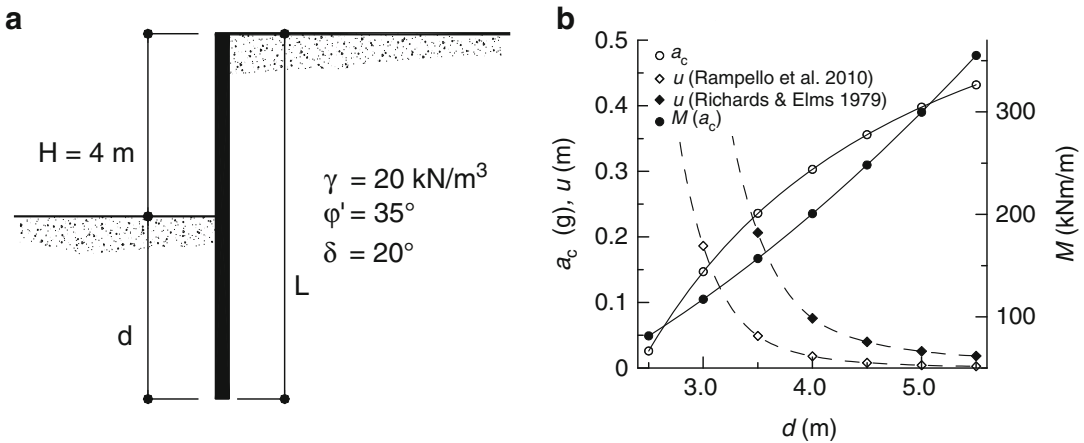
Seismic Design of Earth-Retaining Structures, Fig. 11 (**a**) Example of an embedded retaining wall; (**b**) critical acceleration, displacement, and maximum bending moment plotted as a function of embedded length

the active and passive pressures. The results are very similar to those developed for the example retaining wall of Fig. 9, showing that increasing the embedded length of the wall the critical acceleration increases, and therefore the permanent displacement decreases but the internal forces in the wall become larger. For these wall types, it is recommended that the internal forces be evaluated considering the critical value of the acceleration, that is, considering the activation of the plastic mechanisms depicted in Fig. 10.

When the structure has several additional constraints, the attainment of the soil strength is not sufficient for the activation of a plastic mechanism: if the design requires that the capacity of the wall itself and of the constraints cannot be reached, then these should be classified as a non-displacing wall and the earth pressure should be evaluated accordingly, as explained in a preceding section. In this perspective, no consideration of plastic mechanisms is needed, and the objective of the design is essentially the evaluation of the maximum internal forces in the structural members that should derive necessarily from a study of the soil-structure interaction.

However, the analysis of the soil-structure interaction under seismic conditions is not straightforward and can be carried out at different levels of complexity. A very basic representation of the soil-structure interaction is represented by the subgrade reaction methods that constitute a

standard method for the static design of multi-constrained retaining wall (Fig. 12a). While there is not a general consensus on the use of this method in seismic conditions, a simple strategy for taking into account the inertial force consists of the following steps:

(i) Carry out a static analysis to model the construction sequence (Fig. 12b).
(ii) Evaluate the additional forces due to the earthquake and the reduction of the coefficient of passive resistance, as a function of the maximum acceleration estimated for the site.
(iii) Apply these forces as external loads on the subgrade reaction model, and reduce the coefficient of passive resistance (Fig. 12c).
(iv) Run the program for equilibrium and compatibility (Fig. 12d).

Summary

The current trend in the design of earthquake-resistant structures is based on the evaluation of the seismic performance of the system. For a retaining structure, the seismic performance is expressed by its cumulative displacement at the end of the earthquake. Although it would be desirable to evaluate directly the performance of the structure, it has been shown that a design

S
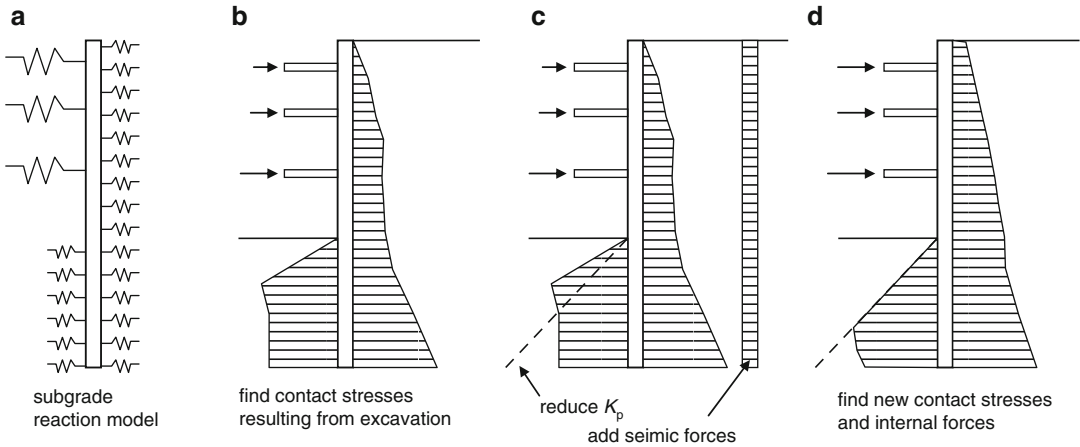
Seismic Design of Earth-Retaining Structures, Fig. 12 Use of a subgrade reaction model for the seismic analysis of a multi-propped retaining wall

requirement limiting the seismic displacements can be transformed into an equivalent requirement on the seismic action that needs to be considered in a conventional design process, which is commonly based on forces. Specifically, the conventional seismic coefficient used in a pseudo-static calculation can be made to decrease as the allowable displacement decreases. Specific cases of non-displacing walls have been evidenced that may call for a different approach, in which the internal forces are evaluated assuming that the soil does not mobilize its strength during the earthquake.

References

Callisto L, Aversa S (2008) Dimensionamento di opere di sostegno soggette ad azioni sismiche. In: Opere geotecniche in condizioni sismiche, MIR 2008. Pàtron, Bologna, pp 273–308

EN 1998-5 (2003) Eurocode 8: design of structures for earthquake resistance – part 5: foundations, retaining structures and geotechnical aspects. CEN European Committee for Standardization, Brussels

Franklin AG, Chang PK (1977) Earthquake resistance of earth and rockfill dams. Report 5, Permanent displacement of earth embankments by Newmark sliding block analysis. MP S-71-17, Soils and Pavements Laboratory, US Army Waterways Experiment Station, Vicksburg

Lancellotta R (2007) Lower-bound approach for seismic passive earth resistance. Géotechnique 57(3):319–321

Matsuzawa H, Ishibashi I, Kawamura M (1984) Dynamic soil and water pressures of submerged soils. J Geotech Eng 111(10):1161–1176

Mononobe N, Matsuo H (1929) On the determination of earth pressure during earthquakes. In: Proceedings of the world engineering congress, Tokio, vol 9, pp 177–185

Newmark NM (1965) Effects of earthquakes on dams and embankments. Fifth Rankine lecture. Géotechnique 15(2):139–193

Okabe S (1924) General theory on earth pressure and seismic stability of retaining wall and dam. J Jpn Civ Eng Soc 10(5):1277–1323

Rampello S, Callisto L, Fargnoli P (2010) Evaluation of slope performance under earthquake loading conditions. Ital Geotech J 44(4):29–41

Rampello S, Callisto L, Masini L (2011) Spinta delle terre sulle strutture di sostegno. Atti delle XXIII Conferenze Geotecniche di Torino

Richards R, Elms DG (1979) Seismic behavior of gravity retaining walls. J Geotech Eng Div ASCE 105(GT4):449–464

Seed HB, Booker JR (1977) Stabilization of potentially liquefiable sand deposits using gravel drains. J Geotech Eng 103(7):757–768

Whitman RV (1990) Seismic design and behavior of gravity retaining walls. In: Proceedings of the design and performance of earth retaining structures, Ithaca, ASCE GSP 25, pp 817–842

Wong CP (1982) Seismic analysis and improved seismic design procedure for gravity retaining walls. MSc thesis, Department of Civil Engineering, M.I.T., Cambridge, MA

Wood JH (1973) Earthquake-induced soil pressures on structures. PhD thesis, California Institute of Technology, Pasadena

Younan AH, Veletsos AS (2000) Dynamic response of flexible retaining walls. Earthq Eng Struct Dyn 29:1815–1844

Seismic Design of Pipelines

Hany El Naggar
Department of Civil and Resource Engineering,
Dalhousie University, Halifax, NS, Canada

Synonyms

Earthquakes; Lifelines; Liquefied soil; PGD; Pipeline breakings; Seismic-induced settlement

Introduction

Pipelines are a cost-effective means for the transportation of water supply and sewers, in addition to commercial fluids such as oil and gas. The designation of pipeline systems as "lifelines" indicates that they should be designed to function and operate at all times for public safety and well-being and also for economic reasons. Engineers must consider the different types of loads that are going to be imposed on the pipeline, the environment that the pipeline will travel in, and the type of material that the pipeline is going to convey. Consequently, pipelines must be designed for different loads such as stresses due to pressure generated by the flow (internal pressure), external pressure generated by the weight of earth and by live loads for buried pipelines, or external pressure generated by fluid if the pipe is submerged underwater. Also, seismic loads (earthquake loads) in medium and high seismicity zones are detrimental and should be considered. In general, pipeline design includes several general steps: (1) load determination, (2) critical performance evaluation (determining the critical stress and/or deformation), (3) comparison of the critical performance with the limiting criteria established by adopted codes and standards, and finally (4) selection of the pipe.

Over the past century, several catastrophic earthquakes caused severe damage to buried pipelines. In the 1906 San Francisco earthquake, one of the main reasons that caused the extensive damage was due to the failure of several water pipelines which hindered firefighting efforts. Several years later, in the 1971 San Fernando earthquake, about 80 % of the reported destruction were in welded steel buried pipelines. Following the 1989 Loma Prieta earthquake, the East Bay Municipal Utilities District identified over 120 water pipeline breaks, and the San Jose Water Company reported another 155 pipe breaks. However, most of the serious damage reported by Pacific Gas and Electric Company (PG&E) occurred to natural gas mains and service lines. During the Northridge earthquake (1994) in California, some natural gas pipelines were severely damaged and the leak of containment fluid caused a large explosion in the Northridge town resulted in several deaths (Lau et al. 1995; O'Rourke and Palmer 1996). In the 1995 Hyogo-Ken Nanbu earthquake (in Japan), the natural gas leakage from buried pipelines resulted in numerous number of fires (531 cases reported) which started primarily due to gas release and electricity sparks affecting areas of over one square kilometer totally burnt (Scawthorn and Yanev 1995). More recently, the 1999 Chi-Chi earthquake (in Taiwan) also caused severe damage to natural gas distribution systems. More than 100,000 customers were affected after the earthquake, and the estimated economic loss of five major natural gas companies was approximately US$ 25 million (Chen et al. 2000). In the 1999 Izmit earthquake (in Turkey), Tupras refinery suffered serious fire damage as the pipeline conveying water from neighbor lake was damaged, so the refinery was dependent on internal water reservoirs which were insufficient. In 2001 two earthquakes occurred (1 month apart) in San Salvador; several pipelines in rural areas experienced severe destruction caused by huge landslides. Till now, in several very recent earthquakes occurred in Chile (2010), New Zealand (Darfield, 2010), Japan (Tohoku, 2011), and Italy (earthquake of Emilia, 2012), serious damages to pipelines have been witnessed.

Accordingly, pipelines should be designed to function and operate during and following design earthquakes for life safety and economic reasons. As mentioned above several earthquakes in the

S
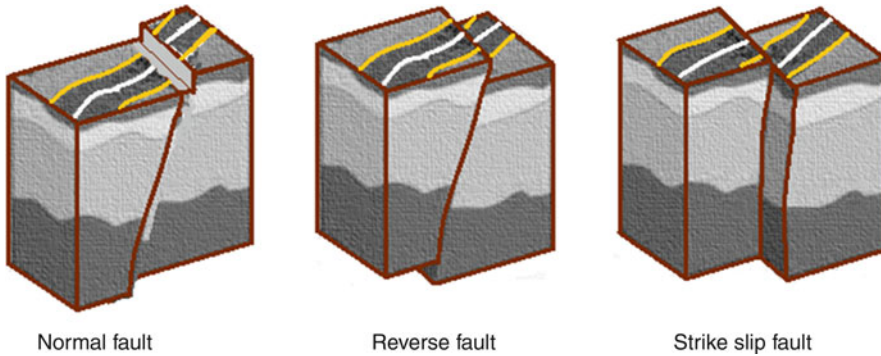
Seismic Design of Pipelines, Fig. 1 Main types of faults

last few decades resulted in too many pipeline breakings, and that extensive pipe breakage has the potential to lead to great economic harm to our urban communities.

The main focus of this chapter will be centered on the seismic analysis and design of pipelines.

Permanent Ground Deformation and Seismic Wave Propagation Hazards

In seismic events, buried pipelines can be damaged mainly by either the permanent ground deformations (PGD) or by the transient seismic wave propagation. PGD movements include faulting, landslides, lateral spreading due to liquefaction, and seismic settlement. Even though PGD risks are usually restricted to small regions within the pipeline, the chances of them causing severe damage is substantial since they impose large deformations. On the other hand, the seismic wave propagation risks typically affect the whole pipeline, but with lower damage rates (as the total deformations are general less and not permanent). For example, only 5 % of the affected area experienced lateral spreading during the 1906 San Francisco earthquake; approximately 52 % of all pipeline breaks occurred within one city block of the lateral spreading (O'Rourke et al. 1985).

Permanent Ground Deformation (PGD)
This section describes in details the four different forms of permanent ground deformation and

presents methods to calculate the amount of PGD as well as the extent of the PGD zone. Equations to quantify the amount of PGD are provided. Also, useful observations and notes to determine the extent of the PGD zone are presented and discussed.

Fault
Stresses in the earth's crust push the two sides of the fault. Eventually enough stress builds up and the rocks slip suddenly releasing energy in waves that travel through the rock to cause earthquake. Accordingly, earthquakes occur on faults. A fault is a thin zone of crushed rock separating blocks of the earth's crust. When an earthquake occurs on one of these faults, the rock on one side of the fault slips with respect to the other. The fault surface can be vertical, horizontal, or oblique to the surface of the earth. If the earthquake magnitude is large enough, the offset along the fault will propagate all the way to the earth's surface causing surface rupture (fault offset). Figure 1 below shows the main types of faults. In normal and reverse faults, the major ground displacement is vertical with a minor horizontal displacement. These ground displacements pose axial tension/compression and bending stresses in the pipeline depending on the direction of movement. On the other hand, in strike slip fault the main deformation (the offset) occurs in the horizontal plane, which poses axial tension/compression and bending stresses in the pipeline depending on the intersection angle of the pipeline and the fault.

The following empirical equations by Wells and Coppersmith (1994) are the most recognized equations to estimate the average fault displacement relative to the size of the considered earthquake:

For the normal fault $\rightarrow \log \delta_f = -4.45 + 0.63 M$

$$(1)$$

For the reverse fault $\rightarrow \log \delta_f = -0.74 + 0.08 M$

$$(2)$$

For the strike slip fault $\rightarrow \log \delta_f = -6.32 + 0.90 M$

$$(3)$$

where δ_f is the average fault displacement in meters and M is the moment magnitude of the earthquake.

Landslides

Seismically induced landslide involves a wide range of downslope mass ground movements, which can occur in offshore, coastal, and onshore environments.

Offshore Landslides In offshore landslides, the response of the pipeline is mainly governed by the orientation of the pipeline to the direction of ground movement. The imposed displacements for undersea slides are so large that pipeline response is likely controlled by the maximum force available at the soil-pipe interface.

Onshore Landslides There are several types of onshore landslides based on soil movements, geometry of the slide, and the types of material involved. The main types are rock falls, rock topples, slides, and lateral spreads. Rock fall and rock topple can cause direct damage to above-ground pipelines by the impact of falling rock fragments. In an earth slide the earth moves relatively as a block; they typically develop along natural slopes and embankments.

Based on Newmark's Block model for landslides (Newmark 1965), the critical acceleration at which the slide will be triggered α_c can then be determined from

$$a_c = g \, (\text{FOS} - 1) \sin \alpha \qquad (4)$$

where g is the acceleration due to gravity, FOS is the factor of safety, and α is the angle of the slope.

Then the displacement of the block can be calculated by double integrating the ground acceleration. Jibson and Keefer (1993) proposed the following equation to estimate the Newmark displacement, δ_s, in centimeters as

$$\text{Log } \delta_S = 1.460 \log I_a - 6.642 \, a_c + 1.546 \quad (5)$$

where α_c is the critical acceleration in gs and I_α is the Arias Intensity in m/s.

Arias Intensity can be calculated using the following relationship developed by Wilson and Keefer (1983):

$$\text{Log } I_a = M - 2 \log R - 4.1 \qquad (6)$$

where M is the earthquake magnitude and R is source distance in kilometers.

Lateral Spreading Due to Liquefaction

Seismic shaking may trigger the liquefaction of saturated loose cohesionless soils. The liquefaction process increases the pore water pressure in the ground to a level where the effective stress approaches zero at which point the soil loses entirely its shear strength, which in turn results in the lateral movement of the soil (lateral spreading). In past major earthquakes, large number of pipelines suffered massive damage caused by liquefaction-induced ground failures (Zhang et al. 2004).

Youd et al. (2002) proposed two empirical equations to approximately estimate the PGD due to liquefaction:

For lateral spreads down gentle ground slopes (GS),

$$\begin{aligned}
\text{Log} \delta_L = {} & -16.213 + 1.532 M - 1.406 \log R^* \\
& - 0.012 R + 0.338 \log S + 0.54 \log T_{15} \\
& + 3.413 \log(100 - F_{15}) \\
& - 0.795 \log(D50_{15} - 0.1)
\end{aligned}$$

$$(7)$$

Seismic Design of Pipelines, Fig. 2 Sketch showing the ground slope and the free face lateral spreads

Ground slope   Free face
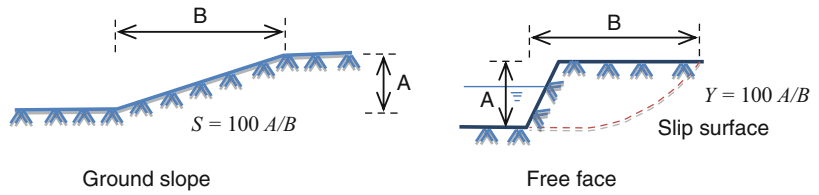
For lateral spreads at a free face (FF),

$$
\begin{aligned}
\text{Log}\,\delta_L = {} & -16.710 + 1.532\,M - 1.406\,\log R^* \\
& - 0.012\,R + 0.592\,\log Y + 0.54\,\log T_{15} \\
& + 3.413\,\log(100 - F_{15}) \\
& - 0.795\,\log(D50_{15} - 0.1)
\end{aligned}
\tag{8}
$$

where δ_L is the PGD in meters, M is the earthquake magnitude, S is the ground slope %, γ is the free face ratio (in % see Fig. 2), T_{15} is the thickness in meters of the saturated cohesionless soil layer with a corrected standard penetration test (SPT value of less than 15), F_{15} is the percentage of average fines contents in T_{15} (in %), $D50_{15}$ is the mean grain size in mm in T_{15}, and R^* is an adjusted distance parameter in kilometers given by

$$
R^* = R + 10\,e^{(0.89M - 5.64)}
\tag{9}
$$

where R is the horizontal distance in kilometers from the site of interest to the nearest bound of the seismic energy source (do not use less than 0.5 km).

Seismic Settlement

Seismic-induced settlement may be caused by densification of cohesionless soils, consolidation of cohesive soils, or consolidation of liquefied soil. In this section we will only discuss liquefaction-induced ground settlement as it can cause larger settlement and hence higher potential for damage to buried pipelines.

Takada and Tanabe (1988) proposed the two following empirical equations to calculate liquefaction-induced settlement at embankments and leveled sites:

For embankments

$$
\delta_{GS} = 0.11\,H_1 H_2\,\frac{a_{\max}}{N} + 20
\tag{10}
$$

For leveled sites

$$
\delta_{GS} = 0.30\,H_1\,\frac{a_{\max}}{N} + 2
\tag{11}
$$

where δ_{GS} is the liquefaction-induced settlement in centimeters, H_1 is the thickness of saturated cohesionless soil layer in meters, H_2 is the height of embankment in meters, N is the SPT N-value in the cohesionless layer, and $\alpha_{\max}$ is the ground acceleration in gals.

Seismic Wave Propagation

For the seismic analysis and design of buried pipelines, the effect of the seismic wave propagation on the pipeline is usually characterized by the induced ground strain and curvature. Newmark (1967) developed a straightforward method to estimate the ground strain. The general form of a traveling wave in Newmark's method is given by

$$
U = f\left(\frac{t}{T} + \frac{x}{\lambda}\right)
\tag{12}
$$

where U is the function of the separation distance between the two points, x, and the speed of the seismic wave; T is the period of the repeating motion; and λ is the wavelength.

For particle motion parallel to the direction of propagation (R-waves), ε_g, the ground strain along the direction of propagation can be calculated as

$$
\varepsilon_g = \frac{V_{\max}}{C_R}
\tag{13}
$$

where $V_{\max}$ is the maximum horizontal ground velocity in the direction of wave propagation and C_R is the propagation velocity of the R-wave.

For particle motion perpendicular to the direction of propagation (S-waves), C_g, the curvature can be calculated as

Seismic Design of Pipelines,
Fig. 3 Kennedy
et al. (1977) model for one
side of the fault

$$C_g = \frac{A_{\max}}{C_S{}^2} \qquad (14)$$

where $A_{\max}$ is the maximum ground acceleration perpendicular to the direction of wave propagation and C_s is the propagation velocity of the S-wave.

Wave propagation with particle motion parallel to the pipeline (R-wave) would induce large axial strain in the pipeline. On the other hand, S-waves propagating parallel to the pipeline with particle motion perpendicular to the pipeline direction would induce only bending strains.

For S-waves traveling in a horizontal plane and at an angle with respect to the pipeline, the S-wave particle motion (perpendicular to its direction of propagation) would have one component parallel to the pipeline inducing axial strain, ε_g, and another component perpendicular to the pipeline inducing bending strain, ε_b, which can be calculated as

$$\varepsilon_g = \frac{V_{\max}}{2\,C_s} \qquad (15)$$

and

$$\varepsilon_b = \frac{\pi D\, V_{\max}}{\lambda\, C_s} \qquad (16)$$

where ε_b is the upper bound bending strain (maximum), $V_{\max}$ is the peak ground velocity, C_s is the apparent propagation velocity of the S-wave, λ is the wavelength, and D is the pipe diameter.

The bending strains induced in a pipeline due to traveling waves (S-waves, L-waves, or the vertical component of R-waves) are generally small compared to the axial strains induced in the ground by the traveling waves.

Pipelines Response to Faulting

This section presents the response of continuous pipelines subject to fault offsets. PGD due to faulting can be resolved into two components: longitudinal PGD (parallel to the pipeline) and transverse PGD (perpendicular to the pipeline) axis. In the case of "normal fault" type (see Fig. 1), the pipeline will be subjected to bending axial tensile force, caused by the transverse and longitudinal components, respectively. In this case, tensile rupture would be the most probable failure mechanism. In the "reverse fault"-type case, the pipeline will be subjected to bending axial compressive force, caused again by the transverse and longitudinal components, respectively. In this case, buckling would be the most likely failure mechanism. In the last case of "strike slip fault," the pipeline can be subjected to either tension or compression depending on the intersection angle between the pipeline and the fault and pipe and the relative movement at the fault.

Kennedy et al. (1977) proposed a simplified method to analyze the tensile and bending behavior of pipelines due to fault movements. Figure 3 shows the Kennedy et al. model for one side of the fault.

According to Kennedy et al. (1977), the total strain in the pipe (bending + tensile) is given by

$$\varepsilon = \varepsilon_a + \varepsilon_b = \frac{\Delta L}{L} + \frac{D}{2R_c} \qquad (17)$$

where ΔL is the total elongation of the pipeline, L is the total length of the pipeline, D is the pipe diameter, and R_c is the radius of curvature of the curved portion. R_c can be estimated by

$$R_c = \frac{\sigma \pi D t}{P_u} \qquad (18)$$

where σ is the axial stress at fault crossing and P_u is the peak lateral pipeline-soil interaction force per unit length which can be calculated using the 1984 ASCE guideline using the following relations for sand and clay, respectively:

For sand

$$P_u = \gamma H N_{qh} D \qquad (19)$$

For clay

$$P_u = c_u N_{ch} D \qquad (20)$$

where γ is the unit weight of the soil, H is the embedment depth of the pipeline, and N_{qh} and N_{ch} are the horizontal bearing capacity factors for sand and clay, respectively. Figure 4 presents

horizontal bearing capacity factors for sand and clay after Hansen (1961).

The total elongation of the pipeline, ΔL, can be estimated using the following equation:

$$\Delta L = \delta_f \, \cos \beta + \frac{\left(\delta_f \, \sin \beta\right)^2}{3 L_c} \qquad (21)$$

where δ_f is the average fault displacement, β is the fault angle, and L_c is the horizontal projection length of the laterally deformed pipeline (see Fig. 3).

L_c can be approximately calculated using the following simplified equation:

$$L_c = \sqrt{R_c \delta_f \sin \beta} \qquad (22)$$

Pipelines Response to Longitudinal PGD

This section presents the response of continuous pipelines subject to longitudinal permanent ground deformations, PGD, where the soil movement is in the same direction as the pipeline.

Seismic Design of Pipelines,
Fig. 4 Horizontal bearing capacity factors for sand and clay (Reproduced after Hansen 1961)
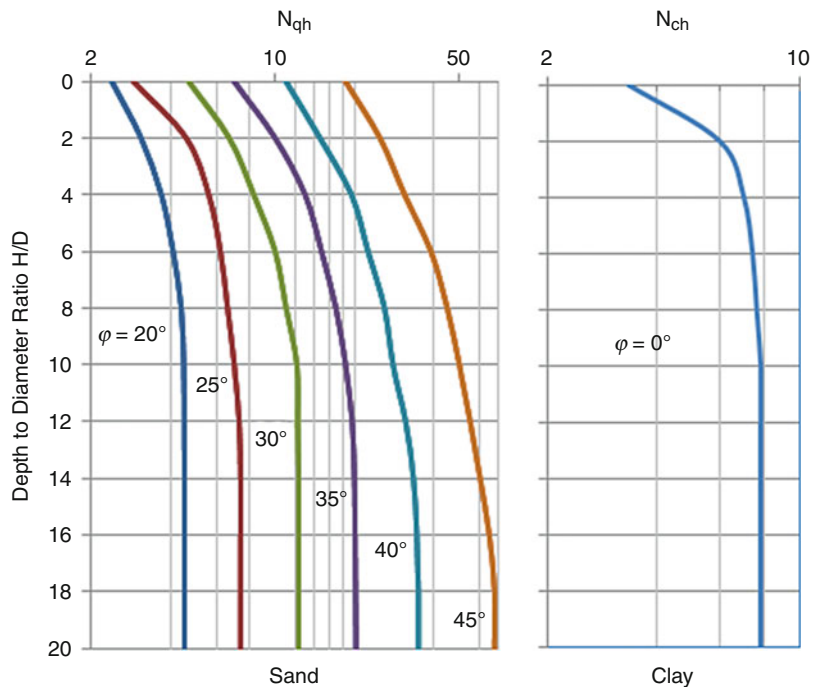
O'Rourke et al. (1995) proposed an inelastic pipeline model to analyze the response of pipelines to longitudinal PGD. Figure 5 shows the considered model. To develop this model O'Rourke et al. utilized an idealized block pattern in which a mass of soil having length L moves down a slight incline. The soil displacement on either side of the PGD zone is zero, while the soil displacement within the zone is a constant value δ.

In O'Rourke et al. model, a block of soil between Points B and D moves to the right pulling the pipe laterally with it with the soil forces acting on the pipeline within the PGD zone to the right. On the other hand, the soil between Points A and B and the soil between Points D and E resist the pipeline movement and these soil restraint forces are directed to the left. The combined pipeline-soil interaction forces result in a region of pipe axial tension between points A and B and a region of axial compression between points D and E. The conditions outlined in Fig. 5 correspond to the case where the PGD, δ, is comparatively large and the length of the PGD zone, L, is comparatively short. In that case, the maximum pipe displacement is less than the ground displacement and the pipe strain is controlled by L.

Figure 6 presents the other possibility of the O'Rourke et al. model where the length of the PGD zone is relatively large while the amount of PGD is comparatively short. Also in this case, there is still axial pipe tension between points A and B and axial compression between points D and E; however, the zone is long enough that the pipe displacement matches that of the ground between Points C and D where the axial force and strain in the pipe are zero.

As it can be seen from Figs. 5 and 6, the axial force in the pipeline in the segment AB is linearly proportional to the distance from Point A. Accordingly, the pipelines strain and displacement can be evaluated using the following Ramberg-Osgood model relations:

$$\varepsilon(x) = \frac{\beta_p x}{E} \left[1 + \frac{n}{1+r} \left(\frac{\beta_p x}{\sigma_y} \right)^r \right] \quad (22)$$

$$\delta(x) = \frac{\beta_p x^2}{E} \left[1 + \frac{2}{2+r} \frac{n}{1+r} \left(\frac{\beta_p x}{\sigma_y} \right)^r \right] \quad (23)$$

where n and r are Ramberg-Osgood parameters (given in Table 1 below), E is the modulus of elasticity of steel, σ_y is the effective yield stress, and β_p is the pipe burial parameter, defined as the friction force per unit length t_u divided by the pipe cross-sectional area A.

The pipe burial parameter β_p can be obtained from the following:

For sand

$$\beta_p = \frac{\tan \varphi \, \gamma \, H}{t} \quad (24)$$

For clay

$$\beta_p = \frac{\alpha \, c_u}{t} \quad (25)$$

where φ is the angle of shear resistance, γ is the effective unit weight of the soil, H is the embedment depth of the pipeline, α is the adhesion factor for clay, c_u is the undrained cohesion of the clay (see Fig. 7 below), and t is the pipe wall thickness.

Wrinkling of the Pipe Wall in Compression

Substituting a critical local buckling strain into Eq. 22, one can obtain the critical length of PGD zone L_{cr}. This can then be used to calculate the critical ground movement δ_{cr} from Eq. 23. The critical strain in compression may be taken as 0.175 t/R.

Pipeline Response to Transverse PGD

This section presents the response of continuous pipelines subject to transverse permanent ground deformations, PGD, where the soil movement is perpendicular to the pipeline. When subjected to transverse PGD, the pipeline will stretch and bend as it attempts to conform to the transverse ground movement profile. In this case, the failure
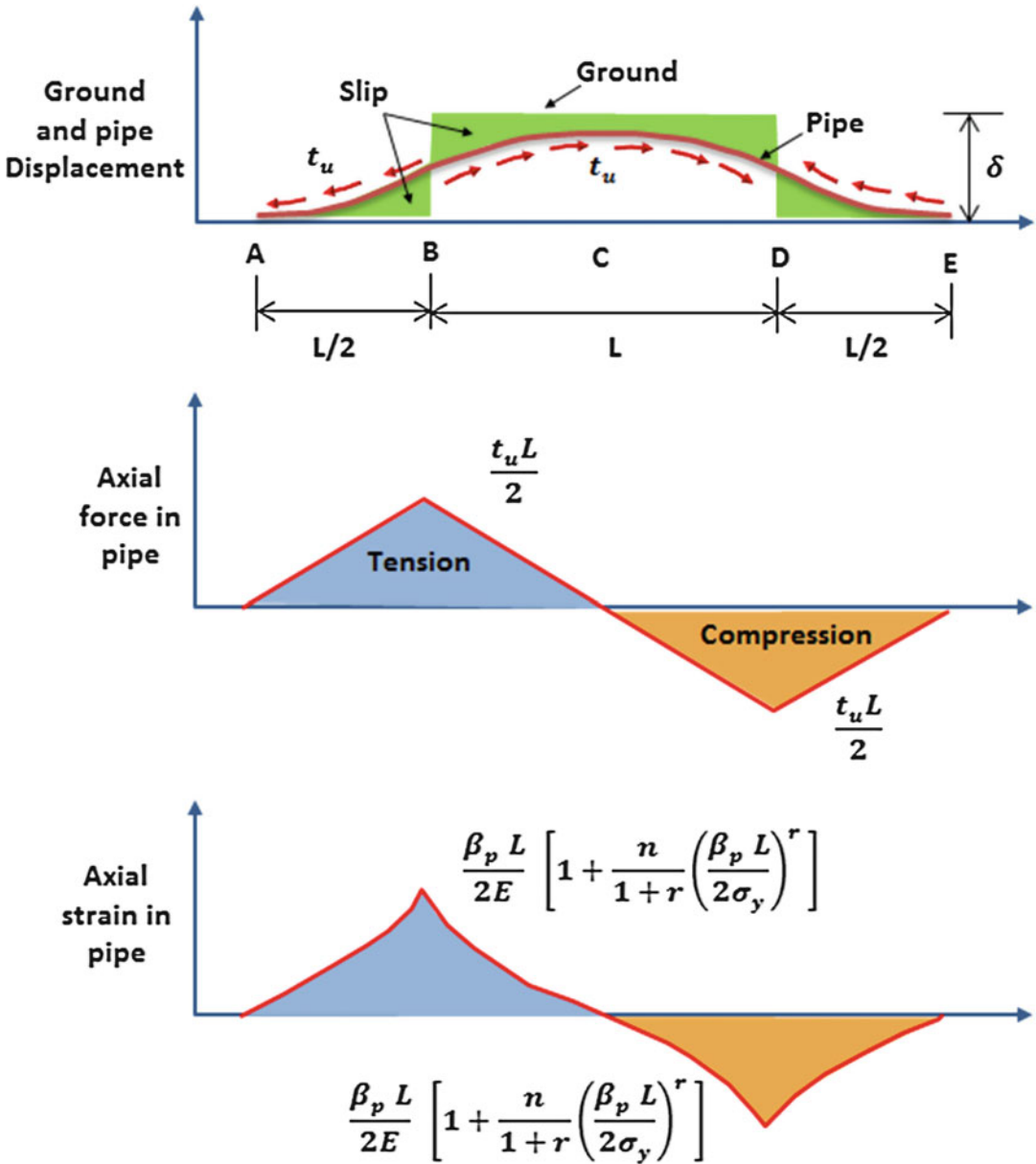
Seismic Design of Pipelines, Fig. 5 O'Rourke et al. model to analyze the response of pipelines to longitudinal PGD

mechanism of the pipeline will be governed by the relative magnitude of the axial tensile strain and the bending strain. If the tensile strain is relatively small, the pipe wall may buckle in compression due to excessive bending. Instead, if tensile strain is relatively large, the pipe may rupture in tension due to the combined effects of the tensile and bending stresses.

In general, the response to transverse PGD is a function of the magnitude of PGD, the width of the PGD zone, and the pattern of ground deformation. Two types of transverse ground deformation patterns are discussed here: the spatially distributed transverse PGD pattern and the abrupt transverse PGD pattern. Figure 8 shows sketch of the considered patterns.
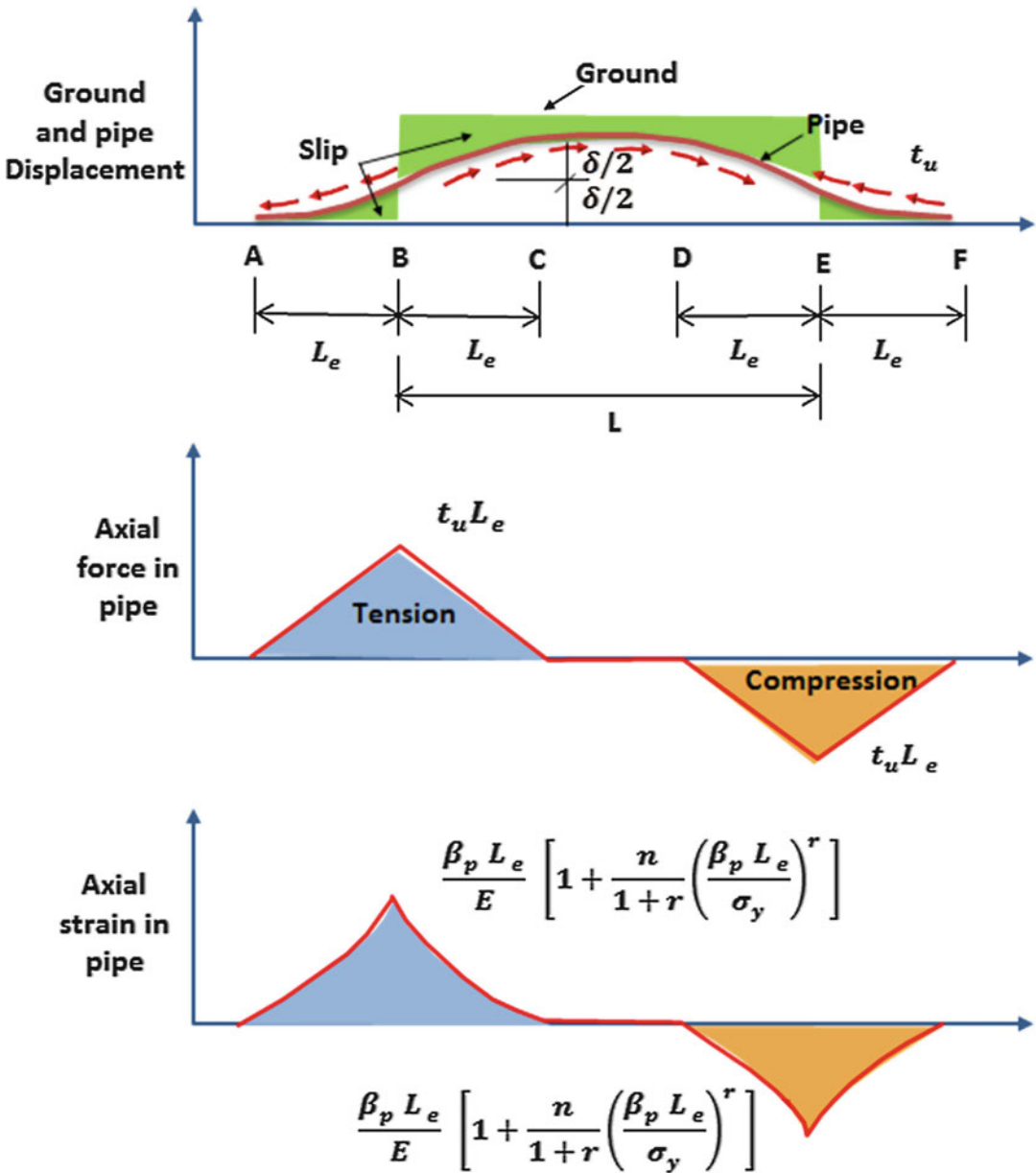
Seismic Design of Pipelines, Fig. 6 O'Rourke et al. model to analyze the response of pipelines to longitudinal PGD (PGD zone is relatively large, while the amount of PGD is comparatively short)

Spatially Distributed Transverse PGD

O'Rourke (1989) proposed a simple model to analyze the response of pipelines to spatially distributed transverse PGD. In this model O'Rourke considered two types of response (wide and narrow width PGD zones) as shown in Fig. 9 below. In the wide width PGD zone, the pipeline is relatively flexible and its lateral displacement is assumed to closely conform to the soil outline. Accordingly, the pipeline strain is expected to be mainly due to the ground curvature (i.e., displacement controlled). On the other hand, for the narrow width PGD case, the pipeline is relatively stiff and the pipeline lateral

displacement is significantly smaller than that of the soil. Hence, the pipeline strain is anticipated to be due to loading at the pipeline-soil interface (i.e., load controlled).

The maximum bending strain, ε_b, in the pipeline is given by the following:

– *For the wide width PGD zone*

$$\varepsilon_b = \pm \frac{\pi^2 \delta D}{W^2} \qquad (26)$$

– *For the narrow width PGD zone*

$$\varepsilon_b = \pm \frac{p_u W^2}{3\pi E t W^2} \qquad (27)$$

And the average axial tensile strain, ε_a, in the pipeline is estimated by the following:

Seismic Design of Pipelines, Table 1 Yield stress and Ramberg-Osgood parameters for mild steel and X-grade steel

	Grade B	X-42	X-52	X-60	X-70	X-80
Yield stress (MPa)	241	290	359	414	483	552
n	10	15	9	10	5.5	16
r	100	32	10	12	16.6	16

– *For the wide width PGD zone*

$$\varepsilon_a = \left(\frac{\pi}{2}\right)^2 \left(\frac{\delta}{W}\right)^2 \qquad (28)$$

– *For the narrow width PGD zone*

The axial tension in this case is small and neglected.

where δ is the magnitude of the PGD, D is the pipe diameter, W is the length of the PGD zone, P_u is the maximum lateral force per unit length at the pipeline-soil interface, E is the elastic modulus of the pipeline material, and t is the pipe wall's thickness.

Abrupt Transverse PGD

Parker et al. (2008) proposed a simple model to analyze the response of pipelines to transverse PGD. Figure 10 presents the geometric and force details of the model. In this model the width of the abrupt transverse PGD is $2W_1$. Within this width the pipeline is subject to a lateral force per unit length P_{u1}. This lateral load is resisted by soil resistance forces P_{u2} over a distance W_2 on each side of the abrupt transverse PGD zone (see Fig. 10). Therefore, from horizontal equilibrium in the direction of pipeline, we get

Seismic Design of Pipelines, Fig. 7 Adhesion factors (Reproduced after Honegger and Nyman 2004)
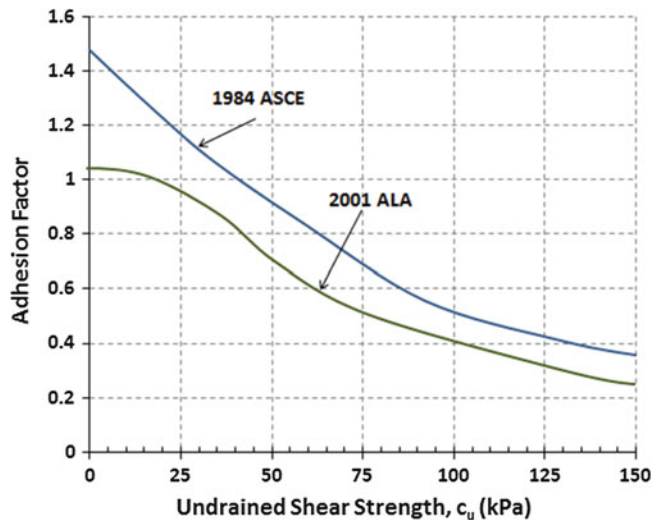
$$P_{u1} W_1 = P_{u2} W_2 \qquad (29)$$

The tensile force in the pipeline is assumed to be a constant value T_o within the PGD zone. Beyond the margins the pipeline axial tension decreases linearly at Points C and E to zero at Points A and G (see Fig. 10).

According to Parker et al. (2008), the total elongation due to the pipeline deformation

Spatially distributed transverse PGD pattern

Abrupt transverse PGD pattern

Seismic Design of Pipelines, Fig. 8 Sketch of the considered patterns

(Point B to Point F) can be estimated using the following equation:

$$\Delta L = \frac{1}{3} P_{u1}^2 W_1^3 \left(1 + P_{u1}/P_{u2}\right) / T_o^2 \qquad (30)$$

Thus, the axial pipe strain is calculated as

$$\varepsilon_a = \frac{\Delta L}{L} \qquad (31)$$

and the bending strain in the pipeline can be evaluated as

$$\varepsilon_b = \frac{D\, P_{u1}}{2\, T_o} \qquad (32)$$

Pipelines in Liquefied Soil

It is common practice that pipelines are buried at depths of 2 m or less from the ground surface. Thus, the top of the susceptible liquefiable soil layer is commonly located below the bottom of the pipeline. However, in some cases when the pipeline is buried at a river bed in saturated sand, for example, the soil surrounding the pipeline may liquefy during strong seismic shaking event. In this case,

Seismic Design of Pipelines, Fig. 9 O'Rourke (1989) model to analyze the response of pipelines to spatially distributed transverse PGD
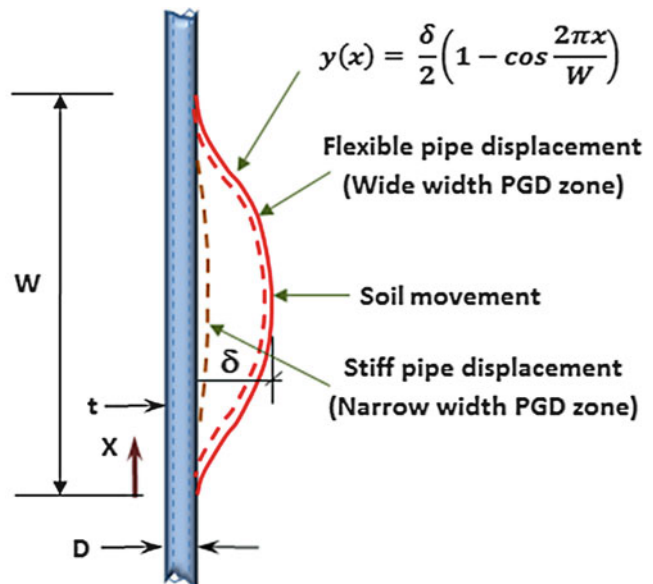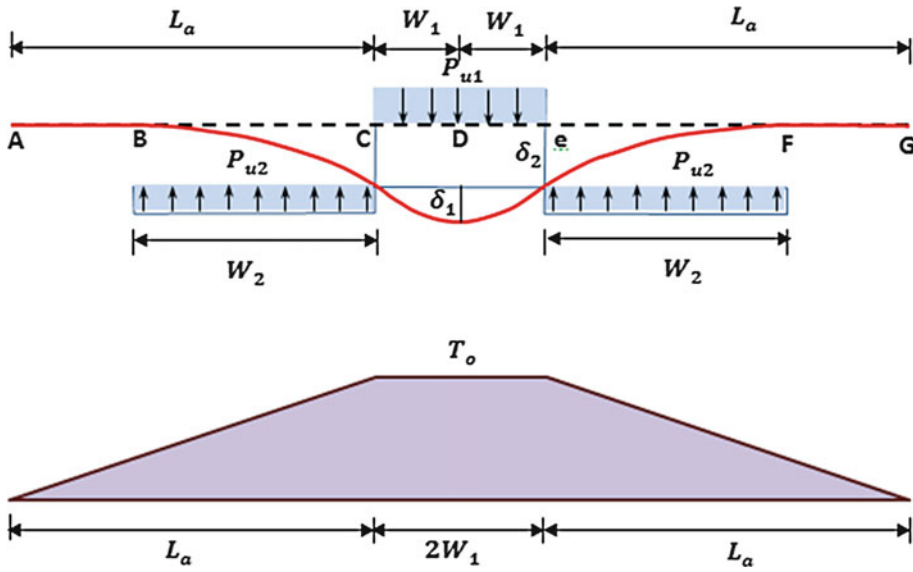
Seismic Design of Pipelines, Fig. 10 Geometric and force details of the Parker et al. (2008) model

the pipeline will probably deform laterally following the flow of the liquefied soil downward a mild slope or move upward due to buoyancy, particularly when something restrains the pipeline at one point or a compressive load acts on the pipeline.

Horizontal Movement

When a pipeline is surrounded by liquefied soil, the pipeline may move laterally due to the flow of liquefied soil downslope. The response of a buried pipe surrounded by liquefied soil subject to spatially distributed transverse PGD can be analyzed using the O'Rourke (1989) method presented earlier.

Vertical Movement

When a pipeline is surrounded by liquefied soil, the pipeline may uplift due to the buoyancy and moves upward. Hou et al. (1990) proposed analytical method to analyze the response of pipelines subjected to vertical movements. According to Hou et al. (1990), the uplifting force per unit length, P_{uplift}, acting on the pipeline within the liquefied zone is given by

$$P_{\text{uplift}} = \frac{\pi D^2}{4} \left(\gamma_{\text{soil}} - \gamma_{\text{contents}} \right) - \pi D t \gamma_{\text{pipe}} \tag{33}$$

where D is the pipe diameter, γ_{soil} is the unit weight of the liquefied soil, γ_{contents} is the unit weight of the contents inside the pipe (water, oil, gas, etc.), γ_{pipe} is the unit weight of the pipe material, and t is the pipe wall's thickness.

The maximum uplift displacement and/or the spacing for pipeline restraints is given by (see Fig. 11 below)

$$A \, \delta_{\max}{}^3 + 16 \, I \, \delta_{\max} - \frac{16 P_{\text{uplift}} W_s{}^4}{E \, \pi^5} \tag{34}$$

and the maximum strain in the pipeline is then given by

$$\varepsilon_{\max} = \pm \frac{\pi^2 \, \delta_{\max} D}{W_s{}^2} + \frac{\pi^2 \, \delta_{\max}{}^2}{4 \, W_s{}^2} \tag{35}$$

where A is the cross-section area, I is the moment of inertia, W_s is the spacing of the restraints, and E is the elastic modulus of the pipe's material.
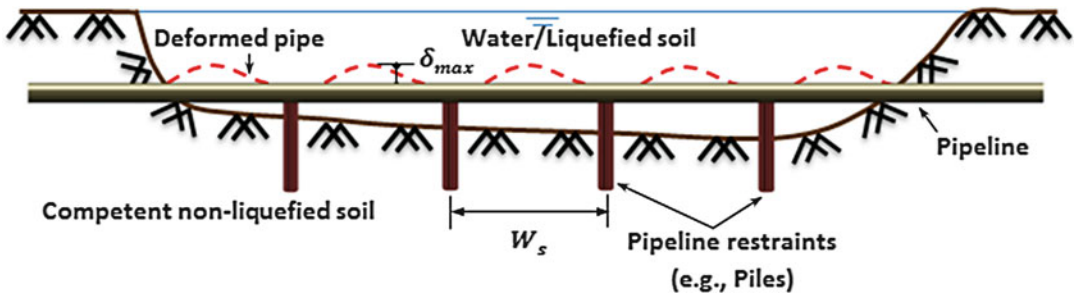
Seismic Design of Pipelines, Fig. 11 Profile of pipeline crossing liquefied zone

Seismic Design of Pipelines, Fig. 12 External forces on the buried pipeline by the adjacent soil during seismic shaking
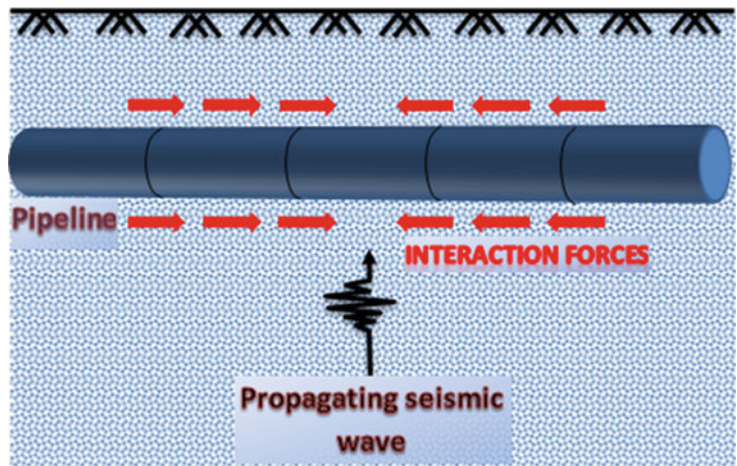
Seismic Design Guidelines and Pipeline-Soil Interaction

The pipeline-soil interaction effect exerts external forces on the buried pipeline by the adjacent soil when seismic motions are provided. The type of soil around the pipeline plays a significant role in its seismic behavior. In general, the soil displacement will produce friction like forces at the pipeline-soil interface (see Fig. 12 below). The overall seismic performance of buried pipeline is strongly related to the force-deformation relationship at the pipeline-soil interface (the p-y curves). For cohesionless soils, the probability of liquefaction becomes higher in loose materials. In cohesive soils, softer soils will undergo greater differential settlement due to consolidation and higher amplification effects and accordingly greater interaction forces.

Pipelines in Competent Non-liquefied Soil

The 1984 ASCE Guidelines suggest for the purpose of analysis idealized elastoplastic models for the force-deformation relationship at the pipeline-soil interface (see Fig. 13 below).

Longitudinal Movement

In this case the relative movement is parallel to the pipeline which results in axial forces at the pipeline-soil interface. The 1984 ASCE Guidelines provide relations for both cohesionless and cohesive soils. For cohesionless materials, the longitudinal resistance is due to the friction in the longitudinal direction at the pipeline-soil interface. The normal pressure which leads to the axial friction is the overburden and the lateral soil pressures. In the 1984 ASCE Guidelines, the normal pressure is taken as the average of the vertical and at rest lateral soil pressures acting

S
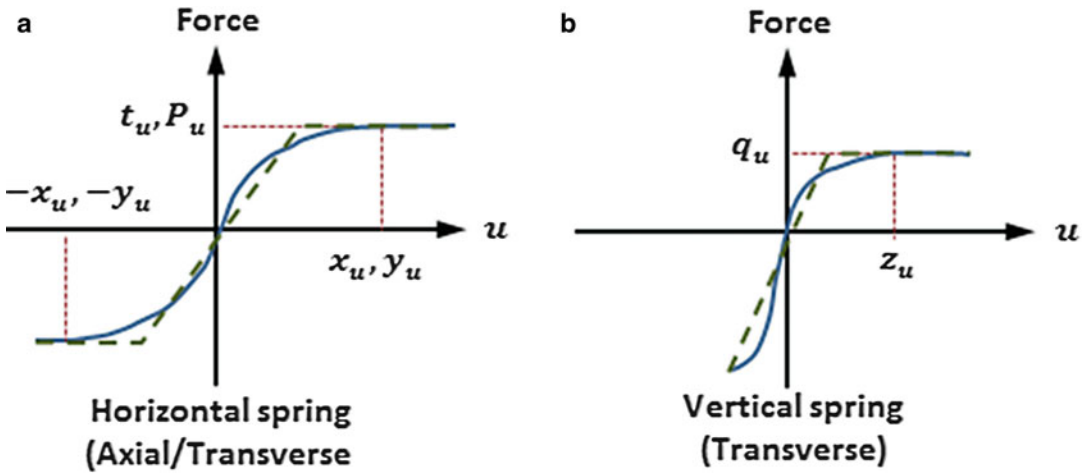
Seismic Design of Pipelines, Fig. 13 Idealized elastoplastic models for the force-deformation relationship at the pipeline-soil interface (Reproduced after ASCE 1984)

on the pipeline. (The 1984 ASCE Guidelines assumes $X_u \cong 2.5 \sim 5.0$ mm).

$$t_u = \pi D \gamma H \left(\frac{1 + k_o}{2} \right) \tan k\varphi \qquad (36)$$

For cohesive materials, the longitudinal resistance is proportional to the adhesion at the pipeline-soil interface:

$$t_u = \pi D \alpha c_u \qquad (37)$$

where D is the pipe diameter, γ is the effective unit weight of the soil, H is the depth of the pipeline, φ is the angle of shear resistance, k_o is the coefficient of lateral soil pressure at rest, k is a friction factor, a is the adhesion factor (given in Fig. 7), and c_u is the undrained shear strength of the soil.

Transverse-Horizontal Movement

In this case the relative movement is perpendicular to the pipeline which results in transverse-horizontal forces at the pipeline-soil interface. The 1984 ASCE Guidelines provide relations for both cohesionless and cohesive soils. For cohesionless materials, the maximum soil resistance in horizontal transverse direction may be calculated using the following equation:

$$P_u = \gamma H D N_{qh} \qquad (38)$$

and the maximum elastic relative displacement in horizontal transverse direction is

$$y_u = \begin{cases} (0.07 \sim 0.10)(H + D/2) \text{ for loose sand} \\ (0.03 \sim 0.05)(H + D/2) \text{ for medium sand} \\ (0.02 \sim 0.03)(H + D/2) \text{ for dense sand} \end{cases} \qquad (39)$$

For cohesive materials, the maximum soil resistance in horizontal transverse direction may be evaluated using the following equation:

$$P_u = c_u N_{ch} D \qquad (40)$$

The maximum elastic relative displacement in horizontal transverse direction is

$$y_u = (0.03 \sim 0.05)(H + D/2) \qquad (41)$$

where γ is the unit weight of the soil, H is the embedment depth of the pipeline, and N_{qh} and N_{ch} are the horizontal bearing capacity factors for sand and clay, respectively (Fig. 4).

Pipelines in Liquefied Soil

The response of continuous pipelines buried in liquefied soil layer is very sensitive to the

stiffness of the soil (Suzuki et al. 1988; Miyajima and Kitaura 1989). Based on several experimental results, it is recommended that the stiffness of liquefied soil ranges from 1/100 to 3/100 of that for non-liquefied soil (Yoshida and Uematsu 1978; Matsumoto et al. 1987; Yasuda et al. 1987; Tanabe 1988). Accordingly, the reduced stiffness will be used at the pipeline-soil interface. Analysis is then performed using the same procedures as in the competent non-liquefied soil case (on the conservative side).

Summary

This chapter focused on the seismic analysis and design of pipelines. In seismic events, buried pipelines can be damaged mainly by either the permanent ground deformations (PGD) or by the transient seismic wave propagation. Hence, different analysis methods for pipelines subjected to permanent ground deformation (PGD) or transient seismic wave propagation hazards were presented and discussed in details. In addition, the response of continuous pipelines subject to fault offsets was discussed. Also, several methods were introduced to predict the response of pipelines to either longitudinal or transverse PGD.

Earthquake shaking may trigger the liquefaction of saturated loose cohesionless soils. The liquefaction process increases the pore water pressure in the ground to a level where the effective stress approaches zero at which point the soil loses entirely its shear strength. A separate section in this chapter presented and discussed methods of analyzing and designing pipelines buried in liquefiable soils.

The overall seismic performance of buried pipelines is strongly related to the pipeline-soil interaction. Consequently, this chapter dedicated a section to provide guidelines for the seismic design of pipelines considering the pipeline-soil interaction effects as it plays a significant role in its seismic behavior.

Cross-References

▶ Earthquake Magnitude Estimation
▶ Earthquake Mechanisms and Tectonics
▶ Earthquake Response Spectra and Design Spectra
▶ Liquefaction: Performance of Building Foundation Systems
▶ Seismic Vulnerability Assessment: Lifelines
▶ Soil-Structure Interaction

References

American Society of Civil Engineers (ASCE) (1984) Guidelines for the seismic design of oil and gas pipeline systems. Committee on Gas and Liquid Fuel Lifelines, ASCE

Chen WW, Shih BJ, Wu CW, Chen YC (2000) Natural gas pipeline system damages in the Ci Ci earthquake (The City of Nantou). In: Proceedings of the 6th international conference on seismic zonation, Palm Springs, California.

Hansen JB (1961) The ultimate resistance of rigid piles against transversal forces. Bulletin 12, Danish Geotechnical Institute, Copenhagen

Honegger, D.G. and Nyman, D. (2004). PRCI guidelines for the seismic design and assessment of natural gas and liquid hydrocarbon pipelines. Pipeline Research Council International, Published by Technical Toolboxes, Houston, Texas

Hou Z, Cai J, Liu X (1990) Response calculation of oil pipeline subjected to permanent ground movement induced by soil liquefaction. In: Proceedings of the China-Japan symposium on lifeline earthquake engineering, Beijing, pp 107–114

Jibson RW, Keefer DK (1993) Analysis of the seismic origin of landslides: examples from the new Madrid seismic zone. Geol Soc Am Bull 105:521–536

Kennedy RP, Chow AW, Williamson RA (1977) Fault movement effects on buried oil pipeline. J Transp Eng Div ASCE 103(TE5):617–633

Lau DL, Tang A, Pierre J-R (1995) Performance of lifelines during the 1994 Northridge earthquake. Can J Civ Eng 22:438–451

Matsumoto H, Sasaki Y, Kondo M (1987) Coefficient of subgrade reaction on pile in liquefied ground. In: Proceedings of the second national conference on soil mechanics and foundation engineering, Tokyo, Japan, pp 827–828 (in Japanese)

Miyajima M, Kitaura M (1989) Effects of liquefaction-induced ground movement on pipeline. In: Proceedings of the Second U.S.-Japan workshop on liquefaction, large ground deformation and their effects on lifelines, Buffalo, Technical report NCEER-89-0032,

S

Multidisciplinary Center for Earthquake Engineering Research, Buffalo, pp 386–400

Newmark NM (1965) Effects of earthquakes on dams and embankments. Geotechnique 15(2):139–160

Newmark NM (1967) Problems in wave propagation in soil and rocks. In: Proceedings of the international symposium on wave propagation and dynamic properties of earth materials, University of New Mexico Press, pp 7–26

O'Rourke MJ (1989) Approximate analysis procedures for permanent ground deformation effects on buried pipelines. In: Proceedings of the second U.S.-Japan workshop on liquefaction, large ground deformation and their effects on lifelines, Buffalo. Technical report NCEER-89-0032, Multidisciplinary Center for Earthquake Engineering Research, Buffalo, pp 336–347

O'Rourke TD, Palmer MC (1996) Earthquake performance of gas transmission pipelines. Earthq Spectra 12(3):493–527

O'Rourke TD, Grigoriu MD, Khater MM (1985) Seismic response of buried pipes. Pressure vessel and piping technology – a decade of progress, ASME, pp 281–323

O'Rourke MJ, Liu XJ, Flores-Berrones R (1995) Steel pipe wrinkling due to longitudinal permanent ground deformation. J Transp Eng 121(5):443–451

Parker E, Moore R, Evans T, Usher N (2008) Evaluation of landslide impact on deepwater submarine pipelines. OTC 19459, Offshore Technology Conference, Houston

Scawthorn C, Yanev PI (1995) Preliminary report 17 January 1995, Hyogoken Nambu, Japanese earthquake. Eng Struct 17(3):14657

Suzuki H (1988) Damage to buried pipes caused by large ground displacement. In: Proceedings of the first Japan-U.S. workshop on liquefaction, large ground deformation and their effects on lifeline facilities, Tokyo, pp 127–132

Takada S, Tanabe K (1988) Estimation of earthquake induced settlements for lifeline engineering. In: Proceedings of the ninth world conference earthquake engineering, vol VII, pp 109–114

Tanabe K (1988) Fundamental study on seismic assessment and design of buried pipelines subjected to ground failure during earthquake. Doctoral dissertation, Kobe University (in Japanese)

Wells DL, Coppersmith KJ (1994) New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bull Seismol Soc Am 84(4):974–1002

Wilson RC, Keefer DK (1983) Dynamic analysis of a slope failure from the 6 August 1979 Coyote Lake, California earthquake. Bull Seismol Soc Am 73(3):863–877

Yasuda S, Saito K, Suzuki N (1987) Soil spring constant on pipe in liquefied ground. Proceedings of the nineteenth JSCE conference on earthquake engineering, Tokyo, Japan, pp 189–192 (in Japanese)

Yoshida T, Uematsu M (1978) Dynamic behavior of a pile in liquefaction sand. Proceedings of the fifth Japan earthquake engineering symposium, Tokyo, Japan, pp 657–663 (in Japanese)

Youd TL, Hansen C, Bartlett S (2002) Revised multilinear regression equations for prediction of lateral spread displacement. J Geotech Geoenviron Eng ASCE 128(12):1007–1017

Zhang G, Robertson PK, Brachman RWI (2004) Estimating liquefaction-induced lateral displacements using the standard penetration test or cone penetration test. J Geotech Geoenviron Eng ASCE 130:861–871

Seismic Design of Tunnels

Youssef M. A. Hashash[1] and
Maria Ines Romero-Arduz[2]
[1]Department of Civil and Environmental Engineering, 2230c Newmark Civil Engineering Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[2]Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Synonyms

Cut-and-cover structures; Earthquakes; Free-field deformations; Immersed-tube tunnels; Nonlinear response; Racking; Soil-structure interaction; Tunnels

Introduction

The complete enclosure in soil or rock makes their seismic behavior different than that of aboveground structures or superstructures. Underground structure seismic response is constrained by seismic response of the surrounding soil and cannot experience free vibrations as is the case for aboveground structures. This section focuses on the seismic analysis and design of large linear underground structures commonly used for metro structures, highway tunnels, and large water/sewage transportation ducts in urban areas and can be grouped into three broad

Cut-and-cover (rectangular)   Cut-and-cover (vertical tubes)   Cut-and-cover (center columns or wall)

Immersed Tube Tunnels (Single and multiple tubes)   Bored Tunnel (Circular)
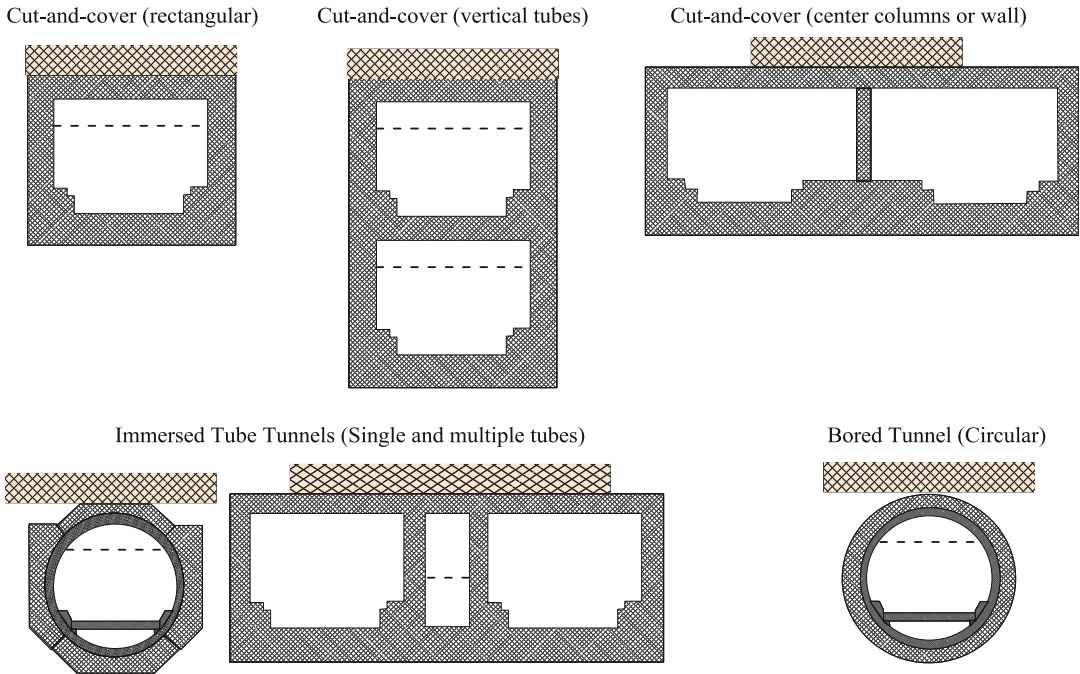
Seismic Design of Tunnels, Fig. 1   Typical cross sections of underground structures

categories: (1) bored or mined tunnels, (2) cut-and-cover tunnels, and (3) immersed tunnels (Fig. 1).

The section starts with a selected review of performance of underground structures during seismic events. This is followed by presentation of a performance-based framework for design and analysis of underground structures considering both permanent and transient deformations. A number of seismic design issues are then discussed including vertical ground shaking and response, interaction of temporary and permanent structures, impact of superstructure and adjacent structures, transitions and tunnel joints, seismic retrofit of existing facilities, design considerations for structural support members, and precast tunnel lining.

Performance of Underground Facilities During Seismic Events

Based on several studies that documented earthquake damage to underground facilities in the past and the behavior of underground facilities

in recent large magnitude earthquakes (e.g., Tohoku, Japan, 2011; Maule, Chile, 2010), underground structures suffer appreciably less damage than surface structures. However, damage or failure of a limited section of an underground structure can be disruptive to post-earthquake recovery or operation in densely populated urban areas as such damage can interrupt the function of an entire system whether it is part of a mass transit or vehicular transportation network or large water or sewage transportation tunnels.

Damage is related to a number of parameters including ground motion intensity, ground conditions, and structural support system. Shallow tunnels tend to be more vulnerable to earthquake shaking than deep tunnels, and those constructed in soil can undergo more deformation than those constructed in competent rock. Circular bored tunnels are less susceptible to earthquake damage than cut-and-cover tunnels. Shaking damage can be reduced by stabilizing the ground around the tunnel and by improving the contact between the lining and the surrounding soil using grouting. Stiffening the lining without stabilizing the

S

Seismic Design of Tunnels, Fig. 2 Street view of Dakai subway station collapse (Iida et al. 1996)

surrounding poor ground may only result in excessive seismic forces in the lining. Damage at tunnel portals may be caused by slope instability (Hashash et al. 2001).

Underground Structures in the United States

The Bay Area Rapid Transit system (BART) in San Francisco, California, sustained the 1989 Loma Prieta earthquake without damage and was operational after the earthquake. It consists of underground stations and tunnels embedded in soft bay mud deposits connected to Oakland via the transbay-immersed-tube tunnel. It was one of the first underground facilities designed with seismic considerations: special seismic joints were designed to accommodate differential movements. Limited displacements were measured at that joint.

During the same earthquake, the Alameda Tubes, a pair of immersed-tube tunnels that connect Alameda Island to Oakland in California, experienced structural cracking on the ventilation buildings and limited water leakage due to liquefaction of loose deposits above the tubes.

The 1994 Northridge earthquake caused no damage to concrete lining of bored tunnels of the metro system in Los Angeles, California (Hashash et al. 2001).

Underground Structures in Japan

The Dakai subway collapse (Fig. 2) during the 1995 Hyogoken-Nambu earthquake in Kobe, Japan, was the first collapse of an urban underground structure due to earthquakes shaking. The collapse experienced by the center concrete columns (Fig. 3) was due to lack of shear reinforcement, leading to a collapse of the ceiling slab and settlement of the soil cover. The 1962 station design did not include specific seismic provisions. However, in the 2011 Tohoku earthquake, the underground subways in Sendai experienced strong shaking with no reports of damage. Only a water distribution plant in Kashima City and a wastewater treatment plant in Itako City were observed to have been damaged by liquefaction. The damages include uplift of buried tanks, offsets in underground tunnels, damage to support utilities, and damage to major trunk lines on and off the site (Ashford et al. 2011).

Underground Structures in Taiwan

The 1999 Chi-Chi earthquake did not cause visible signs of damage in several highway tunnels located in central Taiwan. The main damage occurred at tunnel portals because of slope instability. No damage was reported in the Taipei subway, located 100 km from the ruptured fault
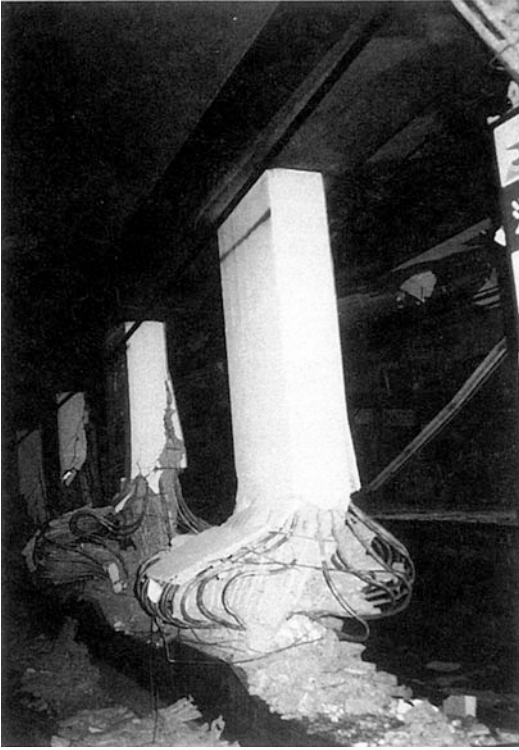
Seismic Design of Tunnels, Fig. 3 Dakai subway station collapse (Iida et al. 1996)

zone (Fig. 4). However, some tunnels in mountainous areas were severely damage due to slope failure (Fig. 5).

Underground Structures in Turkey
The August 17, 1999, Kocaeli earthquake had minimal impact on the Bolu twin tunnels, a 1.5-billion-dollar project under construction at that time. It had an excavated section of 15 m tall by 16 m wide and crossed several minor faults parallel to the North Anatolian Fault. After the earthquake, continuous monitoring showed no movement due to the earthquake. The November 12, 1999, earthquake caused collapse of both tunnels 300 m from its eastern portal, in a clay gauge material in the unfinished section (Hashash et al. 2001).

Underground Structures in Chile
Cut-and-cover highway box structures with three lanes of traffic in one direction and approximately 1 km long constructed in relatively stiff

gravels, the Santiago Metro tunnels and underground stations, and highway tunnels on the southbound of Route 5 appeared to be undamaged by the 2010 Maule earthquake in Chile, as shown in Fig. 6. In the northbound of Route 5, "La Calavera" tunnel near Calera has a rock block dislodged, but the tunnel was old and had problems before the earthquake (Elnashai et al. 2010).

In summary, well-engineered underground structures performed well even under strong shaking in recent earthquakes in different parts of the world. However, underground structures are vulnerable to permanent ground displacements such as liquefaction, slope stability, and fault displacement. There is also vulnerability to transient ground motions when insufficient structural detailing is provided, the underground structure is constructed in loose ground, masonry lining is used, or near field effects are present.

Performance-Based Seismic Evaluation Framework

Underground structures under earthquake effects can undergo permanent deformations and/or transient deformations. Factors influencing these effects include shape, dimensions, and depth of the structure, properties of the surrounding soil or rock, properties of the structure, and severity of ground shaking (Hashash et al. 2001). Table 1 summarizes a performance-based framework for seismic design and analysis of underground structures. The framework consists of three main steps: definition of seismic environment, evaluation of ground response to shaking, and assessment of structure response due to seismic shaking.

Step 1: Definition of Seismic Environment
Seismic analysis of underground structures starts with site-specific definition of its seismic environment. A detailed field and laboratory investigation program is necessary; the field investigation program should include definition of the site stratigraphy and direct measurements of shear wave velocity profiles and cone

S

Seismic Design of Tunnels, Fig. 4 Slope failure at tunnel portal, Chi-Chi earthquake, central Taiwan (Hashash et al. 2001)
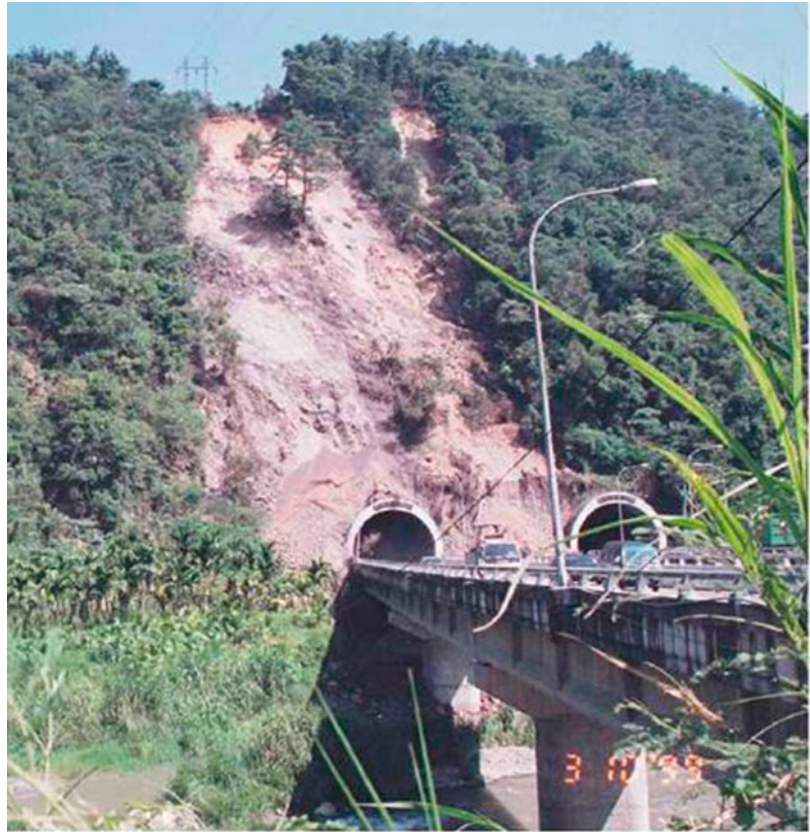
penetration resistance of soft soils as well as assessment of potential geo-hazards including slope instability, fault displacement, and lateral spreading. Appropriate static and cyclic laboratory tests for major soil units are also required.

Site-specific probabilistic and/or deterministic seismic hazard analyses, as well as hazard analyses using conditional (mean) spectra, are needed to define seismic hazard levels for permanent condition of the structure (operational and maximum levels). Increasingly seismic hazard, using a shorter return period, is being considered for temporary conditions during construction.

Seismic performance criteria selection is a crucial aspect in the design of an underground structure. Performance objectives include explicit target performance of the structure and system performance. Performance objectives are not purely technical requirements and should include owner and user requirements, policy considerations, and life-cycle costs. It is an iterative process

based on analysis findings to answer the question of what is feasible and at what cost as illustrated in Fig. 7.

Often a two-level criterion is adopted: operating design earthquake (ODE) and maximum design earthquake (MDE). Those are defined using response spectra developed in the seismic hazard analysis. A suite of three component motions is needed for each of the design earthquake levels for site response analysis and soil-structure interaction modeling. It is preferred to use recorded motions instead of synthetic motions to spectrally match the target spectra. Ground motion spatial incoherence must be taken into account for long structures including (1) wave passage, (2) extended source effects, (3) ray-path effects, and (4) local site effects. One-dimensional equivalent linear and nonlinear site response analyses are conducted to assess how the ground motion is affected by the soil column. One-dimensional site response
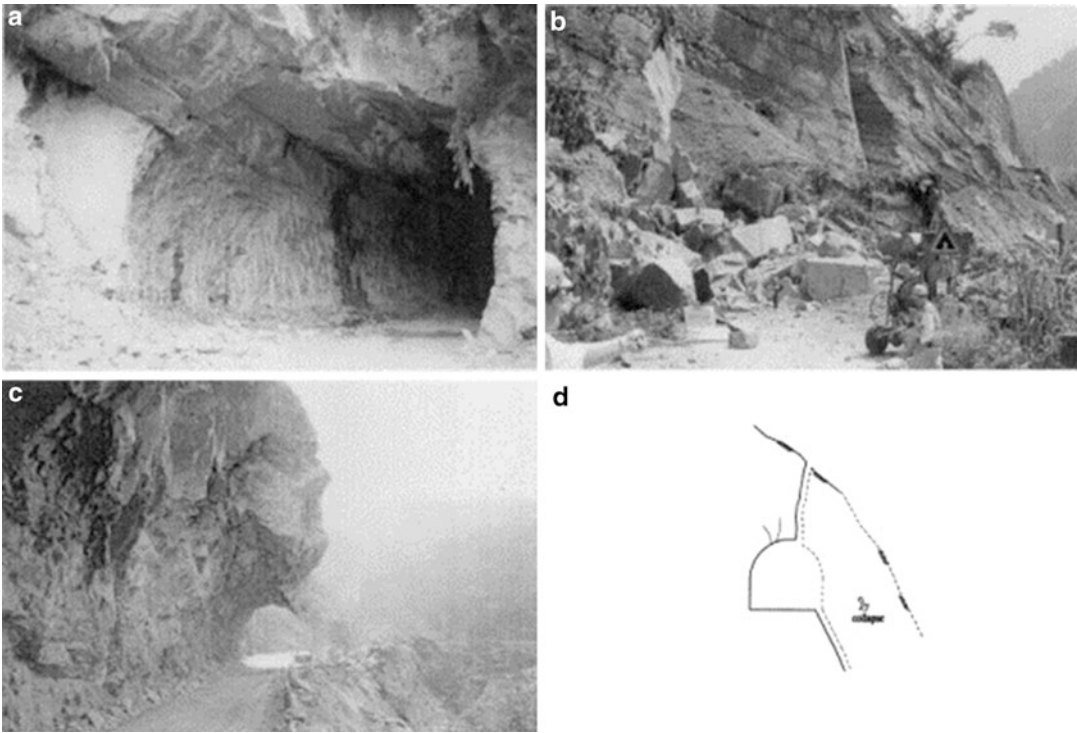
Seismic Design of Tunnels, Fig. 5 Chi-Shue tunnel before and after Chi-Chi earthquake (Wang et al. 2001)

Seismic Design of Tunnels, Fig. 6 *Left*: Highway box structures in Santiago, *Right*: Highway tunnel in Route 5 South

Seismic Design of Tunnels, Table 1 Performance-based framework for design and analysis of underground structures

Step 1: Definition of Seismic Environment	Step 2: Evaluation of Ground Response to Shaking	Step 3: Assessment of structure behavior due to seismic shaking
Field and laboratory investigations	Permanent deformations • Liquefaction • Slope Instability • Fault Displacement	Seismic Design Loading Criteria • Loading Criteria for MDE • Loading Criteria for ODE
Seismic hazard analysis • Deterministic Seismic Hazard Analysis (DSHA) • Probabilistic Seismic Hazard Analysis (PSHA)	Transient deformations • Longitudinal Extension/Compression • Longitudinal Bending • Racking/Ovaling	Underground Structure Response to Ground Deformations • Free-Field Deformation Approach • Soil-structure Interaction Approach
Seismic performance criteria • Maximum Design Earthquake (MDE) • Operating Design Earthquake (ODE)		Additional Seismic Design Issues
Seismic input motions • Acceleration, Velocity, Displacement Amplitudes • Target Response Spectra, Motion Time History • Spatial Incoherence of Ground Motion		
Site response analysis		

analyses are used in the analysis of underground structures to:

(a) Obtain free-field racking deformations (differential sideways movements between the top and bottom elevations of rectangular structures) along the underground structure height which can be used in pseudo-static soil-structure interaction.
(b) Obtain input motions for dynamic soil for dynamic soil-structure interaction analysis.
(c) Obtain strain-compatible soil properties for use in pseudo-static and dynamic soil-structure interaction.
(d) Assess potential liquefaction and ground failure.

Step 2: Evaluation of Ground Response to Shaking

Evaluation of ground response to shaking can be divided into permanent deformations or ground failure and transient deformations or ground shaking.

Permanent deformations or ground failure includes liquefaction, slope instability, and fault displacement. Liquefaction, prevalent in loose sand and fill deposits, can result in generation of sand boils, loss of shear strength, lateral spreading, and slope failure. Tunnels in liquefiable deposits can experience increased lateral pressures, loss of lateral passive resistance, flotation or sinking, lateral displacements if lateral spreading happens, permanent settlement, and compression/tension failure after soil consolidation. A landslide intercepting a tunnel can result in concentrated shearing displacements and collapse of a cross section. The potential for these failures is greatest when a pre-existing landslide intersects the tunnel, in shallower parts of tunnel, and at tunnel portals. An underground structure may have to pass across an active fault zone; in these situations the tunnel must tolerate the expected displacements. The design for permanent deformations is discussed in section "Design for Permanent Deformations."

Transient deformations can be quite complex due to interaction of seismic waves with surficial
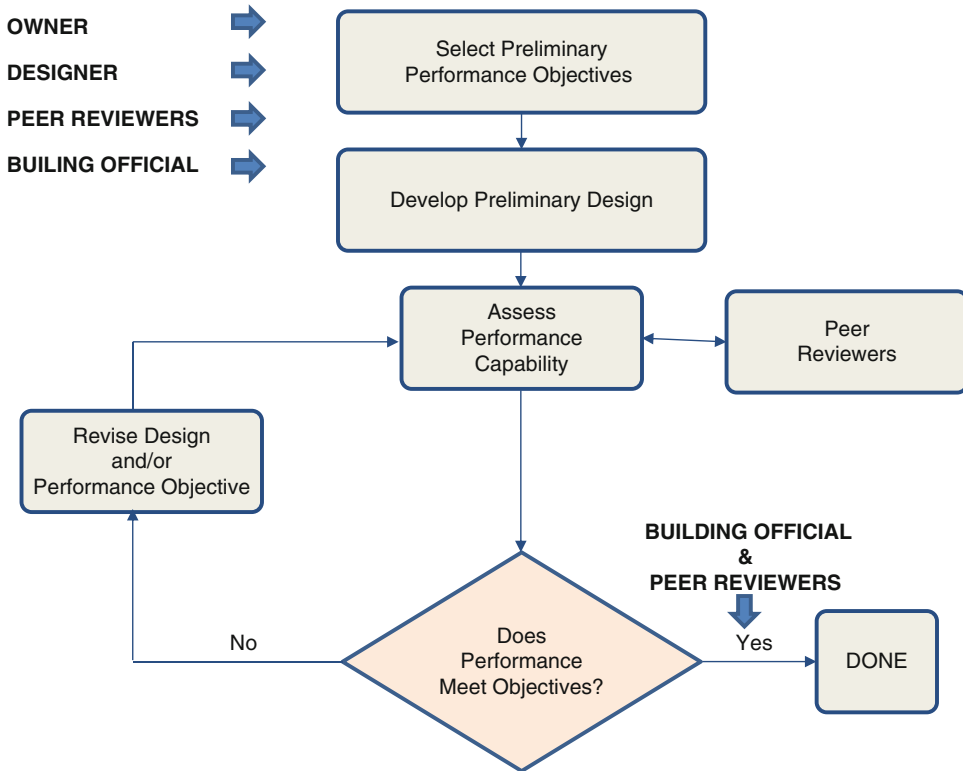
Seismic Design of Tunnels, Fig. 7 Performance-based design procedure after Hamburger and Hooper (2011)

deposits. Underground structures undergo three primary modes of deformation during seismic shaking: compression-extension, longitudinal bending, and ovaling/racking. The design for transient deformations is discussed in section "Performance Evaluation Under Transient Ground Deformations."

Step 3: Assessment of Structure Behavior Due to Seismic Shaking

The evaluation of structure behavior will be primarily a deformation controlled soil-structure interaction problem. Pseudo-static and dynamic soil-structure interaction approaches have been used in the evaluation of structure response and are discussion in section "Performance Evaluation Under Transient Ground Deformations."

Section "Additional Seismic Performance Issues" discusses additional seismic design issues such as vertical ground shaking and response;

interaction of temporary and permanent structures; permanent changes in state of stress of soil; impact of superstructure and adjacent structures; tunnel joints such as portals, stations, and tunnel segment; seismic retrofit of existing facilities; design considerations for structural support members; precast tunnel lining; and seismic design of buried reservoirs.

Design for Permanent Deformations

Designing underground structures for permanent deformations may not be viable, but ground stabilization techniques can help prevent large deformations. Some solutions include ground improvement, drainage, soil reinforcement, grouting, earth retaining systems, or even removing problematic soils or relocating the tunnel alignment.

Underground Structures Crossing Active Faults

In the Century Area Tunneling Safety and Fault Investigations TAP report for the Los Angeles Metro (Cording et al. 2011), some recommendations are provided for consideration when an underground tunnel segment crosses an active fault:

– The segments must be designed to accommodate fault displacement without collapse and with the capability of being repaired.
– The alignments should be selected so that the tunnel crosses at a relatively sharp angle to the fault zone to minimize the length of tunnel that must accommodate fault displacements.
– Methods employed to allow fault displacement on a tunnel lining include excavating to a larger section at the crossing to facilitate realignment, providing the tunnel with a strong but flexible lining like ductile steel segments or articulated joints, and placing crushable backpacking material around the structural lining.
– As there is no precedent of placing metro underground stations on active faults, designing for it will be extremely difficult and cost-prohibitive. If possible, it is advisable to avoid this scenario.

The design approach for tunnels crossing active faults will change depending on the displacement magnitude and the tunnel width over which the displacement is distributed:

(a) If fault displacements are small and/or distributed over a relatively wide zone, providing articulation of the tunnel lining through ductile joints is a possible solution. The closer the joint spacing, the better the tunnel performance will be; this is more viable in soft soils where displacements can be effectively redistributed over the tunnel lining. The tunnel can then deform in an S-shape through the fault zone without rupture. It is always necessary to keep the tunnel watertight when using joints.

(b) If large displacements are concentrated in a narrow zone, retrofit will consist of enlarging the tunnel section across and beyond the displacement zone. The length over which the enlargement is made is a function of fault displacement and permissible curvature of the road or track; the longer the enlarged tunnel, the smaller the post-earthquake curvature (Power et al. 1996). This solution has been implemented in the San Francisco BART system and Los Angeles Metro rapid transit tunnel system. Concrete-encased steel ribs provide sufficient ductility to accommodate distortions with little strength degradation. Under axial displacements, even though compression is more damaging to the tunnel lining than extension, both will result in unacceptable water inflow. A solution for water tightness is flexible couplings (Wang 1993), used for the Southwest Ocean Outfall in San Francisco. Cellular concrete may also be used within the enlarged tunnel, because it has a low yield strength that can minimize the loads on the tunnel liner while also providing adequate resistance for normal soil pressures and other seismic loads.

Estimating fault displacement is a key issue to design tunnels crossing active faults. One option to estimate fault displacement is using empirical relationships that express expected displacements in terms of some source parameter. Deterministic and probabilistic fault displacement hazard analyses can be used to assess fault displacement hazard where a displacement attenuation function is used in a probabilistic seismic hazard analysis (Coppersmith and Youngs 2000; Youngs et al. 2003).

Flotation in Liquefiable Deposits

Liquefaction evaluation is discussed elsewhere. If liquefaction is limited to soil layers above the underground structure, then it is unlikely to influence the racking deformations of the structure. However, if the structure is partially or entirely embedded in liquefiable soil, additional evaluations are required. Underground structures may experience flotation in liquefiable deposits.
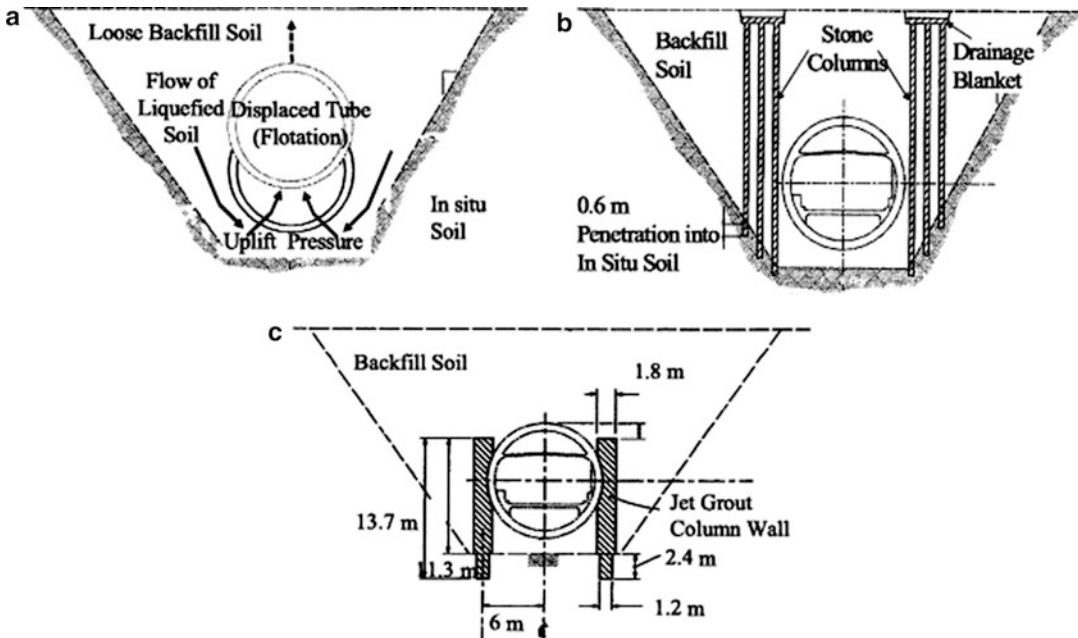
Seismic Design of Tunnels, Fig. 8 Isolation principle, use of cutoff walls to prevent tunnel uplift (Schmidt and Hashash 1999)

As shown in Fig. 8a, when the tunnel experiences uplift due to flotation, the liquefied soil moves underneath the displaced tunnel and lifts it further up (Schmidt and Hashash 1999). Uplift can be prevented through isolation using cutoff walls, such as sheet pile walls; stone columns (Fig. 8b); or jet grout columns (Fig. 8c). Sheet piles with drainage capability can also reduce excess pore water pressure. The rise in excess pore water pressure is prevented at the bottom of the tunnel and in the soil underneath with these barrier walls. With longer barrier walls and a wider structure, uplift is more difficult. After the liquefaction potential is mitigated, flexible joints can be used to allow for differential displacements at tunnel connection joints.

Slope Instability and Lateral Spreading

Stabilizing the soil or changing the alignment is often the most practical way to mitigate slope instability and lateral spreading. It is challenging to design an underground structure to resist or take these movements unless the hazard is localized and the movement is small (Power et al. 1996).

Performance Evaluation Under Transient Ground Deformations

If permanent deformations are not expected, then the underground structure must be designed for transient ground deformations. In this case, the underground structure response is controlled by the ground deformation and the peak ground velocity.

The focus of underground structure seismic design is on free-field deformations of the ground and their interaction with the structure, since the inertia of the surrounding soil is large relative to the inertia of the structure. Figure 9 shows response of underground structures to seismic motions: axial compression and extension, longitudinal bending, and ovaling/racking. Axial deformations are due to seismic waves producing motions parallel to the tunnel axis, bending is due to seismic waves producing particle motion perpendicular to the longitudinal axis, and ovaling/racking is due to shear waves propagating normally to the tunnel axis. Design considerations for axial and bending deformations are generally in the direction along

S

a Axially-propagating wave
Compression - extension
Wave parallel to tunnel axis

b Diagonally-propagating wave
Compression - extension

c Longitudinal bending
Particle motion perpendicular to longitudinal axis

d Normally-propagating shear wave
Ovaling of tunnel section
Shear wave nearly normal to tunnel axis

e Normally-propagating shear wave
Racking of tunnel section
Shear wave nearly normal to tunnel axis

f Normally-propagating wave
Compression of tunnel section
Wave parallel to tunnel axis

Seismic Design of Tunnels, Fig. 9 Deformation modes of tunnels due to seismic waves

the tunnel axis and in the transverse direction for ovaling/racking.

There are four main approaches to tackle the design for transient deformations: free-field deformation methods, pseudo-static soil-structure interaction analyses, dynamic soil-structure interaction finite-element analyses, and dynamic earth pressure methods. The merits and drawbacks of these methods are summarized in Table 2.

Free-field deformation methods assume that the underground structure deformations are identical to those of the surrounding ground. They do not take into account soil-underground structure interaction and are most appropriate when the structure (racking) stiffness is equivalent to that of the surrounding ground.

Pseudo-static soil-structure interaction models account for the kinematic interaction between the soil and the underground structure neglecting inertial interaction. They are often used for practical design purposes when the structure is not too complex (NCHRP 611, Anderson et al. (2008)).

Nowadays, the ease of access to high performance computers makes it possible to perform dynamic soil-structure interaction analyses within a reasonable amount of time. These types of analyses allow problems with complicated tunnel geometry and ground conditions to be solved efficiently. However, the selection of parameters for a complex problem requires expertise; therefore, it is important to always verify the computer model solution with simpler pseudo-static or closed-form solutions.

The presence of a rectangular frame structure in the ground will induce dynamic earth pressures acting upon the structure. Complex shear and normal stress distributions along the exterior surfaces of the structures are expected, but quantifying those distributions require rigorous dynamic soil-structure interaction, since they heavily depend on how the interface is modeled.

In the past, the Mononobe-Okabe method was used to calculate the seismic-induced dynamic earth pressures on underground structures. The method assumes the earthquake load is caused by inertial forces of the surrounding soil and calculates the load using soil properties and a determined seismic coefficient. This method is not applicable in the case of underground structures,

Seismic Design of Tunnels, Table 2 Comparison of seismic design approaches

Approaches	Advantages	Disadvantages	Applicability
Free-field deformation methods	1. Comparatively easy to formulate, many 1D wave propagation programs available	1. Nonconservative for tunnel structure more flexible than ground 2. Conservative for tunnel structure stiffer than ground 3. Overly conservative for tunnel structures significantly stiffer than ground 4. Less precision with highly variable ground conditions	For tunnel structures with equal stiffness to ground
Pseudo-static soil-structure interaction methods	1. Good approximation of soil-structure interaction 2. Comparatively easy to formulate 3. Reasonable accuracy in determining structure response 4. Computationally efficient 5. Sensitivity analysis can be easily performed	1. Ignores inertial effects 2. Less precision with highly variable ground 3. Shear displacement not transmitted uniformly to shallow box structures	Most conditions except for variable soil profile, shallow structures
Dynamic soil-structure interaction finite-element analysis	1. Best representation of soil-structure system 2. Best accuracy in determining structure response 3. Capable of solving problems with complicated tunnel geometry and ground conditions (significant variations in soil stiffness)	1. Computationally demanding 2. Uncertainty of design seismic input parameters may be several times the uncertainty of the analysis	All conditions
Dynamic earth pressure methods	1. Serve as additional safety measures against seismic loading	1. Lack of rigorous theoretical basis 2. Resulting in excessive deformations for tunnels with significant burial 3. Use limited to certain types of ground properties	None

since they will move with the ground and will not form an active wedge.

When designing underground structures for transient deformations, sufficient ductility is needed to absorb imposed deformations without losing the capacity to carry static loads. Care should be exercised in not increasing the stiffness of the structure as this tends to attract additional loads thus increasing the demand on the structure.

Free-Field Deformation Approach

Free-field deformations are the deformations caused by seismic waves on a given soil profile in the absence of structures or excavations. The interaction between the soil and the underground structure is neglected, but provides a first-order estimate of the underground structure deformation. Imposing the free-field deformations directly on the underground structure can underestimate or overestimate the structure deformations.

Closed-Form Elastic Solutions

Initial estimates of strains and deformations in a tunnel can be developed using simplified closed-form solutions. There are many assumptions within the formulation of these methods:

(a) The seismic wave field is considered as a plane of wave with the same amplitude at all locations along the tunnel, differing only in their arrival time. Wave scattering and 3D wave propagation are neglected. The results
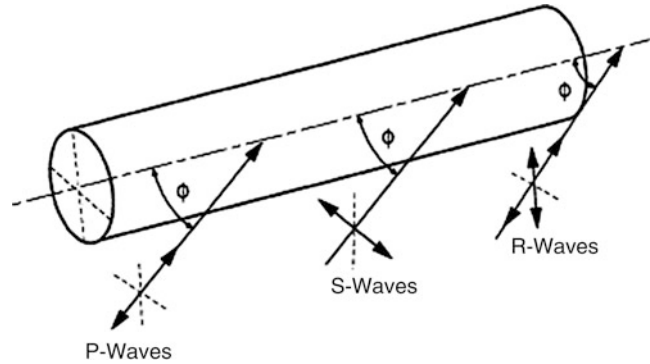
of these analyses should be interpreted with care (Power et al. 1996).

(b) Harmonic wave propagating at a given angle of incidence in a homogeneous, isotropic, elastic medium. The critical incidence angle resulting in the maximum strain is typically used (Newmark 1967). However, the strain order of magnitude estimated by this method is useful as initial design tool and design verification method.

(c) St John and Zahrah (1987) developed free-field solutions for axial and curvature strains due to compression, shear, and Rayleigh waves. Figure 10 shows the seismic waves causing the strains. Treating the tunnel as an elastic beam allows the calculation of combined axial and curvature deformations.

The strain bending component is relatively small compared to axial strains, but if the tunnel radius increases, the curvature contribution increases. Tunnel cracks may open and then close in the lining due to the cyclic nature of the axial strains. As long as the cracks are small, are uniformly distributed, and do not affect the performance of the tunnel, even unreinforced concrete linings are considered adequate. It is important to emphasize that the p- and s-wave velocities used are those of the deep rock. The range for s-wave is between 2 and 4 km/s and p-wave between 4 and 8 km/s (Power et al. 1996).

Ovaling and Racking Deformation

Ovaling deformations, developed by waves acting perpendicular to the circular tunnel lining, are caused predominately by vertically propagating shear waves (Wang 1993). Ground shear distortions can be defined assuming a non-perforated ground or a perforated ground. As shown in Fig. 11, both cases ignore the tunnel lining (soil-structure interaction), where the maximum diametric strain is in terms of the maximum free-field shear strain (γ_{max}) and the Poisson ratio (v_m). The first can be used to approximate the behavior of a tunnel lining whose stiffness is equal to the medium it replaces. The second can be used to approximate the behavior of a tunnel lining whose stiffness can be neglected in comparison with the stiffness of the medium.

A rectangular box structure will undergo transverse racking deformations (Fig. 12) when subjected to earthquake shear distortions. Racking deformations are defined as the differential sideway movements between the top and bottom elevations of rectangular structures. If an initial calculation is needed, it can be calculated based on St John and Zahrah (1987) equations.

Numerical Analysis

Many computer programs are available to estimate free-field shear distortions: SHAKE (Schnabel et al. 1972), FLUSH (Lysmer et al. 1975), D-MOD (Matasovic 1993), and DEEPSOIL (Hashash et al. 2011; Hashash and Park 2001), among others. One-dimensional site response analyses can be used to characterize the change in the propagating ground motions on variable soil profiles, but these analyses only take into consideration vertically propagating shear waves. However, these are the waves that

Seismic Design of Tunnels, Fig. 11 Free-field shear distortions of perforated and non-perforated ground

Continuous soil

$$\frac{\Delta d}{d} = \pm \frac{\gamma_{max}}{2}$$

$\frac{\Delta d}{2}$

Free-Field Shear Strain, γ

Perforated soil

$$\frac{\Delta d}{d} = \pm 2\gamma_{max}(1-\upsilon_m)$$

$\frac{\Delta d}{2}$

Open cavity

Free-Field Shear Strain, γ
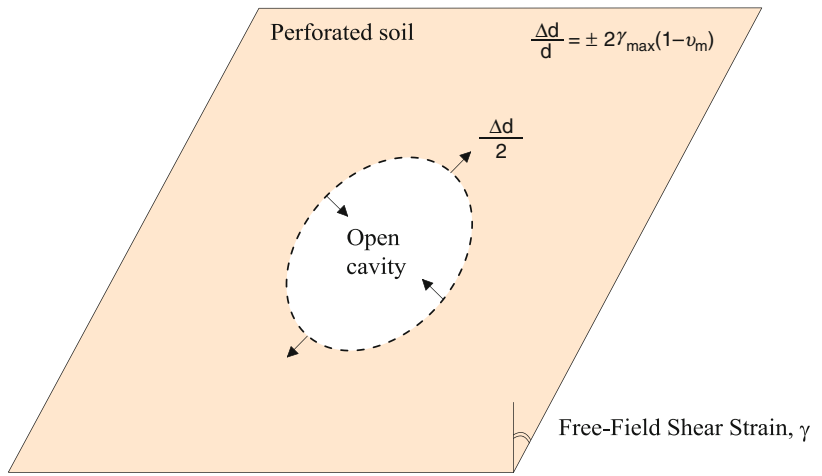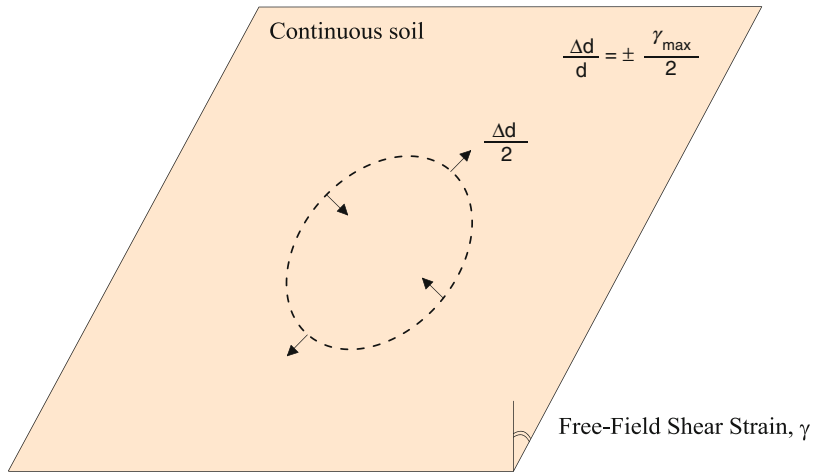
Seismic Design of Tunnels, Fig. 12 Typical free-field racking deformation imposed on a rectangular frame (Wang 1993)

Horizontal Shear Deformation, Δ (ft)

0.1   0.2   0.3   0.4

Δ_B Δ_A

Depth Below Surface (ft)

50

Δ_{diff}

100

150

Soil Deformation Profile

Ground Surface

Δ_{diff}

A

B

$\Delta_{diff} = \Delta_A - \Delta_B$

Δ_{diff}

Racking Deformation of a Box Structure

S
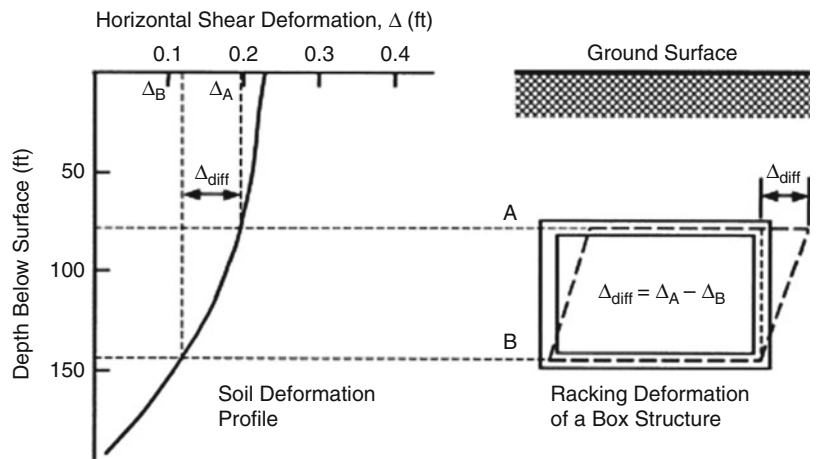
carry most of the seismic energy. The analyses can be performed in equivalent linear frequency domain or nonlinear time domain. The resulting free-field shear distortion can be expressed in the form of shear strain or shear deformation profile with depth.

Applicability of Free-Field Deformation Approach

The free-field deformation approach is a simple and effective design tool when earthquake-induced ground motions are small. However, in structures located within soft soil profiles, the method gives overly conservative designs, because free-field ground distortions in these soils are large. It also neglects the difference in stiffness between the lining and the surrounding soil. The presence of an underground structure modifies the free-field deformations; methods to model this interaction will be described in the following sections.

Pseudo-static Soil-Structure Interaction

In pseudo-static soil-structure interaction analyses, the soil and structure inertia due to seismic shaking is neglected. The problem is simplified to that of a structure in a soil medium subjected to simple shear on horizontal and vertical planes.

The beam-on-elastic foundation approach is used to model soil-structure interaction effects. Both the lining and the medium are assumed to be linear elastic. Wang (1993) presents a summary of closed-form elastic solutions for axial force and moment developed for circular tunnels due to seismic waves propagating along and perpendicular to the tunnel axis. Adding stiffness and strength to the structure may attract more forces, so a better solution would be to add ductility. These solutions are dependent on the estimates of appropriate spring coefficients compatible with anticipated displacements and wave lengths. They are often limited to idealized seismic wave forms.

Most pseudo-static SSI analyses focus on the interaction of vertically propagating shear waves with the transverse section of a tunnel. These analysis approaches are described next.

Transverse: Ovaling Deformations of Circular Tunnels

Peck et al. (1972) proposed closed-form solutions in terms of thrust, bending moments, and displacement under external loading. The lining response was a function of structure compressibility and flexibility ratios, in situ overburden pressure, and at-rest earth coefficient. To adapt to seismic loading, the free-field shear stress replaces the in situ overburden pressure and earth coefficient. The stiffness of the tunnel relative to the ground is measured by the compressibility (C) and flexibility (F) ratios. Those are the extensional stiffness and flexural stiffness of the medium relative to the lining.

Under this framework, Wang (1993) presented solutions for the diametric strain, the maximum thrust, and the bending moment under full-slip conditions, meaning normal force but no tangential shear force are present between the lining and the medium. For most cases the interface condition is between full slip and no slip. Slip interface can only happen in tunnels in soft soils or cases of severe seismic loading and full-slip assumption may lead to underestimation of the maximum thrust. As shown in Wang (1993) and NCHRP 611 (Anderson et al. 2008), for ground Poisson's ratio less than 0.5, thrusts decrease with decreasing compressibility ratio, but for Poisson's ratio of 0.5, the thrust response is independent of compressibility. The normalized lining distortion can be a plotted as function of flexibility ratio, as shown in Fig. 13. When $F < 1.0$, the lining is considered stiffer than the ground and deforms less than the ground. When $F > 1.0$, the lining is expected to deform more than the free field with an upper limit equal to the perforated ground case as described in Table 3. Penzien (2000) provides an analytical procedure to evaluate racking deformations of rectangular and circular tunnels. His solutions for ovaling deformations in terms of thrust and moment are very close to those of Wang (1993) for full-slip condition. However, the value of thrust for no-slip condition is much smaller in Penzien (2000) than in Wang (1993), differing in one order of magnitude.
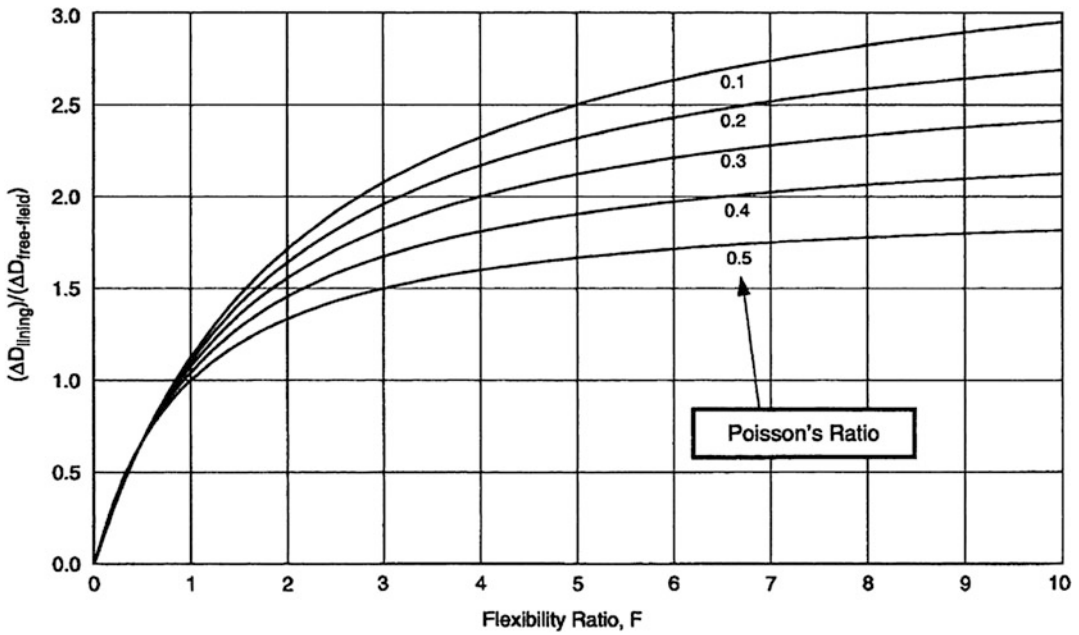
Seismic Design of Tunnels, Fig. 13 Normalized lining deflection vs. flexibility ratio, full-slip interface, circular lining (Wang 1993)

Seismic Design of Tunnels, Table 3 Explanation of flexibility ratio

Flexibility ratio F	Meaning
$F \to 0$	The structure is rigid, so it will not rack regardless of the distortion of the ground
$F < 1$	The structure is stiff relative to the medium and will therefore deform less. Usually soft soil, and racking deformations are large
$F = 1$	The structure and medium have equal stiffness, so the structure will undergo approximately free-field distortions
$F > 1$	The structure racking stiffness is smaller than that of the soil. Usually stiff soil, and racking deformations are small
$F \to \infty$	The structure has no stiffness, so it will undergo deformations identical to the perforated ground

Hashash et al. (2005) compared the two analytical solutions to finite-element method numerical analyses to validate which of the solutions provide the correct solution to this problem.

The results from the numerical analyses agree with Wang (1993) solutions, highlighting the limitations of the other analytical solution. Sedarat et al. (2009) show that interface condition between the tunnel lining and the surrounding soil has an important impact on the computed thrust in the lining but limited impact on computed lining deformation.

Transverse: Racking Deformations of Rectangular Tunnels

Box-shaped cut-and-cover tunnels, common for transportation tunnels, have seismic characteristics different from circular tunnels because the walls and slabs of the box-shaped tunnels are stiffer. They are also often placed at shallower depths compared to circular tunnels. Therefore, it is important to carefully consider the soil-structure interaction due to increased stiffness and the increased seismic ground deformations at shallow depths (Hashash et al. 2001). Numerical analyses are often employed to compute the response of the tunnel structure to deformation of the surrounding soil.
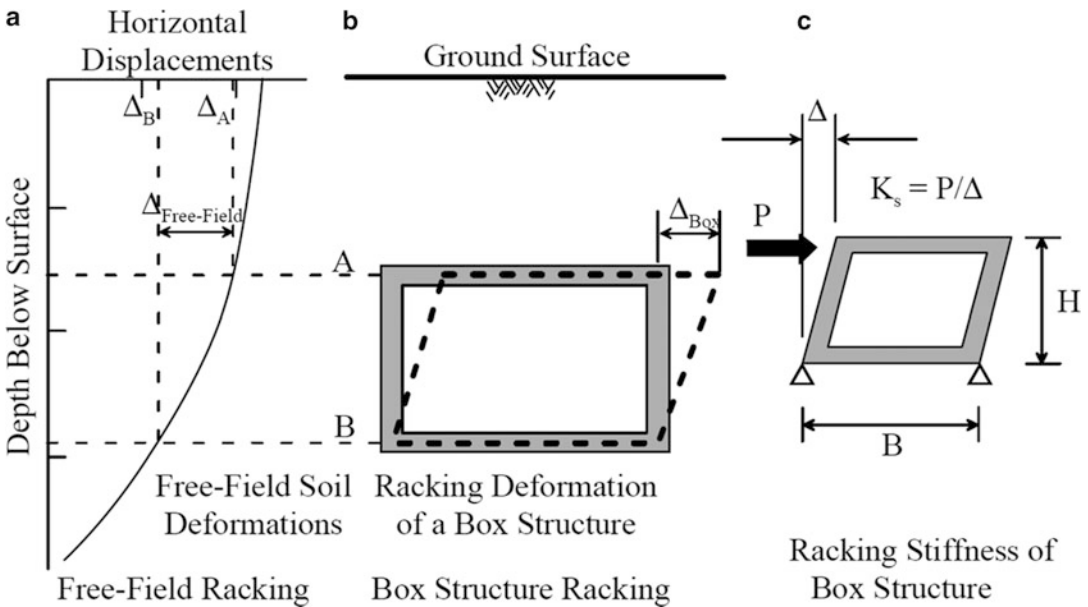
Seismic Design of Tunnels, Fig. 14 Racking ratio, free-field and structure racking (Hashash et al. 2010)

Wang (1993) and Anderson et al. (2008) employed such techniques to develop relationships between racking ratio and flexibility (Fig. 14).

Huo et al. (2006) present an analytical solution for deep rectangular structures with a far-field shear stress. Complex variable theory and conformal mapping were used to develop the solution of structures in homogeneous, isotropic, elastic medium.

The relative stiffness between soil and structure, structure geometry, input earthquake motions, and tunnel embedment depth are factors that contribute to the soil-structure interaction effect. The most important of those is the flexibility ratio (F) that refers to the shear stiffness of the soil relative to the structure that replaces it (Wang 1993):

$$F = \frac{G_m \cdot B}{P \cdot H}$$

where G_m is the shear modulus of soil or rock medium, B is the width of the structure, P is the force required to cause a unit racking deflection of a rectangular frame structure, and H is the height of the tunnel as illustrated in Fig. 14. For a rectangular frame with an arbitrary configuration, the flexibility ratio can be determined by

performing a simple frame analysis; even for simple one-barrel frames, no computer analysis is needed.

The racking ratio (R) is the ratio structure racking distortion to the free-field ground distortion:

$$R = \frac{\Delta_{\text{structure}}}{\Delta_{\text{free-field}}}$$

NCHRP 611 (Anderson et al. 2008) developed a relationship between the racking ratio (R) of rectangular conduits and the flexibility ratio (F), whereby

$$R = \frac{2F}{1 + F}$$

Hashash et al. (2010) describes the procedure to be used for performing 2D pseudo-static racking analysis of rectangular tunnels (Fig. 15). In this type of analysis, the soil is assumed to be massless and the section deforms in pure shear. In a first step, Fig. 15a, the free-field racking deformation time history is computed from 1D site response analyses using hazard-compatible ground motion time histories.
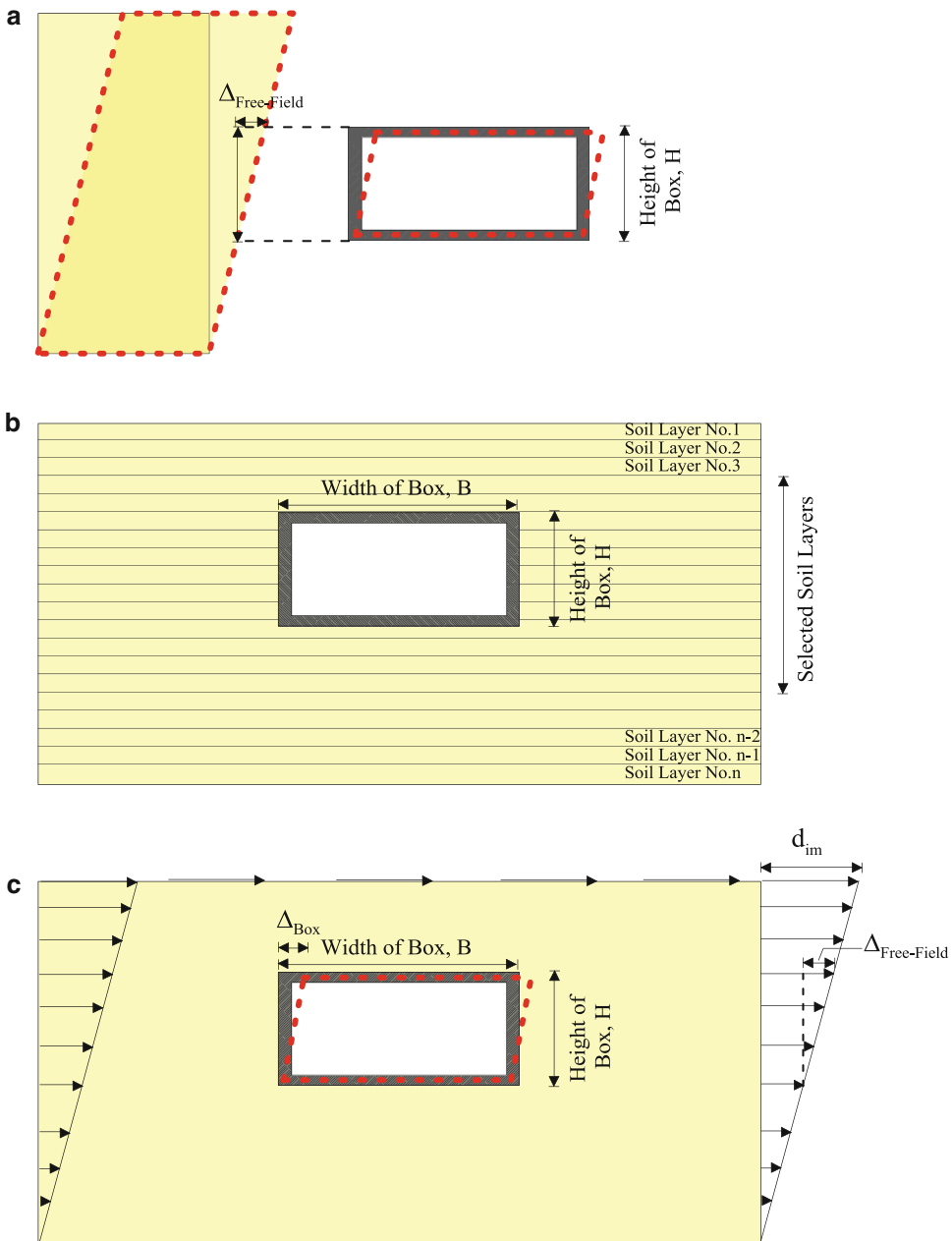
Seismic Design of Tunnels, Fig. 15 2D pseudo-static numerical analysis (Hashash et al. 2010)

To perform the 2D numerical analysis, the second step (Fig. 15b) is to define the elastic properties of a uniform soil medium as the average strain-compatible elastic properties of the selected soil layers (Anderson et al. 2008). As recommended in Hashash et al. (2010), layers 1–3 m above and below the structure should be included. Shear modulus values can be selected using the strain-compatible shear wave velocities from site response analysis in the selected layers, from which the average shear modulus over the selected layers can be calculated. In this step, the

structural member properties are needed: E (stiffness), I (moment of inertia), and A (cross-sectional area).

Finally, in a 2D numerical analysis, the lateral displacement (d_{im}) time histories obtained from 1D site response analysis are applied at the left, right and top boundaries of the model to impose the free-field racking calculated in the first step on the model, as shown in Fig. 15c. With the numerical analysis, the soil medium will transmit shearing deformations to the box structure and the box racking deformation time history can be obtained. This is used to obtain the racking ratio (R).

Two-dimensional pseudo-static numerical analyses can be a very useful tool, but they have some limitations. The ground surface shear displacements for shallow box structures cannot be transmitted uniformly. The model can be artificially extended to address this problem. Racking deformations are assumed to vary uniformly over the height of the box structure. The response of individual layers is not represented. This becomes a problem when layers with very different stiffness are part of the soil profile.

Dynamic Soil-Structure Interaction

The complex soil-structure interaction of underground structures during seismic loading can be simulated using numerical analysis tools which include lumped mass/stiffness methods and finite-element/difference methods.

Lumped mass/stiffness methods are useful to analyze the 3D behavior of a tunnel lining in a simplified manner. Many parameters for the springs that represent the structure stiffness and the soil stiffness must be defined to have a realistic model.

In finite-element/difference models, the tunnel structure is discretized and the soil surrounding the tunnel can be either discretized or represented by springs. 2D and 3D models can also be used to analyze the inelastic sections of the tunnel cross section. Discrete element models are useful when weak planes in the soil profile exist, since in this method the soil/rock mass is represented by an assemblage of distinct blocks (rigid or deformable) with a prescribed constitutive relationship.

One of the advantages of these dynamic soil-structure interaction analyses is that the soil layers are modeled to reflect the idealized site stratigraphy; each soil layer can be either modeled as a linear elastic material with strain-compatible shear modulus and damping values or characterized via soil constitutive models that represent soil nonlinearity and hysteretic response at small strains. However, the use of the nonlinear constitutive models requires careful selection of input parameters and thus more advanced testing to define those input parameters. The nonlinear behavior and the frequency content of the free-field environment contribute to the structural racking behavior.

Hashash et al. (2010) provides a simplified 2D dynamic soil-structure interaction procedure that makes computational effort manageable for design purposes of transverse response of rectangular tunnels. The first step is to perform a 1D site response analysis to obtain the acceleration and displacement time history throughout the soil profile and then obtain the strain-compatible shear wave velocities and damping ratios for the 2D model layers.

The numerical analysis involves applying the displacement time history at the model base, then propagating the ground motion through the soil to simulate the soil-box interaction. The displacement time histories must be obtained at four monitored points, as shown in Fig. 16, to calculate the relative box displacement $\Delta_{\text{box}} = \max\left[abs\left(\delta_{h,C} - \delta_{h,D}\right)\right]$ and the free-field relative displacement $\Delta_{ff} = \max\left[abs\left(\delta_{h,A} - \delta_{h,B}\right)\right]$ to then calculate the racking ratio R $= \frac{\Delta_{\text{box}}}{\Delta_{ff}}$. The complete soil profile does not need to be included; a limited thickness of soil that captures the characteristics of wave propagation is sufficient. In the dynamic soil-structure analysis, the soil profile represented in the analysis reflects the actual soil profile, as there is no need to use an average soil layer needed in pseudo-static analyses.
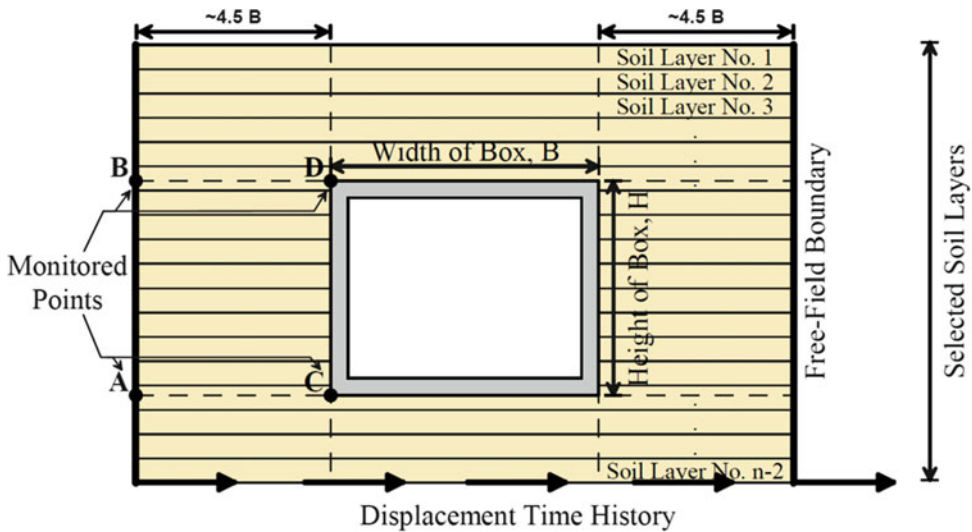
Seismic Design of Tunnels, Fig. 16 2D dynamic soil-structure interaction (Hashash et al. 2010)

Comparison of Pseudo-static and Dynamic Racking Soil-Structure Interaction Analyses

Hashash et al. (2010) performed a series of pseudo-static and dynamic soil-structure interaction analyses of single- and double-box structures in stiff and soft soil profiles using equivalent linear and nonlinear site response analysis with 14 ground motion time histories. The results are shown in Fig. 17. The results of the study found that for $F < 1$ the dynamic analysis and pseudo-static analysis appear to be quite similar and slightly above NCHRP 611 (Anderson et al. 2008). For $4 < F < 9$, the dynamic results show more scatter and higher racking ratios compared to pseudo-static results, both above NCHRP611 (Anderson et al. 2008). Dynamic analyses for $10 < F < 13$ show slightly lower racking ratios than pseudo-static analysis, both below NCHRP611 (Anderson et al. 2008). Some of the main conclusions from Hashash et al. (2010) are: (a) numerical approaches provide results and trends that are consistent with results obtained from simplified closed-form solutions, (b) there is a need to account for variability in the input ground motions and site response analysis methods as they affect the flexibility ratio (F), and (c) dynamic analyses must be performed to verify and supplement the results of pseudo-static soil-structure interaction.

Performance Evaluation for Immersed-Tube Tunnels

Pseudo-static longitudinal 3D models can be used to analyze axial and bending deformations in immersed-tube tunnels. In a lumped mass analysis approach, the tunnel lining is divided into individual segments with different masses and stiffnesses. The masses are then connected by springs that represent the axial, shear, and bending stiffness of the tunnel as shown in Fig. 18 (Hashash et al. 1998). Free-field displacement time histories that consider effects of wave passage/phase shift and incoherence are calculated at selected locations along the tunnel's length. The computed free-field displacement time histories are then applied at the ends of springs to represent soil-tunnel interaction in a quasi-static analysis. If a dynamic analysis is needed, appropriate damping factors need to be incorporated into the structure as well as springs to represent the soil.

Recent work from Anastasopoulos et al. (2007) focuses on nonlinear response to
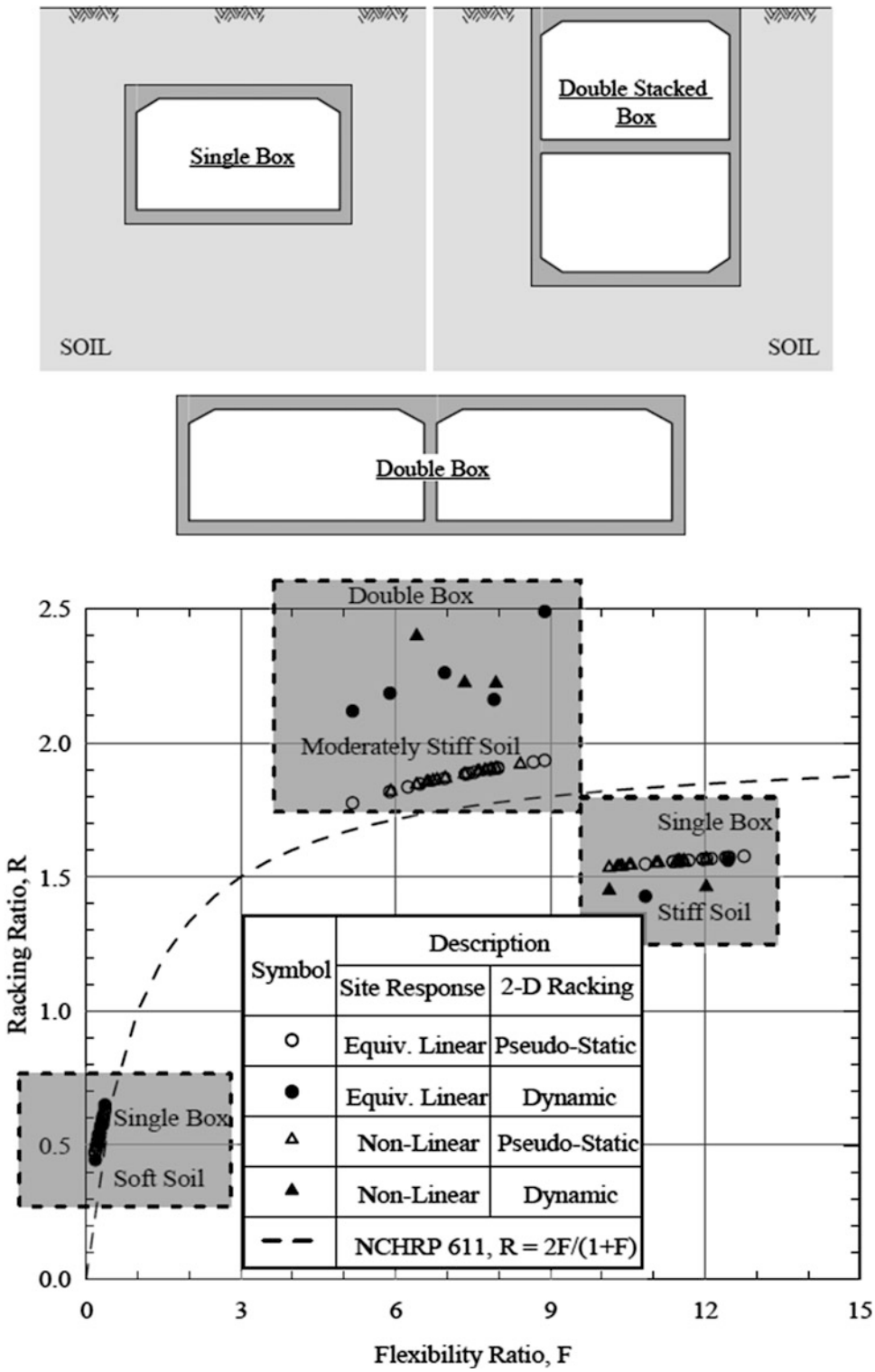
Seismic Design of Tunnels, Fig. 17 Results from comparison of 2D pseudo-static and dynamic numerical analyses (Hashash et al. 2010)
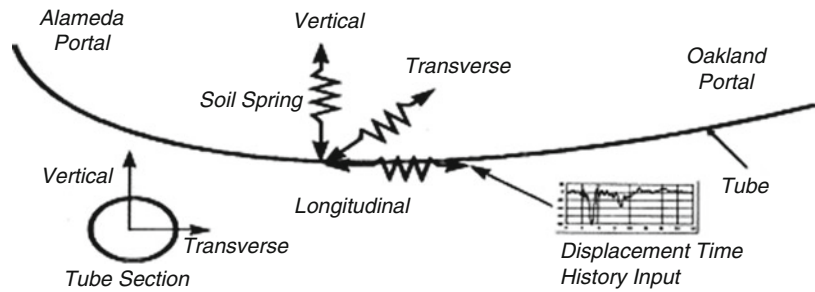
strong seismic shaking of deep immersed tunnels ($\approx$70 m). The free-field acceleration time histories are computed at the base of the tunnel through 1D wave propagation analysis using equivalent linear and nonlinear analyses. The computed free-field acceleration time histories are then imposed on the supports of the tunnel in the form of excitation. The tunnel is modeled as a multi-segment beam connected to the ground through calibrated springs, dashpots, and sliders. Wave passage effects are taken into account using Eurocode 8 (EC8 2002); however, the geometric incoherence was not considered because it did not make a difference when added to wave passage. The soil is assumed to be uniform along the tunnel.

A finite-element analysis is used to perform a nonlinear dynamic transient analysis of the tunnel. Tunnel segments are modeled using beam elements that take into account shear rigidity. The joints are modeled with nonlinear hyperelastic elements. The bored tunnels at the end of both segments are incorporated in the analysis as beams on viscoelastic foundation. Influence of segment length and joint properties was then investigated parametrically.

The results from Anastasopoulos et al. (2007) show that seismic response of immersed tunnel correlates better with PGV than PGA consistent with prior studies. There are some key conclusions applicable to immersed tunneling projects. First, a properly designed immersed tunnel can resist near-fault soil-amplified excitation with a PGA as large as 0.6 g and PGV as large as 80 m/s and containing long period pulses. Also, the net tension and excessive compression between segments can be avoided by suitable design of joint gaskets and relatively small segment lengths. However, it is important to note that this study did not examine the heterogeneous nature of the soil conditions and ground motion incoherency and did not investigate the time-dependent stress relaxation on the rubber gasket or the effect of tectonic displacements from fault rupturing.

Additional Seismic Performance Issues

A number of additional items must be considered when evaluating the seismic performance of box structures. The detailed treatment of all these items is beyond the scope of this document, but a brief overview of some of them is presented.

Response to Vertical Ground Shaking

Significant vertical ground motions often due to near-fault effects can impose vertical loads on the roof of box structures. Figure 19 shows a schematic of the vertical loading on relatively shallow box structure (Hashash et al. 2010). Two types of analyses can be performed to assess the vertical loads. The vertical acceleration near the ground surface can be estimated as part of the seismic hazard analysis and then be used to computer pseudo-static inertial load of the soil on the tunnel roof. Alternatively, 2D dynamic soil-structure interaction analysis representing both the underground structure and the soil can be performed to compute the vertical inertial loading on the roof. There remain significant uncertainties in selecting appropriate soil properties for propagating vertical ground motion in such a model.

S

Seismic Design of Tunnels, Fig. 19 Vertical ground shaking effects (Hashash et al. 2010)
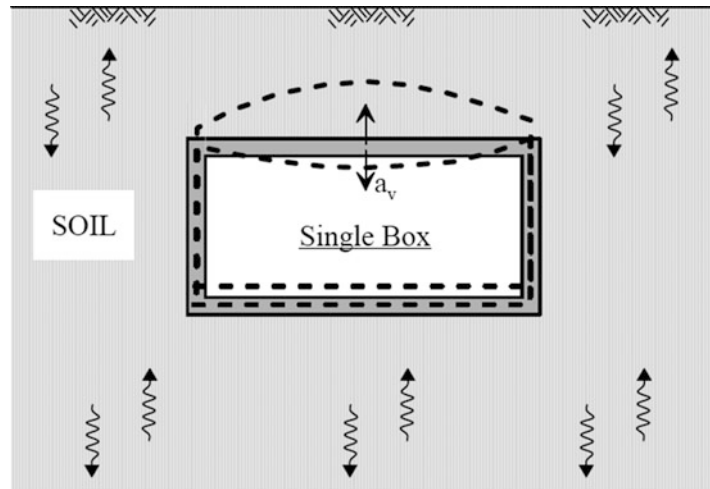
Interaction of Temporary and Permanent Structures

Braced excavations are used when space is needed to construct a shallow underground structure. Temporary excavations in highly seismic urban areas are also being seismically designed though for a lower level of seismic shaking than a permanent structure. Dynamic soil-structure interaction analyses are preferable to analyze racking in temporary structures, because it allows the dynamic load increments and levels of deformation on the temporary system to be estimated without significant simplifications.

It is customary to neglect the contribution of the temporary shoring wall on the permanent tunnel box. This is based on the assumption that the presence of the temporary walls will enhance the performance of the system and neglecting it is prudent. The effect of the presence of shoring walls on the seismic response of permanent box structures was studied in Hashash et al. (2010). For many projects only a few feet of the top of shoring is system is cut off after completion of the box construction while the rest of the wall remains in place. The study analyzed a single box surrounded by soft clay using both 2D pseudo-static and dynamic methodologies, as discussed earlier. Three cases were considered in Fig. 20a.

Figure 20b shows that modeling the shoring wall over the height of the structure only slightly reduces the racking ratio of the tunnel structure, and hence neglecting the presence of the temporary walls is acceptable. An increase between 15 % and 20 % in the racking ratios is estimated when the temporary shoring wall is modeled above and below the permanent structure. The shoring walls act as extended wings, transferring soil loads to the structure from above and below the structure and therefore creating more racking deformations (Hashash et al. 2010). Therefore, neglecting the contribution of the shoring walls in this case will lead to underestimating of the racking deformation and is thus not advisable.

Impact of Superstructure and Adjacent Structures

Shallow underground structures for public transportation are a key component of sustainable cities. In dense urban environments, underground structures are built in close proximity to high-rise building foundations. Tall buildings have the potential to change the ground motions in the foundation soil and therefore transmit significant forces and base shear to adjacent underground structures. It is therefore important to evaluate the impact of transmitted forces from the superstructure to the soil to the underground structure under earthquake loading.

If aboveground structures are built over existing underground structures, the interaction between the structures must be evaluated.
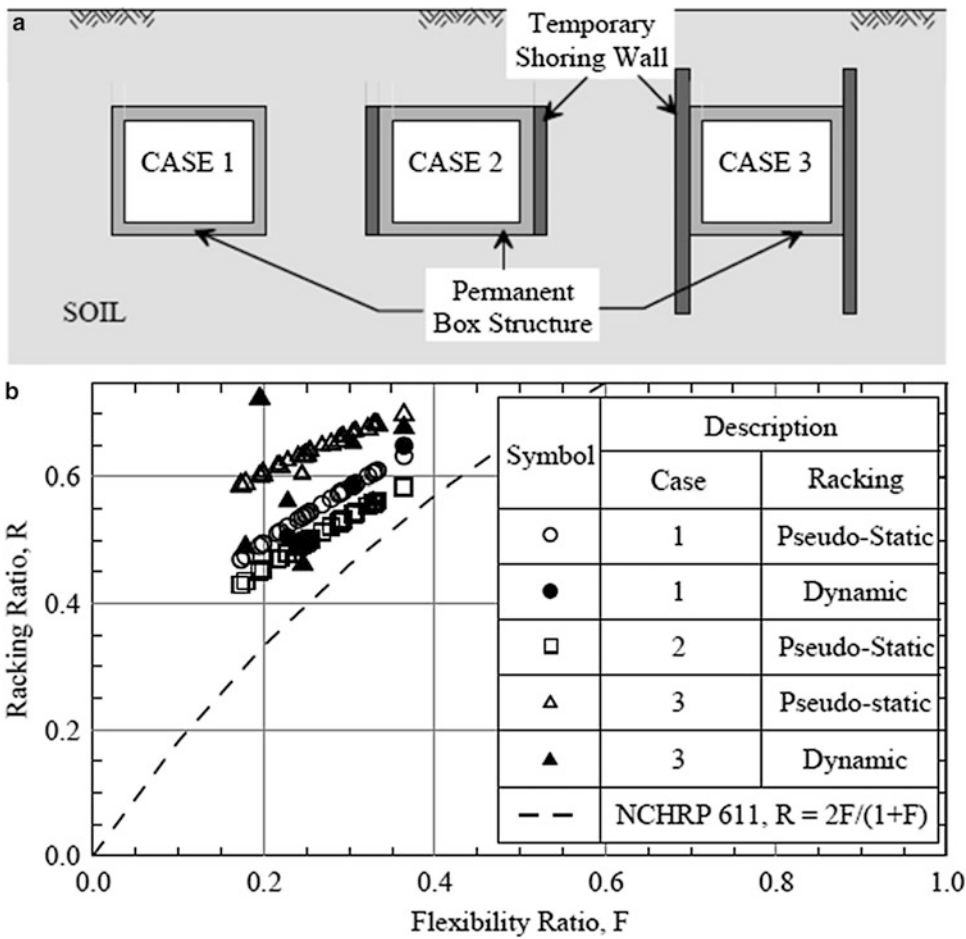
Seismic Design of Tunnels, Fig. 20 Effect of shoring walls on permanent structures (Hashash et al. 2010)

The global system performance can be evaluated by means of a numerical analysis of seismic soil-box-structure interaction.

A related item is that the placement of wide box structures in deep excavations in soft soils may lead to significant changes in the soil properties beneath the structure. The influence of those changes on the long-term dynamic response of the structure should be taken into account (Hashash et al. 2010).

Tunnel Joints: Portals, Stations, and Tunnel Segments

Stiffness differences in the tunnel structure may generate differential movements and stress concentration. Some examples include tunnel-building or tunnel-station connections, tunnel junctions, tunnels crossing distinct geologic media, and local restraints of any type (Hashash et al. 2001).

The most common solution to these interface problems is the use of flexible joints. The first step in the design process is to use closed-form solutions or numerical analyses to determine the required allowable rotation and differential movements in the longitudinal, transverse directions. The joints also must be designed to support static and dynamic earth and water loads before and during the earthquake while remaining watertight. If a continuous design is implemented, very large forces and moments are likely to be generated (Hashash et al. 2001).

Tunnel portals and vent structures have part of their structure above ground; therefore, the
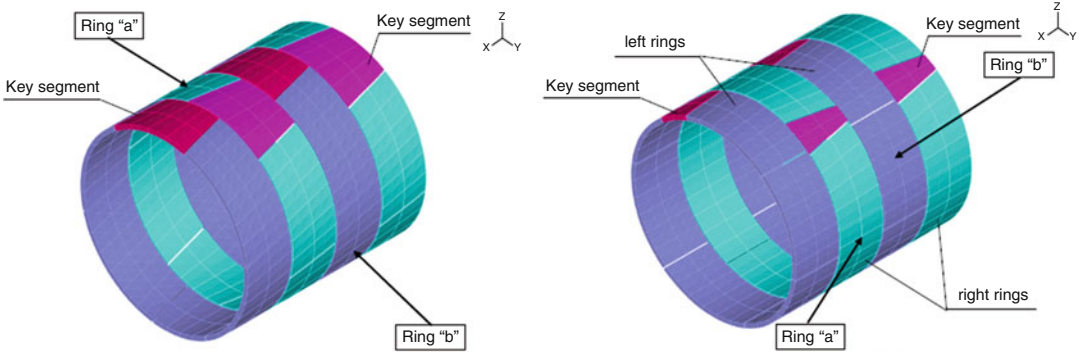
Seismic Design of Tunnels, Fig. 21 FEM lining models for trapezoidal and rectangular geometric configurations boundary conditions (Kramer et al. 2007)
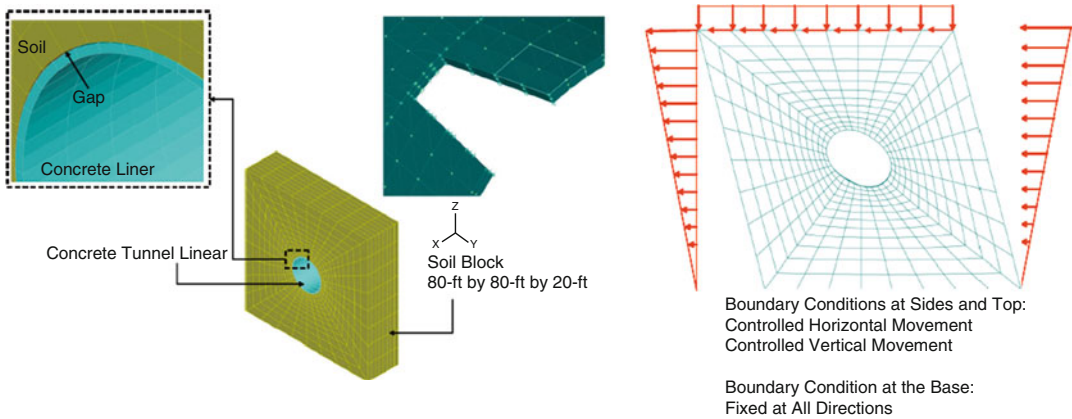
Seismic Design of Tunnels, Fig. 22 Details of mesh development and boundary conditions (Kramer et al. 2007)

seismic design should account for inertial effects or potential pounding. It is always preferable that those structures are isolated from the tunnel through flexible joints (Hashash et al. 2001).

For soil-rock transitions, it is recommended that the tunnel not be cast directly against the rock. Instead, there should be some over-excavation filled with soil or aggregate. If that is not possible in bored tunnels, then a flexible lining can be installed (Hashash et al. 2001).

Precast Tunnel Lining

When a tunneling machine is used to excavate a tunnel, the lining usually is erected in segments. Single pass precast concrete segmental lining systems are often employed. The segment joint connections must therefore be designed to accommodate anticipated deformations. The joint behavior can be kept elastic or allow inelasticity if proper detailing is performed. Numerical analyses are often used to evaluate the seismic performance of segmental lining. Kramer et al. (2007) performed 3D finite-element analyses, Fig. 21, to compute radial and circumferential joint behavior during seismic ovaling and wave propagation for the Silicon Valley Rapid Transit (SVRT) Project in San Jose, California. These analyses incorporate inelastic constitutive soil behavior, cracked concrete properties, and no-tension, frictional segment joint surfaces (Fig. 22). Kramer et al. (2007) provide guidelines for precast lining design in high seismicity zones.

Seismic Retrofit of Existing Facilities

When considering seismic retrofit of an existing tunnel structure, the retrofit strategy depends on the structure damage mode. If there is a gross stability problem, the seismic retrofit strategies involve strengthening the structure itself or the surrounding geologic materials.

One concern for a circular tunnel is the contact quality between the liner and the surrounding geologic media. Strengthening this interface includes replacing the lining, increasing the lining thickness by adding reinforced concrete, or adding reinforcing bars or internal steel liner. Increasing lining thickness is not always a good solution, as it tends to attract more load. Methods that increase ductility as well as strength are more effective (Power et al. 1996). Adding circumferential joints along the tunnel axis can also reduce the stresses and strains induced by longitudinal propagating waves. However, the value of adding joints must be weighed against the expected performance of the liner without joints. It is important to verify in the retrofit design that there will be no additional water leakage and that the joints will not become weak spots (Power et al. 1996).

In cut-and-cover structures, some seismic retrofit strategies include increasing ductility of reinforced concrete lining, adding confinement at existing columns, and adding steel plate jackets at joints.

Summary

The performance of the underground facilities during recent seismic events showed that the underground structures have suffered appreciably less damage than surface structures. However, the failure of even one of these underground structures can be detrimental to the proper post-earthquake operation of a tunnel network in urban areas. Thus, evaluation of underground structures to seismic shaking is necessary.

A performance-based approach is recommended for seismic evaluation of underground structures. Underground structures seismic response is controlled by the deformation of the soil or rock medium in which they are embedded and their seismic evaluation is different from aboveground structures.

Seismic evaluation of underground structures includes evaluation for permanent and transient ground deformations. Pseudo-static and dynamic analysis approaches can be used to estimate the deformation of underground structures due to transient ground deformations and are presented.

Additional seismic design issues were discussed, including vertical ground shaking and response, interaction of temporary and permanent structures, impact of superstructure and adjacent structures, tunnel joints, and seismic retrofit of existing facilities.

Cross-References

▶ Conditional Spectra
▶ Dynamic Soil Properties: In Situ Characterization Using Penetration Tests
▶ Liquefaction: Countermeasures to Mitigate Risk
▶ Nonlinear Dynamic Seismic Analysis
▶ Nonlinear Finite Element Analysis
▶ Nonlinear Seismic Ground Response Analysis of Local Site Effects with Three-dimensional High-fidelity Model
▶ Performance-Based Design Procedure for Structures with Magneto-Rheological Dampers
▶ Probabilistic Seismic Hazard Models
▶ Seismic Actions Due to Near-Fault Ground Motion
▶ Site Response for Seismic Hazard Assessment
▶ Site Response: 1-D Time Domain Analyses
▶ Site Response: Comparison Between Theory and Observation
▶ Soil-Structure Interaction

References

Anastasopoulos I, Gerolymos N, Drosos V, Kourkoulis R, Georgarakos T, Gazetas G (2007) Nonlinear response of deep immersed tunnel to strong seismic shaking. J Geotech Geoenviron Eng 133(9): 1067–1090. doi:10.1061/(ASCE)1090-0241(2007)133:9(1067)

S

Anderson DG, Martin GR, Lam I, Wang JN (2008) Seismic analysis and design of retaining walls, buried structures, slopes and embankments, vol NCHRP 611: National Cooperative Highway Research Program, Transportation Research Board

Ashford SA, Boulanger RW, Donahue JL, Stewart JP (2011) Geotechnical quick report on the Kanto Plain Region during the March 11, 2011, Off Pacific Coast of Tohoku Earthquake, Japan GEER Association report no. GEER-025a: Geotechnical Extreme Events Reconnaissance (GEER)

Coppersmith KJ, Youngs RR (2000) Data needs for probabilistic fault displacement hazard analysis. In: Cello G, Tondi E (eds) The resolution of geological analysis and models for earthquake. Pergamon Press, Oxford

Cording EJ, Martin G, Parker H (2011) Century city area tunneling safety and fault investigations: tunnel advisory panel. Los Angeles County Metropolitan Transportation Authority, Los Angeles

EC8 (2002) Eurocode 8: design of structures for earthquake resistance. European Committee for Standardization (CEN), Brussels

Elnashai AS, Gencturk B, Kwon O-S, Al-Qadi IL, Hashash Y, Roesler JR, et al. (2010) The Maule (Chile) earthquake of February 27, 2010 consequence assessment and case studies: Mid-America Earthquake Center

Hamburger RO, Hooper JD (2011) Performance-based seismic design. Mod Steel Constr 4:36–39

Hashash YMA, Park D (2001) Non-linear one-dimensional seismic ground motion propagation in the Mississippi embayment. Eng Geol 62(1–3):185–206

Hashash YMA, Tseng WS, Krimotat A (1998) Seismic soil-structure interaction analysis for immersed tube tunnels retrofit. Paper presented at the proceedings of the 1998 conference on geotechnical earthquake engineering and soil dynamics III. Part 2 (of 2), Seattle, 3–6 Aug 1998

Hashash YMA, Hook JJ, Schmidt B, Yao JI-C (2001) Seismic design and analysis of underground structures. Tunn Undergr Space Technol 16:247–293

Hashash YMA, Park D, Yao JIC (2005) Ovaling deformations of circular tunnels under seismic loading, an update on seismic design and analysis of underground structures. Tunn Undergr Space Technol 20(5):435–441. doi:10.1016/j.tust.2005.02.004

Hashash YMA, Karina K, Koutsoftas D, O'Riordan N (2010) Seismic design considerations for underground box structures. Paper presented at the 2010 earth retention conference – earth retention conference 3, Bellevue, 1–4 Aug 2010

Hashash YMA, Groholski D, Musgrove M, Park D, Phillips C, Tsai C-C (2011) DEEPSOIL V5.0, manual and tutorial. Board of Trustees of University of Illinois at Urbana-Champaign, Urbana

Huo H, Bobet A, Fernández G, Ramírez J (2006) Analytical solution for deep rectangular structures subjected to far-field shear stresses. Tunn Undergr Space Technol 21(6):613–625

Iida H, Hiroto T, Yoshida N, Iwafuji M (1996) Damage to the Daikai subway station. Soils and Foundations 283–300. Special issue on geotechnical aspects of the January 17, 1995 Hyogoken-Nambu Earthquake, Tokyo, Japan

Kramer GJE, Sedarat H, Kozak A, Liu A, Chai J (2007) Seismic response of precast tunnel linings. Paper presented at the proceedings of the 2007 rapid excavation and tunneling conference, Toronto. Publisher and copyright holder: Society for Mining, Metallurgy & Exploration (SME). www.smenet.org

Lysmer J, Udaka T, Tsai CF, Seed HB (1975) FLUSH: a computer program for approximate 3-D analysis of soil-structure interaction problems. Earthquake Engineering Research Center, Berkeley

Matasovic N (1993) Seismic response of composite horizontally-layered soil deposits. PhD thesis, University of California, Los Angeles. Retrieved from http://wwwlib.umi.com/pqdd2/search/do?query=au%28Matasovic%2C%20Neven%2C%29%20and%20da%281993%29

Newmark NM (1967) Problems in wave propagation in soil and rock. Paper presented at the international symposium on wave propagation and dynamic properties of earth materials, Albuquerque, NM

Peck RB, Hendron AJ, Mohraz B (1972) State of the art in soft ground tunneling. Paper presented at the rapid excavation and tunneling conference, New York

Penzien J (2000) Seismically induced racking of tunnel linings. Earthq Eng Struct Dyn 29(5):683–691. doi:10.1002/(sici)1096-9845(200005)29:5<683::aid-eqe932>3.0.co;2-1

Power MS, Rosidi D, Kaneshiro J (1996) Volume III Strawman: screening, evaluation, and retrofit design of tunnels. National Center for Earthquake Engineering Research, Buffalo

Schmidt B, Hashash YMA (1999) Preventing tunnel flotation due to liquefaction. Paper presented at the proceedings of the second international conference on earthquake geotechnical engineering, Lisbon

Schnabel PB, Lysmer JL, Seed HB (1972) SHAKE: a computer program for earthquake response analysis of horizontally layered sites. Earthquake Engineering Research Center, Berkeley

Sedarat H, Kozak A, Hashash YMA, Shamsabadi A, Krimotat A (2009) Contact interface in seismic analysis of circular tunnels. Tunn Undergr Space Technol 24:482–490

St John CM, Zahrah TF (1987) Aseismic design of underground structures. Tunn Undergr Space Technol 2(2):165–197

Wang JN (1993) Seismic design of tunnels: a state-of-the-art approach, Monograph 7. Parsons Brinckerhoff Quade & Douglas, New York

Wang WL, Wang TT, Su JJ, Lin CH, Seng CR, Huang TH (2001) Assessment of damage in mountain tunnels due

to the Taiwan Chi-Chi earthquake. Tunn Undergr Space Technol 16(3):133–150

Youngs RR, Arabasz WJ, Anderson RE, Ramelli AR, Ake JP, Slemmons DB, et al. (2003) A methodology for probabilistic fault displacement hazard analysis (PFDHA). Earthq Spectra 19(1):191–219. doi:10.1193/1.1542891

Seismic Design of Waste Containment Systems

Edward Kavazanjian Jr.
School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ, USA

Synonyms

Equivalent linear analysis; Geosynthetic; Landfill; Performance-based design; Site response; Solid waste; Waste containment systems

Introduction

Maintaining the integrity of waste containment systems subject to seismic loading is not only important with respect to protecting human health and the environment from wastes and waste by-products, it is also important with respect to the availability of facilities essential for post-earthquake recovery, i.e., of landfills for proper disposal of the large volumes of earthquake-generated waste. Important considerations in the seismic design of waste containment systems include the response of the waste mass itself to seismic excitation, the stability of the waste mass (including its foundation), and the integrity of the engineered components of the waste containment system. Engineered components of waste containment systems include base and side-slope liner (barrier) systems, leachate and gas collection and removal systems, and final cover systems. Liner and cover systems include both natural soil and geosynthetic elements that serve as barriers to advective transport

of liquids and gas, as drainage layers for landfill leachate and surface water infiltration, and as protection for other components of the waste containment system. Geosynthetic waste containment system elements, including geomembrane and geosynthetic clay liner (GCL) barrier layers, geonet drainage layers, and geotextile filter, separation, and cushion (geomembrane protection) layers, are typically only found in containment systems constructed within the last 25 years, appearing with increasing frequency over that time span.

Figure 1 illustrates the components of municipal solid waste (MSW) landfill waste containment systems. MSW landfills are by far the most common type of engineered waste containment system, followed (with respect to frequency of occurrence) by cover systems (caps) for uncontrolled dumps and hazardous waste sites and then by hazardous waste landfills. Caps and containment systems for hazardous waste landfills employ similar elements to MSW landfills. The performance of ancillary facilities at a waste disposal site subject to seismic loading, including leachate and gas treatment facilities, surface water control systems, access roadways, and landfill monitoring systems, is also an important consideration but will not be addressed herein.

Following a brief review of the performance of landfills in earthquakes, this entry summarizes the state of knowledge on waste mass seismic response and the properties of the waste and geosynthetic materials relevant to seismic analysis and design. Current approaches to the seismic design of landfills are then discussed, culminating in a discussion of the value of performance-based seismic design for these important facilities.

Seismic Performance of Waste Containment Systems

In general, the seismic performance of waste containment systems subject to strong ground motion from earthquakes has been acceptable, i.e., has not resulted in a harmful discharge of waste or waste by-products to the environment.

S
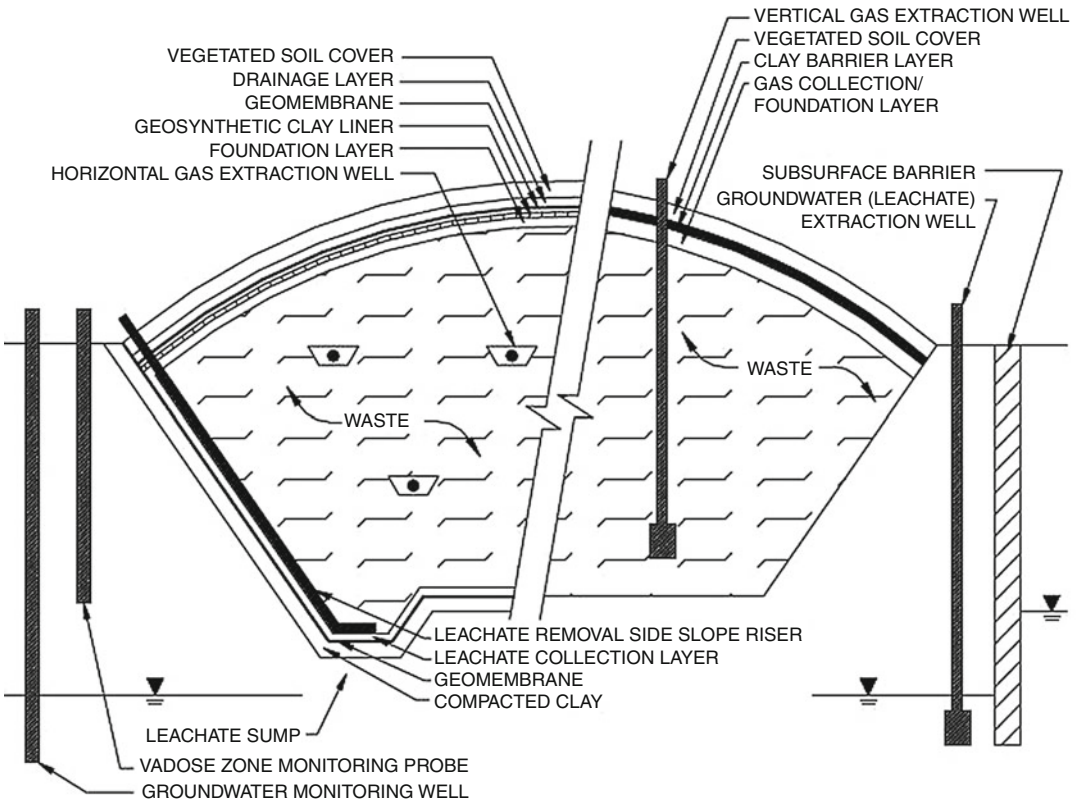
Seismic Design of Waste Containment Systems, Fig. 1 Containment system components for modern (*left*) and older (*right*) landfills (Kavazanjian et al. 1998)

In the epicentral region of the 1989 M 6.9 Loma Prieta earthquake, several MSW landfills with 2H:1V slopes rising up to 60 m above grade withstood earthquake motions with peak ground accelerations estimated to be on the order of 0.6 g with only minor cracking in the soil cover, e.g., observed cracks on the order of a 50 mm or less (Buranek and Prasad 1991; Johnson et al. 1991; Orr and Finch 1990). However, there were observations of more significant cracking in the soil cover on the very steep north slope of the Operating Industries, Inc. (OII) Landfill (average inclination, 1.5H:1V; maximum inclination, 1.3H:1V) in the 1987 M 6.1 Whittier Narrows earthquake and the 1994 M 6.7 Northridge earthquake when they were subject to peak ground accelerations estimated to be less than half the 0.6 g level experienced by landfills in the Loma Prieta event (Kavazanjian et al. 2013b; Matasovic et al. 1998). Furthermore, neither the

OII Landfill nor any of the landfills subject to strong ground shaking in the Loma Prieta earthquake employed geosynthetic containment system elements (Kavazanjian et al. 2013b; Matasovic et al. 1998), Tensile rupture of a geomembrane barrier layer was observed at two separate locations at the Chiquita Canyon Landfill in the Northridge earthquake, one location at the crest of a lined side slope and a second location near the crest of a lined side slope (EMCON 1994). Figure 2a shows the cover soil cracking observed at the OII Landfill, and Fig. 2b shows one of the geomembrane tears observed at the Chiquita Canyon Landfill following the Northridge earthquake.

Geosynthetic waste containment system elements are of particular concern with respect to seismic loading because they create the potential for planes of weakness along which slippage and instability can occur (due to a relatively low

Seismic Design of Waste Containment Systems, Fig. 2 Damage to landfills observed following the Northridge earthquake: (**a**) cracking in soil cover on the benches of the OII Landfill (Courtesy or Raymond B. Seed); (**b**) tear in the side-slope liner at the Chiquita Canyon Landfill (Courtesy of Robert M. Koerner)

interface or in-plane shear resistance) and because of their potential for tensile rupture which could result in breaching of barrier, filter, and protective layers and disruption of drainage layers. Stability failures of landfills have occurred under static loading due to the low interface shear strength of a geosynthetic waste containment system element (Koerner and Soong 2000; Mitchell et al. 1990; Seed et al. 1990) and the tensile rupture of the side-slope geomembranes at the Chiquita Canyon Landfill were cited previously. Furthermore, there are no case histories of cover systems with geomembrane elements subject to free-field ground motions in excess of 0.2 g, and cover systems are of particular concern due to the potential for amplification of seismic motions by the waste mass. There is also a potential for hidden damage to containment system elements beneath the surface of the landfill, where damage cannot be directly observed or readily detected by the monitoring systems employed in current practice.

Waste Mass Seismic Response

Overview: The seismic response of the waste mass is an important consideration in waste containment system seismic design as it controls the loading on the containment system elements as well as the global stability of the waste mass. Because there is often a significant impedance contrast between the waste mass and the underlying foundation material (soil or rock), landfill seismic response is susceptible to amplification of free-field motions. Amplification of earthquake ground motions is reported to have occurred at the one landfill at which strong ground motions have been recorded to date, OII Landfill (Kavazanjian et al. 2013b), and Bray and Rathje (1998) have conducted analyses suggesting that the amplification potential of municipal solid waste landfills is similar to the amplification reported by Harder (1991) between the base and crest of earth dams in the transverse direction. Figure 3 compares the amplification of peak ground acceleration from the base to the

S
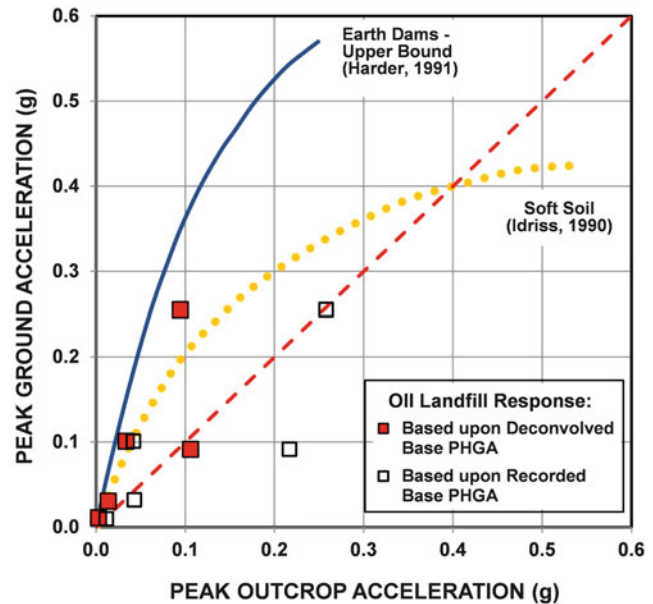
crest of the OII Landfill (as determined by back analysis) to the amplification of the transverse peak ground acceleration from the base to the crest of earth dams as reported by Harder (1991).

Waste Properties: Seismic response of a waste mass is controlled by waste mass properties, the waste thickness, and the impedance contrast between the base of the landfill and the waste mass (which depends upon the properties of the waste mass and the underlying foundation material) as well as the characteristics of the free-field ground motion. For landfills with geosynthetic elements in the liner and cover system, slip along the interface of a geosynthetic element may also affect the seismic response. However, slip at geosynthetic interfaces is usually ignored in practice when evaluating landfill seismic response. Analyses in which the potential for slip at the interface is ignored are termed "decoupled" analyses as landfill seismic response is decoupled from the geosynthetic interface behavior and potential displacement at the interface. A decoupled analysis is generally considered conservative with respect to predicting landfill seismic performance, though there are relatively rare situations where this has been shown not to be the case (Bray and Rathje 1998).

Most seismic response analyses in landfill practice are conducted using the equivalent linear method. Waste mass properties of importance in an equivalent linear seismic response analysis include the waste total unit weight (or mass density), the small-strain shear modulus, and equivalent linear shear modulus reduction and damping curves. Poisson's ratio may also be of interest if two-dimensional equivalent linear response analyses are to be conducted. Small strain shear modulus is generally established based upon shear wave velocity and unit weight or mass density. Waste mass shear strength and the interface shear strength of geosynthetic elements (or in the case of a geosynthetic clay liner, the in-plane shear strength) are essential properties in seismic analysis of waste containment systems for waste mass stability and cover system assessments. The properties of natural soil materials used in landfill construction are also important in a landfill seismic stability assessment but are not discussed herein.

Figure 4, from Zekkos et al. (2006), provides a family of typical unit weight profiles for MSW landfills that depend upon the amount of cover soil and compaction effort provided by the landfill operator during waste placement. Due to difficulties in measuring the waste unit weight in the

Seismic Design of Waste Containment Systems, Fig. 4 Typical MSW total unit weight vs. depth profiles (Zekkos et al. 2006)
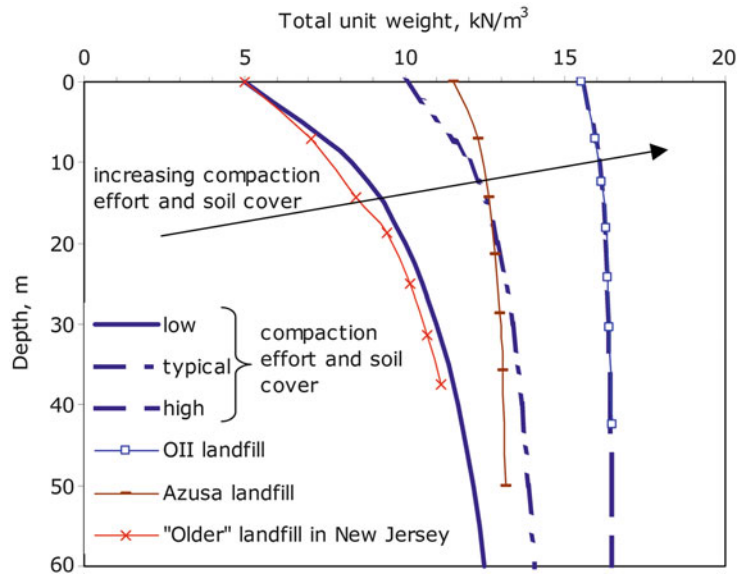
field, Fig. 4 or some other relationship for typical unit values is usually used in landfill seismic analysis and design rather than direct measurement, even for existing facilities.

Shear wave velocity and the small-strain modulus of MSW have been measured in the field and in the laboratory. Zekkos et al. (2008) and Yuan et al. (2011) report shear wave velocity and small-strain modulus values measured in laboratory testing of MSW reconstituted at three different ratios of refuse (MSW particles >20 mm, i.e., retained on a 20-mm sieve) to soil-sized (MSW particles <20 mm, i.e., passing a 20-mm sieve) material (note that these investigators used the same MSW and same composition ratios but different laboratory devices). However, unlike unit weight, waste mass shear wave velocity is readily measured in the field, e.g., using noninvasive surface wave techniques. Field data is more reliable than laboratory testing on reconstituted specimens, and case histories of site-specific shear wave velocity measurements for landfills are being reported in the literature with increasing frequency. Figure 5 provides a summary of shear wave velocity profiles for MSW landfills worldwide as compiled by Ramaiah (2013).

Equivalent linear shear modulus reduction and damping curves for MSW are available from

three sources: (1) from back analysis of strong ground motion records from the OII Landfill alone (e.g., Idriss et al. 1995), in some cases supplemented with cyclic laboratory testing of waste (Matasovic and Kavazanjian 1998); (2) from large-diameter cyclic triaxial testing by Zekkos et al. (2008) on reconstituted waste specimens from the Tri Cities Landfill for three different composition ratios; and (3) from large-scale laboratory simple shear testing by Yuan et al. (2011) on reconstituted waste specimens from the Tri Cities Landfill for the same three composition ratios employed by Zekkos et al. (2008). Figure 6 presents the equivalent linear shear modulus reduction (Fig. 6a) and damping (Fig. 6b) curves from Zekkos et al. (2008) from large-scale triaxial testing, termed the UCB data, and from Yuan et al. (2011) from large-scale simple shear testing, termed the ASU data. These two data sets are relatively consistent with each other and with the OII field data and represent the best available information on these property relationship.

Poisson's ratio values for MSW have been derived from field measurements of shear and compressional wave velocity at the OII Landfill (Matasovic and Kavazanjian 1998) and from laboratory measurements of axial and radial strain (Zekkos 2005). Poisson's ratio can also be
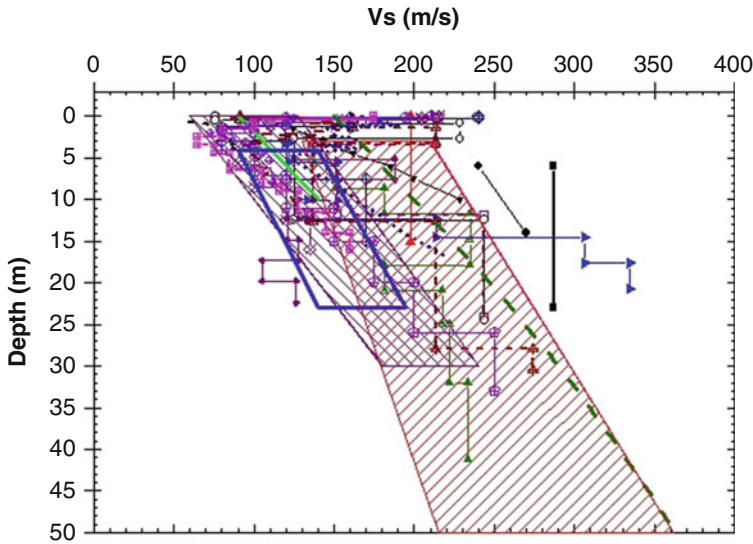
Legend	Landfill Name	Location	Method	Reference
■	Puente Hills landfill	California, USA	Down-hole	Earth Technology 1988
●	Puente Hills landfill	California, USA	Cross-hole	Earth Technology 1988
▲	Richmond landfill	California, USA	Down-hole	Sharma et al. 1990
▼	North west Regional Landfill Facility	Arizona, USA	Down-hole	Houston et al. 1995
◆	Live Oka landfill	Georgia, USA	SASW	Haker et al. 1997
▲	Sanifil landfill	Georgia, USA	SASW	Haker et al. 1997
▶	Bolton Raod landfill	Atlanta, USA	SASW	Haker et al. 1997
□	Olympic View Sanitary Landfill (OVSL)-Old	Washington, USA	SASW	Matasovic & Kavazanjian 2006
◇	OVSL-West	Washington, USA	SASW	Matasovic & Kavazanjian 2006
▲	OVSL -South	Washington, USA	SASW	Matasovic & Kavazanjian 2006
⊕	Villalba landfill	Madrid, Spain	SASW	Cuellar et al. 1998
⊕	Valdemingomez landfill	Madrid, Spain	SASW	Pereira et al. 2002
✳	Lasdehasas landfill	Madrid, Spain	SASW	Mateos 2006
- - -	Maldegem landfill	Ghent, Belgium	SASW	Van Impe & Bouzza 1996
- - ··	Site 1 (6-8 years old)	Melbourne, Australia	CSSASW	Bouzza & Kavazanjian 2000
——	Site 2 (7-8 years old)	Melbourne, Australia	CSSASW	Bouzza & Kavazanjian 2000
——	Spadina landfill	Saskatoon, Canada	MASW	Fleming et al. 2011
⊞	Gazhipur landfill	Delhi, India	SASW	Ramaiah et al. 2013
⊕	Okhla landfill	Delhi, India	SASW	Ramaiah et al. 2013

- - -	Recommended Vs profile in absence of site specific data	Kavazanjian et al. 1995
▨	Range of Vs for Southern California Landfills	Kavazanjian et al. 1996
▨	Vs range for Tri-Cities, Altamont and Redwood landfills	Zekkos 2005
▭	Vs range for Michigan landfills (13 profiles from 4 landfill sites) Sahadewa et al. 2011	

Seismic Design of Waste Containment Systems, Fig. 5 Shear wave velocity at MSW landfills worldwide (Ramaiah 2013)

Seismic Design of Waste Containment Systems, Fig. 6 Equivalent linear property relationships for waste specimens reconstituted at different composition ratios (Yuan et al. 2011): (**a**) shear modulus reduction; (**b**) damping
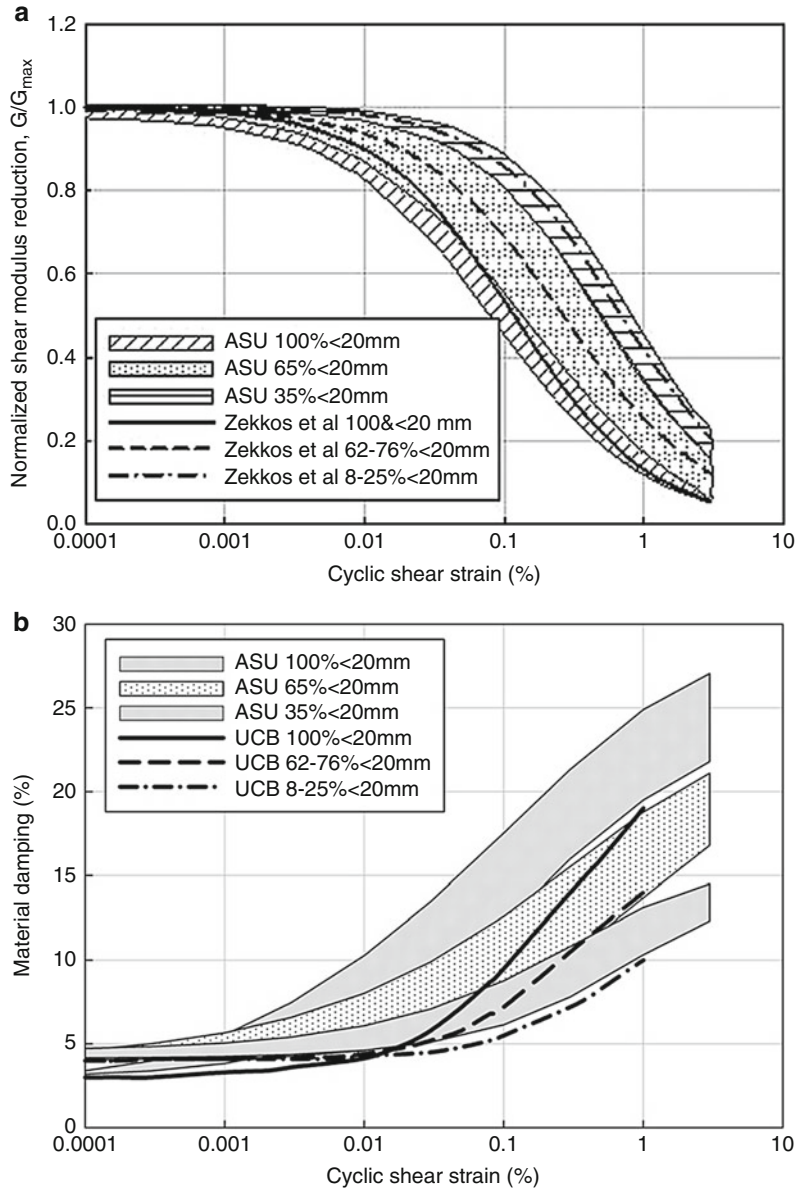
inferred from the field measurements of the lateral earth pressure at rest in MSW reported by Dixon et al. (1999). The data on Poisson's ratio from these studies is scattered, both within and among the studies, but it appears that values on the order of 0.2–0.4 are appropriate for unsaturated MSW.

Due to difficulties associated with conventional laboratory and field testing, studies of MSW shear strength have employed back

analyses of landfill failures under static loading to develop shear strength envelopes. Kavazanjian et al. (1996) employed back analysis of both landfill failures and steep stable landfill slopes to develop a bilinear strength envelope in which MSW was assigned a cohesion of 25 kPa at low normal stresses and a friction angle of 33° at higher confining pressures. Based upon back analyses of failed landfill slopes, Eid et al. (2000) proposed that typical MSW shear strength could

be characterized by a linear failure envelope with a friction angle of 35° and a cohesion intercept and that the lower- and upper-bound shear strength should be characterized using zero and 50-kPa cohesion, respectively. Based upon laboratory triaxial and direct testing on relatively large-sized specimens (e.g., 150 mm-diameter triaxial specimens), Bray et al. (2009) suggested that MSW shear strength was dependent on the normal stress and characterized by a cohesion of 15 kPa and a friction angle that was equal to 36° at a normal stress less than or equal to one atmosphere and that decreased by 5° for every log cycle of normal stress. Both Bray et al. (2009) and Kavazanjian et al. (2013b) have noted the anisotropic structure and shear strength of MSW, with the strengths reported above applicable to the horizontal plane, the weakest plane in a waste mass due to the preferred alignment of the long axes of waste constituents in this direction, and with substantially greater shear strength on non-horizontal planes that cut across the long axes of the constituents. Based upon analysis of landfill performance in the Northridge earthquake, Augello et al. (1995) suggested that the shear strength of MSW subject to seismic loading was greater than the static strength. Bray et al. (2009) suggest that the dynamic shear strength of MSW is at least 20 % greater than the static shear strength. However, in most seismic evaluations, the static shear strength has been found to be sufficient to demonstrate landfill seismic stability.

The data on material properties for waste materials other than typical MSW is rather sparse. Hazardous waste in modern landfills is generally stabilized and containerized prior to placement in a landfill. The containers are then stacked and backfilled with soil. In older hazardous waste landfills, the waste may have been containerized, typically in partially filled steel drums, without stabilization. Cone penetrometer testing at a 20–30-year-old hazardous waste landfill where the waste was placed in partially filled steel drums without stabilization showed no evidence that the drums were still intact or of voids (as would be expected at the top of a partially filled drum) (Matasovic et al. 2005).

Furthermore, hazardous waste disposed of in landfills is often a hazardous material (e.g., pesticide residue, organic solvents) or mixed waste (e.g., potentially radioactive construction and demolition debris) in a soil matrix. Hence, it is common to assume that the properties of the waste mass at a hazardous waste site are the same as the properties of the backfill soil (or of the native soil at a hazardous material spill site) (Kavazanjian and Matasovic 2001).

Geosynthetic Properties: The interface shear strength of geosynthetic materials and the in-plane internal strength of GCLs can vary widely, depending upon the type of geosynthetic material, the nature of the material in contact with the geosynthetic material, and for GCLs the internal reinforcement and moisture content (Zornberg et al. 2005; McCartney et al. 2009). Furthermore, many geosynthetic materials have significantly different peak and large-displacement shear strengths.

Hydrated unreinforced or poorly reinforced GCLs can have extremely low in-plane shear strengths, in some cases as low as the hydrated strength of their sodium bentonite core (e.g., a friction angle, ϕ, equal to 4° at high normal stresses). However, such materials are rarely used in landfill construction. Hydrated reinforced GCLs can have a substantial peak internal shear resistance, sometimes in excess of $\phi = 30°$, and still maintain a relatively high strength when subject to large deformation (i.e., $\phi > 15°$), depending on the quality of the reinforcement.

Smooth geomembranes in contact with cohesive soils (e.g., compacted clay) can have very low shear strength ($\phi =$ on the order of 10°), but textured geomembranes can provide substantially higher interface strengths, even when subject to a large displacement (ϕ sometimes on the order of 16–20° for the large displacements). Granular materials in contact with smooth geomembranes may have an interface strength exceeding 70 % of the shear strength of the granular soil, while the shear strength of a granular material in contact with a textured geomembrane may approach the shear strength of the soil. Nonwoven geotextile placed against a smooth geomembrane will have a relatively low shear

strength but can have a peak interface strength exceeding that of most natural soils if placed against a textured geomembrane. Woven geotextile placed against a smooth geomembrane will typically have a higher interface strength than a nonwoven geotextile against a smooth geomembrane. Woven geotextiles are not typically placed against textured geomembranes. Geonets placed upon a geomembrane (textured or smooth) may have a low interface strength ($\phi = 10°$) but can be heat bonded to a nonwoven geotextile and then placed against the geomembrane if a higher interface strength is desired.

Minimum required interface strengths are typically established by design analyses and then specified in design documents. Laboratory testing is often conducted during design to demonstrate that the required strengths are achievable with available materials. The geosynthetic materials used in construction are then generally subject to conformance testing prior to the start of construction and to quality assurance testing during construction to demonstrate that the specific materials used in construction can provide the specified interface shear strength.

Other geosynthetic properties of importance to seismic design include the tensile strength and the strain at yield and at failure of the geosynthetic elements of the containment system. For advanced analyses, the shear and axial stiffness of geosynthetic material elements may be required.

Design Analyses

Decoupled Analysis: Most seismic analyses of landfill systems conducted in practice today employ a decoupled approach in which the response of the landfill to the seismic loading is considered separately from integrity assessments for the waste mass and containment system components. The seismic response analysis is conducted assuming that the containment system components and waste mass maintain their integrity, that there is no relative displacement at geosynthetic interfaces, and that no instabilities develop. The most common type of seismic response analysis is a one-dimensional equivalent linear analysis of a representative column or columns for the waste mass. For important projects, nonlinear one-dimensional analysis and two-dimensional equivalent linear analyses are sometimes performed. Nonlinear two-dimensional analyses are primarily used for research purposes. The integrity assessment is usually based upon the calculated seismic response, permanent seismic displacement calculated in a Newmark-type seismic displacement analysis, and empirical rules on allowable displacement for the various elements of the containment system. The Newmark displacement analysis is typically conducted using shear stress time histories from the seismic response analysis and a yield acceleration calculated using conventional limit equilibrium analysis. The standard of practice is to use the large-displacement static shear strength of any interface engaged in the limit equilibrium analysis. If more than one interface is aligned with a slip surface engaged in the limit equilibrium analysis, the large-displacement static strength of the interface with the lowest static peak strength is employed.

Due to the interrelationship between the yield acceleration, shear stress time history, and calculated seismic displacement, multiple potential sliding mechanisms must be checked in a landfill seismic stability assessment. Potential sliding mechanisms include veneer failures of the landfill cover along interfaces with containment system elements (both geosynthetic and natural soil) and between the cover soil and the waste, circular surfaces through the waste mass and/or the foundation, and sliding block surfaces along planes of weakness in waste, including along interfaces with geosynthetic and natural soil containment system elements and other potential planes of weakness. Figure 7 provides an example of a landfill seismic stability analysis in which multiple potential sliding surfaces were evaluated (Kavazanjian et al. 2013b). In this case, the horizontal planes of weakness were continuous layers of daily and interim covers hypothesized to be present within the waste mass. The calculated seismic displacement on each of these failure surfaces is considered independently in the integrity assessment.
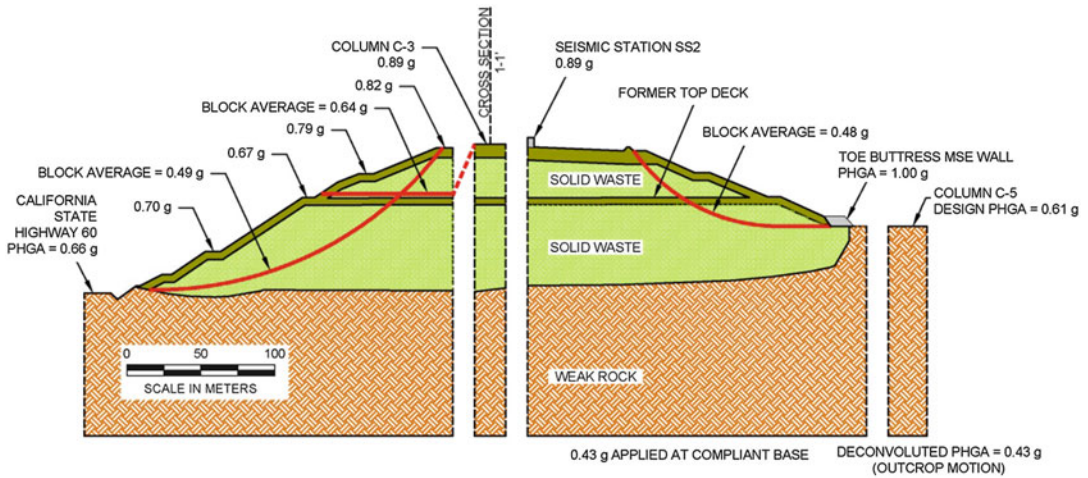
Seismic Design of Waste Containment Systems, Fig. 7 Potential sliding mechanisms assumed in seismic stability assessment for the OII Landfill: g-values represent the peak average acceleration for each potential sliding mechanism (Kavazanjian et al. 2013b)

Due to the conservatism associated with the decoupled approach and the use of large-displacement static shear strength, the calculated permanent seismic displacement from this type of analysis is considered merely an index of seismic performance and not the expected permanent seismic displacement along the hypothesized sliding surface. A calculated seismic displacement of 150 mm or less is typically considered indicative of minimal permanent seismic displacement in the design earthquake, and calculated displacements of up to 1 m are considered indicative of limited displacement. Calculated displacements of greater than 1 m are considered indicative of the potential for large, possibly uncontrolled displacements in the design earthquake.

Table 1 summarizes typical values of acceptable calculated seismic displacement for non-geosynthetic components of landfill containment systems from a decoupled analysis, along with the anticipated duration for the interim and final repairs of these components. For the geosynthetic elements of waste containment systems, 150 mm of calculated displacement is generally considered to be indicative of no damage, and up to 3 m of displacement may be considered acceptable for geosynthetic elements in which

damage can be readily detected in a post-earthquake inspection and readily repaired. It should be noted that these general guidelines are based primarily upon back analyses of the performance of slopes, embankments, and landfills in earthquakes using conventional decoupled analysis. If a more sophisticated seismic analysis is conducted, e.g., a two-dimensional nonlinear coupled analysis of seismic response and displacement, these guidelines may not be appropriate, particularly with respect to damage to the geosynthetic elements of the containment system.

Coupled Analysis: Coupled analysis of landfill seismic response and waste containment system integrity is rare and complex due to difficulties in modeling slip at interfaces and geosynthetic element stress–strain behavior. Bray and Rathje (1998) developed a one-dimensional model for landfill response that included slip at the interface. Kavazanjian et al. (2012) describe a fully coupled two-dimensional model for performance-based seismic design of waste containment systems that accounts for both slip at an interface and geosynthetic element stress–strain behavior. The interface model allows for a post-peak decrease in the interface shear strength. Geosynthetic

Seismic Design of Waste Containment Systems, Table 1 MSW landfill seismic design criteria and performance standards for decoupled analysis (Kavazanjian et al 1998)

Cover system component	Design criteria and performance standard	Interim remediation to restore compliance	Repair to pre-earthquake condition
Final cover			
Soil monocover on side slopes	150 mm of soil deformation; partial failure contained on site	3 months to strip vegetation and regrade and recompact areas of cracking	12 months to restore vegetation
Landfill gas control			
Collection wells	Up to 25 % of wellheads broken	1 month to route headers around broken wellheads	12 months to repair/ replace broken wellheads
Headers	Up to 25 % of header pipes cracked or broken	1 month to bypass broken header pipes	3 months to repair/ replace broken headers
Vacuum pumps	Power loss; no structural damage	None required	1 month to restore off-site power
Leachate transmission pipes	Acceptable breakage of pipes with double containment	1 month to bypass broken pipes	3 months to repair broken pipes
Surface water management			
Conveyance systems (bench channels, down drains, culverts)	Cracking and up to 300 mm of displacement	2 months to completely restore surface pathways	9 months to replace/ rebuild surface pathways
Sedimentation basin	Minor cracking of concrete	2 weeks to 1 month to patch the cracks	9 months to rebuild the basin (if needed)
Access roads	300 mm displacement (cracking)	2 months to patch the cracks	12 months for full repair

elements of the waste containment system are modeled as structural beam elements with zero moment of inertia (to allow for buckling) and with a parabolic stress–strain model, allowing for explicit calculation of forces and strains in the geosynthetic elements of the containment system. Parametric studies using this model suggest that in some cases the geosynthetic elements of a waste containment system may not sustain significant damage even though a decoupled Newmark analysis results in a calculated displacement of greater than 150 mm. In particular, when the interface shear strength on the top of a geosynthetic element is less than the interface shear strength on the bottom of the geosynthetic element and there are no irregularities (penetrations, seams) in the element, the interface may be able to sustain large sliding displacements without incurring damage to the geosynthetic element, i.e., without exceeding its

tensile strength. However, a back analysis of the geomembrane tears observed at two separate locations at the Chiquita Canyon Landfill (cited previously) has suggested that damage can occur to geosynthetic waste containment system elements when the calculated Newmark displacement is less than 150 mm and when there are irregularities in the containment system, e.g., due to strain concentrations around patches in the geomembrane at locations where coupons were removed for destructive testing as part of the construction quality assurance (CQA) program (Kavazanjian et al. 2013).

Performance-Based Design: Figure 8 illustrates the strain in the geomembrane for a landfill side-slope liner as predicted by the performance-based seismic design model described by Kavazanjian et al. (2012). While this model has yet to be fully validated, back analysis of the performance of the Chiquita Canyon Landfill in
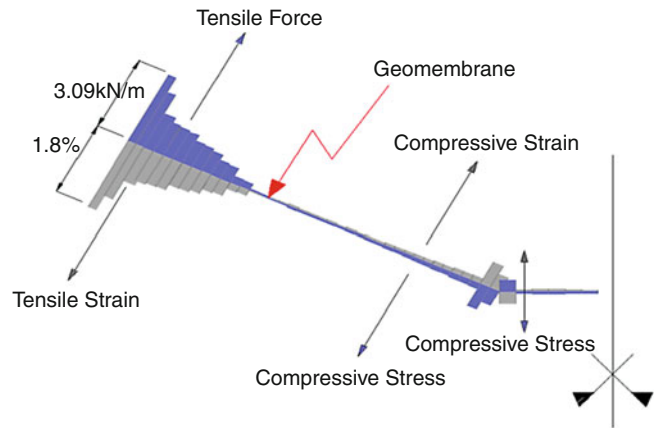
the Northridge earthquake (Kavazanjian et al. 2013a) indicates that strains due to the anchoring of the geomembrane at the crest of the slope and strain concentration factors due to seams in the geomembrane (Giroud 2005) at locations where coupons were removed for CQA testing must be considered in order to predict the tears observed in the liner following the earthquake. These recent findings suggest that, at a minimum, new guidelines need to be developed for acceptable locations for removal of coupons for CQA testing and for minimizing other irregularities in the geosynthetic elements of the containment system. However, ultimately performance-based design using fully coupled models that consider interface and geosynthetic material behavior to explicitly predict strain and forces in geosynthetic elements of modern waste containment systems is likely to become the standard of practice, at least for critical facilities in areas of high seismicity.

Summary

Seismic design of waste containment systems is an important consideration with respect to both maintaining the protection afforded by the containment system for human health and the environment and providing safe and secure capacity for waste disposal during earthquake response and recovery operations. Important engineering considerations with respect to the seismic design of waste containment systems include the seismic response of the waste mass and methods for evaluating the integrity of containment system components subject to seismic loading. Observations of landfill performance in earthquakes indicate that the waste mass may amplify free-field ground motions and that containment system elements, including geosynthetic liner system components and soil cover, are susceptible to earthquake-induced damage. Waste mass seismic response depends to a large extent upon the waste mass material properties, including the small-strain stiffness, strain-softening behavior, and internal damping in the waste. For landfills with geosynthetic liner and cover systems, the interface and in-plane shear strength of the geosynthetic components is also an important consideration with respect to seismic response. Most containment system performance analyses employed in engineering practice are decoupled analyses in which waste mass seismic response is calculated using the equivalent linear method and the performance of liner system components is assessed based upon the calculated seismic displacement from a Newmark-type analysis. The seismic performance of containment system components is then assessed based upon empirical rules for allowable calculated displacement. While more sophisticated fully coupled performance-based analyses in which the forces, strains, and displacements induced by the earthquake in containment system components are explicitly calculated are possible, such analyses have yet to make their way into engineering practice.

Cross-References

References

Augello AJ, Matasovic N, Bray JD, Kavazanjian E Jr, Seed RB (1995) Evaluation of solid waste landfill performance during the Northridge earthquake. In: Earthquake design and performance of solid waste landfills. Geotechnical special publication no. 54. ASCE, Reston, VA, pp 17–50

Bray JD, Rathje EM (1998) Earthquake-induced displacements of solid-waste landfills. J Geotech Geoenviron Eng ASCE 124(3):242–253

Bray JD, Zekkos D, Kavazanjian E Jr, Athanasopoulos GA, Reimer M (2009) Shear strength of municipal solid waste. J Geotech Geoenviron Eng ASCE 135 (6):709–722

Buranek D, Prasad S (1991) Sanitary landfill performance during the Loma Prieta earthquake. In: Proceedings of the second international conference on recent advances in geotechnical earthquake engineering and soil dynamics. University of Missouri, Rolla, pp 1655–1660

Dixon N, Jones DRV, WhittleRW (1999) Mechanical properties of household waste: in situ assessment using pressuremeters. In: Proceedings of the Sardinia'99: seventh international waste management and landfill symposium, vol III, pp 453–460

Eid HT, Stark TD, Douglas WD, Sherry PE (2000) Municipal solid waste slope failure. 1: waste and foundation properties. J Geotech Geoenviron Eng, ASCE 126 (5):397–407

EMCON (1994) Northridge earthquake seismic evaluation, Chiquita Canyon landfill. Technical report prepared for Laidlaw Environmental Systems by EMCON Associates, San Jose

Giroud J P (2005) Quantification of geosynthetic behavior. Geosynt Int 12(1):2–27, Special Issue on the Giroud Lectures

Harder LS Jr (1991) Performance of earth dams during the Loma Prieta earthquake. In: Proceedings of the second international conference on recent advances in geotechnical earthquake engineering and soil dynamics. University of Missouri, Rolla, pp 1673–1691

Idriss IM, Fiegel G, Hudson MB, Mundy PK, Herzig R (1995) Seismic response of the operating industries landfill. Earthquake design and performance of solid waste landfills. Geotechnical special publication no. vol 54. ASCE, Reston, VA, pp 83–118

Johnson ME, Lundy J, Lew M, Ray ME (1991) Investigation of sanitary landfill slope performance during strong ground motion from the Loma Prieta Earthquake of October 17, 1989. In: Proceedings of the second international conference on recent advances in geotechnical earthquake engineering and soil dynamics. University of Missouri, Rolla, pp 1701–1706

Kavazanjian E Jr, Matasovic N (2001) Seismic design of mixed and hazardous waste landfills. In: Proceedings of the fourth international conference on recent advances in geotechnical earthquake engineering and soil dynamics. University of Missouri-Rolla, Rolla (on CD ROM)

Kavazanjian E Jr, Matasovic N, Stokoe KH II, Bray JD (1996) In-Situ shear wave velocity of solid waste from surface wave measurements. In: Proceedings of the 2nd international congress environmental geotechnics, Osaka, Balkema, vol 1, pp 97–104

Kavazanjian E Jr, Matasovic N, Caldwell J (1998) Damage criteria for solid waste landfills. In: Proceedings of the sixth US national conference on earthquake engineering, EERI (on CD ROM)

Kavazanjian E Jr, Arab MG, Matasovic N (2012) Performance based design for seismic design of geosynthetics-lined waste containment systems. In: Proceedings of the second international conference on performance-based earthquake geotechnical engineering, State of the Art – State of the Practice Volume, Keynote Lecture No. 19, University of Catania, p 23 (on CD ROM)

Kavazanjian E Jr, Arab M, Matasovic N (2013a) Performance of Two Geosynthetics-Lined Landfills in the Northridge Earthquake. In: Proceedings of the seventh international conference on case histories in geotechnical engineering, Missouri University of Science and Technology, p 14 (on CD ROM)

Kavazanjian E Jr, Matasovic N, Bachus RC (2013b) The 11th Peck lecture: pre-design geotechnical investigation for the OII superfund site landfill. J Geotech Geoenviron Eng. ASCE 139(10): 1839–1863

Koerner RL, Soong T-Y (2000) Stability assessment of ten large landfill failures. In: Advances in transportation and geoenvironmental systems using geosynthetics. Geotechnical special publication no. 103. ASCE, Reston, VA, pp 1–38

Matasovic N, Kavazanjian E Jr (1998) Cyclic characterization of oII landfill solid waste. J Geotech Geoenviron Eng ASCE 124(3):197–210

Matasovic N, Kavazanjian E Jr, Anderson RL (1998) Performance of solid waste landfills in earthquakes. Earthquake Spectra EERI 14(2):319–334

S

Matasovic M, Kavazanjian E Jr, De A, Dunn RJ (2005) - CPT-based seismic stability assessment of a hazardous waste site. Int J Soil Dynam Earth Eng 26(2):201–208

McCartney JS, Zornberg JG, Swan RH (2009) Analysis of a large database of GCL-Geomembrane interface shear strength results. J Geotech Geoenviron Eng ASCE 134(2):209–223

Mitchell JK, Seed RB, Seed HB (1990) Kettleman hills waste landfill slope failure. I: liner-system properties. J Geotech Eng ASCE 116(4):647–668

Orr WR, Finch MO (1990) Solid waste landfill performance during Loma Prieta Earthquake. In: Geotechnics of waste fills: theory and practice, vol 1070, Special technical publication. ASTM International, Conshohocken, pp 22–30

Ramaiah BJ (2013) Static and seismic stability assessment of municipal solid waste landfills. Ph.D dissertation, Department of Civil Engineering, Indian Institute of Technology Delhi, Hauzkhas

Seed RB, Mitchell JK, Seed HB (1990) Kettleman hills waste landfill slope failure. II: stability analyses. J Geotech Eng 116(4):669–690

Yuan P, Kavazanjian E Jr, Chen W, Seo B (2011) Compositional effects on the dynamic properties of municipal solid waste. Waste Manag 31:2380–2390

Zekkos Zeccos D P (2005) Evaluation of static and dynamic properties of municipal solid-waste. PhD thesis, Department of Civil and Environmental Engineering, University of California, Berkeley

Zekkos D, Bray JD, Kavazanjian E Jr, Matasovic N, Rathje E, Reimer M, Stokoe KH III (2006) Unit weight of municipal solid waste. J Geotech Geoenviron Eng ASCE 132(10):1250–1261

Zekkos D, Bray JD, Riemer MF (2008) Shear modulus and material damping of municipal solid waste based on large-scale triaxial testing. Can Geotech J 45:45–58

Zornberg JG, McCartney JS, Swan RH Jr (2005) Analysis of a large database of GCL internal shear strength results. J Geotech Geoenviron Eng ASCE 131(3):1–14

Seismic Event Detection

Jens Havskov[1] and Gerardo Alguacil[2]
[1]Department of Earth Science, University of Bergen, Bergen, Norway
[2]Instituto Andaluz de Geofísica, University of Granada, Granada, Spain

Synonyms

Earthquakes; Seismic events; Seismic signals; Seismogram; Volcanic events

Introduction

Seismic stations record signals continuously; however, the main interest for the user is the recordings of seismic events. The events will normally be contained in a very small fraction of the actually recorded signal (Fig. 1).

As it is seen in this 24 h seismogram, there are only a few sections with real seismic events and a lot of other wiggles which is just noise. The more or less continuous noise seen throughout the seismogram is the earth's background noise (see below and Figs. 2 and 4). In the old days when all the signals were recorded on paper (on seismograms), the events were detected easily by eye, particularly for small local events with high-frequency content. This method is sometimes also used with today's digital recordings which can be displayed on screen or printed on paper. But the majority of event detection is done automatically, either in real time to limit the amount of data recorded or off-line when large data sets have been collected, e.g., from field experiments. Over the years a few simple and very many sophisticated algorithms have been developed to automatically detect a "real" signal in the presence of noise. Few, if any, have been developed that can beat the human eye, which still is used to evaluate the efficiency of the detector. The use of a trigger or a detector can have two purposes:

1. Determine the onset of real seismic events in a large data set. These onset times will be used for further processing.
2. Extract out the waveform time segments of a larger data set which contains seismic events, either in a field recorder to limit the amount of data recorded or to extract out just the useful events for processing.

In this entry, the most common event detection methods will be described. An overview of detection methods is found in Sharma et al. (2010).
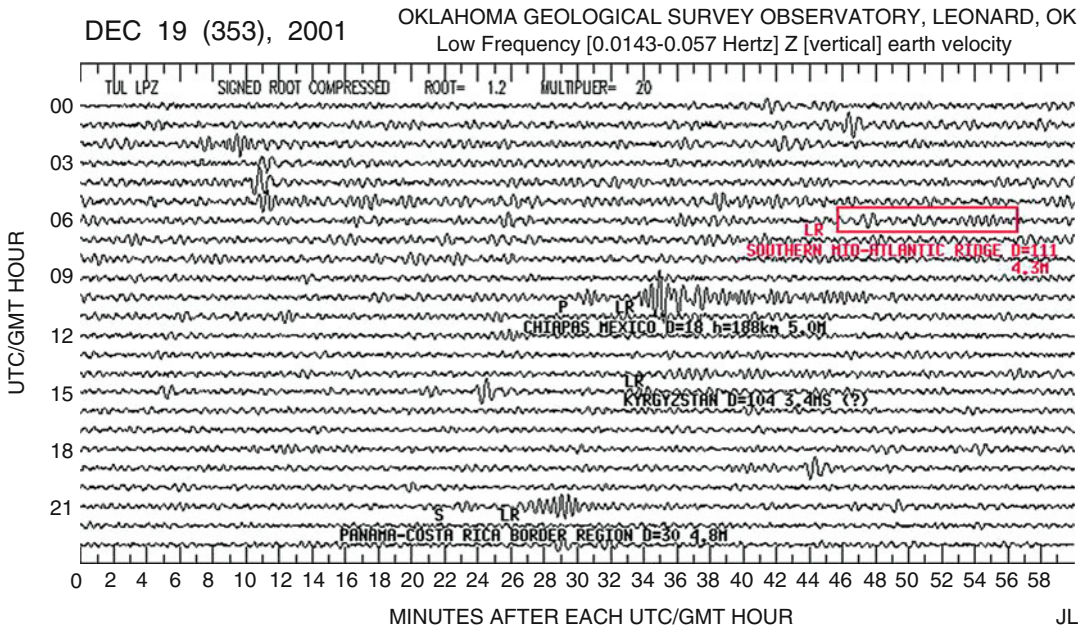
Seismic Event Detection, Fig. 1 A seismogram for a 24 h period. Detected events are indicated with the name of the region where the earthquake occurred, *D* is the distance from the station to the event in degrees, *h* is the depth of the earthquake, and *M* is the magnitude (Figure from www.okgeosurvey1.gov/seis/20011219.gif)
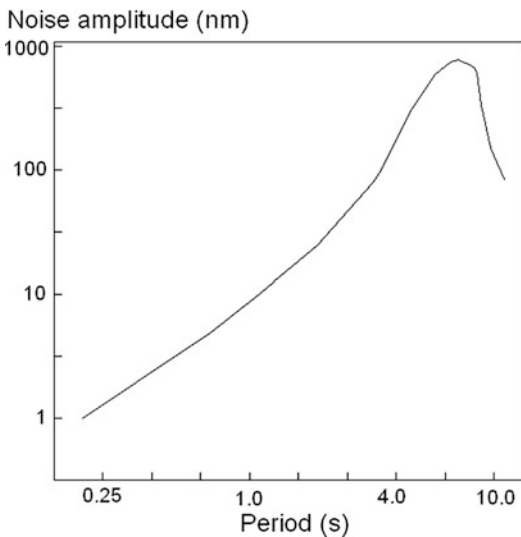
Seismic Event Detection, Fig. 2 Typical amplitude of the microseismic background noise as a function of period(s) (The data for the plot are taken from Brune and Oliver 1959)

Examples of Seismic Signals

Detecting real seismic events in a continuous data stream can be quite a challenge, particularly if there is nonstationary noise like noise from cars passing by or wind gusts and in some cases it is virtually impossible to distinguish between man-made noise and real seismic events.

In addition, there is an ever present so-called microseismic background noise (see Figs. 4 and 1). This noise is caused by standing waves in the ocean and wave action near the coasts. It is present all over the world, strong near the coast, and weaker far from the coast. It has a maximum amplitude at a period of about 4–6 s. Figure 2 shows a typical curve with the noise as a function or period.

Seismic events can have quite a range of appearances and frequency content depending on the type of seismic sensor used and the

**Seismic Event Detection,
Fig. 3** Seismogram of
a local event (M = 1.5) at
5 stations. Sensors are short
period and the traces are not
filtered. Each trace is auto
scaled. The first arrival
phase *P* and second arrival
phase *S* are indicated. The
epicentral distances
(distance from the
earthquake to the station
measured along the
surface) are given above the
traces to the *right* (The
figure is from Havskov and
Ottemöller 2010)
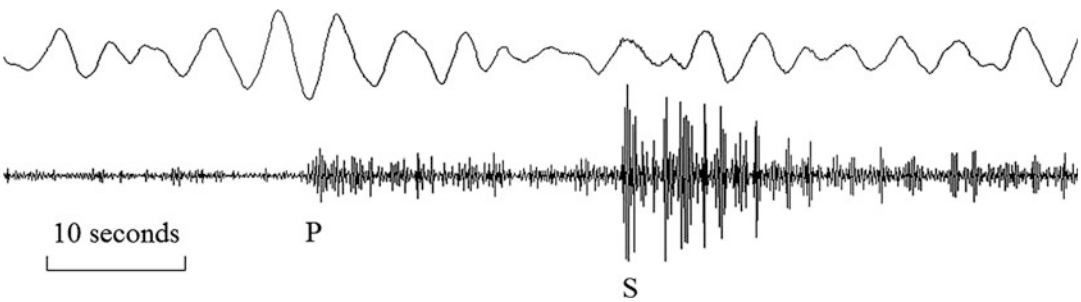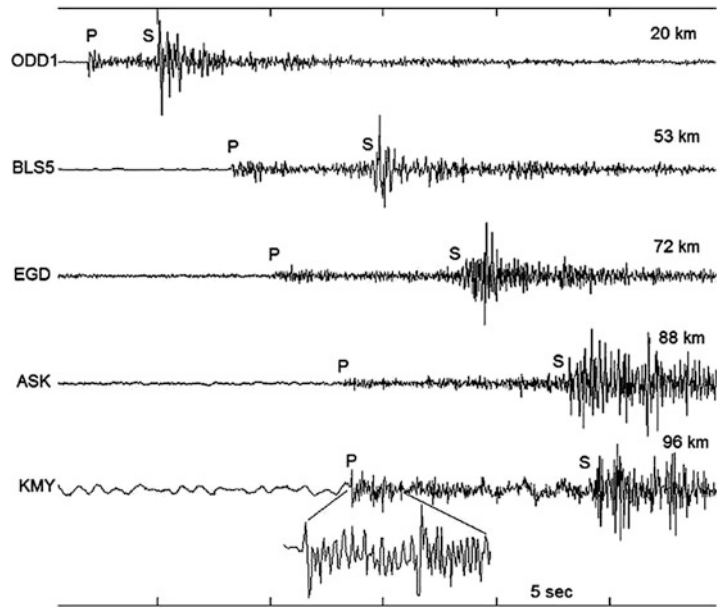
Seismic Event Detection, Fig. 4 A magnitude 1.5 event recorded on a broadband station. The *top* trace is the original trace and the *bottom* trace has been filtered from 5 to 10 Hz

distance to the event and its magnitude. Essentially there are three types of seismic sensors: short-period measuring ground velocity (frequencies above 1 Hz), broadband measuring ground velocity (frequencies typically from 0.01 to 50 Hz), and accelerometers measuring ground acceleration from 0 to 100 Hz. Some examples will be given of seismic events.

Local small earthquake recorded with short-period sensors (Fig. 3). The signals are high frequency, and P- and S-phases are clearly seen although the P-phase is getting close to the noise level for station ASK.

Small event recorded on a broadband station (Fig. 4). The high background noise level, the

microseismic noise, is the harmonic signal dominating the record. It has a period of 4.5 s. A small high-frequency event will "drown" in this noise and can practically not be seen on the original record. By filtering the signal from 5 to 10 Hz, the local event is clearly seen. This illustrates that event triggering often will have to work on filtered traces. So how did humans detect this kind of event before digital recording? Simply by recording the signals in two different frequency bands, short period above 1 Hz and long period below 0.1 Hz, and thereby filtering out the noise around the microseismic peak (Fig. 2).

Distant event recorded on a broadband station (Fig. 5). This magnitude 6.2 event is recorded at

Seismic Event Detection, Fig. 5 Recording of an earthquake at a large distance (4,200 km). The insert above *left* shows a zoom of the noise before the *P*
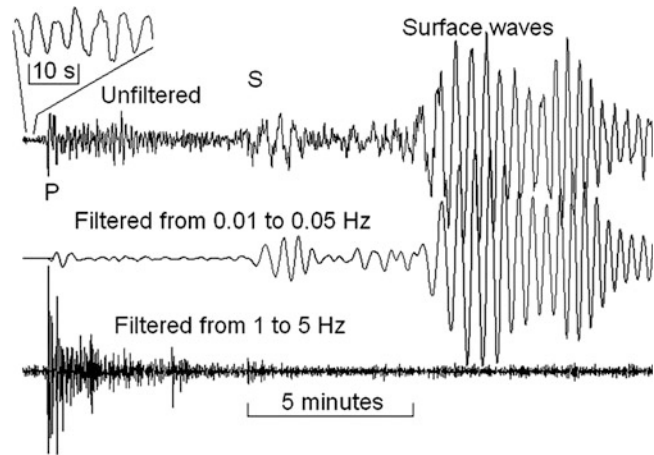
a distance 4,200 km. It is seen that a large distant event has much lower frequencies than small local events, partially due the source nature of earthquakes and partly due to filtering out the high frequencies along the path.

The unfiltered trace shows, before the P, the microseismic noise. For smaller events the P might drown in the noise. The microseismic noise is filtered away in the middle trace and the phases appear very clear. The bottom trace is filtered so the signal looks like it would have been recorded on a short-period seismograph. In this case a detector would not be able to find the true extent of the seismic event since the long-period S- and surface waves are not seen.

Volcanic events (Fig. 6). Volcanoes present a special problem since the signals from the events often are very emergent. Figure 6 shows the most common types.

The tectonic event is like any other earthquake and is easy to detect. The shallow events are very harmonic and monochromatic and of sufficient amplitude to be easy to detect. The surface event can be very difficult to detect automatically since it is very emergent and might be mistaken for noise. The harmonic tremor is a continuous series of harmonic events and cannot be detected.

Many events close in time: Figure 7 shows swarm activity of small earthquakes in a 24 h period. The human eye can easily detect the individual events while an automatic trigger would have problems with separating the events.
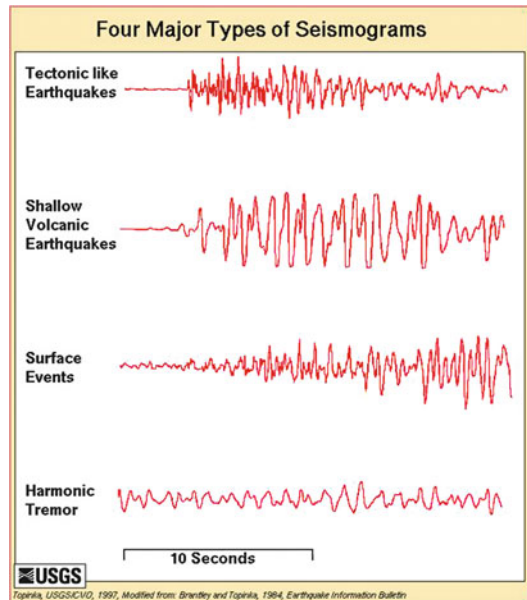
Seismic Event Detection, Fig. 6 Major types of seismic signals from volcanoes (From http://vulcan.wr.usgs.gov/Imgs/Gif/Monitoring/Seismic/quakes.gif)

Trigger Methods

A trigger or detector is an algorithm that checks the signal for variations that could indicate an event. There are several trigger algorithms available, some very sophisticated using neural networks (e.g., Romeo et al. 1995) and pattern recognition (e.g., Köhler et al. 2010). In the hands of an expert, they can significantly improve the event detections/false trigger ratio,

IG.AEP2..HHZ
IG.AEP2..HHZ
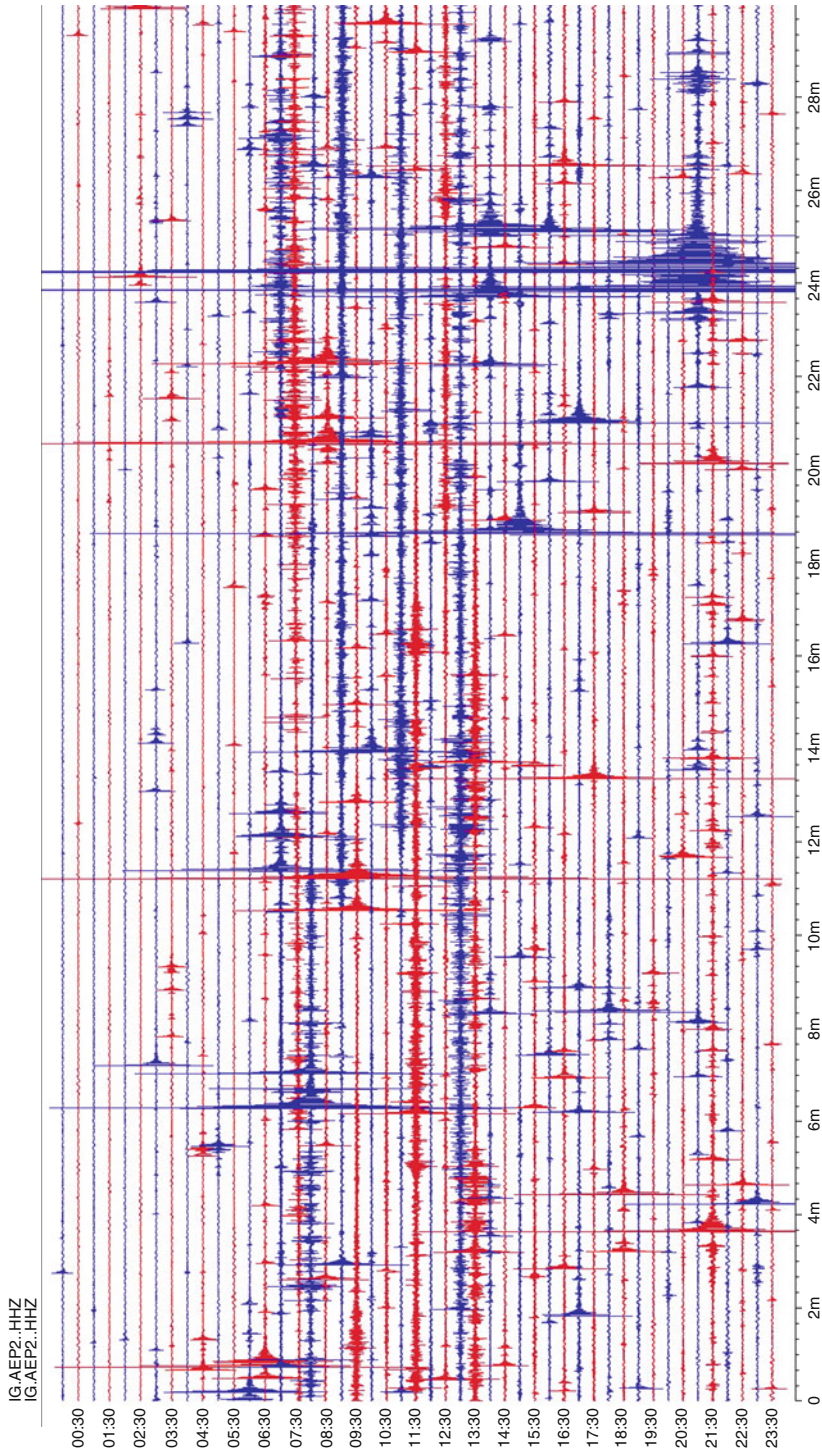
Seismic Event Detection, Fig. 7 Swarms of earthquakes. A 24 h record on 5 February 2013 of a temporary broadband station during the seismic swarm of Torreperogil (Southern Spain)

Seismic Event Detection,
Fig. 8 An example of level
trigger. The plot shows
a horizontal component of
the accelerogram of a small
local earthquake. The
trigger level is marked with
two horizontal lines. In this
case the de-trigger level is
the same as the trigger
level. Usually the triggers
of the three components are
combined, and an event is
declared if at least two or
three channels are triggered
within a given time window
(e.g., 1 s)
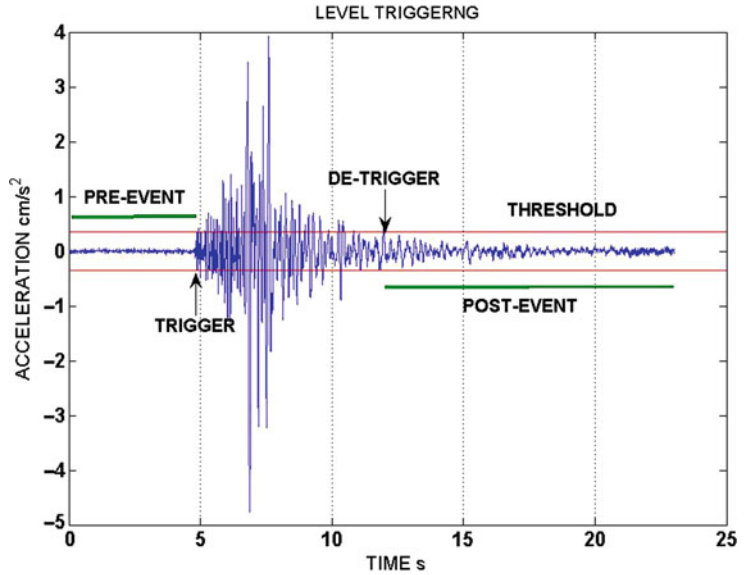
particularly for a given type of seismic events. However, the sophisticated adjustments of operational parameters to the actual signal and seismic noise conditions, at each seismic site, which these triggers require, have often been proven unwieldy and subject to error. Hence, for practical purposes, only two types of triggers are widely used, the level trigger and the *STA/LTA* trigger (see below), and only those will be described here.

The level trigger: This trigger is also called the amplitude threshold trigger and simply searches for any amplitude exceeding a preset threshold. The start of the event is declared when this threshold is reached and the end of the event is declared when the level reaches the de-trigger level, usually smaller than the trigger level. When this type of trigger is installed in a recorder, that only shall record the real events, the recording window will be extended with a pre-event time before the trigger and a post-event time after the trigger (Fig. 8). This is to ensure that the whole event is recorded.

The level trigger has some obvious problems. If DC (a constant voltage) is present in the signal and not removed before triggering, it can prevent triggering or keep the instrument in continuous trigger mode. If no averaging or filtering is done,

the level trigger will also trigger on a single high value (a spike), which obviously is not a seismic signal.

This algorithm is often used in strong-motion seismic instruments (recording large ground accelerations), where high sensitivity is not an issue and where consequently man-made and natural seismic noise is not critical. The level trigger is now largely replaced by the *STA/LTA* trigger (see next section), since computer power no longer is an issue. However, many instruments still have the level option.

The short-term average – long-term average trigger (STA/LTA) is the most frequently used trigger algorithm. A single channel of seismic signal is typically processed as follows: The signal is band-pass filtered, and the absolute average *STA* (short-term average) over the STA time window is determined. Typically, the *STA* time window is 0.5 s for local earthquakes. A short *STA* time makes the trigger more sensitive to short-term variations in the signal, while a longer *STA* is better at averaging out short-term fluctuations and it can be compared to a low-pass filter. If there are spikes in the signal, a long *STA* must be used in order to average out the interference. This will of course reduce the sensitivity to short-lasting signals, but if a longer-lasting signal is the

objective, *STA* can be as long as the duration of the main P-wave train. The most common mistake in setting the *STA* is to use a too short value.

The same filtered signal is also used to calculate the *LTA* (long-term average) over the *LTA* time window, which is typically 50–500 s. Thus, *LTA* will give the long-term background signal level, while the STA will respond to short-term signal variations. The ratio between *STA* and *LTA* is constantly monitored, and once it exceeds a given threshold, the trigger level, the start of an event is declared for that trace. The *LTA* should be short enough to adapt to slow changes of the background noise and long enough to avoid reducing the sensitivity to triggering on low-amplitude emergent signals. This can be a problem with very emergent signal where the STA might increase fast enough to prevent triggering (see the surface event in Fig. 6). Since the natural background noise usually changes very slowly, it is generally better to use a too long than a too short *LTA* time. If *LTA* is not frozen during triggering (see next section), the *LTA* time must be at least as long as the event duration to prevent premature de-triggering. Using a too short *LTA* time combined with a high trigger ratio might even prevent triggering completely since the *LTA* is adjusted upward so fast that *STA/LTA* never exceeds the trigger ratio.

Once the event starts, the *LTA* is usually frozen, so that the reference level is not affected by the event signal. The end of the event is declared when the *STA/LTA* ratio reaches the de-trigger level. Trigger levels and de-trigger levels of 4.0 and 2.0, respectively, are typical values. In order to get the complete event, the recording will start pre-event time before the trigger time. Likewise, to avoid the truncation of the signal, recording continues some time after the de-triggering, the post-event time. The maximum time of recording can usually be limited by a maximum recording time setting, and it is also sometimes possible to discard triggers lasting less than a certain time, the minimum trigger time.

STA and *LTA* must be calculated on signals without a DC component in order to reflect the real signal. The DC is normally removed by filtering, which is desirable also for other reasons (see below). The *STA/LTA* trigger algorithm is well suited to cope with fluctuations of natural seismic noise, which are slow in nature. It is less effective in situations where man-made or natural seismic noise of a bursting or spiky nature (e.g., wind gusts) is present. At sites with high, irregular seismic noise, the *STA/LTA* trigger usually does not function well, meaning there are too many false triggers.

A typical implementation will now be described. The trigger must be running continuously in the computer, so the *STA* and *LTA* values are calculated as running averages:

$$STA_i = STA_{i-1} + \frac{|x_i| - STA_{i-1}}{NSTA}$$

$$LTA_i = LTA_{i-1} + \frac{|x_i| - LTA_{i-1}}{NLTA} \quad (1)$$

$$R_i = \frac{STA_i}{LTA_i}$$

where x_i is the signal (filtered or unfiltered); *STA*, the short-term average; *LTA*, the long-term average; *R*, the *STA-LTA* ratio; and *NSTA* and *NLTA*, the number of points in the *STA* and *LTA* windows, respectively. In this case, *x* is the unfiltered signal. This works if the signal has no DC component, but if *x* has a large DC component, this will be transferred to the *STA* and *LTA* and they might be so large that *R* never reaches the trigger ratio.

There are two common variations of the trigger algorithm Eq. 1. One is to use the squared x_i instead of the absolute value. Using the squared x_i, makes the trigger more sensitive to changes. The other is the Allen trigger (Allen 1978), in which the absolute value of x_i is replaced by the characteristic function c_i. In a simplified form, c_i is calculated as

$$c_i = x_i^2 + k \cdot (x_i - x_{i-1})^2 \quad (2)$$

where k is a constant. The second term is proportional to the squared first time derivative of the signal. The characteristic function is sensitive to

Seismic Event Detection,
Fig. 9 *Below*:
A spectrogram of the signal plotted above, with three local earthquakes closely spaced in time. It is apparent that the amplitude of high frequencies decays faster than lower frequencies after the first arrival of each event. The colors indicate spectral amplitudes
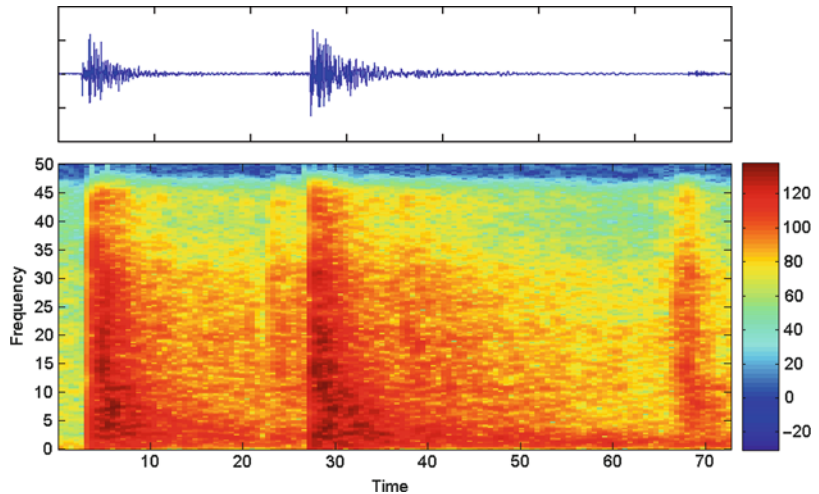
both amplitude variations and frequency changes.

All the above described trigger methods have in common that they are *not-looking-forward* algorithms, i.e., the present value of the trigger parameters depends only on present and past values of the signal. This makes them easy to implement with simple recursive, real-time operations.

If the seismic events are very close in time (Fig. 7), it is obvious that the STA/LTA would produce many failures. If one event is not finished before the next event starts, the first event would simply be prolonged so correct start times of events would only be obtained when a new trigger occurs after the previous has finished. A more sophisticated trigger might also monitor the frequency content of the signal. Since the end of the event usually has a lower-frequency content than the initial P-wave, this could be used to declare the start of a new event before the previous event is finished. This can be illustrated with a spectrogram which is a time-frequency representation of the spectral content of a signal. Figure 9 shows an example with three events occurring within 70 s.

A human analyst works in a quite different way: He detects an event looking at the entire record and evaluates the past, present, and subsequent values of the trace (or something like the envelope of it) to decide the presence of an event

and its onset time. So the human analyst would have no problem detecting the events in Fig. 7.

Figure 10 shows an example of how the algorithm *STA/LTA* works for the standard trigger using the absolute value.

The original signal is a small high-frequency earthquake with a medium-level microseismic background noise, which is reflected in the slow variations of the *STA* and more smoothed in the *LTA*. At about time 12 s, the ratio gets above the threshold of 3 (1. trigger). It immediately gets below again, but since the de-trigger ratio is 2, the trigger remains activated. In the above example, the *LTA* is frozen once an event has been declared, which can be seen from the flat *LTA* level. A few seconds after the de-triggering, the ratio again rises above the trigger level and remains in trigger state for about 1 s (2. trigger). This signal has a good signal-to-noise ratio, so why has the trigger not been activated much longer? The first reason is simply that the DC has not been removed before calculating *STA* and *LTA*. Since *STA* and *LTA* before triggering are about 220, the event will de-trigger when STA is 440 and the maximum ratio is 7 (Fig. 10 and Table 1). The DC of this signal is only 200 as compared to the maximum value of 3,400; yet if it is not removed, it will have a large influence on the trigger performance.

When the DC is removed, the figure corresponding to Fig. 10 will look almost the

S

Seismic Event Detection,
Fig. 10 The seismic
trigger. *Top* trace is the
original signal; the second
trace, the *STA*; the third
trace, the *LTA*; and the last
trace, the *STA*/*LTA* ratio.
The *STA* window is 0.4 s
and the *LTA* window is 50 s.
The trigger ratio is 3 and the
de-trigger ratio 2. All
amplitudes are counts (the
raw number generated by
the recorder) (Figure from
Havskov and Alguacil
(2010))
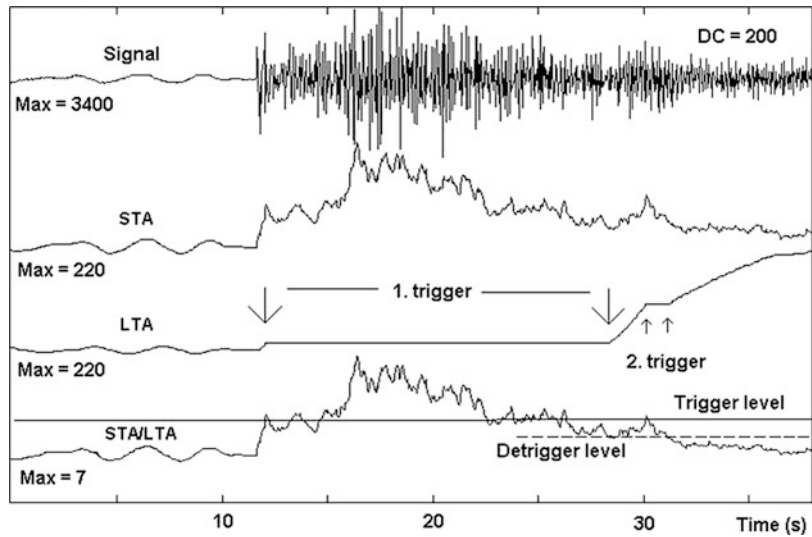
Seismic Event Detection, Table 1 Effect of DC and
filter on STA, LTA, and STA/LTA. The units of the
numbers are in counts

	DC not removed	DC removed	Filtered 5–10 Hz
STA and LTA before trigger	220	40	3
Maximum STA	1,500	1,450	600
Maximum STA/LTA	7	36	200
Duration of trigger (s)	15	50	105

same, but since both noise, *STA* and *LTA*, are
smaller due to the removal of the DC, STA/LTA
will be much higher for the earthquake signal,
and the event will remain in trigger state for 50 s
instead of 15 s (Table 1). The performance can be
further improved by filtering. The signal has
a typical microseismic background noise
superimposed, and this gives the main contribu-
tion to the *STA* and *LTA* before triggering. By
using a band-pass filter from 5 to 10 Hz, the
performance is further improved (Table 1) and
the trigger lasts for 105 s.

So, the conclusion is that every trigger must
have an adjustable band-pass filter in front of the
trigger algorithm, which has both the function of
removing the DC and making the trigger algo-
rithm sensitive to the frequency band of interest.

This is particularly important with broadband
seismometers, where small earthquake signals
often are buried in dominant 0.2–0.3 Hz seismic
noise and a trigger without a filter simply would
not work except for the largest events (Fig. 4).
Some recorders allow several trigger parameter
sets to be used simultaneously. This is needed if,
e.g., a broadband station has to trigger on micro-
earthquakes, teleseismic P-waves, and surface
waves which each requires separate filters, *STA*,
and *LTA*.

For more detail of how to set and use
STA/LTA triggers, see Trnkoczy, information
sheet 8.1 in New Manual of Seismological Obser-
vatory Practice (NMSOP) (Bormann 2012).

Improved Triggering Using Several Stations, the Network Trigger

Whatever trigger method is used with single
channels or single three-component stations,
there is always likely to be quite a lot of false
triggers, all depending on the local conditions and
how sensitive the trigger has been set up. When
several stations within a single network are
processed together, the number of false triggers
can be significantly reduced setting up
a requirement of concurrent triggering. This
means that for a network trigger to be declared,
there must be a minimum number of triggers
within a given time window, the array

Seismic Event Detection, Fig. 11 A volcanic event recorded on an array of vertical sensors. The *bottom* trace (*black*) is the sum of the individual channels (color), once they have been time aligned using the P-wave relative propagation times. The improvement on the signal-to-noise ratio is apparent
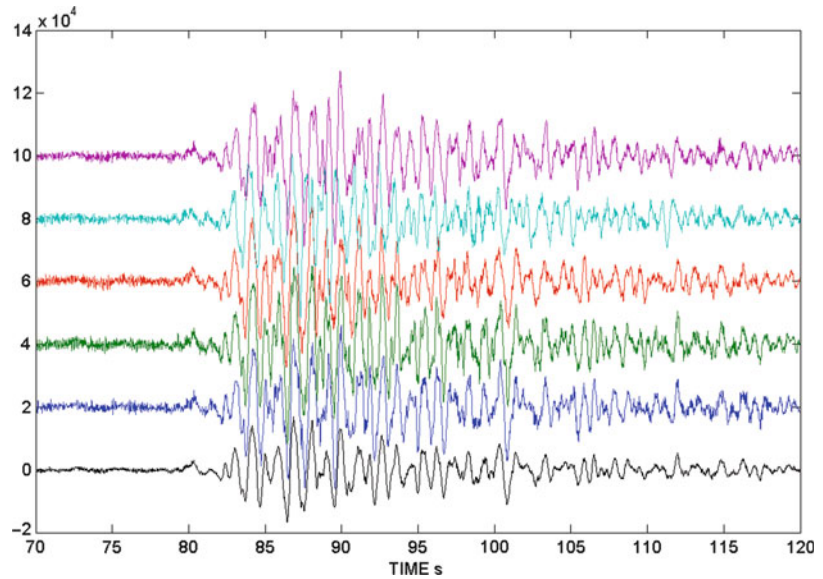
propagation window. The array propagation window is calculated as the time the waves take to propagate across the network. If, e.g., the network is 100 km in diameter and assuming the case that triggering takes place on the S-waves only and the event is outside the network, then the array propagation window should be $100/V_s$ where V_s is the S-wave velocity. If some stations trigger on P and others on S, the array propagation window should be even larger by the S-P time so in general the array propagation time is set to a value larger than $100/V_s$. In case there is swarm activity (Fig. 7), a network trigger might not help a lot since the events come so close in time that it would be hard to associate events, that is, if the events are correctly triggered in the first place.

Triggering Using a Seismic Array

A seismic array consists of multiple sensors in a small enough area that signals are correlated, and the seismic array is therefore also sometimes called a seismic antenna.

The signal-to-noise ratio of a seismic signal is improved by summing the signals of each station in the array using a delay corresponding to ray travel time across the array so the signals get lined up. In this way the coherent seismic signal will be enhanced and the supposedly incoherent noise will cancel out. Running a seismic trigger algorithm on the summed signal will therefore be more sensitive and more reliable than on any of the single stations. In a way this is similar to the network trigger where the same event must be present in all channels. An example is shown in Fig. 11.

Triggering Using Three Channels

Most seismic stations today have three channel sensors oriented vertically, north–south and east–west. As a first small improvement in the triggering, it can be required that all three channels trigger on the seismic event. Since external disturbances affect all three channels, this requirement might only improve the detection marginally. With three channels, it is possible to make a polarization analysis or in other words determine the particle motion of the ground (see Fig. 12).

For simplicity, the polarization is only shown in the north-vertical plane. It is seen that the P-wave is not polarized in the noise and highly polarized in the P-wave signal. By continuously calculating the degree of polarization, it is possible to distinguish a P-wave from background noise which would have no polarization.

S
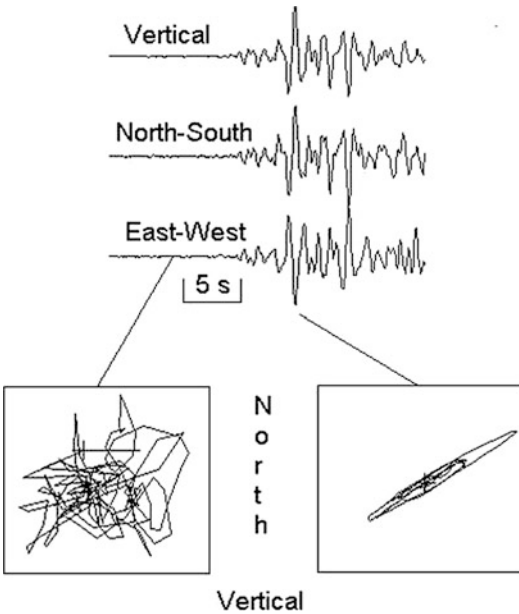
Seismic Event Detection, Fig. 12 *Top*: the three components of ground velocity for a local earthquake. *Bottom*: the particle motion in the north-vertical plane at two different times as indicated on the figure. To the *left* is the analysis of the noise and to the *right* the analysis of the P-wave signal

This property has been used to improve the detection of P- and S-waves (Roberts et al. 1989; Withers et al. 1998).

Triggering Using a Known Waveform

Many seismic events have very similar waveforms if they originate from the same or nearby sites. This could also be the case for volcanic events even though they might not be from exactly the same area. A "known" waveform shape can then be used to detect unknown arrivals by cross correlating the known wavelet with the signal of unknown events. An example of similar events is seen in Fig. 13. Despite the large distances, the signal shapes are nearly identical for the two events.

The only trigger parameter would then be the minimum correlation needed to declare an event. Using this technique, it is often possible to detect events which otherwise could not be detected. When using a seismic array, this method is particularly powerful (Gibbons and Ringdal 2006).
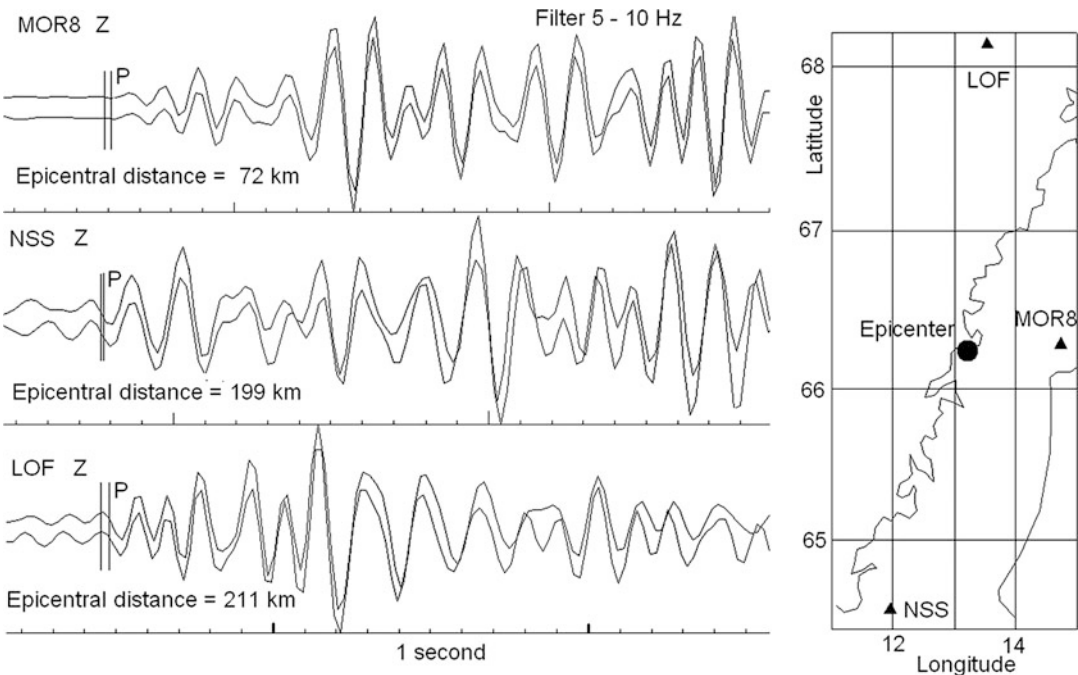
Seismic Event Detection, Fig. 13 *Left*: the first couple of seconds of the P-waves for the two events occurring in the same area as recorded with three different stations. The *top* trace event has magnitude 1.8 and the *bottom* trace event magnitude 1.6 (Figure from Havskov and Ottemöller (2010))

Summary

Today, huge quantities of continuous seismic data are recorded, and automatic detection of the seismic events is indispensable. Over the years very many algorithms have been developed to detect "real" seismic events; however the overall most used and reliable detectors are based on the standard STA/LTA method. Combined with a network trigger, this becomes a simple and robust event detector, in particular considering that the number of stations is continuously increasing so more use of the network trigger can be done, making the requirement for the single-channel detector less stringent.

Cross-References

▶ Passive Seismometers
▶ Principles of Broadband Seismometry
▶ Seismometer Arrays
▶ Seismic Noise

References

Allen RV (1978) Automatic earthquake recognition and timing from single traces. Bull Seismol Soc Am 68:1521–1532

Bormann, P. (ed) (2012) New Manual of Seismological Observatory Practice (NMSOP-2), IASPEI, GFZ German Research Centre for Geosciences, Potsdam; nmsop.gfz-potsdam.de

Brune JN, Oliver J (1959) The seismic noise at the earth's surface. Bull Seismol Soc Am 49:349–353

Gibbons SJ, Ringdal F (2006) The detection of low magnitude seismic events using array-based waveform correlation. Geophys J Int 165:149–66

Havskov J, Alguacil G (2010) Instrumentation in earthquake seismology, 3rd edn. Springer, Dordrecht, 358 pp

Havskov J, Ottemöller L (2010) Routine data processing in earthquake seismology. Springer, Dordrecht, 347 pp

Köhler A, Ohrnberger M, Scherbaum F (2010) Unsupervised pattern recognition in continuous seismic wavefield records using self-organizing maps. Geophys J Int 182:1619–1630

Roberts RG, Christofferson A, Cassedy F (1989) Real time event detection, phase identification and source location using single station 3 component seismic data and a small PC. Geophys J 97:471–480

Romeo G, Mele F, Morelli A (1995) Neural networks and discrimination of seismic signals. Comput Geosci 21:279–288

Sharma BK, Kumar A, Murthy VM (2010) Evaluation of seismic event detection algorithms. J Geol Soc India 75:533–538

Withers M, Aster R, Young C, Beiriger J, Harris M, Moore S, Trujillo J (1998) A comparison of select trigger algorithms for automated global seismic phase and event detection. Bull Seismol Soc Am 88:95–106

Seismic Fragility Analysis

Murat Altug Erberik
Civil Engineering Department, Middle East
Technical University, Ankara, Turkey

Synonyms

Damage probability matrix; Fragility curve; Fragility surface; Limit state; Numerical simulation; Seismic capacity; Seismic demand; Uncertainty

Introduction

Seismic fragility can be defined as the proneness of a structural component or a system to fail to perform satisfactorily under a predefined limit state when subjected to an extensive range of seismic action. In accordance with the above definition, seismic fragility analysis can be regarded as a probabilistic measure for seismic performance assessment of structural components or systems. There are two different end products of seismic fragility analysis: damage probability matrix and fragility curve.

Damage probability matrix (DPM) is a table that provides discrete values of damage state probabilities for specified levels of ground motion intensities. Each column of DPM stands for a constant level of ground motion intensity whereas each row of DPM denotes the probability of being in a predefined damage state. Hence, any element of DPM represents the probability of

S

experiencing a certain damage state for a specific level of ground motion intensity.

Fragility curves are continuous functions that represent the probability of exceeding predefined limit (or performance) states for specific levels of ground motion intensity. Fragility curves can be developed for structural and nonstructural components, or they can be generated for structural systems and assemblies. In the field of earthquake engineering, the most common application is to generate fragility curves for building structures. Depending on the characteristics of the building stock in the region of interest, the fragility curves can be derived for many different structural types (i.e., steel, reinforced concrete, or masonry buildings; low-, mid-, or high-rise construction; frame, wall, or mixed structures; etc.). Other than building structures, fragility curves have also been generated for infrastructures (like bridges, dams), lifeline systems (like transportation, power, water supply networks), unrestrained equipment, etc.

Fragility curves can be used for different purposes. The main purpose is to determine the seismic performance of new or existing structural systems. In the former case, the fragility information is obtained for optimal design and in the latter case for condition assessment. Fragility curves can also be employed in order to assess the efficiency of different intervention techniques on existing structural systems. When the curves are derived to represent a certain class of structural systems, they can also be used for regional seismic damage and loss estimation studies. Such studies are very popular nowadays since they can be employed for pre-earthquake mitigation efforts and postearthquake decision-making processes.

There is a close relationship between fragility curves and DPMs so that the information obtained from fragility curves can be converted to construct DPMs or vice versa. Figure 1 demonstrates how to generate DPM from a given set of fragility curves. The vertical axis, simply named as "probability," actually stands for the probability of exceeding a limit state. The horizontal axis represents the complete range of ground motion intensity levels (ILs). Different

ground motion intensity parameters can be used in the generation of fragility curves. As seen from the figure, four limit states (LSs) have been assigned, which in turn means that there exist five damage states (DSs). Attainment of limit states is an important phase of fragility curve generation and is discussed in following sections. The columns of DPM are obtained by intersecting the fragility curve set with vertical lines (dashed lines in Fig. 1) at specific intensity levels and calculating the damage state probabilities, which are corresponding portions between any two limit states in these vertical alignments.

Different Approaches for Seismic Fragility

There are four main resources to generate the seismic fragility of a structural system or component: expert opinion, field data, experimental data, and analytical simulations.

One of these main data sources for the derivation of fragility curves is based on the opinion of the experts from the field of earthquake engineering that are invited to make a prediction about the probable damage state of a structure or a class of structures under different levels of seismic action. The fragility curves generated by this judgment-based data are called as "judgmental fragility curves." Until recently, expert opinion was the main source to generate fragility information especially for a wide range of structures and performance limits in the United States by the introduction of the ATC-13 document (Applied Technology Council 1985) and HAZUS earthquake loss estimation methodology (National Institute of Building Sciences 1999). However, the main issue that impairs the use of expert opinion is the subjectivity and bias of the collected data since it depends on the unique experience of the experts involved and it is very difficult to quantify this subjectivity or eliminate the bias.

Fragility curves can also be derived by using the structural damage data obtained from field observations after earthquakes. The fragility curves obtained by this approach are called as
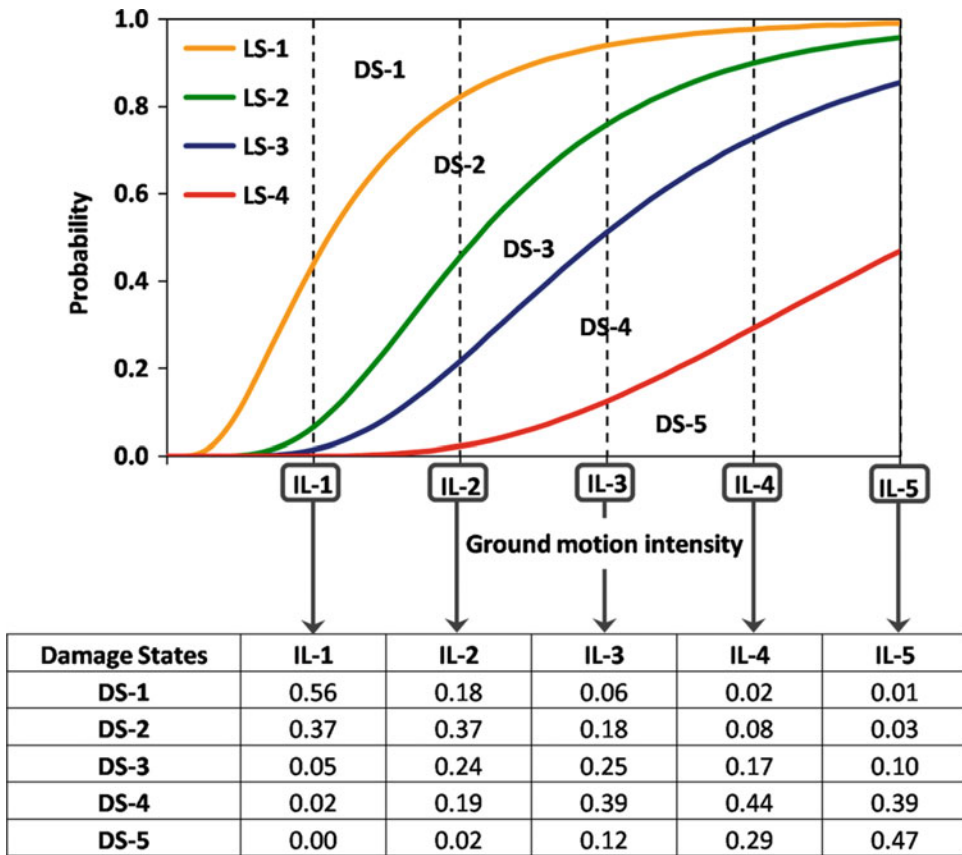
Seismic Fragility Analysis, Fig. 1 Conversion from a set of fragility curves to DPM (Prepared by using the data from Erberik and Elnashai [2004])

Damage States	IL-1	IL-2	IL-3	IL-4	IL-5
DS-1	0.56	0.18	0.06	0.02	0.01
DS-2	0.37	0.37	0.18	0.08	0.03
DS-3	0.05	0.24	0.25	0.17	0.10
DS-4	0.02	0.19	0.39	0.44	0.39
DS-5	0.00	0.02	0.12	0.29	0.47

"empirical fragility curves." Although it is difficult to construct empirical fragility curves from field observations, the obtained fragility curves are invaluable in the sense that they reflect the exact behavior of a structure or a class of structures. The difficulty comes from the fact that for empirical fragility curve generation, an extensive field data from many earthquakes with different magnitudes should be gathered. For instance, Rossetto and Elnashai (2003) gathered a huge database that is composed of 340,000 buildings inspected after 19 different earthquakes and locations in order to construct the empirical fragility curves for European-type reinforced concrete frame buildings. Furthermore, it is not an easy task to estimate the spatial distribution of earthquake intensity at different locations where the buildings under consideration reside. Another issue is the nonuniformity of field data in terms of ground motion intensities and damage state definitions since it is being collected from many different resources. In this case, it becomes crucial to convert all field data to a standard format in order to use in the generation of empirical fragility curves.

Experimental data can be employed in the generation of fragility curves by using a similar approach as in the case of empirical fragility curve generation based on field data. However, in order to obtain sufficient data for fragility curve generation, it may be required to conduct a large number of laboratory tests with different levels of loading intensities and material properties. Hence, this way of generating fragility curves is generally an expensive solution and therefore not popular when compared to the

generation of fragility curves based on analytical simulations.

The most commonly used data source for quantification of seismic fragility comes from the analytical simulations since most of the issues encountered in the case of judgmental, empirical, and experimental-based approaches are handled by the use of numerical data obtained through analytical modeling. It is possible to generate response statistics for any type of structure or class of structures subjected to a wide range of ground motion intensities. However, in this case, there are two main issues: computational effort and limitations in analytical modeling due to idealization of actual structures. Although the computing technology is very powerful nowadays, generation of analytical fragility curves can take considerable time and effort depending on the level of detail in analytical modeling. In some cases, it can even become unfeasible such as when it is required to run a dynamic analysis thousands of times on a detailed finite element model of a special structure. Hence, this means that it is inevitable to introduce some simplifications and idealizations to the analytical model. But then, the issues arising from limitations of analytical modeling begin to play an important role in seismic fragility analysis. This may cause the deviation of the response of the analytical model from the actual behavior of the structure under concern. As a result, analytical approach to quantify seismic fragility always possesses a trade-off between accuracy and computational effort. In this regard, there are different alternatives to be used as the analysis method for seismic fragility analysis. These can be listed as linear static, linear dynamic, nonlinear static, and nonlinear dynamic methods from the simplest to the most complex approach. Among these, linear static method is not generally used in seismic fragility analysis due to the dynamic nature of the actual response. However, nonlinear static analysis is an alternative when it is difficult to carry out series of dynamic analyses to generate fragility functions. Linear dynamic analysis can sometimes be used, especially when the structural system under concern exhibits brittle behavior under seismic action and does not go far

beyond the elastic range of response, like unreinforced masonry structures. The best method is always nonlinear dynamic approach, but in practice, this may not be the case due to computational difficulties as discussed above.

Components of Seismic Fragility Analysis

Basically, seismic fragility analysis is the comparison of seismic capacity and demand and to estimate whether the seismic capacity is exceeded for a well-defined performance level when the structural system is subjected to specified levels of ground motion intensity. Due to the probabilistic nature of seismic fragility analysis, both seismic capacity and demand are defined by probability functions in terms of certain random variables to quantify all the uncertainties involved in the process.

In the field of earthquake engineering, there is no consensus on the best method for seismic fragility analysis. However, the general framework of seismic fragility analysis can be more or less presented as shown in Fig. 2. Each element of the flowchart defines a different component of seismic fragility analysis procedure. The elements enclosed by boxes with dashed lines can be regarded as a single component in the analysis. Each component of seismic fragility analysis is explained in detail in the following paragraphs.

The first component is the development of the analytical model and generation of structural simulations to account for structural variability. At this stage, the major challenge is the selection of the analytical model. Depending on the level of idealization and simplification, a wide range of models from equivalent single-degree-of-freedom (SDOF) systems to detailed finite element (FE) meshes can be employed in order to carry out the structural simulations (Fig. 3). Since a huge number of analyses are required to construct the response statistics and the fragility curves, the use of a simple model (like a SDOF model) becomes very advantageous in terms of computational effort. However, it should be noted that such a simple model can only realize
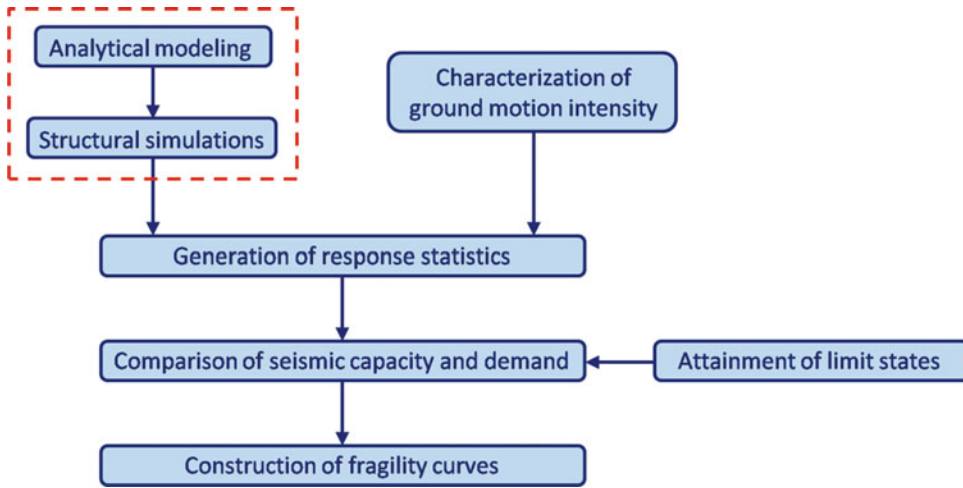
Seismic Fragility Analysis, Fig. 2 General framework of seismic fragility analysis
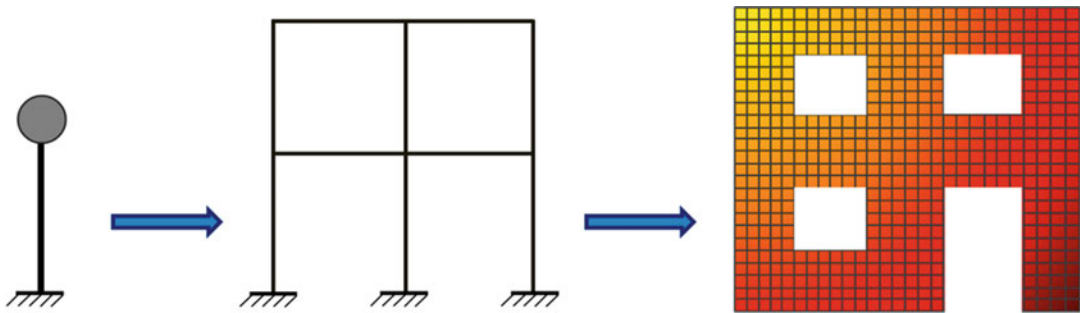
Seismic Fragility Analysis, Fig. 3 Different levels of idealization for analytical modeling of structural systems from the simplest to the most sophisticated

the global behavior of the structural system to some extent and it is blind to local structural characteristics or construction details. Hence, such models are generally used to generate the fragility curves of a class of structures (for instance, mid-rise reinforced concrete frame structures) in order to predict regional damage or loss. On the other hand, it is not possible to use simplified models for special structures, in which the structural and construction details are important and should be reflected in the seismic fragility analysis. A good example for this case is historical masonry structures, which are unique and possess their own structural characteristics. In such cases, finite element mesh modeling seems to be the best solution, but then the problem is to carry out numerous analyses (especially

if they are dynamic in nature) on this finite element mesh within a reasonable period of time. Briefly, the problem of selecting the appropriate modeling strategy is multidimensional, and one has to make the decision by considering all the pros and cons.

After the selection of the analytical model, the next step is to generate the structural simulations. Due to the probabilistic nature of seismic fragility analysis, some of the major structural parameters within the analytical model are considered as random variables with appropriate probability density functions assigned to them. Normal or lognormal distributions are commonly used for convenience. These can be mechanical properties like stiffness or strength to account for the material variability or geometric properties like

length, height, or cross-sectional dimensions. In order to generate the population of analytical simulations, a sampling method is required. The most popular method is Monte Carlo sampling (Metropolis and Ulam 1949). It is a robust and a straightforward method to generate structural simulations, but the disadvantage is that it requires a very large sample size, which makes the method unfeasible to use in the case of detailed and complex analytical models. This shortcoming opens a door to constrained sampling methods to reduce the sample size, like the Latin hypercube model (McKay et al. 1979).

Since dynamic analysis is an indispensable tool for seismic fragility analysis, characterization of ground motion intensity becomes an important component to generate fragility information. To conduct dynamic analyses, there are two options: to employ (1) actual or (2) synthetic ground motion records. If it is possible to find a sufficiently large number of records from different earthquakes in the region of interest, then the first option seems to be the best solution to reflect the effect of regional seismicity and local geologic characteristics. Ground motion records are generally selected to cover the whole range of intensity levels. However, this may induce the need to scale ground motion records to fill the gaps of seismic intensity levels, especially for high intensities. Scaling ground motion records is a different research topic in earthquake engineering, and there are many different methods that one can employ. The second option, using synthetic ground motion records, becomes crucial if there exists an insufficient number of actual ground motion records in the study region. Generating synthetic ground motion records is another huge area of research in earthquake engineering, and there are many different approaches in the literature. However, it should be noted that in order to generate reliable synthetic ground motion records in a region, geologic and seismological parameters should be well studied and known beforehand.

Another challenge in seismic fragility analysis is the selection of the ground motion intensity parameter since it is difficult to determine a unique parameter that fully characterizes earthquake ground motions. Descriptive parameters like Modified Mercalli Intensity (MMI) had been used in the past, but they are no more popular due to the fact that they are judgmental and subjective measures. In addition, they do not correlate well with damage. The most commonly used group of parameters are single-valued (or peak-valued) parameters like peak ground acceleration (PGA) and peak ground velocity (PGV). These parameters can be easily obtained from the ground motion time history. However, single-valued parameters are insensitive to frequency content and duration of the ground motion record, which is extremely important for obtaining the seismic performance of a structural system in some cases. To overcome the shortcomings of single-valued parameters, spectral parameters like spectral acceleration (S_a) and spectral displacement (S_d) can also be employed since these are enhanced parameters which are functions of both the ground motion characteristics and SDOF system properties. Although rare, there are also cases where involved parameters like effective peak acceleration, Housner's spectral intensity, and input energy have been used as the intensity parameter in seismic fragility analysis.

After generating the structural simulations and selecting the ground motion records, the next step is to carry out a large number of analyses in order to build up the response statistics. These are scattered plots in which the vertical axis stands for a response parameter like displacement, drift, or force, whereas the horizontal axis represents the ground motion intensity parameter (see Fig. 4a). All the uncertainty and the variability arising from the ground motion characterization and the analytical simulation is reflected in this plot. Although scattered, the general trend in this plot is that as the ground motion intensity increases, seismic demand characterized by the response parameter also increases.

The response statistics can be employed to generate fragility curves if some limit states are defined as a function of the response parameter under consideration. This is another important stage in seismic fragility analysis because the definition of a limit state directly affects the
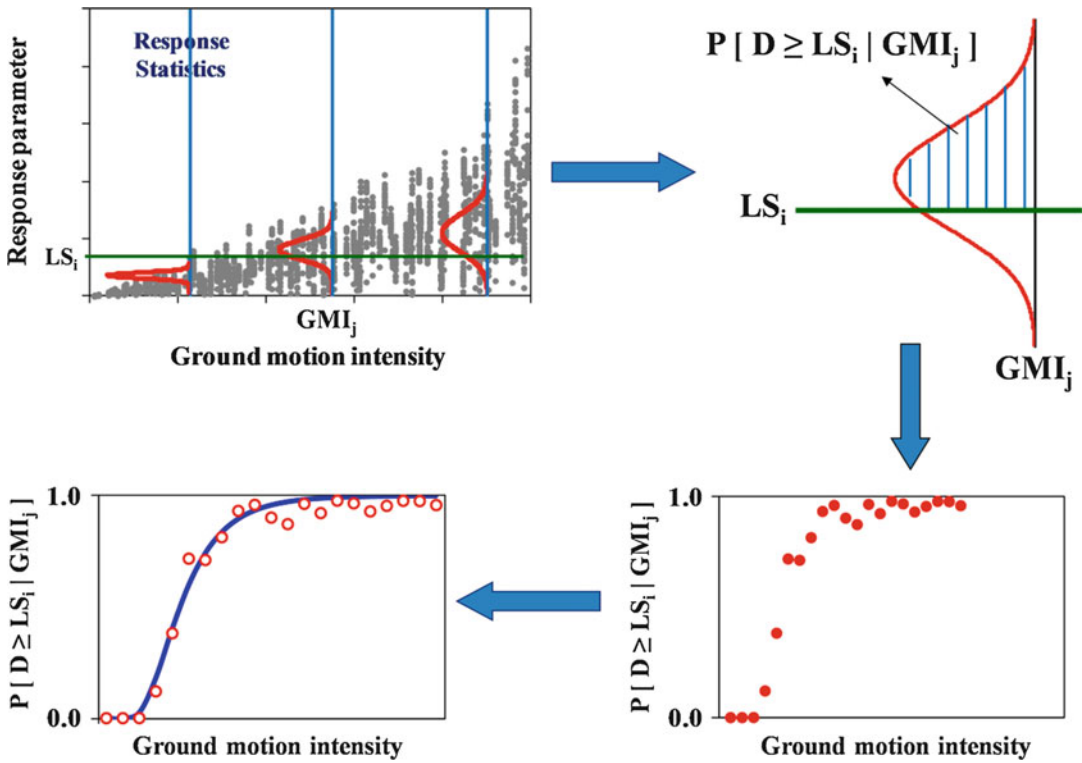
Seismic Fragility Analysis, Fig. 4 Schematic representation of the fragility curve generation procedure (Adopted from Ugurhan et al. 2011) (**a**) response statistics, (**b**) probability of exceeding a limit state at a specific ground motion intensity, (**c**) discrete fragility information for a certain limit state, (**d**) continuous fragility function

corresponding fragility curve. There are many different ways to attain the limit states. Limit state values can be directly obtained from the previous studies in the literature if the structural type under consideration had been studied before. However, in most of the cases, the local and special characteristics of the structural system requires the attainment of new limit states. In this case, the best method is to rely on field observations or experimental findings if any. In the absence of such data, analytical approaches can be used to determine the limit states. For instance, nonlinear static analysis, in particular pushover method, is a good candidate for this task. The instantaneous capacity of the structural system and the progression of damage can be monitored in this type of analysis.

Generally, two or more limit states are defined in seismic fragility analysis. For instance, in

FEMA-273 document (Federal Emergency Management Agency 1997), three limit states were defined as "Immediate Occupancy," "Life Safety," and "Collapse Prevention" in terms of interstory drift. The limit states are generally considered as deterministic parameters. However, there are cases in literature in which limit states are also considered in a probabilistic manner to account for the variability in seismic capacity.

Each vertical line of scattered data in response statistics stands for a specific ground motion intensity level with its own statistical distribution. If the limit states are also added to the same plot, then it becomes possible to determine the probabilities of exceeding the limit states at specific ground motion intensity levels. This is illustrated in Fig. 4b, in which LS_i represents the ith limit state and GMI_j represents the jth intensity

level. Hence, the shaded area in the given distribution denotes the probability of exceeding the ith limit state at the jth ground motion intensity level, i.e., $P[D \geq LS_i|GMI_j]$. The value obtained represents one of the data points in Fig. 4c. If this process is repeated for different intensity levels for the whole range, the discrete fragility information shown in Fig. 4c can be obtained for a certain limit state. Then, repeating the process for different limit states reveals a set of fragility information. The last step is to convert this discrete information into a continuous function by curve fitting. The continuous fragility functions are generally represented by cumulative lognormal distribution (Fig. 4d), although other forms of equations (like exponential function) have also been used in the literature. The shape of the fragility curve is dependent on the overall uncertainties involved in the generation process. Less variability means a steeper fragility curve. In the limit, the fragility curve with no uncertainty (i.e., deterministic case) turns out to be a step function.

Additional Information on Seismic Fragility Analysis

The topic of seismic fragility analysis has been a very popular and active research area in the field of earthquake engineering for the last two decades. Hence, there are many additional issues to be discussed related to this topic. Some of these are mentioned in the following subsections.

Different Analysis Approaches in the Literature
In the previous section, a fragility curve generation procedure that makes use of time history analyses was introduced. However, it should be stated that this is not the only alternative to generate fragility curves. For instance, instead of using time history analysis, incremental dynamic analysis (IDA) can also be used for the generation of fragility curves. In this approach, the structural system under consideration is subjected to a series of nonlinear time history analyses, in which the intensity of the ground motion is

increased until collapse at each run by using scaling methods (Vamvatsikos and Cornell 2002). Hence, it becomes possible to monitor the complete system response under seismic action from linear elastic behavior to collapse by using IDA approach, which makes it a good candidate for seismic fragility analysis.

In literature, there are two different nonlinear static procedures that have replaced dynamic analyses in the generation of fragility curves, named as capacity spectrum method (CSM) and displacement coefficient method (DCM). The CSM was first developed by Freeman et al. (1975) and then became more publicized thanks to the introduced in the ATC-40 document (Applied Technology Council 1996). It has been very popular since then in performance based earthquake engineering. The CSM method is based on the idea to compare a nonlinear capacity curve (obtained from a pushover analysis) in acceleration-displacement response spectrum (ADRS) format with an elastic demand spectrum that is reduced to account for equivalent damping. The intersection point is obtained by trial and error and it is called as "performance point". The second method, DCM, was first introduced by the FEMA-273 document (Federal Emergency Management Agency 1997). Just like the CSM, it is based on constructing the pushover curve of the structural system under concern. Then, this information is converted into equivalent SDOF response. The final step is to modify the linear elastic response of this equivalent SDOF system by the help of some empirical coefficients that account for MDOF response, inelastic behavior, degradation, and P-Δ effect. The final product is an estimate of the maximum displacement of the structural system, called as "the target displacement." By using both methods, it is much easier to generate fragility curves since it is only a matter of repeating the pushover analysis rather than carrying out complicated nonlinear time history analyses.

There is also a different class of analysis methods, called as response surface methods, which are employed to reduce the large number of computations in seismic fragility analysis. In these methods, a response function is generated

in terms of the major random variables (ground motion intensity parameter, geometric or mechanical properties) to predict the seismic response of structural systems without the need to carry out complex dynamic analyses. The other advantage of these methods comes from the fact that both seismic demand and capacity can be directly incorporated with their inherent joint probability distributions in the response surface, which can only be achieved by making qualitative physical assumptions in the other analysis approaches.

Quantification of Uncertainty in Seismic Fragility Analysis

Seismic fragility analysis requires the treatment of uncertainties involved in the process since both seismic demand and capacity possess different sources of uncertainty. The uncertainties may either arise from inherent randomness in nature (aleatoric uncertainty), which cannot be reduced, or from lack of knowledge or data (epistemic uncertainty). From another point of view, the uncertainties in seismic fragility analysis may belong to ground motion characteristics and the demand on the structural system or the structural capacity under seismic action, performance level identification, and the limitations in analytical modeling. All sources of uncertainty should be quantified and reduced (if possible) in the generation of fragility curves.

Most of the researchers have stated that the uncertainties arising from ground motion characteristics outweigh the uncertainties due to structural system characteristics and analytical modeling. But, of course, this may not be always the case, depending on the specific properties of seismic action and the structural system under concern.

The reliability of the generated fragility curves can be tested by plotting the confidence intervals for selected levels of confidence with upper and lower bounds. If the uncertainties arising from the components of seismic fragility analysis are significant, the confidence intervals will eventually be large, indicating a poor estimate of performance that is represented by the fragility information.

Combined Fragility Curves

Fragility curves for individual structures can be combined together to yield the fragility information of a class of structures with similar properties. This is especially required for regional damage and loss estimation, to obtain the fragilities of different construction types such as reinforced concrete frame structures, unreinforced masonry structures, etc. In this case, the combined fragility curve can be obtained by using (Shinozuka et al. 2000)

$$F_C(X) = \sum_{i=1}^{M} P_i F_i(X) \qquad (1)$$

where

$$P_i = \frac{N_i}{\sum_{i=1}^{M} N_i} \qquad (2)$$

In these equations, F_C stands for the combined fragility, whereas F_i represents the fragility curve of the ith individual structure and X denotes any ground motion intensity parameter used in the generation of the fragility curve. Parameter N_i is the number of structures that the fragility information F_i belongs to, and the summation in the denominator of Eq. 2 represents the total number of structures in that specific class for which the information will be combined.

Fragility Surfaces

In some cases, it becomes difficult to obtain the seismic fragility of a structural system by using a single ground motion intensity parameter since there is no such unique parameter that is perfectly correlated with seismic damage as mentioned above. Then, it is possible to use two different ground motion intensity parameters together in the development of fragility information, which is called as fragility surface rather than fragility curve. In this case, there are two horizontal axes that define the ground motion intensity and one vertical axis that defines the probability of exceeding a certain limit state. This means the response statistics obtained can be plotted as a surface, not a curve. The dual ground motion intensity parameters for fragility surfaces can be

S

selected as magnitude and distance, PGA and PGV, spectral acceleration or displacement at two different vibration periods of the structural system (for instance, the first and the second vibration mode), etc.

Summary

Seismic fragility analysis can be regarded as a probabilistic measure for seismic performance assessment of structural components or systems. This analysis concept is very popular since it can be used in different research areas of earthquake engineering. Due to its probabilistic nature, it requires the use of random variables to characterize both seismic demand and capacity. In addition to this, there exist challenges such as the idealization of structure through an analytical model, selection of the ground motion intensity parameters and the analysis approach, attainment of the limit states, and quantification of the involved uncertainties. That is why the researchers in this field have used many different approaches and methods to obtain their own seismic fragility information. But the main goal is always to predict the performance of structural systems with a quantified level of confidence under different levels of seismic action.

Cross-References

▶ Analytic Fragility and Limit States [P(EDP| IM)]: Nonlinear Dynamic Procedures
▶ Analytic Fragility and Limit States [P(EDP| IM)]: Nonlinear Static Procedures
▶ Empirical Fragility
▶ Estimation of Potential Seismic Damage in Urban Areas
▶ Seismic Loss Assessment

References

Applied Technology Council (1985) Report ATC-13: earthquake damage evaluation data for California. Applied Technology Council, Palo Alto
Applied Technology Council (1996) Report ATC-40: seismic evaluation and retrofit of concrete buildings. Applied Technology Council, Redwood City
Erberik MA, Elnashai A (2004) Fragility analysis of flat-slab structures. Eng Struct 26:937–948
Federal Emergency Management Agency (1997) NEHRP guidelines for the seismic rehabilitation of buildings. FEMA Publication 273. Federal Emergency Management Agency, Washington, DC
Freeman SA, Nicoletti JP, Tyrell JV (1975) Evaluations of existing buildings for seismic risk – a case study of Puget Sound Naval Shipyard, Bremerton, Washington. In: Proceedings of the US National conference on earthquake engineering, Berkeley, pp 113–122
McKay MD, Conover WJ, Beckman RJ (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 221:239–245
Metropolis N, Ulam S (1949) The Monte Carlo method. J Am Stat Assoc 44(247):335–341
National Institute of Building Sciences (1999) HAZUS technical manual. Prepared for Federal Emergency Management Agency, Washington, DC
Rossetto T, Elnashai A (2003) Derivation of vulnerability functions for European-type RC structures based on observational data. Eng Struct 25:1241–1263
Shinozuka M, Feng MQ, Lee J, Naganuma T (2000) Statistical analysis of fragility curves. J Eng Mech 126(12):1224–1231
Ugurhan B, Askan A, Erberik MA (2011) A methodology for seismic loss estimation in urban regions based on ground-motion simulations. Bull Seismol Soc Am 101(2):710–725
Vamvatsikos D, Cornell CA (2002) Incremental dynamic analysis. Earthq Eng Struct D 31(3):491–514

Seismic Fragility of Aging Offshore Platforms

Viswanath Kammula[1] and Srinivas Sriramula[2]
[1]Structural Engineer-in-Training, Soscia Engineering Limited, Toronto, Canada
[2]Lloyd's Register Foundation (LRF) Centre for Safety and Reliability Engineering, School of Engineering, University of Aberdeen, Fraser Noble Building, Aberdeen, UK

Synonyms

Aging infrastructure; Fragility analysis; Offshore structures; Probabilistic assessment; Seismic risk

Introduction

The importance of infrastructure to both the fabric of society and its economy is becoming increasingly more apparent. The robustness of infrastructure systems can be judged by their capacity to accommodate change over time. However, the rapid rate at which our infrastructure is aging requires new solutions for providing a resilient infrastructure that can last for future generations.

Earthquakes are paramount among the natural hazards impacting the infrastructure. The past several decades have witnessed a series of costly and damaging earthquakes. The events would have even more devastating effects if a major earthquake were to hit an old and unprepared infrastructure. The challenge that engineers today face is not only on how to construct new infrastructure with proper resistance to earthquakes but rather to update older structures that do not have the proper defense should a natural disaster occur. Failing to modernize older infrastructure could result in very serious consequences.

Offshore platforms are one of the key infrastructure systems housing the facilities for exploration and production of oil and gas. In addition to acting as accommodation for offshore workers, these platforms contain a range of equipment from power generation to oil processing. Assessing the reliability of these structural systems is highly complicated due to the complexity of interaction between various elements. It is possible to classify the offshore platforms further as floating and fixed offshore platforms (Sadeghi 2007), with usage depending on the depth of water at the installation location. These structures are popular in the waters of California, the Gulf of Mexico, Nigeria, the North Sea, and the Persian Gulf.

This entry will focus on a very specific and popular type of offshore platform, the fixed steel jacket platform (FSJ). According to Ferreira (2003), there are around 7,500 FSJ platforms distributed around the world. These platforms have been popular because of their relatively simple design and the availability of oil at shallow waters in the initial days of exploration.

However, as the resources in shallow water have started to extinguish, the exploration activities have moved to deeper waters. Nonetheless, there are many fixed steel jacket platforms that are still in service.

When the FSJ platforms were originally designed, the design life was considered as 25 years. However, due to the continued availability of hydrocarbons at some locations, the operators are looking at various options to extend the design life. For example, in the United Kingdom Continental Shelf (UKCS) sector of the North Sea, out of the 288 offshore installations, about 50 % of the fixed platforms have exceeded their design life (Stacey et al. 2008). When it comes to extending the life of these structures, there are various key limit states one needs to consider, depending on the geographical location. For example, the effects of wave fatigue on offshore structures are a key consideration in one zone, whereas the structural response under earthquakes is a serious limit state at a different location.

Considering the complexity and popularity of these structures, this entry looks at ways to characterize the probabilistic response of FSJ platforms under earthquake loads. The seismic vulnerability curve or seismic fragility curve is one of the most important elements in the assessment of seismic damage and the evaluation of the performance of different structural systems. The fragility curves are used to represent the probabilities of structural damage due to earthquakes as a function of ground motion indices such as peak ground acceleration (PGA), spectral acceleration (S_a), and spectral displacement (S_d). The characterization of earthquake excitation and the identification of different damage levels have an important role in developing the fragility curves and understanding the seismic performance of these structures.

This entry provides details of the development of fragility curve for a typical FSJ offshore platform using the SAC-FEMA technique. The following sections provide mathematical details of the SAC-FEMA technique, an overview on the reference structure and establishment of dynamic characteristics. Then, the ground motions used in

fragility analysis and the analysis results are discussed. The entry concludes with suggestions to formulate a better inspection and maintenance framework based on the summary of findings.

Seismic Fragility Analysis of Structures

Theory of Fragility Analysis

Introduction
Seismic damage and loss estimations are essential for disaster planning and formulating damage-reduction and insurance policies. Regional loss estimation can be mathematically described through Eq. 1, as shown below (Kwon 2007):

$$P[Loss] = \sum_s \sum_{LS} \sum_d P[Loss \,|D = d].P[D = d|LS].P[LS|IM = s].P[IM = s] \quad (1)$$

where $P[\bullet]$ = probability of events in the brackets. The term *Loss* refers to direct or indirect losses from a seismic event; *IM* is an intensity measure of a seismic hazard such as spectral acceleration or peak ground acceleration (PGA); and s is a realization of the intensity measure. $P[LS|IM = s]$ is a conditional probability of attaining structural limit state (LS), and $P[D = d|LS]$ is a conditional probability of attaining damage (minor, moderate, or major), which is a qualitative measure. The loss of a region is determined based on the damage state. Among these elements, the term, $P[LS|IM = s]$, refers to strong, motion-shaking severity influencing the probability of reaching or exceeding a specified performance limit state. This term is referred as *vulnerability* or *fragility*. The strong, motion-shaking severity may be expressed as an intensity, peak ground parameters (PGA, PGV, or PGD), or spectral ordinates (S_a, S_v, or S_d) corresponding to a fundamental structural period.

Analysis Techniques
Among others, some practical approaches to perform fragility analysis are the SAC/FEMA method, effective fragility analysis method, and methods based on response surface (Pinto et al. 2004). The SAC-FEMA method, which provides the basis for the FEMA-350 (2000) for the seismic design and assessment of steel moment-resisting frames, is the most practical approach among these options. The SAC-FEMA method is efficient when the response is predominantly dominated by the first mode and when failure is influenced by a single scalar variable, for instance, the interstory drift. However, when failure involves several random variables, then one would need to use methods based on effective fragility analysis techniques or response surfaces.

Mathematical Details of SAC-FEMA Technique
The SAC-FEMA method (Pinto et al. 2004) is one of the popular methods to perform vulnerability analysis. In the SAC-FEMA approach, the seismic hazard $P[IM = s]$ is defined in terms of spectral acceleration ordinates, calculated at the fundamental period of the structure. In this method the failure occurs when the maximum demand over the duration of the seismic excitation exceeds the corresponding capacity. The seismic hazard is combined with the drift demand to define drift hazard as follows:

$$H_D(d) = \int P[D \geq d \,|S_a = x] \,| \, dH(x) \,| \quad (2)$$

where $|dH(x)|$ means the absolute value of the derivative of the site's spectral acceleration hazard curve times dx. The seismic demand is related to the hazard as follows:

$$\hat{D} = aS_a{}^b \quad (3)$$

where $\hat{D}$ is the median value of demand D. The constants, a and b, are determined from a regression analysis of structural demands calculated from the dynamic analyses. The term $\beta_{D,}$ standard deviation of the natural logarithm of demand D, is calculated about the median value of demand D. Ideally, at each spectral acceleration value, there would be multiple demand values, as two ground motions with same spectral acceleration could result in different drifts. Hence, the

median value over each set of demands is corresponding to a spectral ordinate value that needs to be calculated. Also every set is associated with a unique dispersion. To account for the dispersion, a random variable ϵ, with unit median and dispersion equal to β_D, is introduced. Equation 4 then updates to:

$$D = aS_a{}^b\epsilon \tag{4}$$

However, for simplicity, only one demand is used for a given spectral acceleration, and it is assumed that this value would be the median demand value, and the dispersion is assumed to be the same across all the spectral acceleration values. The demand D is assumed to be log-normally distributed about the median, with standard deviation of the natural logarithm equal to β_D. For a given intensity level, the probability of reaching a structural demand, d, can be calculated using cumulative distribution function of log-normal distribution, given as:

$$P[d \leq D \,|\, S_a = x] = 1 - \Phi\left(\frac{\ln(d/ax^b)}{\beta_{D|S_a}}\right) \tag{5}$$

Equations 4 and 5 result in a closed form solution of drift hazard as shown below:

$$H_D(d) = P[D \geq d] = H\left(S_a^d\right) \exp\left[\frac{1}{2}\frac{k^2}{b^2}\beta_{D|S_a}^2\right] \tag{6}$$

The drift hazard should be combined with the drift capacity, C, to evaluate the annual probability of the performance level not being met, P_{PL}.

$$P_{PL} = \int P[C \leq d] \,|dH_D(d)| \tag{7}$$

The drift capacity C is assumed to follow log-normal distribution with a median value of $\hat{C}$ and dispersion β_C. Therefore, the probability of the capacity, C, being lower than the demand, d, is expressed as

$$P[C \leq d] = \Phi\left(\frac{\ln(d/\hat{C})}{\beta_C}\right) \tag{8}$$

Substituting Eqs. 6 and 8 into 7, the probability of annual failure is given as

$$P_{PL} = H\left(s_a^{\hat{C}}\right)\exp\left[\frac{1}{2}\frac{k^2}{b^2}\left(\beta_{D|S_a}^2 + \beta_C^2\right)\right] \tag{9}$$

Previous studies have shown that P_{PL} is dictated by the hazard and not by the uncertainties in the demand and capacity. The above procedure, summarized in Eq. 9, fully couples hazard, demand, and capacity. However, as highlighted in Kwon (2007), the derivation of vulnerability curves is more advantageous than calculating fully coupled risks.

Application to a Typical Fixed Steel Jacket Platform

Analytical Modeling

Fragility analysis was performed on a typical FSJ offshore platform that was designed according to the API guidelines and presented by Nordal et al. (1987). It is assumed that all the lateral forces are transferred to the foundation mainly by the diagonal braces. Furthermore, at each level, the horizontal X-braces transfer the forces among the four vertical frames. In this study, the performance of the platform is estimated through fragility analysis when subjected to seismic forces. A 2-D model of the X-braced vertical frame on grid line A is developed in ETABS (2008); the section sizes are available in Nordal et al. (1987). A screenshot of this analytical model is shown in Fig. 1.

The numerical model of the structure is developed based on the following assumptions:

- All the connections are assumed to be pinned connections.
- All the members are modeled as truss elements.
- Rigid diaphragm condition is assumed only for the level 4, where the mass is defined.
- All supports are assigned pin conditions.

Fig. 1 Analytical model in
ETABS (**a**) 3D view (**b**)
elevation
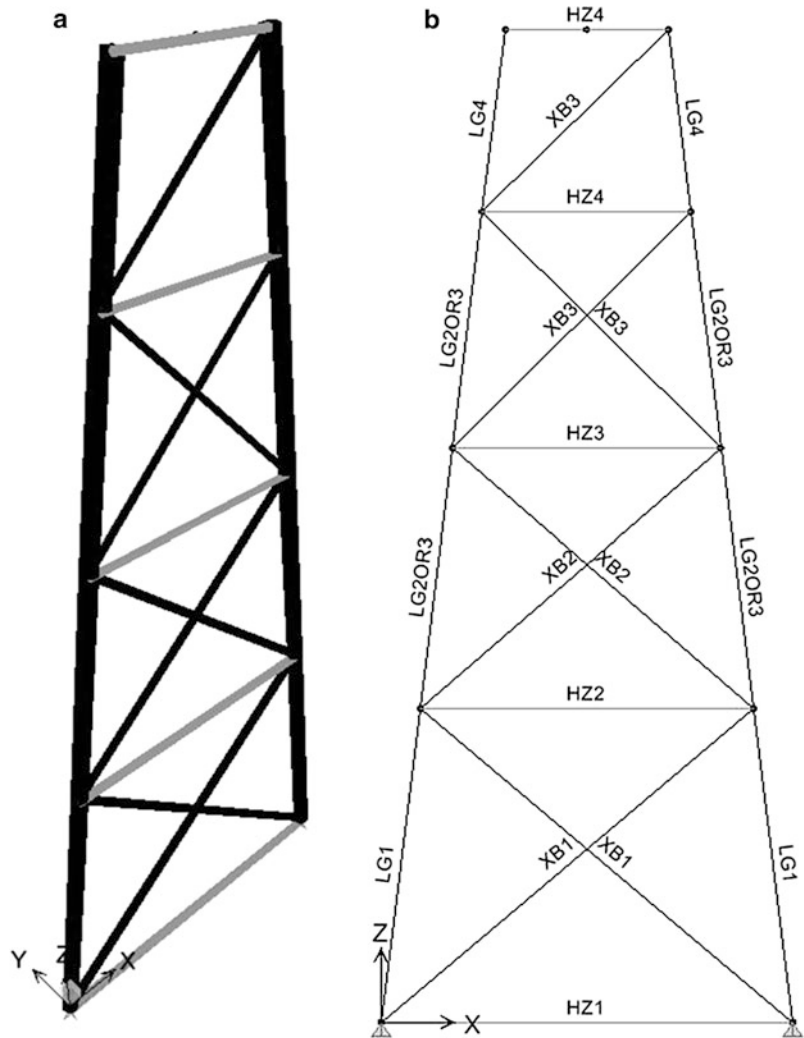
- The primary lateral force-resisting members are the braces.
- Nonlinearity effects are not modeled.

Based on the work of Stear and Bea (1997), the total deck load being supported is taken as 51 MN, which would be shared by the four frames in the *y*-direction. Hence, a mass of 12.75 MN (12,750 KN) is assigned to the top node of the frame. In addition, as discussed in the modeling assumptions, rigid diaphragm condition (all the nodes in a given floor are constrained together) is assigned for the top level.

Evaluation of Dynamic Properties

Using the above defined mass and section sizes, a modal analysis is performed to obtain the fundamental period. First two mode shapes of this structure along with their corresponding periods are shown in Fig. 2. The fundamental period was also determined using static pushover analysis to verify the period obtained using the modal analysis. The fundamental period obtained using both the analyses are in good agreement.

Selection of Ground Motions

For demonstrative purposes, ground motions are selected from the Pacific Earthquake Engineering
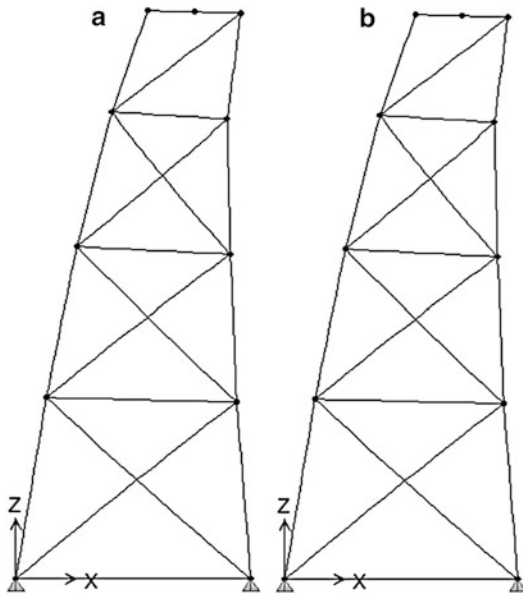
Seismic Fragility of Aging Offshore Platforms, Fig. 2 Mode shapes of the frame (**a**) Mode 1 ($T = 3.71$ s) (**b**) Mode 2 ($T = 0.123$ s)

Center (PEER) New Generation Attenuation (NGA) database. However, it is recommended to select the ground motions based on the site location, if available. The database is composed of 3,551 ground motion recordings that represent over 160 seismic events (including aftershock events) ranging in magnitude from M4.2 to M7.9 (Chiou et al. 2008). A total of 30 far-field ground motions are selected from the PEER NGA database. The PGA range for the final selected suite of ground motions, listed in Table 1, is 0.2 g to 1.8 g with an average of 0.5 g. The ground motion suite includes earthquakes that occurred from 1971 to 1999. The spectral acceleration at the fundamental period of the structure, 3.71 s corresponding to 5 % damping, is obtained by constructing the elastic response spectra of each ground motion (Seismosoft 2011). For instance, for the first ground motion in Table 1, Loma Prieta earthquake (A01090), the pseudo-acceleration elastic spectra is shown in Fig. 3.

Calculation of Story Drift

For a linear elastic system, the peak story drift can be obtained from the elastic pseudo-spectral

acceleration using the following equation (Chopra 2010):

$$A = \omega_n^2 D = \left(\frac{2\pi}{T_n}\right)^2 D \qquad (10)$$

where ω_n is the natural frequency of the structure, D is the peak deformation of the structure due to earthquake, T_n is the natural period of the structure, and A is the pseudo-spectral acceleration corresponding to T_n. The peak displacement of the top node, obtained above, is divided with total story height to obtain the peak drifts. The graph of peak drift versus spectral acceleration is shown in Fig. 4 for the ground motions listed above.

In general, the pseudo-spectral acceleration is very low because of the higher fundamental period. When compared to buildings, the percentage increment in the mass must be higher than the percentage increment in stiffness for the offshore structures. Having obtained the seismic hazard expressed in spectral acceleration and the demand expressed in interstory drift, a reliability analysis can now be performed. The limit states are discussed in the following section.

Formulation of Limit States

The performance limit states are of prime importance in the fragility analysis and need to be carefully determined. The three performance levels that were used for the offshore steel-braced frame are given in Table 2. However, more investigation is required to ascertain the applicability of these limit states to offshore structures. The limits corresponding to each level are also given in Table 2.

Fragility Analysis Results

Using the outlined procedure in section "Mathematical Details of SAC-FEMA Technique," the fragility curves are developed for the three limit states discussed in the previous section and are shown in Fig. 5. It can be inferred that the probability of reaching the higher limit state, characterized with drift 2 %, is almost 100 % for a seismic hazard as low as 1.2 g, which signifies the non-robust seismic design of the structure.

Seismic Fragility of Aging Offshore Platforms, Table 1   Considered ground motions (PEER: NGA Database 2008)

No	Name	PGA, g	Earthquake name	Year	Station name
1	A01090	0.294	Loma Prieta	1989	Foster City – APEEL 1
2	ABRD130	0.313	Whittier Narrows-01	1987	Brea Dam (Downstream)
3	AING000	0.299	Whittier Narrows-01	1987	Inglewood – Union Oil
4	AND250	0.244	Loma Prieta	1989	Anderson Dam (Downstream)
5	ATAR090	0.644	Whittier Narrows-01	1987	Tarzana – Cedar Hill
6	CH10270	0.269	Loma Prieta	1989	Oakland – Outer Harbor Wharf
7	CHY014W	0.229	Chi-Chi, Taiwan	1999	CHY014
8	CHY080N	0.218	Chi-Chi, Taiwan-03	1999	CHY080
9	G03090	0.367	Loma Prieta	1989	Gilroy Array #3
10	Go2000	0.363	Loma Prieta	1989	Gilroy Array #2
11	Go2090	0.321	Loma Prieta	1989	Gilroy Array #2
12	H-Z14090	0.274	Coalinga-01	1983	Parkfield – Fault Zone 14
13	HDA165	0.269	Loma Prieta	1989	Hollister Diff. Array
14	HDA255	0.279	Loma Prieta	1989	Hollister Diff. Array
15	HEC090	0.337	Hector Mine	1999	Hector
16	LAC180	0.316	Northridge-01	1994	LA – City Terrace
17	CHY028-E	0.653	Chi-Chi, Taiwan	1999	CHY028
18	CHY028-N	0.795	Chi-Chi, Taiwan	1999	CHY028
19	H-PVY045	0.592	Coalinga-01	1983	Pleasant Valley P.P. – yard
20	PET090	0.662	Cape Mendocino	1992	Petrolia
21	TAR360	0.990	Northridge-01	1994	Tarzana – Cedar Hill A
22	TCU074-E	0.597	Chi-Chi, Taiwan	1999	TCU074
23	ORR090	0.568	Northridge-01	1994	Castaic – Old Ridge Route
24	ORR291	0.268	San Fernando	1971	Castaic – Old Ridge Route
25	ORR360	0.514	Northridge-01	1994	Castaic – Old Ridge Route
26	PEL090	0.231	San Fernando	1971	LA – Hollywood Stor FF
27	PEL180	0.174	San Fernando	1971	LA – Hollywood Stor FF
28	PEL360	0.358	Northridge-01	1994	LA – Hollywood Stor FF
29	TCU071-N	0.380	Chi-Chi, Taiwan-03	1999	TCU071
30	TAR090	1.779	Northridge-01	1994	Tarzana – Cedar Hill A

Seismic Fragility of Aging Offshore Platforms,
Fig. 3 Pseudo-acceleration elastic response spectra ($\zeta = 5\,\%$)
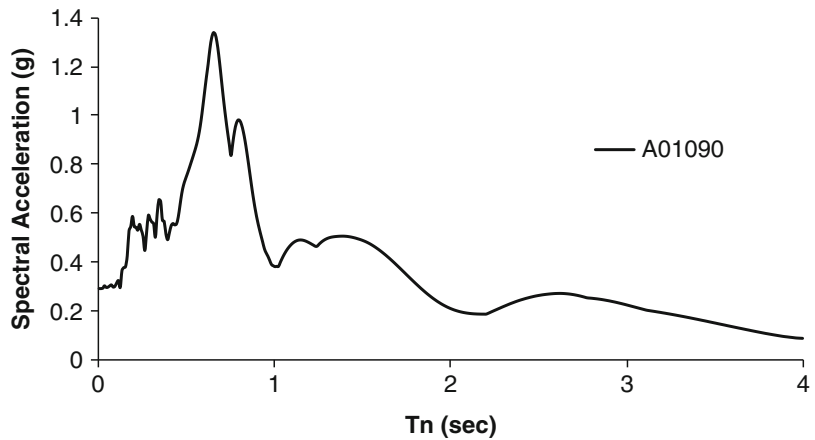
Seismic Fragility of Aging Offshore Platforms, Fig. 4 Peak drift versus spectral acceleration
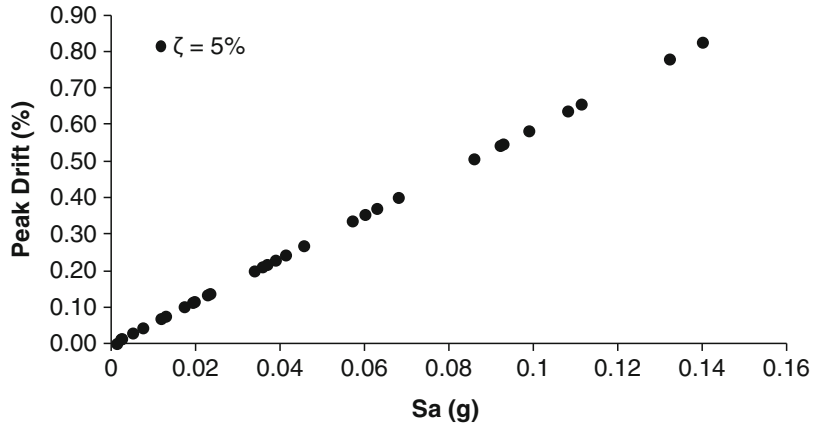
Seismic Fragility of Aging Offshore Platforms, Table 2 Structural performance levels and damage, expressed in terms of drift (Source: ASCE 2007)

Elements	Type	Structural performance levels		
		Collapse prevention	Life safety	Immediate occupancy
Braced steel frames	Primary	Extensive yielding and buckling of braces Many braces and their connections may fail	Many braces yield or buckle but do not totally fail Many connections may fail	Minor yielding or buckling of braces
	Secondary	Same as primary	Same as primary	Same as primary
	Drift	2 % transient	1.5 % transient	0.5 % transient
		or permanent	0.5 % permanent	negligible permanent

The slope of the fragility curve decreases with the limit state which implies that relatively higher hazard increment is necessary to reach the probability of failure as the limit state is increased.

Summary

In this entry, the seismic performance of a typical fixed steel jacket platform is evaluated through fragility analysis. The analysis has shown that the considered offshore platform cannot withstand even the seismic hazard corresponding to the collapse prevention performance level. This undermines the robustness of such structures which are aging, for the ever-increasing seismic hazard. Even though the results were shown on a simple elastic system, the authors believe that the implications could be more serious when an inelastic system was used, as the drift demand would be much higher albeit with a reduction in force demand. Though the fragility curves are derived where the demand is expressed in terms of interstory drift, similar curves can be derived when the demand is expressed in terms of acceleration, residual drift, etc.

The procedure described herein to perform fragility analysis is simple and practical. Hence, when assessing the performance of similar FSJ platforms, practicing engineers can implement it easily. The above results may also be used to prioritize the retrofitting strategy of aging platforms. Fragility curves provide a quick, yet, sufficient insight in to the performance robustness of the structures and help to determine the members that need the utmost attention. This is also ideal for planning the inspection and maintenance strategies for these structures. It is to be noted that other significant limit states also influence such decisions for FSJ platforms; the presented approach is recommended for characterizing the seismic effects along with appropriate consideration of other limit states.

S

Seismic Fragility of Aging Offshore Platforms,
Fig. 5 Seismic fragility curves for interstory drift
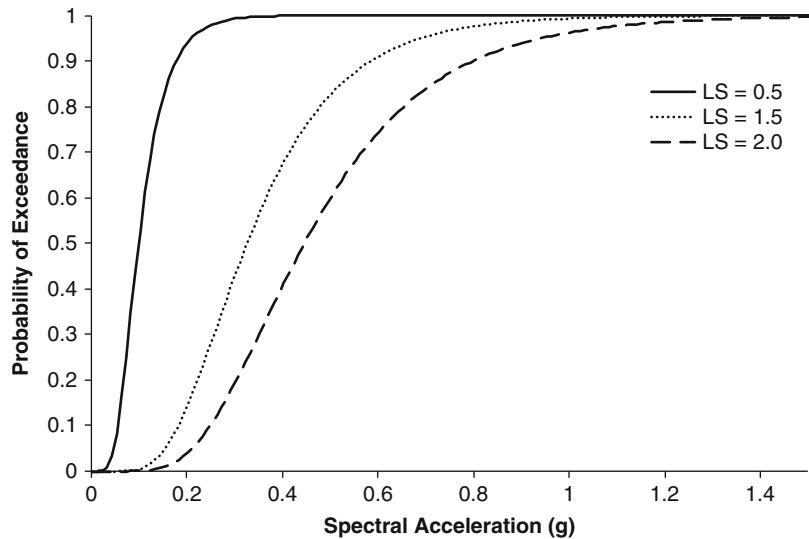
Cross-References

▶ Analytic Fragility and Limit States [P(EDP|IM)]: Nonlinear Static Procedures
▶ Modal Analysis
▶ Performance-Based Design Procedure for Structures with Magneto-Rheological Dampers
▶ Seismic Collapse Assessment
▶ Seismic Fragility Analysis
▶ Seismic Risk Assessment, Cascading Effects

References

ASCE (2007) Seismic rehabilitation of existing buildings. ASCE/SEI Standard 41-06, Reston

Chiou B, Darragh R, Gregor N, Silva W (2008) NGA project strong-motion database. Earthq Spectra 24(1):23–44

Chopra AK (2010) Dynamics of structures–theory and applications to earthquake engineering. Prentice-hall, NewJersey 07458, U.S.A.

ETABS "Version 9.5.0" (Software) (2008) Computers and Structures, Berkeley. www.csiberkeley.com

FEMA 350 (2000) Recommended seismic design criteria for new steel moment-frame buildings. NEHRP Program

Ferreira DD (2003) Fiscal treatment decommissioning and bonds: anticipating impacts of financial assurance requirements for offshore decommissioning, a decision model for the oil industry. Doctoral dissertation, State University Campinas, Brazil, pp 164–168

Kwon O (2007) Probabilistic seismic assessment of structure, foundation, and soil interacting systems. PhD dissertation, University of Illinois at Urbana-Champaign

Pacific Earthquake Engineering Research Center (PEER): NGA Database (2005) The Regents of the University of California (Regents). http://peer.berkeley.edu/nga/. Accessed 10 Nov 2013

Nordal H, Cornell CA, Karamchandani A (1987) A structural system reliability care study of an eight-leg steel jacket offshore production platform. In: Proceedings of The Marine Structural Reliability Symposium, SNAME, Arlington, VA, October 1987.

Pinto PE, Giannini R, Franchin P (2004) Seismic reliability analysis of structures. IUSS Press, Pavia

Sadeghi K (2007) An overview of design, analysis, construction and installation of offshore petroleum platforms suitable for cyprus oil/gas fields. GAU J Soc Appl Sci 2(4):1–16

SeismoSignal "Version 4.3.0" (Software) (2011) Seismosoft, Pavia. www.seismosoft.com

Stacey A, Birkinshaw, M, Sharp JV (2008) Life extension issues for ageing offshore installations. In: American Society of Mechanical Engineers, proceedings of the international conference on offshore mechanics and arctic engineering, vol 5, pp 199–216

Stear JD, Bea RG (1997) Earthquake analysis of offshore platform, screening methodologies. Technical Report to Joint Industry Project Sponsors, Project Phase III, Department of Civil Engineering, University of California at Berkeley

Seismic Instrument Response, Correction for

Jens Havskov[1] and Gerardo Alguacil[2]
[1]Department of Earth Science, University of Bergen, Bergen, Norway
[2]Instituto Andaluz de Geofísica, University of Granada, Granada, Spain

Synonyms

Accelerometer; Amplitude response; Digitizer; Fourier transform; Phase response; Poles and zeros; Response function; Seismometer

Recording seismic signals, whether in digital or analog form, gives a number in a computer or an amplitude on paper (Fig. 1). For velocity sensors this number is proportional to ground velocity above a certain frequency, and for accelerometers the number is proportional to ground acceleration.

What is really needed is the true motion of the ground in displacement, velocity, or acceleration. So considering Fig. 1, the signal to the right is recorded, and the signal to the left is what is needed. They are obviously different. Seismologists most often use displacement while engineers prefer acceleration. In most cases the correction is not as easy as just multiplying with a constant to get the ground motion for two reasons:

– The seismic sensor is not recording the desired ground motion magnitude: displacement, velocity, or acceleration and a transformation must be done.
– The response of the seismic sensor to the ground motion, as well as the response of the recorder, is in general frequency dependent, so a frequency-dependent correction must be made (Fig. 1).

It is thus essential to know how the seismic sensor and recorder modify the input signal, the ground motion, to produce the recorded output signal. If, e.g., for given harmonic ground amplitude displacement $X(\omega)$, the output amplitude $Y(\omega)$ can be calculated as

$$Y(\omega) \,=\, X(\omega) \cdot A(\omega) \qquad (1)$$

where $A(\omega)$ is the displacement amplitude response, then the displacement, $X(\omega)$, of the ground can simply be calculated as

$$X(\omega) \,=\, A(\omega)/Y(\omega) \qquad (2)$$

$A(\omega)$ has traditionally been called the magnification, since for an analog recorder $A(\omega)$ gives how many times the signal is magnified at different frequencies.

In this document, a description of the response function for different kinds of sensors will be described, as well as methods used for making a correction for the instrument response to arrive at the true ground motion, whether in time domain or frequency domain.

The Elements of a Seismic Station

A seismic station consists of a sensor connected to a recorder. The sensor has an output in volts proportional to the ground velocity (velocity sensor) or to the ground acceleration (an accelerometer), and this output is generally only constant in a specific frequency range. The proportionality constant is called the generator constant G and has units of $V/(m \cdot s^{-1})$ and V/g (g is the gravity acceleration) for the two types of sensors, respectively. The sensor output signal enters the digital recorder which will have some filter built in at high frequency to avoid aliasing. The digitizer will give out a number (called counts) proportional to the input voltage, and this defines the digitizer sensitivity in terms of counts/V.

Frequency Response of a Linear System

A seismic sensor and associated electronics in the seismic recorder is assumed to behave as a linear

Seismic Instrument Response, Correction for, Fig. 1 The input is the ground moving with a maximum displacement 0.02 mm. This "signal" goes through an accelerometer and becomes a new signal with a maximum number (count) of 1,250,000. Note also that the frequency content is different for input and output

system. The linearity means that there is a linear relationship between input signal and output signal. If the input signal is $x(t)$ and the output $y(t)$, then multiplying $x(t)$ with a constant will result in an output signal multiplied with the same constant. For example, if the ground velocity is doubled, then the output from the seismometer is also doubled. If two signals of different frequency and amplitude are input, then also two signals with the same frequencies (with different amplitude and phase) are output. The frequency-dependent relation between the output and the input is called the frequency response function. For a more complete description, see Scherbaum (2007).

A very simple linear system is the RC filter consisting of a resistor R and a capacitor C. An RC low-pass filter is seen in Fig. 2.

This circuit lets low frequencies pass without attenuation, while high frequencies are attenuated due to the frequency-dependent impedance of the capacitor. For a sine wave signal of frequency f, the capacitor impedance is the relation between the voltage amplitude at the capacitor terminals and the amplitude of the current flowing through the capacitor.

The impedance Z_c of a capacitor seen by the input for a sine wave signal of frequency f is

$$Z_c = \frac{1}{2\pi f C} = \frac{1}{\omega C} \tag{3}$$

where C is the capacitance (F) and $\omega = 2\pi f$. The voltage over the capacitor is delayed by one fourth cycle (90°) with respect to the current through it. Considering that the RC filter is a frequency-dependent voltage divider, using a monochromatic signal of angular frequency ω and amplitude $X(\omega)$, $x(t) = X(\omega)\cos(\omega t)$, the
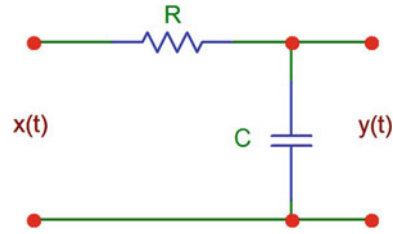
Seismic Instrument Response, Correction for, Fig. 2 RC low-pass filter. R is the resistor and C the capacitor. Input is $x(t)$ and output $y(t)$

output signal amplitude $Y(\omega)$ can be written as (e.g., Havskov and Alguacil 2010)

$$Y(\omega) = \frac{1}{\sqrt{1 + \omega^2 R^2 C^2}} X(\omega) = \frac{1}{\sqrt{1 + \omega^2 / \omega_0^2}} X(\omega) \tag{4}$$

where $\omega_0 = 1/RC$. ω_0 is also called the corner or cutoff frequency of the filter, and for $\omega = \omega_0$, the amplitude has been reduced to $1/\sqrt{2} = 0.707$ (see Fig. 4). From Eq. 4 the amplitude frequency response function of the filter $A(\omega)$ (since only amplitudes are considered) is

$$A(\omega) = \frac{1}{\sqrt{1 + \omega^2 / \omega_0^2}} = \frac{Y(\omega)}{X(\omega)} \tag{5}$$

If $A(\omega)$ is completely known, then the amplitude of the harmonic input signal $X(\omega)$ can be calculated from the measured signal as

$$X(\omega) = Y(\omega) / A(\omega) \tag{6}$$

The amplitude response can be measured very simply as shown in Fig. 3. By varying the frequency, both input and output amplitudes can be
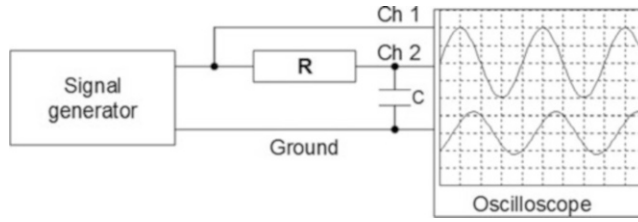
Seismic Instrument Response, Correction for, Fig. 3 Measuring the amplitude and phase response function of an RC filter. The signal from a signal generator goes directly to channel 1 (Ch1, *top trace*) on the oscilloscope and to channel 2 (Ch2, *bottom trace*) through the filter so both input and output is measured

measured at different frequencies to produce the amplitude response function.

From Fig. 3, it is seen that the output signal not only has been changed in amplitude but also has been delayed a little relative to the input signal; in other words, there has been a phase shift; see definition below.

In this example, the phase shift is negative (see definition in Eq. 7). The complete frequency response of the filter therefore consists of both the amplitude response function and the phase response function $\Phi(\omega)$; see Fig. 4. Considering a general input harmonic waveform $x(\omega,t) = X(\omega) \cdot \cos(\omega t)$ at frequency ω, the output can be written as

$$y(\omega, t) \ = \ X(\omega) \cdot A(\omega) \ \cdot \ \cos(\omega t \ + \ \Phi(\omega)) \tag{7}$$

The phase shift is here defined as a quantity being added to the phase as seen above. Thus comparing Fig. 4 and Eq. 7, it is seen that the phase shift is negative. This is the most common way of defining the phase shift, but the opposite sign is sometimes seen, and it may then be called phase delay or phase lag. So it is very important to know which definition has been used.

Using Eq. 7 to correct the amplitude and phase of a single monochromatic signal in time domain is easy, but to make the correction for a whole signal with different frequencies is not simple.

This can be considerably simplified if the complex representation of harmonic signals is used. Instead of writing $\cos(\omega t)$, the real part of the exponential function can be used:
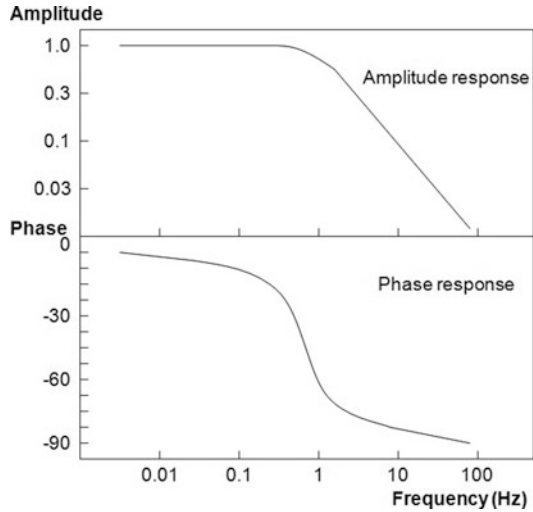
Seismic Instrument Response, Correction for, Fig. 4 Amplitude and phase response of an RC filter with a corner frequency of 1 Hz. The phase response (or phase shift) is given in degrees

$$e^{i\omega t} \ = \ \cos(\omega t) \ + \ i \ \sin(\omega t) \tag{8}$$

Equation 7 can now be written as

$$\begin{aligned} y(\omega, t) &= X(\omega)A(\omega)e^{i(\omega t + \Phi(\omega))} \\ &= X(\omega)A(\omega)e^{i\omega t}e^{i\Phi(\omega)} \end{aligned} \tag{9}$$

$y(\omega,t)$ is now a complex number of which the real part is the actual output. This can be further simplified considering that any complex number Z can be written as

$$Z = a + ib = \sqrt{a^2 + b^2}e^{i\Phi} = |Z|e^{i\Phi} \tag{10}$$

S

where $\Phi = \tan^{-1}(b/a)$ which also follows from Eq. 8. The complex frequency response $T(\omega)$ can then be defined as

$$T(\omega) = A(\omega)e^{i\Phi(\omega)} = |T(\omega)|e^{i\Phi(\omega)} \quad (11)$$

and Eq. 9 can be written as

$$y(\omega, t) = X(\omega)T(\omega)e^{i(\omega t)} \quad (12)$$

Now only one complex function, $T(\omega)$, that includes the phase shift can therefore completely describe the instrument frequency response. So far only monochromatic signals have been used and the corrections can easily be made with both the real and complex representations Eqs. 7 and 12, so it might be hard to see why the complex representation of the response function is needed. Later it is going to be shown how an observed seismogram is corrected, not only for one amplitude at a time, but actually for the whole signal (containing a range of frequencies), which is possible with digital data. Obviously this cannot be done with one amplitude at a time, but will have to be done with the amplitude and frequency content of the whole signal, in other words, making spectral analysis. It will hopefully then be clear why a complex representation is needed.

Frequency Response of Seismic Sensors

There are two important types of seismic sensors:

1. Velocity sensor (seismometer). This sensor has an output proportional to velocity in a limited frequency band; see Fig. 5. This is both the case for the traditional passive seismometers like geophones (for frequencies down to about 1 Hz) and modern broadband seismometers (down to frequencies of 0.003–0.03 Hz).
2. Accelerometer. This sensor usually has an output proportional to ground acceleration from 0 Hz (DC) to some cutoff frequency (see Fig. 6). In some cases there will also be a lower cutoff, e.g., at 0.1 Hz.

Velocity Sensor

The complex response function for a velocity sensor for displacement, velocity, and acceleration is the following (Havskov and Alguacil 2010).

Displacement

$$T_d(\omega) = \frac{Z(\omega)}{U(\omega)}G = \frac{i\omega^3 G}{\omega_0^2 - \omega^2 + i2\omega\omega_0 h} \quad (13)$$

Velocity

$$T_v(\omega) = \frac{\omega^2 G}{\omega_0^2 - \omega^2 + i2\omega\omega_0 h} \quad (14)$$

Acceleration

$$T_a(\omega) = \frac{-i\omega G}{\omega_0^2 - \omega^2 + i2\omega\omega_0 h} \quad (15)$$

where G is the generator constant, ω_0 is the natural frequency of the sensor (the frequency at which it will oscillate if no damping), and h is the damping coefficient (how the motion of the mass is damped; see Fig. 5). It is seen that the only difference between the respective three types of response curves is the factor $i\omega$.

For the velocity response, the amplitude and phase response are (Fig. 5)

$$A_v(\omega) = |T_v(\omega)| = \frac{G\omega^2}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4h^2\omega^2\omega_0^2}} \quad (16)$$

$$\Phi_v(\omega) = a\tan\left(\frac{\text{Im}(T_v(\omega))}{\text{Re}(T_v(\omega))}\right) = a\tan\left(\frac{-2h\omega\omega_0}{\omega_0^2 - \omega^2}\right) \quad (17)$$

It is seen that as the damping decreases, the sensor response to velocity becomes constant for frequencies above the natural frequency. It is also seen that if there is little damping, the sensor will have a high gain at the natural frequency, so damping is usually set to 0.707.

Seismic Instrument Response, Correction for, Fig. 5 The amplitude and phase response functions for a velocity seismometer with a natural frequency of 1 Hz. Curves for various levels of damping *h* are shown. It is assumed that the generator constant is 1
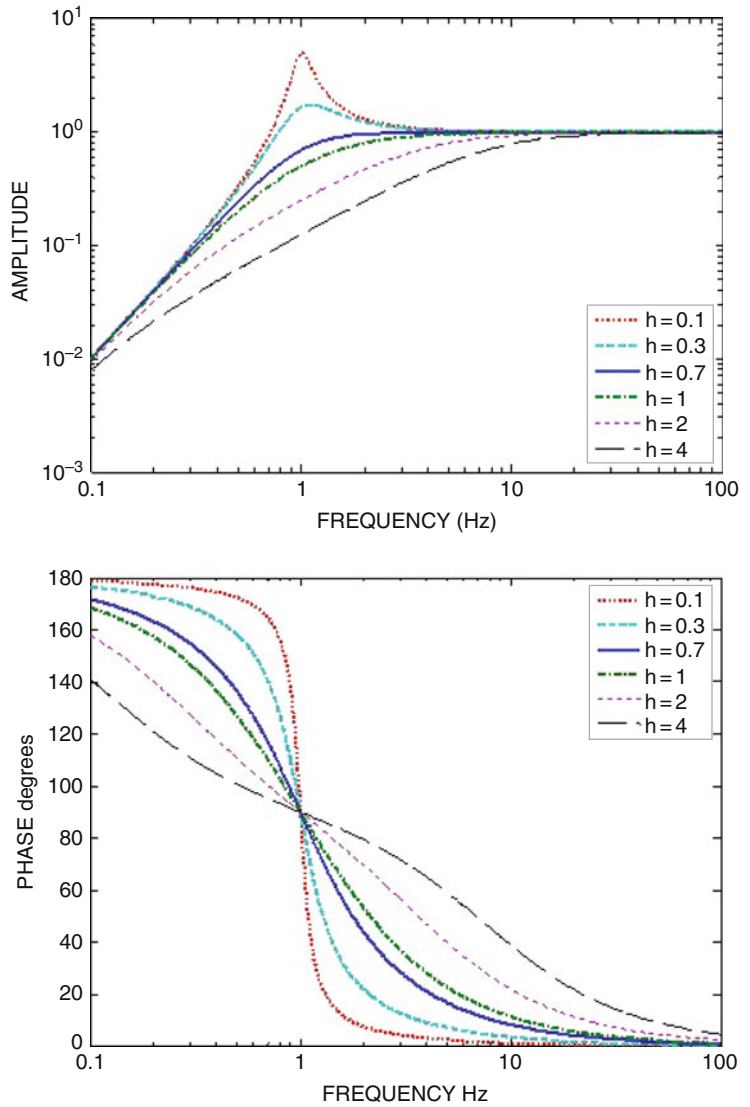
Accelerometer

The acceleration response curve for an FBA accelerometer is (Havskov and Alguacil 2010)

$$T_a(\omega) = \frac{G/\omega_f^2}{\omega_f^2 - \omega^2 + i2\omega\omega_f h} \quad (18)$$

where ω_f is the effective natural (angular) frequency with feedback. The amplitude of this response has a flat value of G from DC up to this frequency (usually above 50 Hz). The instrument correction is then very easy since it is assumed that the feedback natural frequency of the sensor is very high compared to the frequencies of interest so the response for acceleration is simply approximated by multiplying with the generator constant

$$T_a(\omega) = G \quad (19)$$

The velocity and displacement response is then:

Velocity

$$T_v(\omega) = G \cdot i\omega \qquad (20)$$

Displacement

$$T_d(\omega) = -G \cdot \omega^2 \qquad (21)$$

So for engineering purposes, where acceleration is needed, it is very easy to correct for the accelerometer response since it is just a question of multiplying with a constant, while seismologist wanting displacement must make the transformation (corresponding to integrating the signal twice).

In practice, many accelerograms, especially – but not only – those recorded with earlier accelerographs, are affected by offset or baseline shifts that may influence the estimation of characteristic parameters of strong ground motion (Boore 2001). The source of this long-period noise may be true ground motions (rotation, tilt, local deformation) or instrumental (mechanical nonlinearity, electronic effects) and is not yet fully understood (Akkar and Boore 2009). Prior to any integration to compute the ground velocity or displacement, the accelerogram should be corrected from these effects.

In many practical cases, this correction may consist simply in applying a high-pass filter with a suitable cutoff frequency, as it will be shown below. In some cases, a particular processing scheme adapted to an individual record is required. In fact, most accelerogram databases (e.g., Internet Site for European Strong-motion Data, www.isesd.hi.is) include a "processed" version of each accelerogram, which usually consists of baseline correction, band-pass filtering, and conversion to ground motion units.

It is interesting to note that whether a velocity sensor or an accelerometer is used, it is possible to calculate displacement, velocity, and acceleration from both types of sensors; however, as is shown later, in practice, this will be frequency limited due to noise in the system.

Complete Frequency Response

For passive sensors, the above equations represent exactly the seismometer response, and for many broadband sensors, this is also the case up to a particular frequency. A recording of a seismic signal also involves the seismic recorder which also has a frequency response. Modern recorders have little amplitude change or phase shift in the main frequency band used, so the frequency dependence of the recorder is often ignored. Some recorders might modify the signal enough for a correction to be made. In order to completely describe the response for active sensors and possibly also include the response of the recorder, a more general representation of the response function must be used. It turns out that $T(\omega)$ for all systems made from discrete mechanical or electrical components (masses, springs, coils, capacitors, resistors, semiconductors, etc.) can be represented exactly (e.g., Scherbaum 2007) by rational functions of $i\omega$:

$$T(\omega) = c \frac{(i\omega - z_1)(i\omega - z_2)(i\omega - z_3).....}{(i\omega - p_1)(i\omega - p_2)(i\omega - p_3).....} \qquad (22)$$

where c is the combined normalization constant for nominator and denominator polynomials, z are the zeros (or roots) of the nominator polynomial, while the zeros of the denominator polynomial (poles) are p. Using Eq. 22 to represent $T(\omega)$ is the so-called poles and zeros representation, which has become the most standard way.

For example, for the seismometer response function for velocity, $T_v(\omega)$ can be written as

$$T_v(\omega) = \frac{-(i\omega - 0)(i\omega - 0)}{(i\omega - p_1)(i\omega - p_2)} \qquad (23)$$

where

$$p_1 = -\omega_0\left(h + \sqrt{h^2 - 1}\right) \qquad (24)$$

$$p_2 = -\omega_0\left(h - \sqrt{h^2 - 1}\right)$$

So in addition to the poles p_1 and p_2, the seismometer response function has a double zero at $z = 0$ and the normalization constant is -1. Note that since h usually is smaller than 1, the poles are usually complex. Complex poles always appear as conjugate pairs. For the displacement response, the equation is

$$T_d(\omega) = \frac{(i\omega - 0)(i\omega - 0)(i\omega - 0)}{(i\omega - p_1)(i\omega - p2)} \qquad (25)$$

and there is thus one more zero. For the standard accelerometer, the displacement response is simply $(i\omega)^2$ which corresponds to two zeros.

For active sensors, there might be more poles and zeros representing the filtering inherent in active sensors.

The frequency dependence of a recorder can equally be represented by poles and zeros, so it is just a question of adding to the equation in the pole and zero representation.

In addition to the frequency-dependent elements, a complete seismograph or accelerograph also has the frequency-independent elements, the generator constant, and the digitizer gain which has to be multiplied with the normalization constant.

Example

A seismometer has a generator constant G = 300 V/ms^{-1} and the digitizer has a sensitivity of 100,000 counts/V. The velocity sensitivity of the seismograph in the passband (flat for velocity) is then 100,000 counts/V · 300 V/ms^{-1} = 3 · 10^7 counts/ms^{-1}. This number must then be multiplied with the normalization constant for the rest of the system which is the sensor and possibly some filters. For displacement, the amplitude displacement increases with frequency (Fig. 6), but that is
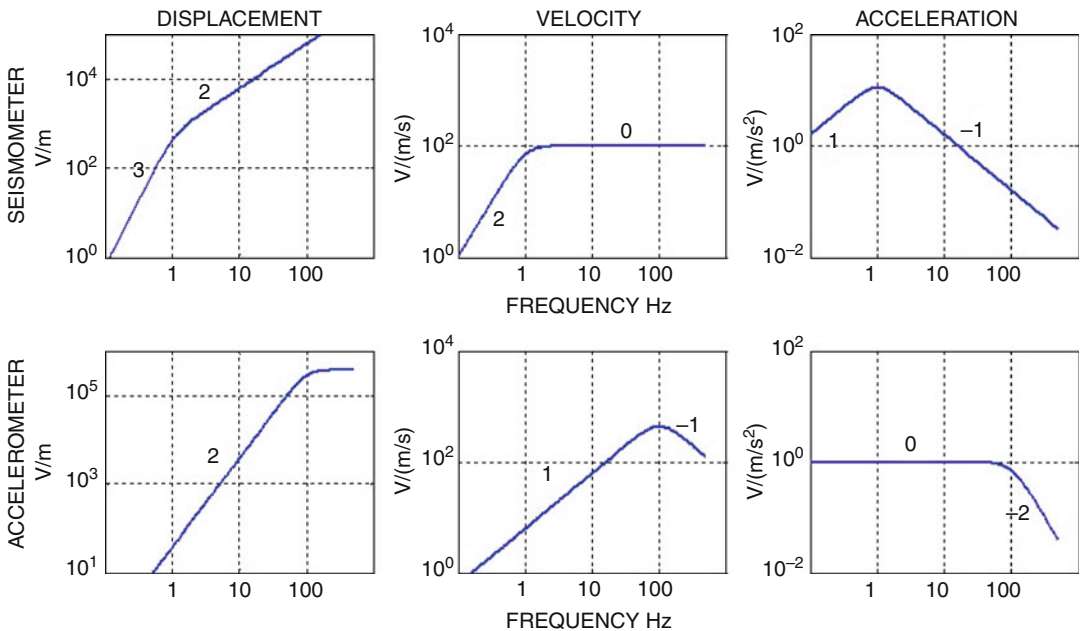
Seismic Instrument Response, Correction for, Fig. 6 Amplitude frequency response of a 1 Hz velocity sensor (*top*) and an accelerometer (*bottom*) with a natural frequency of 100 Hz. From *left* to *right*, the figures show the sensor response for ground displacement, velocity, and acceleration, respectively. The axes are logarithmic and the horizontal axes show frequency in Hz. The asymptotic slope for each segment is indicated. Note how one curve translates into another by just adding or subtracting one unit of the slope corresponding to a multiplication or division by frequency. The response curves assume a generator constant G = 100 V/(m/s) for the seismometer and G = 1 V/(m/s^2) for the accelerometer. Please note that the vertical scales are different
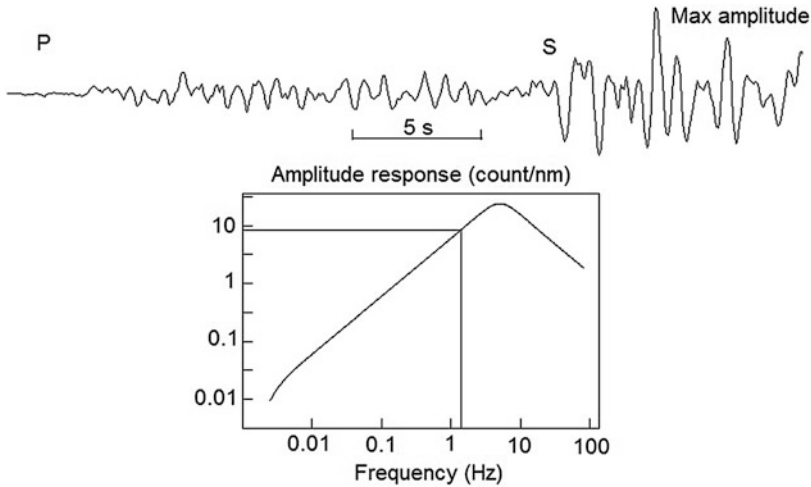
Seismic Instrument Response, Correction for, Fig. 7 *Top*: Seismogram of a local earthquake; the position of the maximum amplitude is indicated. The frequency of the maximum amplitude is 1.4 Hz. The signal is recorded on a broadband station and the amplitude is therefore proportional to ground velocity. *Bottom*: The amplitude displacement response function. The gain corresponding to the frequency of 1.4 Hz is indicated

taken care of with the $i\omega$ factor so the normalization constant still has to be multiplied with the same number. For example, at 10 Hz, the gain for displacement will be $2\pi \cdot 10\text{s}^{-1} \cdot 3 \cdot 10^7$ counts/ms^{-1} = $1.88 \cdot 10^9$ counts/m.

How to Correct for the Response

In the simplest case, there is only a need to correct for a single amplitude reading. In Fig. 7, a signal from a local earthquake ($M = 5.2$) recorded on a broadband station is seen. Figure 7 shows the corresponding amplitude response curve for displacement. Notice that around 8 Hz, there is a filter cutting off the high frequencies. This filter is part of the sensor electronics and will be represented by poles and zeros in the response function. In the seismogram the maximum amplitude is 1,105,000 counts at a frequency of 1.4 Hz. From the response curve the gain of the system is 8 counts/nm or $8 \cdot 10^9$ counts/m. So the maximum ground displacement amplitude is 1,205,000/8 = 138,000 nm. In practical seismology, nm is mostly used for ground motion since it is a more convenient unit than m. In this example, there was no need for a complex response function and the phase shift was not considered since only one particular amplitude was used.

In the more general case, it is desirable to make instrument correction for the complete seismogram. The seismogram seen in Fig. 7 is proportional to ground velocity in its passband, and a general procedure is therefore needed to calculate the corresponding displacement seismogram. Obviously it is not possible to deal with one frequency at a time and the amplitude and frequency content of the whole signal must be used. In other words, a spectral analysis must be used.

The complex frequency response function $T(\omega)$ is now assumed known and the complex spectra of the input and output signals $x(t)$ and $y(t)$ can be defined as $X(\omega)$ and $Y(\omega)$, respectively, using the definition for the Fourier spectrum (e.g., Oppenheim and Schafer 1989). Knowing the complete complex output spectrum, the complete complex input spectrum can be obtained:

$$X(\omega) = Y(\omega)/T(\omega) \qquad (26)$$

If $T(\omega)$ is a seismic instrument response, the instrument-corrected ground motion spectrum is now obtained, and since $X(\omega)$ is complex, this also includes the correction for phase.

The separate amplitude and phase spectra of the seismic signal can then be obtained as

$$A_{\text{signal}}(\omega) = \sqrt{\text{Re}(X(\omega))^2 + \text{Im}(x(\omega))^2} \quad (27)$$

$$\Phi_{\text{signal}}(\omega) = \tan^{-1}\left(\frac{\text{Im}(X(\omega))}{\text{Re}(X(\omega))}\right) \quad (28)$$

Thus in practice, the ground displacement spectrum would be calculated by taking the complex Fourier spectrum of the output signal and dividing it with the complex displacement response function and finally taking the absolute part of the complex ground displacement spectrum. This is basically what is done to calculate source spectra for earthquakes (Havskov and Ottemöller 2010).

All of this could of course easily have been done with the corresponding noncomplex equations since, so far, the phase information has not been used and usually phase spectra for earthquake signals are not used. However, the next step is to obtain the true ground displacement time domain signal instead of the frequency domain signal. With the knowledge about Fourier transforms, this is now easy, since, when all the frequency domain coefficients or the spectrum has been calculated, the corresponding time domain signal can be generated by the inverse Fourier transform. Since the corrected signal consists of a sum of cosine signals with different phases, each of them delayed differently due to the instrument response, the shape of the output signal will depend on the phase correction. So now life is easy, since the corrected complex displacement spectrum is already corrected for phase and the ground motion can be obtained as

$$x(t) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} Y(\omega)/T(\omega) \cdot e^{i\omega t}d\omega \quad (29)$$

While it is possible to use only half of the positive frequencies for making the amplitude spectrum, both positive and negative frequencies must be used for the inverse transformation. Normalization constants of the Fourier transform can be defined in different ways, but if the same routine

is used for both forward and inverse, the normalization constants will cancel out. In practice, since the signal consists of discrete numbers, the spectrum also is done with discrete numbers; see, e.g., Oppenheim and Schafer (1989).

Response Correction in Practice

In theory it is possible to recover the ground motion (displacement, velocity, or acceleration) at any frequency knowing the instrument response. In practice, one has to be careful to only do this in the frequency band where the instrument records real ground motion and not just electronic noise, since the instrument correction then becomes unstable and the output has nothing to do with the real seismic signal. Figure 8 shows an example.

The figure shows the influence of filtering, when estimating the ground displacement signal. In the frequency band 1–10 Hz, the signal looks very much like the original signal although a bit more low frequency, since it is converted to displacement and can nearly be considered an integration of the original signal. In the 0.1–10 Hz range, the earthquake signal almost disappears in the microseismic background noise. Why is this signal considered seismic noise and not instrumental generated noise? First, the earthquake signal has about the same amplitude as above; second, it "looks" like seismic background noise; and third, the amplitude at 1,290 nm is at a period of 5 s (peak amplitude of microseismic noise) which looks reasonable compared to worldwide observations. Note that this is how the earthquake signal would have looked being recorded on a broadband sensor, hardly noticeable. The last trace shows the calculation of the displacement without filtering so the lowest frequency used is $1/T$, where T is the length of the window, here 80 s (only 65 s shown) so f = 0.0125 Hz. The amplitude is now more than 50,000 nm and the signal looks "funny." The large amplitude obviously cannot be right since the microseismic noise has its largest amplitude around 3–8 s and it was 1,290 nm. So this is a clear case of trying to make a displacement signal at frequencies lower than where seismic signals exist in the data.
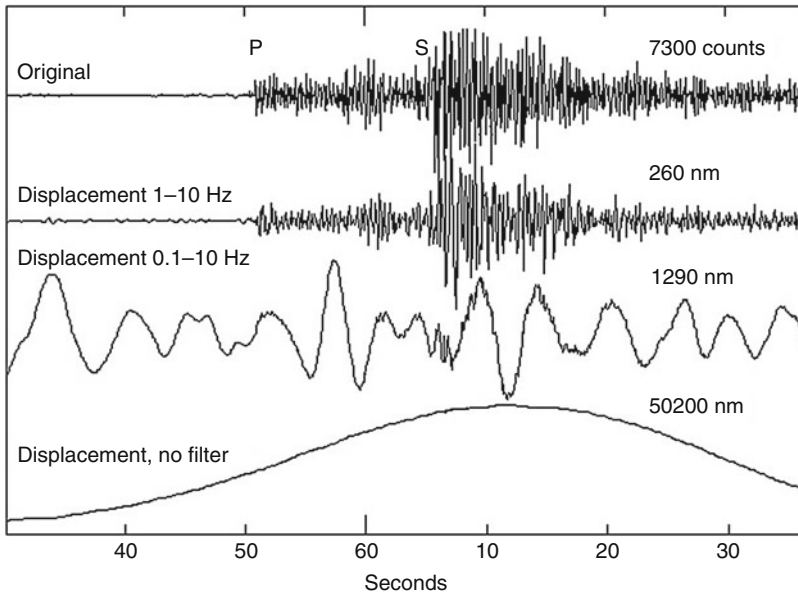
Seismic Instrument Response, Correction for, Fig. 8 Instrument corrections in different filter bands. The top trace is the original recording of a small earthquake with a 1 Hz seismometer. The three bottom traces have been converted to displacement using different filters. The third trace from the top shows harmonic waves at a frequency of about 0.2–0.3 Hz which is typical microseismic background noise originating in the ocean. The amplitudes to the right are maximum amplitudes (Figure from Havskov and Alguacil (2010))
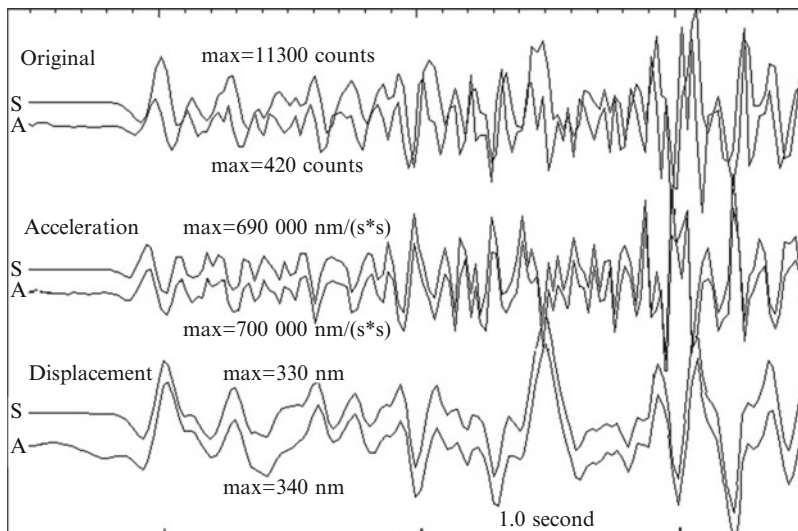
Seismic Instrument Response, Correction for, Fig. 9 Acceleration and displacement. The seismogram in the figure is the first few seconds of a P-wave. On the site there is also an accelerometer installed (A) next to the seismometer (S). The top traces show the original records in counts. The signal from the seismometer is similar to the accelerometer signal but of lower dominant frequencies, and the amplitudes are different. The middle traces show the two signals converted to accelerations and the bottom traces converted to displacement (frequency band 1–20 Hz) (Figure from Havskov and Alguacil (2010))

The ratio of the displacement gain for a 1 Hz seismometer at 1 Hz and 0.0125 Hz is $1/0.0125^3 = 5 \cdot 10^5$. In other words, if the gain at 1 Hz is 1.0, the signal must be multiplied by 1.0 to get the displacement, while at 0.0125 Hz, it must be multiplied by $5 \cdot 10^5$. So any tiny amount of instrumental noise present at low frequencies will blow up in the instrument correction. In the above example, it seems that the displacement signal can be recovered down to 0.1 Hz with a 1.0 Hz sensor.

Engineers often think accelerometers are the only instrument to use and seismologist normally will only use the weak motion velocity sensor (seismometer). However, knowing the response curve and the limitation in the data, it does not really matter which instrument is used; see Fig. 9. In this example, a seismic signal has been recorded on a 1 Hz seismometer and an accelerometer at the same site, and using the respective response function, both signals have been converted to displacement and acceleration. The only limitation is that a seismometer may get saturated by a strong motion and an accelerometer will not be sensitive enough for recording weak motions (e.g., from distant earthquakes).

The corrected signals are now very similar and of the same amplitude. This example clearly demonstrates that, with modern instruments and processing techniques, both accelerometers and seismometers can be used to get the same result.

Summary

Recorded seismic signals represent a number proportional with the ground motion. The most common seismic stations have a velocity sensor with an output proportional to ground velocity in some frequency band and/or an acceleration sensor with output proportional with ground acceleration in combination with a digital recorder. All these elements will have some frequency dependence and it has been shown:

– How to calculate the system frequency response
– How to correct for both the sensor and the recorder in order to obtain the true ground motion, whether in displacement, velocity, or acceleration
– How the noise influences the correction making it necessary to limit the correction to a particular frequency range

Cross-References

▶ Passive Seismometers
▶ Principles of Broadband Seismometry
▶ Recording Seismic Signals
▶ Seismic Accelerometers
▶ Sensors, Calibration of

References

Akkar S, Boore D (2009) On baseline corrections and uncertainty in response spectra for baseline variations commonly encountered in digital accelerograph records. Bull Seism Soc Am 99:1671–1690. doi:10.1785/0120080206

Boore D (2001) Effect of baseline corrections on displacements and response spectra for several recordings of the 1999 Chi-Chi, Taiwan, Earthquake. Bull Seism Soc Am 91:1199–1211

Havskov J, Alguacil G (2010) Instrumentation in earthquake seismology. Springer, Dordrecht, 358 pp

Havskov J, Otemöller L (2010) Routine data processing in earthquake seismology. With sample data, exercises and software. Springer, Dordrecht, 347 pp

Oppenheim AV, Schafer RW (1989) Discrete-time signal processing. Prentice-Hall, Englewood Cliffs, 870 p

Scherbaum F (2007) Of poles and zeros, fundamentals of digital seismology, revised 2nd edn. Springer, Dordrecht, 271 pp

S

Seismic Loss Assessment

Nilesh Shome
Model Development, Risk Management
Solutions, Newark, CA, USA

Synonyms

Damage; Insurance; Loss; Nonstructural component; Risk; Structural component; Vulnerability

Introduction

It is observed that earthquakes, particularly those of large magnitudes, can cause significant damage to structures over a large region surrounding the rupture zone of earthquakes. The top three costliest earthquakes in the world during 1980–2012 are the Mw 9.0 Tohoku-Oki earthquake (2011), Mw 6.8 Kobe earthquake (1995), and Mw 6.7 Northridge earthquake (1994) causing overall economic losses when occurred about $210B, $100B, and $44B respectively (III 2013). The structures those are damaged due to earthquakes are not only buildings, bridges, etc., but also utilities like water, gas and sewer lines, and lifelines like train tracks and roads. The earthquakes also cause landslides due to movement of grounds or liquefaction due to loss of strength of the soil underneath the structures. In addition, earthquakes can cause fire and tsunami leading to significant damage to structures. All the historical losses illustrate that the potential for earthquake loss can be significant in a specific region or to portfolio of buildings for insurance companies. Hence it is important for the federal and state governments or for the insurance companies to estimate accurately the seismic losses to the built environment or to the insured properties to manage earthquake risks.

The performance of buildings nowadays is expressed in terms of monetary loss due to damage, downtime, and number of injuries including fatalities due to earthquakes. See, for example, performance-based earthquake engineering (PBEE) framework developed by Pacific Earthquake Engineering Research (PEER) center (Cornell and Krawinkler 2000) or that developed by ATC-58 (2011). The loss assessment of buildings is often carried out using a vulnerability function (also known as damage function), which defines the distribution of the ratio of the building loss to the building value as a function of ground-motion (GM) intensities. In this chapter, different approaches of development of building vulnerability functions (VF) will be discussed. This will be followed by calculations of the risk of building loss due to earthquakes by integrating the seismic hazard results with the building VFs.

This provides monetary losses to buildings in terms of the average annual loss (AAL) and losses at various return periods.

In 1985, Applied Technology Council (ATC) first published a comprehensive report (ATC-13 1985) to estimate the earthquake loss of existing buildings by developing a suite of VFs based on expert opinions. ATC considered this approach since very limited earthquake damage or loss data were available at that time. The report developed earthquake loss estimates as a function of Modified Mercalli Intensity (MMI) for different facility classes (e.g., low-rise wood frame). Recently, Global Earthquake Model (GEM) project has also followed a similar approach (Jaiswal et al. 2013) for developing VFs.

Professor Karl Steinbrugge based on engineering investigations, data collection, and field studies of earthquake-damaged buildings developed the first, well-accepted VFs providing an estimate of the functional relationships between building damage and earthquake intensities (Steinbrugge 1982). These functions provided a firsthand insight to the damageability of buildings to the insurance companies in the 1990s. Since engineers were interested primarily in the life safety of buildings till the end of the twentieth century, there was limited interest among the engineers to compile the damage data of all the slightly and moderately damaged buildings for assessing performance of buildings in earthquakes. Northridge earthquake (1994) was the first earthquake that provided a large number of insurance claims data, which are used for performance assessment of single-family low-rise wood-frame buildings in high seismic zones of California, for example, Los Angeles and San Francisco (Wesson et al. 2004).

In 1999, the Federal Emergency Management Agency (FEMA) first published a comprehensive methodology to calculate *analytically* earthquake damage of model buildings for a given peak response to estimate potential losses from earthquakes (HAZUS 2003). Later on, the assembly-based vulnerability (ABV) was developed by Porter et al. (2001), and this was adopted by the Consortium of Universities for Research in Earthquake Engineering (CUREE) project (2001)

for developing vulnerability functions for wood-frame buildings. This is a significantly improved analytical approach over the HAZUS. The CUREE approach, however, did not consider more realistic representations of structures (e.g., 2D/3D nonlinear FEM model as has been considered lately in the latest PEER or ATC projects), has limited set of fragility functions compared to those considered lately in the ATC-58 project (2011), and also did not consider the correlation of response or damage between different components of buildings for accurate estimation of variability of total loss of buildings.

Recently, researchers are focused on developing analytical VFs based on detailed nonlinear analysis of buildings. For example, PEER (2011) considered 3D models of 40+ story tall buildings (Shome et al. 2014; Ramirez and Miranda 2009; Shome and Bazzurro 2009), and many others have considered 2D plane frame model for loss estimation of buildings. In addition, ATC-58 (2011) has developed a number of approaches – simple to very detailed and complex – for loss assessment of buildings. Majority of the researchers has adopted a *story-based loss assessment procedure* to take into account the variation in loss over the height, which is particularly important for mid- to high-rise buildings. All these researchers have considered a large number of ground-motion records for nonlinear response analysis of buildings, and a few of those made appropriate correction to the response results because of possible introduction of bias due to scaling of those records for nonlinear analysis (Haselton 2009). It is observed that standard approach of scaling does not take into account dependence of building response on the spectral shape and this introduces bias in the response results. In this context, the framework developed by Jayaram et al. (2012) will be discussed in detail. This approach considered story-based loss assessment, took into account the bias in the response results due to spectral shape, and considered correlation of response, damage, and repair cost of different components of buildings at different stories.

It is well known that there is significant uncertainty in the estimation of loss results and this

uncertainty can be divided into two different categories: (1) aleatory and (2) epistemic (Benjamin and Cornell 1970). Aleatory uncertainty is the inherent variability in the physical system; it is stochastic and cannot be reduced by improving the current approach of loss estimation. Epistemic uncertainty, on the other hand, is associated with lack of knowledge; it can be subjective and is reducible with additional information. There are significant epistemic uncertainties associated with the loss estimates due to uncertainty in the estimation of different parameters, and this should be considered in the decision-making process. FEMA-355F (2000) and ATC-58 (2011) have recommended typical uncertainties of a large number of parameters based on extensive research, and Jayaram et al. (2012) have developed a comprehensive approach to estimate epistemic uncertainty of VFs based on the uncertainty of those parameters.

Methodology

The seismic loss estimation of buildings is dependent on the lateral load resistance characteristics of buildings. Since building codes are being improved over the years, load resistance is dependent on the age of buildings. In addition, the resistance also depends on: (1) construction class (e.g., steel moment-resisting frame), (2) occupancy class (e.g., residential), and (3) region (e.g., high seismic regions in California). Although lateral resistance ideally should not depend on the occupancy class, it is generally observed that commercial buildings are generally well designed and constructed than residential buildings leading to better lateral resistance capacity of those buildings. The loss estimation of portfolio of buildings uses this type of generic information (e.g., HAZUS 2003) by grouping the buildings with similar load resistance capacities. The loss estimation of individual buildings (e.g., ATC-58 2011), on the other hand, uses the detailed building information to estimate accurately the lateral load resistance capacity of buildings for calculating seismic losses. The seismic loss estimation approaches that will be discussed
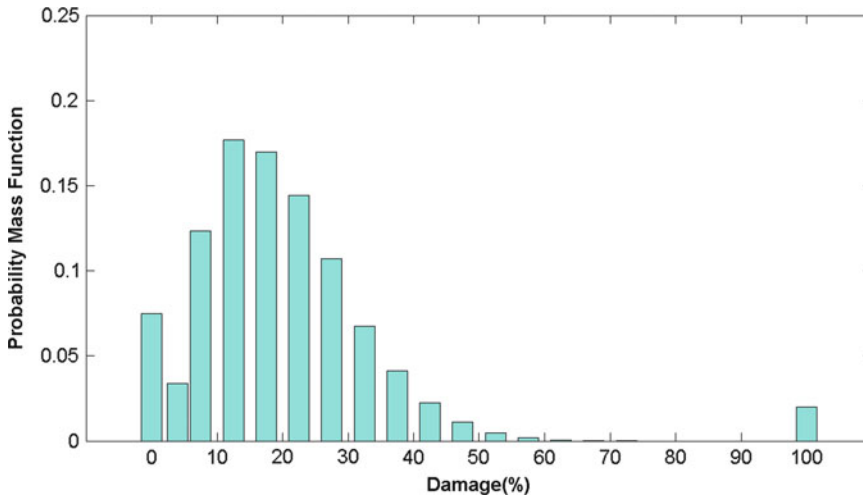
Seismic Loss Assessment, Fig. 1 Illustrative five-parameter Beta distribution

in this chapter can be applied easily for loss estimation of both portfolios and individual buildings. The VFs are developed based on mainly three approaches: (1) empirical approach which is generally expert-opinion-based development, (2) observation-based approach which is based on building loss data, and (3) analytical approach. All these approaches as well as the advantages and disadvantages will be discussed here.

Earthquakes not only cause damage to structural and nonstructural components of buildings but also damage building contents, such as furnitures, and cause closure of buildings. The closure incurs losses due to the expenses to the occupants for temporary rents or hotel stay (called additional living expenses, ALE) or loss of income to businesses (called business interruption, BI, loss). The losses to contents and due to closure of buildings will not be discussed in this chapter. This chapter will focus only on the loss due to *property damage of buildings from shaking of earthquakes*.

Loss Distribution

Since there is significant uncertainty (aleatory and epistemic) in estimating building losses at a given intensity of GM, VFs provide the loss distribution parameters as a function of intensities. It is generally assumed that the loss distribution can be represented by parametric distribution model, and Beta distribution is widely used for building loss calculations following ATC-13 recommendation (1985). The parameters of the distribution can be estimated from the mean and variance of the sample losses (Benjamin and Cornell 1970). The insurance claims data, however, show that the Beta distribution combined with Dirac delta functions at 0 % loss and 100 % loss to model the probabilities of zero (F_0) and complete (F_1) loss fits the observed loss distribution data better. In addition, insurance companies generally pay to completely replace a property (i.e., 100 % loss) beyond a particular threshold (α) of loss (e.g., 70 % loss). Hence, more accurate representation of distribution requires three additional parameters, which are F_0, F_1, and α, in addition to the conventional mean and variance. The 5-parameter distribution for losses is shown in Fig. 1 for illustration. Note that this type of distribution may be more appropriate to model the insurance losses. The methodology for estimating the five parameters of the distribution from loss data will be discussed in detail later on.

Seismic Risk

Seismic risk assessment of buildings requires estimating the probability of losses for all the possible future earthquakes over a specified period of time or estimating losses for a scenario earthquake. The scenario earthquake can be a

historical earthquake, or maximum credible earthquake (MCE), or standard design basis earthquake (DBE), or more frequent service level earthquake (SLE). The standard risk management of buildings entails estimation of average annual loss (AAL) for all the possible earthquakes to determine the insurance premium of buildings. This is the amount of modeled premium that an insurer needs to collect in order to cover the average loss over time. The other risk quantity that is commonly employed is estimation of probability of exceeding certain loss for all the possible earthquakes. When this is estimated for different amount of losses, the result is known in the insurance industry as the exceedance probability or EP curve. The scenario loss for a given earthquake, on the other hand, is useful for deterministic seismic risk management by estimating the Probable Maximum Loss (PML). PML is the largest possible loss which may occur, in regard to a particular risk, given the worst combination of circumstances. Traditionally, PML is the expected shake damage loss given the maximum size earthquake. There is, however, considerable uncertainty over the severity of loss that might arise under a broad variety of events. So an improved PML is a conservative deterministic upper bound, which is typically 90 % upper confidence bound on the portfolio loss from a risk-based MCE, for example, 2,475-year return period event on a fault. All these risk calculations require distribution of building loss as a function of intensity of GM and this is obtained by developing the VFs of buildings. The following section will discuss different approaches of development of VFs.

Empirical Approach: Expert-Opinion-Based Vulnerability Functions

There is significant advantage in developing expert-opinion-based VFs. This approach can be followed to develop VFs relatively quickly and easily without requiring any loss data or developing any analytical approach. Researchers over the years have followed primarily two approaches for developing VFs: (1) Delphi method which was followed by Applied Technology Council (ATC) in the ATC-13 (1985) and (2) Cook's approach followed recently in PAGER and GEM project (Jaiswal et al. 2013). Although the information generated from these approaches may not be very reliable, this type of approach will continue to be used in other projects in the future because of ease in developing VFs when information or resource is limited.

ATC-13 (1985) Approach

One of the first systematic attempts of developing building vulnerability was carried out by ATC-13 for assessing seismic risk in the state of California for the Seismic Safety Commission (SSC). The project was funded by the Federal Emergency Management Agency (FEMA). Since very limited earthquake loss data was available at that time and there was no acceptable analytical approach available for estimating building loss calculations, ATC derived VFs based on experience and judgment of the experts – structural engineers, builders, and the like – by estimating the expected percentage of damage of typical buildings for a specific construction type when these buildings are subjected to a Modified Mercalli Intensity (MMI). Based on personal knowledge and experience, the experts responded to a formal questionnaire with their best estimates of percent physical damage or damage ratio.

In order to develop the VFs, ATC established an advisory project engineering panel (PEP) providing necessary feedback in order to develop consensus of damage or loss estimates of different facility classes. This was modeled after the Delphi method for expert-opinion solicitations, which was originally developed for the Air Force by RAND Corporation in the 1950s. The tasks in Delphi method consist of formulating questionnaires, getting answers to those from experts, iterating the answers one or more times, and aggregating the responses by following a suitable statistical approach. The questionnaire asked each expert to provide low, best, and high estimates of the damage ratio for different construction classes or engineering facilities at different MMI. In addition, the experts were also asked to provide their level of experience with the

S
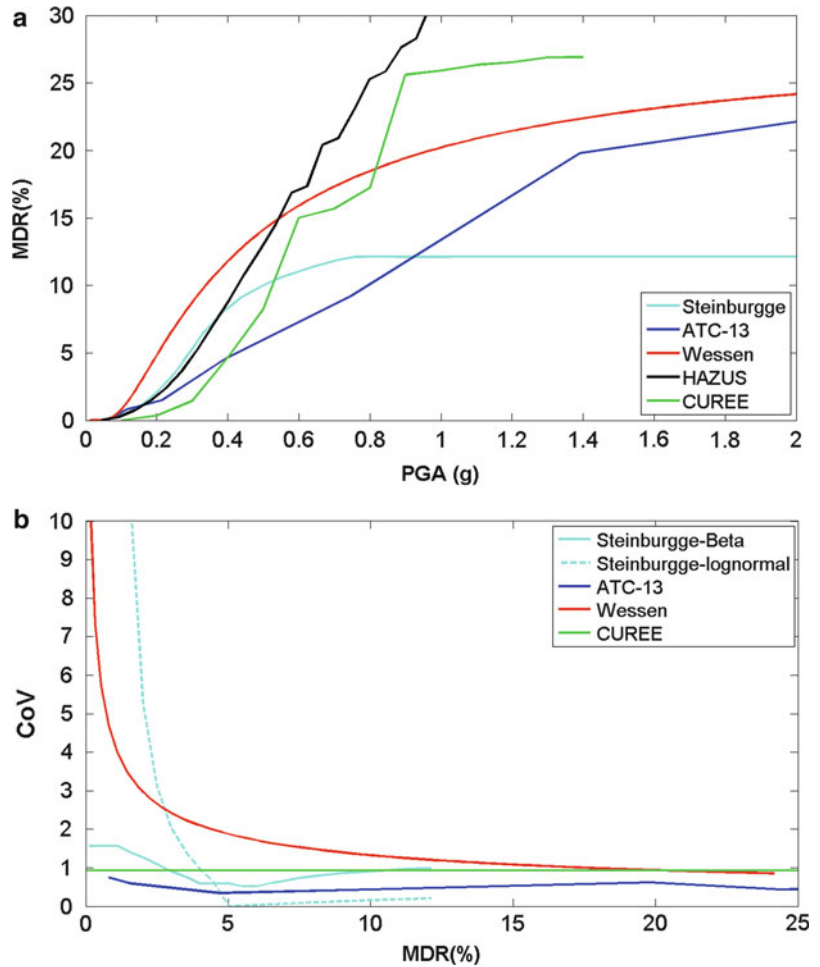
construction class and the degree of uncertainty in their estimates. Subsequently, two other sets of questionnaires were developed in order to reach a consensus on the estimation of different damage estimates. The questionnaire results were processed to develop damage probability matrices (DPM) by fitting Beta distribution to those results. In DPM, the damage to a construction class for a given intensity is described by a series of damage states (DS) and the probability that the DS will occur at that intensity. The VF of single-family wood-frame buildings in California following this approach is shown in Fig. 2 for illustration.

Because it was the first systematic attempt to develop VFs, ATC-13 quickly became the standard for assessing earthquake risk of buildings.

Catastrophe modelers and risk engineers adopted the ATC-13 damage curves until the 1994 Northridge earthquake. It will be shown later on that there is significant difference in the results based on ATC-13 and that observed in the Northridge earthquake. In addition, the ATC-13 approach is subjective. The damage functions from this approach are based exclusively on the opinion of experts. For this reason, it is difficult to calibrate or modify those in order to incorporate new data or new design (e.g., adjustment for new seismic code or other seismic regions). These functions can also be biased as observed in the Northridge earthquake. Hence, in the recent years, researchers have focused on developing VFs either based on observed building-level damage data (Wesson et al. 2004) or from

analytical approaches. This approach combines the experimental damage data and the observed damage data of different components of buildings (Jayaram et al. 2012).

GEM/PAGER Approach

Recently, US Geological Survey (USGS) has adopted Cook's model for developing VFs for the PAGER system and for the GEM project (Jaiswal et al. 2013). This approach involves identifying experts, constructing seed and target questions, eliciting expert judgments, and combining multiple judgments. In this approach, unlike the Delphi approach (adopted by ATC-13), the judgment of scoring is calibrated objectively and classical statistical test is employed for hypothesis testing of the scores. This involves assigning weight to an expert based on two separate scores: (1) *calibration score* which measures the accuracy of the scores by the experts based on background reference density functions of the scores and (2) *information score* which represents the degree with which the distribution of the expert is concentrated with respect to some background distribution. The product of these two scores is the overall score of each of the experts. Since the weights can be recomputed objectively with new observation or addition of new experts, the weights can be recomputed in the future for updating the VFs.

Observation-Based Approach: Vulnerability Functions from Building Damage Data

The most direct approach of developing the VFs would be based on building damage data observed in historical earthquakes. This approach requires damage data of large number of buildings for different construction classes, height, age, and other characteristics, which are important in defining lateral load resistance of buildings. If the damage data is based on detailed inspection of buildings following earthquakes, VFs can be developed accurately. But this requires significant amount of time and resources

and so this type of data is not available in sufficient quantity for developing the VFs. On the other hand, insurance claims data (e.g., those from Northridge earthquake) can provide a large quantity of damage data, which can be used for estimating VFs. But these data typically contain limited information about the buildings (typically provides construction class – primarily materials of construction, e.g., steel, etc., height, and age).

The earliest attempt in developing VFs following this approach was by Steinbrugge (1982). The VFs were developed based on limited damage data for a handful of earthquakes. Since insurance industry was not active in collecting damage data before the 1990s and engineers were focused on studying a few important damaged buildings following earthquakes, Steinbrugge applied his experience and judgment in order to develop the functions based on limited damage data that he collected over the years from small number of earthquakes. The developed VF for single-family wood-frame buildings is shown in Fig. 2. The results show that the vulnerability of residential wood-frame buildings is significantly higher than that predicted by ATC-13 at lower intensities. This function was used extensively by the insurance companies in the 1980s–1990s as a guideline for managing earthquake risks. Later on, this function will be compared with that based on extensive damage data from Northridge earthquake.

In this section, Northridge insurance claims data, which lacks detailed information on the lateral load resistance characteristics of buildings, will be used for developing the VF. The other type of data, which would be a preferred one, is the damage data that provides the detailed lateral load resistance characteristics of buildings. An example on this would be the data of fractured connections of 167 welded steel moment-frame (SMF) buildings collected in the SAC joint venture project. This data was collected in a very detailed and systematic way following Northridge earthquake, and those data were used for developing empirical VFs for building loss calculations in FEMA-351 Rapid Method (Bonowitz and Maison 2003). The project inspected about 18,000 connections and

found that 13 % of those had some kind of damage. Although this approach provided very accurate information about the performance of welded connections of SMF buildings, seismic loss estimation requires repair cost of those connections as well as other structural and nonstructural components of those buildings. Note that the value of the nonstructural components can be more than 70 % of total building value and so the repair cost of these components is significant, particularly at low intensities. Since damage of nonstructural components were not collected, the empirical VF developed in SAC project is primarily for the structural components, and the VF will provide a low estimate of damage at low intensities. In addition, the uncertainty in the damage or loss of buildings at a given intensity cannot be estimated from this limited data, and this information is essential for seismic risk calculations.

Since insurance claims data lacks detailed building information, the VFs developed from this data are for a broad category of buildings, like low-rise wood-frame buildings constructed in the era of moderately developed building code (e.g., 1950–1975 in high seismic regions in California). This would lead to higher uncertainty in the estimation of losses relative to the losses for a specific building. The other issue with the insurance claims data is that the data captures losses only after applying the deductibles (known as gross loss). Hence the building loss is partially known from insurance claims data (formally known as "censored" data), and this missing information is critical for accurate estimation of building vulnerabilities at low intensities.

Development of Vulnerability Functions Based on Claims Data

In this section, insurance claims data will be used for developing VFs, which estimates the probability distribution of building loss as a function of ground-motion intensity. This approach is illustrated here by using the Northridge claims data, which is even today the best source of a large number of high-quality insurance claims data for earthquakes. The data were available only by zip codes and provided the claim amount, value of

the structure, year of construction of the buildings, and the construction class. The scatter plot of insurance loss data for single-family wood-frame dwellings (SFD) constructed after 1975 is shown in Fig. 3a, and the same for steel moment-frame (SMF) buildings based on detailed inspection of beam-column connections is shown in Fig. 3b for visual comparison of the trend and uncertainty in these two data sets. The figure illustrates that the uncertainty conditioned on earthquake intensity (here PGA for SFD and $S_a(T = 2$ s) for SMF) is very high. The vertical stripe of data in the Fig. 3a represents the variation of loss ratio in a specific zip code since Northridge data is available only at the zip code level. Note that although the insurance claims data provides losses for content and additional living expenses (ALE), those are not considered in this section. The building loss data here includes damage to the structural as well as nonstructural components of buildings.

Since the Northridge losses are available by zip code, Wesson et al. (2004) estimated loss distribution by zip code by fitting a 2-parameter Gamma distribution to the claims data. Since insurance data does not capture the losses below deductible, the maximum likelihood approach was followed for estimating the parameters with the constraint that the fraction of losses below deductible must be equal to the observed fraction of no loss. The parameters of the distribution for each of those zip codes are then fitted independently against the intensity measure to get the distribution of loss as a function of intensity measure. The VFs based on this approach are shown in Fig. 2. It is observed that there is significant difference in the results between the vulnerability functions based on Northridge insurance claims data and that estimated earlier in ATC-13 following expert opinion or by Steinbrugge based on limited damage data.

A more direct approach for the estimation of the distribution parameter would be to estimate those directly from the data by grouping those as a function of intensity. In the earlier approach, since each of the parameters of the distribution was fitted independently, the correlation between the parameters was not preserved. In addition, the

Seismic Loss Assessment, Fig. 3 Distribution of damage ratio (ratio of claims amount or repair cost and building replacement cost) distribution as a function earthquake intensity as observed in Northridge earthquake. (**a**) Post-75 single-family wood-frame buildings. (**b**) Steel-frame buildings inspected in SAC project
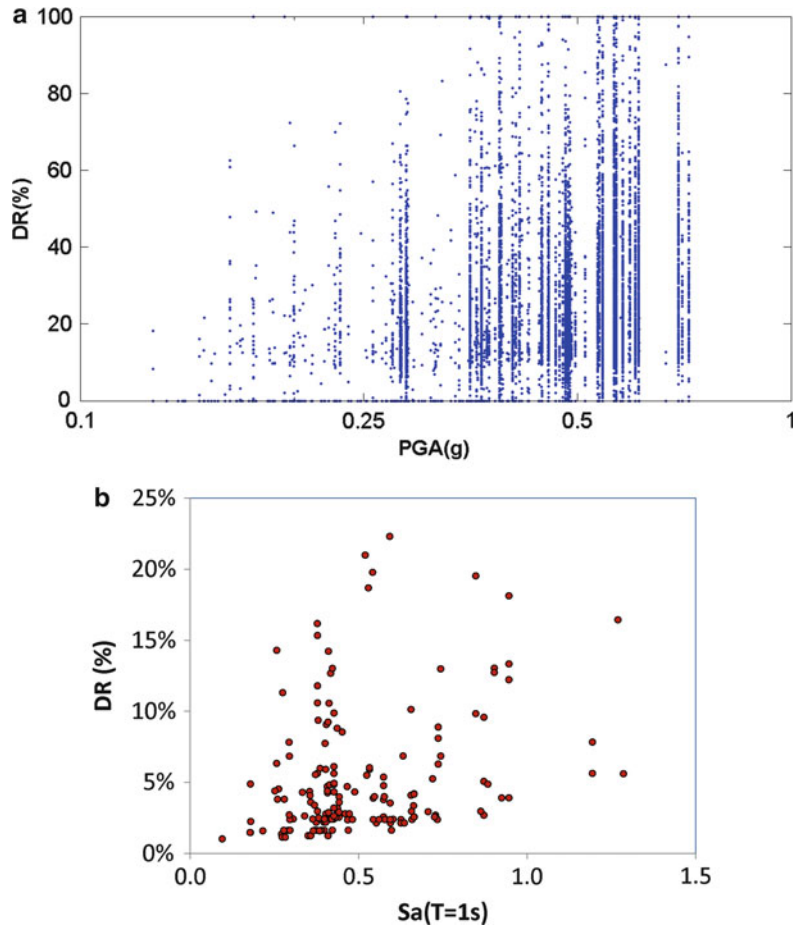
regression analysis gave equal weight to all the parameters for each of the zip codes irrespective of the number of data points in the zip codes. Since Beta distribution is commonly employed for describing the earthquake loss distribution and a five-parameter Beta distribution is deemed to be the most suitable for modeling seismic loss distribution, the fitting of five-parameter distribution model to the loss data will be discussed here. The five parameters of the distribution are: mean, variance, probability of no loss (F_0), probability of complete loss (F_1), and the threshold (α) beyond which losses are considered to be 100 % by insurance companies. The parameters of the probability distribution can be estimated by using two methodologies, namely, the method of moments and the maximum likelihood approach. Since the insurance loss data is available after applying the insurance limits and deductibles,

the method of moments is not suitable for this type of data. Hence *maximum likelihood approach* will be followed to estimate the parameters (see Shome et al. 2012 for details).

In general, the building VFs based on observed damage/loss data from one region cannot be directly extrapolated to other regions since the building design and construction practice varies among different regions. Since the frequency of large-magnitude earthquakes is small and the occurrence of those events near large population center is even rarer, it is not possible to get a large number of damage data for different construction classes in order to develop VFs. Also this approach needs high-quality data, i.e., building-level loss results like those obtained in 1994 Northridge earthquake. The number of data should be high for estimating VFs with high confidence.

Maximum Likelihood Estimation (MLE) of Parameters

The objective here is to estimate the parameters of the five-parameter Beta distribution for the actual loss paid by the building owners for repairing buildings (i.e., distribution for ground-up (GU) loss). The GU loss is computed by adding the gross loss, the deductible, and the losses beyond the limits if applied. But this cannot be done for the cases when the gross loss equals zero. This is a typical case of left-censored data set where the ground-up losses are accurate above deductible, but the data below deductible is not available and is reported as zero. If F_0 is simply estimated as the fraction of data points with zero ground-up loss, F_0 will be severely overestimated when the deductibles are large particularly at low intensities. Note that the data is right censored as well because losses beyond the insurance policy limits are truncated to the limit. This is, however, much rarer since it only occurs at very large damage ratios induced by very high-intensity GM.

The density function for damage ratio $f(D_i)$ at a given intensity equals the following:

$$f(D_i) = \begin{cases} F_0 & \text{for } Di = 0 \\ F_0 + (1 - F_0 - F_1)B_{PDF}(D_i, a, b) & \text{for } 0 < D_i < \alpha \\ [F_1 + (1 - F_0 - F_1)(1 - B_{CDF}(\alpha, a, b))] & \text{for } D_i \geq \alpha \end{cases} \quad (1)$$

where:

α = The threshold beyond which losses are treated as 100 %. Note that this is governed by the underwriting policy of an insurance company.

$B_{CDF}(D_i, a, b)$ = Beta cumulative distribution function (CDF) corresponding to sample damage ratio D_i (=the ratio of the loss to the building's value).

$B_{PDF}(\cdot)$ = Beta probability density function (PDF).

a and b = Parameters for the Beta distribution of the sample damage ratios for a GM intensity bin.

Hence the likelihood of observing damage ratios D_is equals:

$$L = \left[\prod_{D_i=0} F_0 + (1 - F_0 - F_1)B_{CDF}(ded_i, a, b) \right]$$
$$\left[\prod_{D_i \geq \alpha} F_1 + (1 - F_0 - F_1)(1 - B_{CDF}(\alpha, a, b)) \right]$$
$$\left[\prod_{D_i \epsilon (0, \alpha)} F_0 + (1 - F_0 - F_1)B_{PDF}(D_i, a, b) \right]$$
$$(2)$$

where the parameters, F_0 and F_1, correspond to a range of GM intensities for which loss data is accumulated for fitting and these parameters are estimated by maximizing L. The first term in the likelihood estimate corresponds to data points where the observed damage ratio equals zero. This includes likelihood of damage ratio less than deductible ded_i, but reported as zero in the claims dataset, and the true zero damage. The second term corresponds to data points where the observed damage ratio equals 1. This corresponds to the likelihood of observing complete loss and the losses above the threshold, α. The third term corresponds to the rest of the data points where the damage ratio is more than deductible, but less than α. The parameters of the distribution F_0, F_1, a, and b are obtained by maximizing the likelihood of observing all the damages, D_i, which is shown in Eq. 2 as the product of likelihood of the damages.

The results of fit of the five-parameter distribution to the Northridge loss data at high-intensity GM are shown in Fig. 4. The claims data used for fitting the distribution is from a number of insurance companies and the information of the threshold (α) is not available from those companies. Hence for simplicity, it is assumed that α equals to 100 % damage.

Seismic Loss Assessment, Fig. 4 Comparison of distribution of loss at high-intensity ground motion between the data and MLE fit
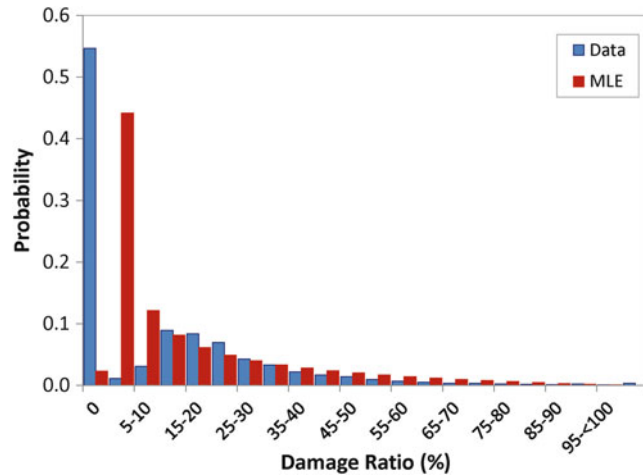
The results show that the fitting captures well the objective of probability mass at zero and 100 % damage. Note that it would be difficult to estimate reliably all the parameters of the distribution when data is limited. In those cases, it is advisable to bring in other sources of data to define the loss distribution. For example, one can use building collapse results – analytical and/or experimental – to define F_I of a specific building or a class of buildings.

Analytical Approach of Development of Vulnerability Functions

There is limited building damage data from earthquakes and that too from small number of construction classes. In addition, building damage characteristics based on the damage data from one region may not be extrapolated to other regions. Hence, VFs for majority of construction classes in most of the regions in the world can be developed only from an analytical approach. Hence HAZUS (2003) first developed a comprehensive analytical methodology to estimate the building seismic losses. HAZUS adopted capacity spectrum method to estimate response of *generic buildings* and estimated only *mean* loss based on damage of different building components. Later on, CUREE-Caltech (2001) developed the assembly-based vulnerability (ABV) methodology for developing *building-specific* VFs. In this

approach, the damage is calculated for the individual building assemblies, e.g., partition wall, based on the response results from *nonlinear time-history analysis* of the model buildings to estimate *mean and variance* of damage of the model buildings as a function of GM intensity. The individual buildings in this approach were represented by deterministic simplified structural models, where building components were idealized as springs and masses. Recently, Jayaram et al. (2012) have developed a systematic simulation approach based on the PEER-PBEE loss assessment framework to develop building-specific VFs from component-level damage at each story of buildings. In this approach, the total building loss is estimated by using the response from detailed nonlinear time-history analysis of buildings. The approach took into account the dependence of response on S_a at multiple periods, the correlation of losses between different components to improve the estimation of uncertainty of losses, and on captured the effects of epistemic and aleatory uncertainties of different parameters considered in loss calculations, such as GMs, structural response parameters, loss costs, etc. All these steps help to get the unbiased loss due to earthquakes and quantify accurately the uncertainty in the loss estimates.

The analytical loss calculation in general consists of (1) repair cost of different building components associated with a damage state or cost

due to demolition, (2) damage state given a demand parameter (EDP), and (3) EDP given an intensity of GM. A brief description of these parameters is given below and the details can be found in HAZUS (2003).

Building Components

In order to carry out the seismic loss calculations, building components are divided broadly into two categories: structural and nonstructural. The loss estimation approach either estimates damage of the individual components (e.g., CUREE-Caltech 2001) or damage of a group of components or subsystems (e.g., HAZUS 2003). The structural component (SD) resists gravity, earthquake, wind, and other types of loads. This includes components like columns, beams, load-bearing walls, etc. For buildings planned by design professionals, the structural components are typically designed and tailored to building-specific configurations and loading conditions. The nonstructural components include a large variety of architectural, mechanical, and electrical components, which can be further divided into as either "drift-sensitive" (NSD) (e.g., non-load-bearing partition walls, exterior curtain walls, etc.) or "acceleration-sensitive" (NSA) components (e.g., suspended ceilings, HVAC systems, etc.). These components are usually neither designed by the professionals nor tailored to specific buildings even when designed. These are tested under some generic load conditions to satisfy the acceptability criteria of the building codes under some idealized conditions. Hence the uncertainty in the performance of structural components is relatively low when designed and constructed by professionals compared to those of the nonstructural components. Drift is used to predict the damage states of the SD and the NSD subsystems or components. Acceleration, on the other hand, is used to predict the damage states of the NSA components.

Damage States

The damage state of each subsystem or component represents a consequence in terms of repair for the type and the severity associated with that state. This is predicted based on the fragility functions of the subsystem or component. The damage states are defined in HAZUS (2003) and these are none, minor (DS1), moderate (DS2), extensive (DS3), and complete damage (DS4). For example, moderate structural damage of concrete shear-wall buildings is defined by the diagonal cracks on most of the shear-wall surfaces and large diagonal cracks with spalling of concrete at the wall ends of some shear walls. The moderate damage of the NSD components or subsystems represents large and extensive cracks of the partition walls requiring repair and repainting as well as replacement of a few walls. The moderate damage state of the NSA subsystem or component represents extensive falling of the suspended-ceiling tiles with disconnected and/or buckled ceiling support framing (T-bars) at a few locations.

Repair Costs

The loss from earthquakes is associated with the repair due to damage of different components or subsystems as well as building replacement cost associated with demolition and collapse. The uncertainty in repair cost arises primarily due to lack of information about the component. The cost of repair works is generally significantly higher than the cost of new construction, and this is more significant for the structural components than the nonstructural components. For example, the repair of beam-column connections of the structural component requires removing the suspended ceilings, inspection of the joints which may require non-destructive testing, closure of part of the floor of the building, etc. All these expenses have to be incurred on top of the repair cost of the connections. Insurance claims data also suggest relatively higher uncertainty in repair cost at low intensities. In order to estimate the repair cost, it is generally assumed that the insurer will pay the cost of repair or replacement of the damaged buildings with the materials of like kind and quality without any deduction for depreciation, which is called *replacement cost method* for valuation.

HAZUS-Based Vulnerability

The HAZUS methodology is developed to estimate losses for different federal, state, and local agencies for planning earthquake risk mitigation, emergency preparedness, and many other uses. Although the methodology can be used to estimate losses for individual buildings, it is developed to estimate losses of portfolio of buildings distributed over a large region. The methodology provides an estimate of the probability of mean loss for none, slight, moderate, extensive, and complete damage of building components at a given level of ground shaking. The building damage characteristic is defined by the following curves: (1) fragility curves that describe the probability of reaching or exceeding different states of damage given the peak-building response and (2) building capacity (pushover) curves that are used to determine the peak-building response for the demand spectrum.

Fragility Functions

The fragility curve is defined by a median value of the engineering demand parameter (EDP), which is, for example, spectral displacement in HAZUS and maximum interstory drift ratio (IDR) in Jayaram et al. 2012, and by the variability associated with the estimation of that damage state for that EDP. The median EDP corresponds to the threshold of the damage state. The curve is assumed to follow the cumulative lognormal distribution function as shown below:

$$P(DS \geq ds_i | EDP) = \Phi\left(\frac{\ln(EDP/\mu)}{\beta}\right) \quad (3)$$

where μ denotes the median value of EDP at which building component or subsystem reaches the threshold of the damage state ds_i, β denotes the corresponding dispersion, and $\Phi(\cdot)$ is the cumulative density function for standard normal distribution. The dispersion considers (1) uncertainty in the damage state threshold, (2) variability in the capacity (response) of buildings, and (3) uncertainty in building response or EDP due to the spatial variability of ground motion.
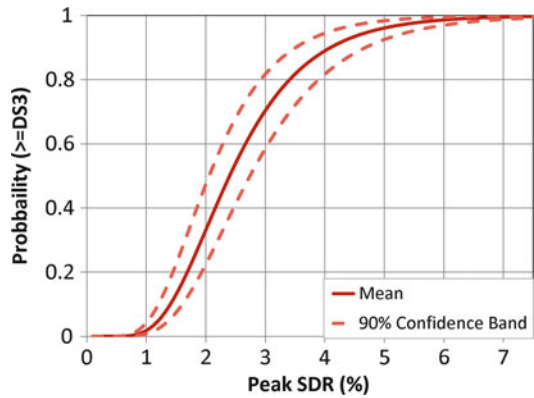
Seismic Loss Assessment, Fig. 5 Mean fragility functions and its uncertainty for extensive damage of nonstructural drift-sensitive subsystem (NSD)

The component fragility functions are estimated from the results of laboratory experiments or from observed damage data. Since there is limited observed damage data or experimental data, past studies recommended fragility functions primarily based on engineering judgment (e.g., HAZUS). Recently, ATC-58 (2011) has made significant effort to develop a suite of fragility functions so that loss calculations can be carried out in day-to-day building design work. Also Network for Earthquake Engineering Simulation (NEES) Program is actively pursuing experimental research (http://www.nees-nonstructural.org/) to develop fragility functions for different nonstructural components of buildings. All these research provide a range of fragility functions giving an estimate of the mean and the standard deviation of the fragility functions. The mean fragility function for extensive damage of the generic NSD subsystem, which is an assembly of a number of components such as partitions, exterior glazing systems, etc., is shown in Fig. 5. The epistemic uncertainty in the estimation of the fragility function for that damage state is shown by the 90 % confidence band.

Building Response Calculations

The building responses in HAZUS are estimated based on the *capacity spectrum method* (CSM). The building capacity curve (also known as a

Seismic Loss Assessment,
Fig. 6 Example building capacity curve, demand spectrum, and building response for wood-frame buildings in California for input spectrum from M7 earthquake on Hayward fault at a site close to the fault
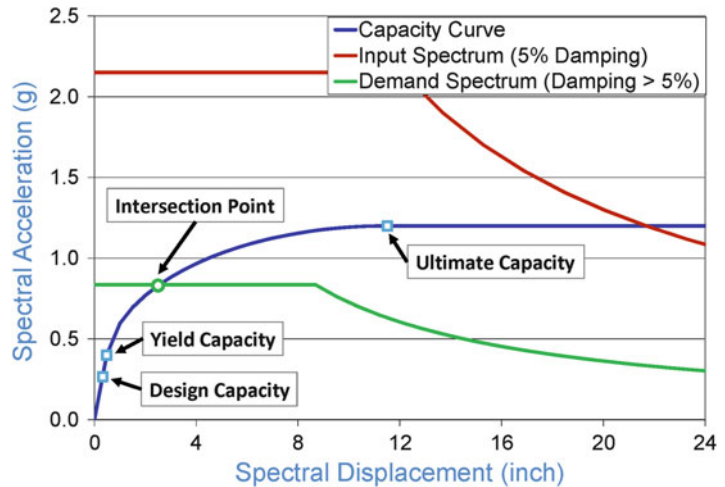
pushover curve) describes the lateral load resistance of a building as a function of its characteristic lateral displacement. The curve is derived from the results of static-equivalent base shear and roof displacement and is defined by three control points: (1) design capacity, (2) yield capacity, and (3) ultimate capacity as shown in Fig. 6. It is assumed that the building capacity will have a range of possible properties and those are lognormally distributed as a function of the ultimate capacity. The peak-building response is the intersection of the building capacity curve and the response spectrum, which defines the seismic demand. Since buildings go beyond its elastic limit, the demand spectrum defined by the elastic damping is reduced due to the higher effective damping. The effective damping includes both the elastic damping and the hysteretic damping associated with the post-yield cyclic response of buildings. Figure 6 illustrates the capacity curve, demand spectrum curve, and the peak-building response at the point of intersection of those curves. The main drawback of CSM is that the MDOF structures are idealized by an equivalent SDOF system using an increased period and damping in order to estimate the seismic demand of the MDOF structures by the capacity spectrum approach. In addition, this method is a nonlinear static procedure, whereas the nonlinear time-history analysis generally provides a more accurate estimation of building

response particularly at higher intensities. Hence this approach may not be suitable for mid-rise and high-rise buildings and for types of buildings at high intensities.

Building Loss Calculations
The fragility curves define all the four damage states for each of the three building subsystems based on the EDPs estimated from the intersection of the capacity curve and the seismic demand curve. It is assumed that the nonstructural damage states are independent of structural damage states. The repair cost is estimated based on the damage of different subsystems of buildings. Note that although the results from HAZUS loss estimation methodology are intended to assess the socioeconomic impact of earthquakes in a region, the result has a number of limitations. Earthquake impact on a society is a very complex process and the economic consequences cannot be considered in a simplistic approach. The dollar losses are calculated for all the three coverages: (1) building repair and replacement costs, (2) contents losses, and (3) additional living expenses (ALE) for residential occupancies or business interruption (BI) losses for commercial or industrial occupancies. In this chapter, only building repair loss calculations are discussed for brevity. This loss considers repair and replacement due to both structural and nonstructural damages, and this is calculated as the product of the probability

of a component in a given damage state and the mean repair costs for that damage state, summed over all the building subsystems and the damage states as given below:

$$L_i = RC_i \cdot \sum_{k=1}^{3} \sum_{j=2}^{5} P_{ijk} \cdot DR_{ijk} \qquad (4)$$

where RC_i is the replacement cost of an entire model building i (e.g., wood frame), P_{ijk} is the probability of model building i and subsystem k (e.g., NSD) being in damage state j (e.g., moderate), and DR_{ijk} is the mean damage ratio for building i and subsystem k in damage state j. The ratio of loss of building L_i and the replacement cost of RC_i is the mean damage ratio (MDR) of the building i for the intensity of GM at which the losses are calculated. This calculation is repeated at different intensities to develop the VF(s) for seismic loss calculations of an individual building or a portfolio of buildings. Note that although HAZUS considers only mean repair cost, simulation-based approach, which will be discussed later on, considers the uncertainty in repair cost as well to calculate the uncertainty in the VFs. Figure 2 shows the VF following HAZUS approach of residential light wood-frame buildings constructed during the period of 1940–1975 following a modern building code (e.g., Uniform Building Code) in the high seismic zone of California for a suite of response spectrum for different magnitude of earthquakes at different distances based on the disaggregation of USGS hazard results in Los Angeles.

Assembly-Based Vulnerability (ABV)/CUREE-Caltech Approach

In this approach (CUREE 2001), the nonlinear building response was calculated based on nonlinear time-history analysis of model buildings representing the best estimate of mass, stiffness (load-deformation characteristics),damping, and building properties required for the analysis. The buildings are modeled as a collection of assemblies of components, such as a gypsum wallboard partition, a floor diaphragm, beams, etc., and each of the assembly is represented by a fragility function to estimate the probability of that assembly in different damage states as a function of structural response. Test data from the CUREE Project, experimental data from other projects, and observation of damage from different earthquakes are used to develop the fragility functions. The losses were estimated for repair of damage of different assemblies following standard construction cost-estimation principles. This approach estimates mean and uncertainties (which is also improvement over HAZUS) in loss estimate by capturing the uncertainties in: ground motion, structural characteristics, component damageability, and repair costs. This approach followed a simulation approach in order to quantify the (aleatory) uncertainty in the loss results. The losses are calculated for individual components of the buildings instead of a group of components like in HAZUS to get the building loss.

The structural response which is calculated from the building simulation results is used as an input to the fragility function to simulate the damage state of the component assemblies. The components are idealized as an assembly of springs and masses to calculate the response. The masses represent the building components that have significant weight, such as the walls, ceiling, floor, etc., and the springs represent building elements, such as the walls, beams, etc., that resist deformation. The idealization took into account for strength and stiffness degradation of the components, along with many other complex nonlinearities, to represent accurately the spring element. The response parameters used in this approach are maximum interstory drift ratio (IDR), peak floor acceleration (PFA), peak transient horizontal shear strain, peak member forces, and residual drift ratio.

The total repair cost is calculated by summing the cost of repairing the damaged assemblies. The cost of repairing a damaged assembly includes labor costs, material requirements, debris removal, and equipment rental. The cost calculation includes the cost of repair of the undamaged assemblies (or line of sight) due to damage of an assembly. The cost also considers the case of collapse of the building. The cost is simulated from the distribution of the cost for repair of each damaged assembly. The simulation

provides total repair or replacement cost at a given level of seismic shaking intensity. The process is repeated many times at a given intensity to estimate the probability distribution of cost for calculating the building-specific VF. The plot of MDR and coefficient of variation (CoV) of typical wood-frame buildings of average size as a function of PGA based on CUREE-Caltech project is shown in Fig. 2. The plot shows that the prediction of MDR and uncertainty are significantly less than those observed in Northridge earthquake.

PEER-PBEE-Based Approach

Recently, Jayaram et al. (2012) have developed a generic procedure for developing the building VFs by following a systematic simulation approach based on the PEER performance-based earthquake engineering (PBEE) framework. The steps involved in the procedure are quantifying ground-motion hazard using a *vector of spectral accelerations*; predicting building response parameters such as story drifts, floor accelerations, and residual drifts under the quantified hazard; estimating structural collapse and demolition; and calculating *story-wise losses* based on the responses. The total building loss is calculated by summing the story losses. The procedure captures the effects of epistemic and aleatory uncertainties in different parameters considered in loss calculations, such as ground motions, structural response parameters, loss costs, etc., in order to quantify the uncertainty in the final loss estimate.

The mathematical framework of the development of the mean vulnerability function based on this approach is given below:

$$E(L|S_a(T_0)) = \iiiint L.dG(L|DM)||dG(DM|EDP)||dG(EDP|S_a(\boldsymbol{T})||dG(\boldsymbol{S_a}(\boldsymbol{T}))|S_a(T_0)| \quad (5)$$

where L is the building loss; $S_a(T_0)$ is the spectral acceleration at a reference period; $E(L|S_a(T_0))$ is the mean loss at $S_a(T_0)$; $dG(L|DM)$ denotes the derivative of the probability of exceedance of the building loss given a damage measure (DM); $dG(DM|EDP)$ is the derivative of the probability of exceedance of the DM given an EDP (e.g., story drift ratio); $dG(EDP|S_a(\boldsymbol{T}))$ is the derivative of the probability of exceedance of the EDP given a vector of spectral acceleration, $S_a(\boldsymbol{T})$; and $dG(S_a(\boldsymbol{T})|S_a(T_0))$ is the derivative of the probability of exceedance of spectral accelerations at multiple periods, $S_a(\boldsymbol{T})$, given $S_a(T_0)$. Note that the reference period, denoted here T_0, is not necessarily the fundamental period of the analytical

model of a structure. The fundamental period is a good choice when the structural performance is reasonably elastic and dominated by the first mode, but it is not a good choice for tall buildings in high seismic regions. The response of these buildings is nonlinear when damaged during earthquakes, and the response is determined by multiple frequencies. The equation shows that the structural response is predicted based on a vector of spectral accelerations, but the vulnerability is represented as a function of scalar intensity measure for simplicity of use, but without sacrificing any accuracy.

The variance of the vulnerability function can be estimated as follows:

$$Var(L|S_a(T_0)) = \iiiint (L - E(L|S_a(T_0))^2|dG(L|DM)||dG(DM|EDP)||dG(EDP|S_a(\boldsymbol{T}))||dG(S_a(\boldsymbol{T}))|S_a(T_0)|$$

$$(6)$$

The above integrals Eqs. 5 and 6 can be evaluated conveniently by using Monte Carlo Simulation (MCS). When the number of variables is large, particularly when the intensity measure is of

Fig. 7 Response spectrum conditioned on $S_a(T_0) = 0.22$ g at $T_0 = 2.5$ s. (**a**) Exponential of the conditional mean of $\ln S_a(\mathbf{T})$. (**b**) Conditional standard deviation of $\ln S_a(\mathbf{T})$
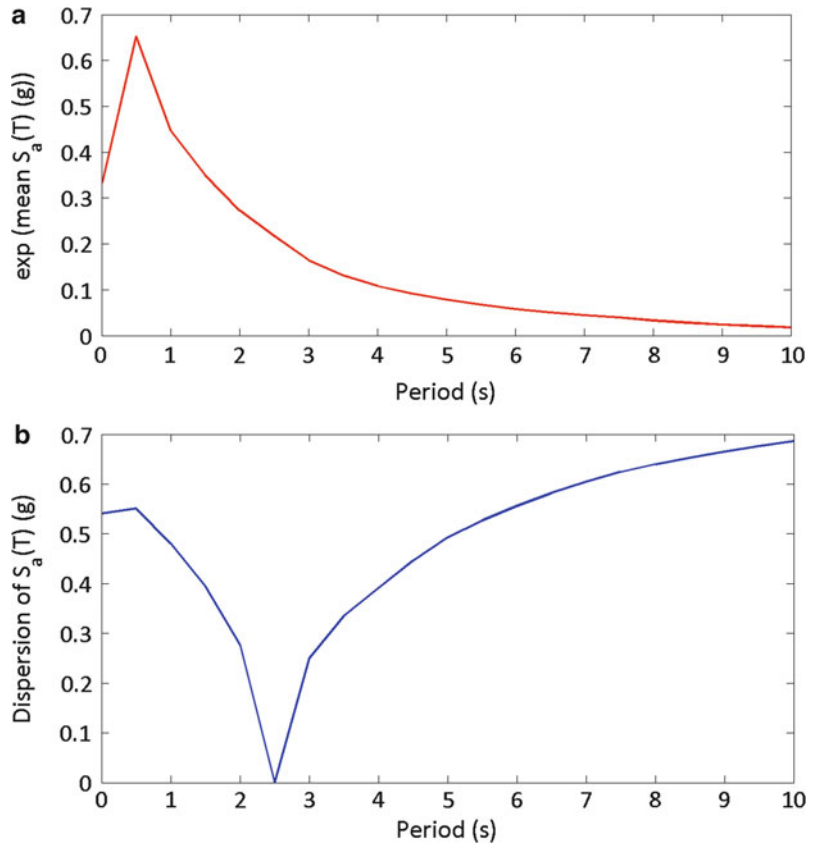
large dimension (i.e., vector of S_a), it is efficient and convenient to use MCS. The MCS approach involves simulating all the random variables in Eqs. 5 and 6 (which are $S_a(\mathbf{T})$, *EDP*, *DM*, and *L*) and then computing the mean and the variance of *L* for a wide range of $S_a(T_0)$ values. The MCS approach provides flexibility and allows consideration of accurate but complex models for different parameters as well as correlation between these parameters. It also allows for the rigorous consideration of collapse and demolition.

Conditional Spectral Accelerations ($S_a(T)|S_a(T_0)$)
Although it is quite common to scale a suite of ground-motion records to $S_a(T_0)$ for nonlinear analysis of buildings for estimating EDPs, studies have shown that this introduces bias in the results – particularly for mid-rise to high-rise buildings (e.g., Haselton 2009). This is due to the fact that the response of these buildings depends on S_a at periods other than T_0. Hence it

is important to select the records in such a way that the distribution of $S_a(\mathbf{T})$ at all the periods conditioned on $S_a(T_0)$ of the scaled records is appropriate. The mean and variance of a typical conditional spectrum in Los Angeles are shown in Fig. 7. It is observed that variance is zero at the conditioning period T_0, but the variance increases rapidly as the period of S_a moves further from T_0. In general, the risk of exceeding $S_a(T_0)$ at a site is caused by earthquakes of several possible combinations of M-R. Since the shape of response spectrum depends on the M-R of earthquakes, there are different possible shapes of response spectrum conditioned on $S_a(T_0)$ due to the different combinations of M-R of earthquakes that can strike a site. The distribution of *M-R* conditioned on $S_a(T_0)$ at a building site can be obtained by deaggregation of hazard for the site, and this different shaped response spectra should be considered in the simulation of building response to improve accuracy. In the simulation of loss

Fig. 8 Response spectra
for an event ($M_W = 7$ and
$R = 10$ km) scaled to the
median $S_a(T_0) = 0.12$ g of a
set of 98 records and the
median spectrum for the
set. The median is
compared with the average
of the four NGA
attenuation models
(Bozorgnia et al. 2014).
The median spectra are
closely matched by
carefully selecting the set of
records
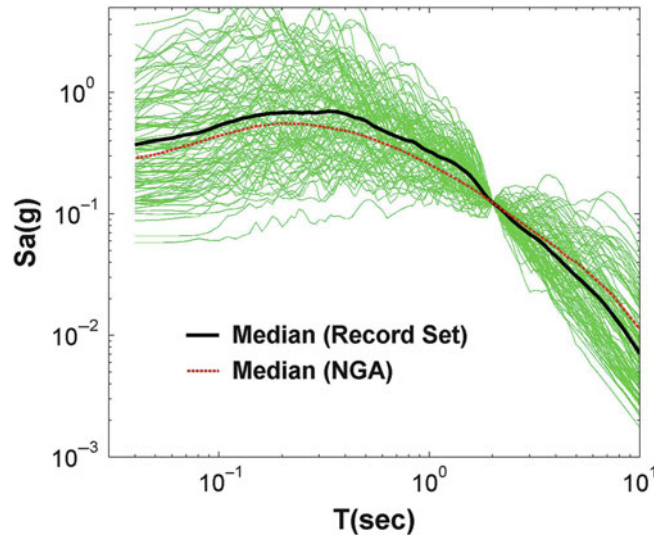
results, first the M and R pairs are sampled based on the distribution from the deaggregation of hazard results, and subsequently, spectral acceleration at multiple periods, $S_a(T)$, conditioned on $S_a(T_0)$, M, and R, is simulated. An example of simulated response spectrum is shown in Fig. 8.

Calculation of EDP

The EDPs required in this approach for loss calculations are interstory drift ratio (SDR), peak floor accelerations (PFA), and residual drift ratio (ResDR). The SDR values are used to predict the damage of SD and NSD components at each story level, the PFA values are used to predict those of the NSA components, and ResDR is used to determine if a building would have to be demolished after an earthquake. The EDPs are first calculated based on nonlinear time-history analysis of buildings for a suite of ground-motion records. But it is nearly impossible to select records that satisfy conditional distribution of S_a at multiple periods (which is can be more than ten for tall buildings) (see, e.g., Jayaram et al. 2012). Hence it is efficient and also accurate to fit regression models to predict independently the different EDPs at different stories and use these fitted models subsequently to estimate the EDPs for the simulated response spectra, which have the right conditional distribution of $\mathbf{S_a(T)}$.

Hence the purpose of computing these EDPs by nonlinear time-history analysis for different recorded ground motions is to develop regression models to predict EDPs as functions of $\mathbf{S_a(T)}$ while carrying out simulations.

During earthquakes, there is a possibility of collapse of structures. Two different types of structural collapses can be observed during the analysis of buildings. The first type of collapse, *simulated collapse*, is the collapses observed in the analyses. This occurs due to excessive lateral drifts and deterioration that lead to loss of the gravity load resistance of structures. This is modeled using a collapse probability curve, which provides the probability of collapse of buildings as a function of $S_a(T_0)$. The other type is referred to as *virtual collapse*, which captures those cases when a structure in nonlinear analyses undergoes a very large drift without collapse due to limitations in numerical modeling. Both these collapses are considered here for developing vulnerability functions. The collapse probability curve is developed using logistic regression, which is commonly used for regressing binary data (Shome and Cornell 2000), and can be expressed as follows:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \cdot S_a(T_0) + \varepsilon \quad (7)$$

Fig. 9 Collapse fragility curve for a modern 20-story steel moment-frame building in Los Angeles
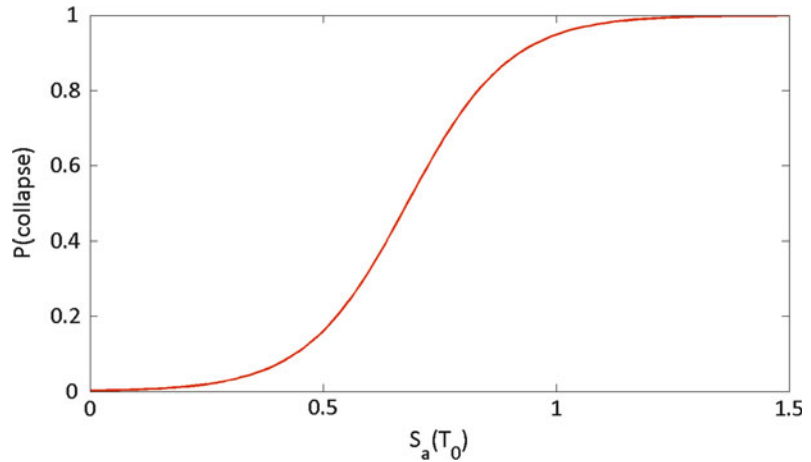
where p = probability of collapse, β_0 and β_1 = regression coefficients, and ε = regression error. The collapse fragility curve for a 20-story modern steel-frame building is shown in Fig. 9.

In addition to fitting the EDPs at each story for estimating the mean and standard deviation of the loss results conditioned on $S_a(T)$, the correlation between different EDPs at the same story and between EDPs at different stories needs to be estimated to jointly simulate the EDPs. The extent of these correlations is computed using the regression residuals. It is assumed that the EDPs follow multivariate lognormal distribution.

Damage Measure and Repair Cost

In this approach, building component losses are simulated based on the repair costs associated with different damage states, which are obtained from the component fragility functions. Since the damage states are defined to be discrete, the simulation of correlated discrete damage states is a difficult task. It is shown by Shome et al. (2014) that the correlation between discrete damage states can be easily established by representing the fragility functions using a "damage-capacity" formulation. In this approach, instead of defining the correlation between the discrete damage states, the correlation between the damage capacity of different components is defined. The damage capacities of all the components can be jointly simulated from a multivariate lognormal

distribution (the multivariate nature is assumed). The distribution is defined by the means and dispersions of the capacity based on the fragility functions and the correlation between those capacities.

The repair costs conditioned on DM are also modeled as a multivariate lognormal random variable with the median and dispersion and correlations of repair costs between different damage states over the height of buildings. The losses of different components across different damage states are expected to be correlated because of common factors involved in repairs, such as the contractor hired for the repairs and the material costs. The parameters for the distribution of repair cost of buildings are based on the study by Ramirez and Miranda (2009) on the cost of contractors for a large number of building projects. Shome et al. (2014) have shown how the information of repair cost of the contractors can be simulated to estimate the lognormal parameters of repair cost for a specific building or a group of buildings.

The total loss is computed by summing up the sample components losses at each story. This is normalized by the building value to get the damage ratio for the building. The scatter plot of simulation of a 20-story modern steel building in Los Angeles, California, is shown in Fig. 10 as a function of $S_a(T_0)$. The mean and the CoV of the damage ratios are then computed by binning the data points into $S_a(T_0)$ bins to obtain the building VF.

Fig. 10 Damage ratios of a modern 20-story SMF in Los Angeles for all the simulated $S_a(T)$ vectors corresponding to a wide range of $S_a(T_0)$ values
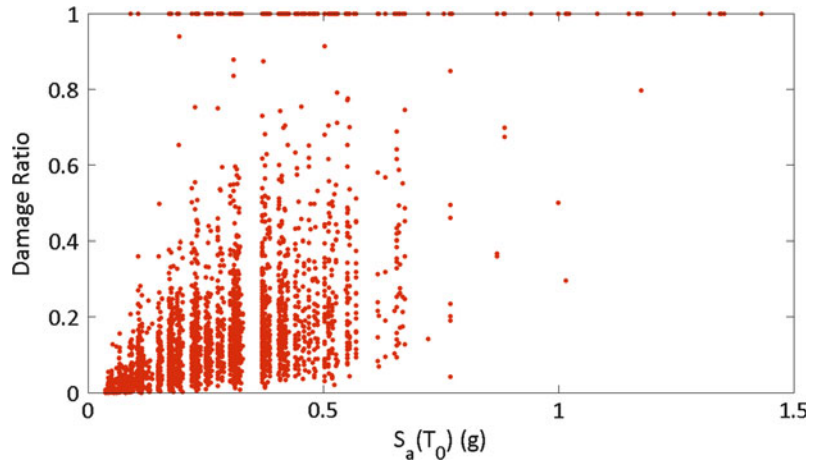
Seismic Loss Assessment,
Fig. 11 Impact of different sources of epistemic uncertainties in the loss results of a modern 20-story steel moment-frame building in Los Angeles. (**a**) Average annual loss (*AAL*). (**b**) 500-year loss
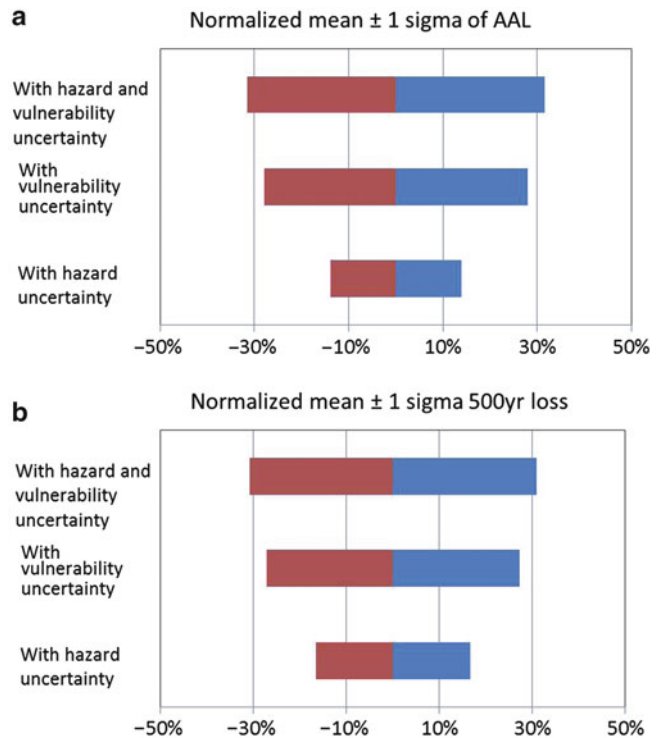
Seismic Loss Results and Uncertainty in Those Results

The VFs developed in the previous sections are used to carry out seismic risk assessments of buildings. The VFs are integrated with the hazard curves to compute the AAL and the return period or EP losses. As discussed, the methodology developed by Jayaram et al. (2012) can be used to account for epistemic uncertainties in the vulnerability functions by developing a suite of functions representing the uncertainty. The losses are then sampled from these functions to compute the epistemic uncertainty of different loss matrices used for building risk assessment like AAL, EP loss, etc. Figure 11 illustrates the normalized

64 % confidence interval (CI) of the AAL and the 500-year return period loss of a modern 20-story steel moment-frame building in Los Angeles. The relative extents of contributions of the epistemic uncertainties in vulnerability and hazard to the total epistemic uncertainty would, however, depend on whether the loss assessment is carried out for a single building or a portfolio of buildings. For a single building, as seen in the figure, the VF uncertainty has a large contribution to the total uncertainty. For a portfolio of buildings, however, the uncertainty in losses across different locations will be mostly independent and will average out. The hazard uncertainty, however, is highly correlated across multiple locations and will not average out. Therefore, for a portfolio, hazard uncertainty will have higher contribution to the total uncertainty.

Summary

Seismic loss assessment of buildings plays an important role in managing earthquake risk for public agencies, insurance companies, as well as individual stakeholders. The loss assessment procedure has evolved over the years – from expert opinion to very rigorous analytical approach. Since the analytical procedure is very complex requiring accurate information of a large number of parameters, there is significant uncertainty in accurate estimation of building losses. Hence loss estimation based on observed damage characteristics of buildings in historical earthquakes only can provide accurate loss results. This chapter has discussed in detail the development of building vulnerability functions (VF) based on observed loss data. The data, particularly those that are from insurance companies, can be censored because of insurance deductibles. This is traditionally dealt with maximum likelihood approach and this approach has been described here. But this approach requires a large number of very high-quality building-specific loss data, which is generally not available. The other disadvantage of data-based vulnerability functions is that the results cannot be extrapolated from one region to the another because of regional building design and construction practices, which are generally unique. Hence this chapter described development of building vulnerability functions based on expert opinion. It is found that this approach can provide significantly biased results and so this is used rarely nowadays. This approach, however, would be used for assessing seismic risk in the future when time and resources are limited.

Over the years, different analytical approaches have been developed and only a few of those have been discussed in this chapter due to limitations in space. This chapter discussed in detail development of vulnerability functions following capacity spectrum method by HAZUS, assembly-based vulnerability, and a comprehensive approach following Jayaram et al. (2012) Analytical approaches have de facto become standard in the recent years for seismic risk calculations. Although these approaches require a large number of parameters, researchers are actively estimating those based on laboratory experiment and observed damage data. It is well known that there is significant uncertainty in the estimation of those parameters vis-à-vis VFs based on those parameters, but the uncertainty in VF can be easily estimated based on the analytical approaches. The importance of consideration of epistemic and aleatory uncertainty and the methodology in order to estimate those in building vulnerability has also been discussed here.

Cross-References

- ▶ Analytic Fragility and Limit States [P(EDP| IM)]: Nonlinear Dynamic Procedures
- ▶ Empirical Fragility
- ▶ Insurance and Reinsurance Models for Earthquake
- ▶ Seismic Risk Assessment, Cascading Effects
- ▶ Seismic Vulnerability Assessment: Masonry Structures
- ▶ Seismic Vulnerability Assessment: Reinforced Concrete Structures

References

ATC-13 (1985) Earthquake damage evaluation data for California. Applied Technology Council, Redwood City

ATC-58 (2011) Guidelines for seismic performance assessment of buildings (75 % draft). Applied Technology Council, Redwood City

Benjamin JR, Cornell CA (1970) Probability, statistics and decisions for civil engineers. McGraw-Hill, New York

Bonowitz D, Maison BF (2003) Northridge welded steel moment-frame damage data and its use for rapid loss estimation. Earthquake Spectra 19(2):335–364

Bozorgnia Y, Abrahamson NA, Atik LA et al (2014) NGA-West2 research project. Earthquake Spectra. doi:10.1193/072113EQS209M

Cornell CA, Krawinkler H (2000) Progress and challenges in seismic performance assessment. PEER Center News 3(2), http://peer.berkeley.edu/news/2000spring/

CUREE (2001) Improving loss estimation for woodframe buildings: CUREE-Caltech woodframe project. In: Porter KA, Beck JL, Seligson HA, et al (eds) Element 4, economic aspects. Consortium of Universities for Research in Earthquake Engineering, Richmond

FEMA 355F (2000) State of the art report on performance prediction and evaluation of steel moment-frame buildings. Federal Emergency Management Agency, Washington, DC

Haselton CB (ed) (2009) Evaluation of ground motion selection and modification methods: predicting median interstory drift response of buildings, PEER report 2009/01. PEER GMSM, Pacific Engineering Research Center, University of California, Berkeley

HAZUS-MH MR3 (2003) Multi-hazard loss estimation methodology: earthquake model. Department of Homeland Security, Washington, DC

III (2013) Earthquakes: risk and insurance issues. Insurance Information Institute. http://www.iii.org/issue_updates/eathquakes-risk-and-insuance-issues.html. Accessed 8 Nov 2013

Jaiswal KS, Wald DJ, Perkins D et al (2013) Estimating structural collapse fragility of generic building typologies using expert judgment. In: 11th international conference on structural safety and reliability (ICOSSAR), New York

Jayaram N, Shome N, Rahnama M (2012) Development of earthquake damage functions for tall buildings. Earthquake Eng Struct Dyn 41(11):1495–1514

PEER-2011/05 (2011) Case studies of the seismic performance of tall buildings designed by alternative means. In: Moehle J, Bozorgnia Y, Jayaram N et al (eds) Task 12 report for the tall buildings initiative. Pacific Earthquake Engineering Research Center Report, Berkeley

Porter KA, Kiremidjian AS, LeGrue JS (2001) Assembly-based vulnerability of buildings and its use in performance evaluation. Earthquake Spectra 17(2):291–312

Ramirez CM, Miranda E (2009) Building-specific loss estimation methods & tools for simplified

performance-based earthquake engineering, report no 171. The John A. Blume Earthquake Engineering Center, Stanford University, Stanford

Shome N, Bazzurro P (2009) Comparison of vulnerability of a new high-rise concrete moment frame structure using HAZUS and nonlinear dynamic analysis. In: 10th international conference on structural safety and reliability (ICOSSAR), Osaka

Shome N, Cornell CA (2000) Structural seismic demand analysis: consideration of collapses. In: 8th ASCE specialty conference on probabilistic mechanics and structural reliability, St Luis

Shome N, Jayaram N, Rahnama M (2012) Uncertainty and spatial correlation models for earthquake losses. In: 15th world conference on earthquake engineering (WCEE), Lisbon

Shome N, Jayaram N, Krawinkler H et al (2014) Loss estimation of tall buildings designed for the PEER tall building initiative project. Earthquake Spectra (accepted for publication)

Steinbrugge KV (1982) Earthquakes, volcanoes and tsunamis – an anatomy of hazards. Skandia America Group, New York

Wesson RL, Perkins DM, Leyendecker EV et al (2004) Losses to single-family housing from ground motions in the 1994 Northridge, California, Earthquake. Earthquake Spectra 20:1021–1045

Seismic Monitoring of Volcanoes

Glenn Thompson
School of Geosciences, University of South Florida, Tampa, FL, USA

Synonyms

Volcano-seismic monitoring

Introduction

Seismicity beneath a volcano usually increases before an eruption because magma and volcanic gas must first force their way up through fractures and passageways. When magma and volcanic gases or fluids move, they will either cause rocks to break or cracks to vibrate. When rocks break, high-frequency earthquakes are triggered. When cracks vibrate, either low-frequency

earthquakes or a continuous shaking called volcanic tremor, which can last from minutes to days, occurs.

Volcanic earthquakes often occur in swarms, which are clusters in time and space of similar earthquakes without an obvious mainshock. Volcano seismologists look for changes in the rate, size, and location of earthquakes and for the occurrence of swarms and tremor to forecast eruptions and to evaluate whether a volcanic eruption is intensifying or ending. Volcanic hazards including explosive eruptions, rockfall, pyroclastic flows, and lahars also cause ground vibrations and can be identified by their seismic signatures. This makes it possible, in principle, to detect them as they happen and issue warnings.

This entry discusses volcano-seismic monitoring from the viewpoint of a scientist leading a seismic monitoring program at a volcano observatory. Seismic monitoring records continuous, high-sample-rate data on the internal state of the volcano. Other monitoring techniques require manual labor to collect or process data, rely on daylight and good weather conditions, only detect volcanic activity once it has reached the surface, or have a low sample rate. For these reasons, seismic monitoring is the backbone of most volcano observatories. A volcano-seismic monitoring program comprises of a seismic network, a telemetry system, data acquisition, and alarm, analysis, and archival systems. At large observatories, personnel may include electronics engineers, volcano seismologists, seismic analysts, software developers, and network administrators. At small observatories, one person may cover most or all of these roles.

A volcano observatory has to be able to process large quantities of streaming data, detect changes in the volcanic system immediately, and respond without having to do a lot of manual analysis. This requires high levels of automation and systems engineering. While top priority is interpreting real-time data using established knowledge from the scientific field of volcano seismology, other important roles are troubleshooting data problems and conducting (or promoting) research that may lead to a better understanding of the volcano or lead to improved monitoring tools.

Much of the equipment and software used for volcano-seismic monitoring were originally developed for regional earthquake monitoring. The main difference is that the software needs to be augmented because of the diverse range of seismic signals recorded at volcanoes, many of which elude standard earthquake detection and location techniques. These signals, and the framework in which they are commonly interpreted, are discussed in the next section. The sections that follow describe volcano observatories including the operations room which is the command center during a crisis, the design of volcano-seismic networks, the real-time seismic monitoring and data analysis tools found in most observatories, and the challenges of maintaining this infrastructure and managing data effectively. In the final section, other monitoring techniques employed at observatories are briefly discussed.

Volcano-Seismic Signals

Volcano seismologists are most interested in signals which are anomalous, because they have high amplitudes, characteristic waveforms, or unusual frequency behavior. Given the variety of volcanoes around the world, with staggeringly different eruptive styles, magma compositions, and viscosities, it is perhaps remarkable that some common types of seismic signals are observed. These signals can be broken down into three broad categories: volcanic earthquakes ("events"), continuous signals, and surface signals. Each of these is discussed below. Any volcano may exhibit some or all of these at some time and perhaps may also exhibit exotic signals unique to that volcano. Classification is important because each type of signal may represent a different physical mechanism. However, classification is problematic.

By analyzing the spatial and temporal patterns between different signal types and volcanic activity, a greater understanding of a particular

S

volcano may be revealed. There are common patterns of behavior, gleaned from analyzing the seismic data from many volcanic eruptions. These are encapsulated in the Generic Volcanic Earthquake Swarm Model (McNutt 1996), a sequence observed at many volcanoes, the main features of which are volcano-tectonic earthquake swarms, followed by low-frequency earthquake swarms, tremor, and eruption.

For a more detailed discussion of volcano-seismic signals and many other topics mentioned in this chapter, Wasserman (2012) is highly recommended.

Volcanic Earthquakes

Volcano-Tectonic Earthquakes
Volcano-tectonic (VT) earthquakes are tectonic earthquakes that occur near active volcanoes. The physical mechanism is shear failure and slip on a fault plane triggered by magma ascent or the relaxation that occurs after magma is erupted. VT earthquakes have clear P- and S-waves (if recorded with a good signal-to-noise ratio) and high-frequency content (>5 Hz). VT earthquakes can be located in the same way as other tectonic earthquakes, i.e., by using the differential travel times between the P and S phases across a seismic network. However, most VT earthquakes are small ($M_L = -0.5$–1.5), and poor signal-to-noise ratio often prevents identification of phases, so in practice many (perhaps most) VT earthquakes cannot be located. Velocity models for volcanoes are often poorly determined, making absolute depths unreliable. Trends in relative depths are a useful diagnostic tool, however, and may indicate the rise of magma toward the surface. VT earthquakes frequently occur in swarms that consist of many similar-sized events and do not occur in mainshock-aftershock sequences where one single event dominates.

Long-Period Earthquakes
Long-period (LP) earthquakes are unique to volcanic regions. They have emergent onsets and a narrow frequency range with a peak frequency typically from 1 to 4 Hz. Tornillos are LP earthquake events with a particularly monotonic appearance and an exponentially decaying tail. LP earthquakes lack discernable P or S phases; consequently, most are not located. Evidence from Soufrière Hills Volcano (Montserrat) that they are often associated with venting from the surface of the dome and often trigger rockfall suggests they may originate at depths of less than 1 km. LP earthquake focal mechanisms reveal a volumetric component which is evidence of a fluid phase. LP earthquake swarms in volcanic systems are often associated with eruptions or intrusions and are believed to be due to processes such as pressure-induced oscillations of fluid-filled cracks in magmatic and hydrothermal systems.

Hybrid Earthquakes
Hybrid earthquakes have a high-frequency P-wave onset, typical of a VT earthquake, followed by a long-period tail. They may represent triggering of an LP earthquake by a VT earthquake. They typically occur in swarms and may be indicative of magma intrusion or extrusion.

Low-Frequency Earthquakes
Collectively, long-period and hybrid earthquakes are referred to as low-frequency earthquakes. There may be a continuum between long-period and hybrid earthquakes.

Deep Long-Period Earthquakes
Some volcanoes produce low-frequency earthquakes (e.g., 20–40 km depth), particularly in the early stages of unrest. These have been called deep long-period (DLP) earthquakes. These have emergent P and S phases, are rich in frequencies below 5 Hz, and are inferred to represent movement of deep-seated magma and associated fluids in the mid-to-lower crust. They look like VT earthquakes with the high frequencies filtered out, perhaps because they occur in a highly attenuating region.

Explosion Quakes and Very-Long-Period Signals
Explosion quakes are signals that accompany Strombolian or other (larger) explosive eruptions.

These signals are identified by the occurrence of an airwave which is caused by expanding gas accelerated at the vent exit. This wave mainly travels through the air with the typical speed of sound (330 m/s at 20 °C).

With the advent of broadband volcanic seismology, very-long-period (VLP) signals have been observed in seismograms from some volcanoes. Many of these are broadband versions of explosion quakes. VLP signals may be produced by the rapid expansion of a large gas volume at shallow depth within the conduit. The gas expansion might result from shallow gas coalescence and expansion or from expansion of a gas slug formed at greater depth.

Continuous Signals

Swarms
Earthquake swarms are sequences of earthquakes closely clustered in space and time without a single mainshock. Volcanic earthquakes often occur in swarms, whereas nonvolcanic earthquakes usually follow a mainshock-aftershock sequence. The Gutenberg-Richter law describes the relative frequency of occurrence of earthquakes of different magnitudes, and this is encapsulated in a parameter called the b-value (see Sanchez et al., "▶ Frequency-Magnitude Distribution of Seismicity in Volcanic Regions," this volume). Volcanic earthquake swarms typically have b-values higher than one (most of the energy is in small earthquakes), whereas mainshock-aftershock sequences typically have $b \leq 1$ (most of the energy is in a single mainshock). Volcanic earthquakes occur in hot, highly heterogeneous material containing many small faults. Nonvolcanic earthquakes tend to occur in more homogeneous material with failure on a single larger fault, which then loads adjacent faults, causing many of them to fail producing aftershocks. Low-frequency earthquake swarms often contain one or more families of repeating earthquakes. Each family is identified by a unique waveform and is the result of the same nondestructive source process being activated repeatedly in the same location.

Benoit and McNutt (1996) examined the reports of over 600 swarms to compile the Global Volcanic Earthquake Swarm Database. They identify three main types of volcanic earthquake swarm. Type 1 swarms begin before an eruption and have a mean (and mode) duration of 8 days. Type 2 swarms begin coincident with, or during, an eruption. Type 3 swarms (39 % of records) were not associated with eruptions and have a mean duration of 3.5 days and a mode of 1.5 days. The most common depth for swarms is 2–3 km. Their study does not distinguish between different types of volcanic earthquakes. They suggest it may also be biased by the underreporting of swarms not associated with eruptions.

Tremor
Volcanic tremor is a narrowband (usually 1–4 Hz), continuous vibration thought to be due to sustained subsurface movement of magma or volatiles and is often observed before explosive eruptions. It may last from a few minutes to months in duration. Tremor has similar spectral characteristics as low-frequency earthquakes. Harmonic tremor shows one or many regularly spaced overtones in addition to a fundamental frequency. Sometimes spectral peaks in harmonic tremor glide upward (or downward) in frequency over as little as a few minutes. Eruption tremor, a continuous vibration coincident with explosive eruptions, has a wider frequency range (0.5–10 Hz). Dome collapses, lahars, weather storms, and telemetry problems can all produce signals that could be confused with tremor.

Tremor may be the result of continuous excitation of the source that produces low-frequency earthquakes. The resonance of a fluid-filled conduit is one model for the origin of tremor, as it is for low-frequency earthquakes. Interface waves traveling along the crack or conduit wall can produce overtones, like an organ pipe. Gliding lines could be caused by a change in the sound speed of the fluid (e.g., due to a change in bubble density) or a change in a length of the section of conduit that is resonating (e.g., a change in nucleation depth).

S

Tremor may also be a superposition of low-frequency earthquakes. There are many observations of low-frequency earthquake swarms merging into volcanic tremor. Modeling has shown that harmonic tremor could appear when overlapping earthquake signals occur at regular time intervals that differ by less than 2 % (Powell and Neuberg 2003). As this regular rate of earthquakes gradually increases (or decreases), gliding spectral lines are produced. In this model, tremor is a low-frequency earthquake swarm.

Surface Signals
Surface signals are those generated by hazardous surface processes such as rockfall, pyroclastic flows, and lahars. Since these processes occur at the surface, they mainly generate surface waves.

Rockfall and Pyroclastic Flow Signals
Actively growing lava domes are highly unstable, and blocks can be observed falling almost continuously, disintegrating into smaller blocks and plumes of hot ash. Most of these rockfalls are small, but many generate detectable seismic signals. If larger parts of the dome collapse, or if there is an explosive component destabilizing the blocks, a pyroclastic flow may be generated. Pyroclastic flows are more energetic, produce vigorously convecting ash clouds, behave more like a fluid (less friction), and are therefore able to travel farther and at higher speeds. There is a continuum from the smallest rockfalls to the largest pyroclastic flows, and they mostly are spawned from the part of the dome that is actively growing. Pyroclastic flows can also be generated by the collapse of an eruption column.

Rockfall signals are emergent, contain a wide range of frequencies (1–10 Hz, peaking 3–4 Hz), and are dominated by surface waves. They have been located by exploiting the amplitude distribution of rockfalls signals across the seismic network. On Soufrière Hills Volcano (Montserrat), some rockfall and many pyroclastic flows have a long-period precursor, and it is possible that this is an explosive degassing signal. At Unzen Volcano (Japan), pyroclastic flow signals are found to comprise three parts. First, a dome collapse is

observed simultaneously with a small LP signal. Second, free-falling blocks impact the slope below (and fragment) simultaneous with a 0.5 Hz signal. Third, the fragmented material generates a rockfall signal.

Dome collapse signals are the superposition of many rockfall events, which may be occurring simultaneously on different flanks of the volcano. Dome collapses may last minutes to hours. From studying dome collapse signals, it is possible to tell how long dome collapses lasted and which phases of the collapse were most energetic. Several major dome collapses at Soufrière Hills Volcano (Montserrat) and Merapi Volcano (Indonesia) have been triggered by intense rainfall.

Lahar Signals
The composition of lahars varies from muddy water to highly erosive, dense mixtures of wet ash, rocks, and boulders that set like concrete. Lahars usually occur during or immediately after heavy rainfall. Barclay et al. (2006) found that about 2 cm per hour of rain falling on unconsolidated materials was sufficient to trigger lahars on Soufrière Hills Volcano (Montserrat). However, hot volcanic material can generate lahars by mixing with crater lakes or with snow and ice on glaciated volcanoes. Volcanic activity may also melt the base of a glacier and cause an outbreak flood that briefly rivals the force of a major river. These are particularly common in Iceland and are known in Icelandic as *jökulhlaups*.

Lahar signals are tremor-like signals that can be distinguished from pyroclastic flow signals by their duration (tens of minutes to hours), higher-frequency content (6–10 Hz, up to 100 Hz in some cases), and slower speeds; they show amplitude peaks on different stations perhaps minutes apart as they travel down valley. To improve the detection of lahar signals, seismic stations, video cameras, and tripwires can be added at various positions adjacent to (or within) the valley.

Difficulties of Event Classification
Classification of volcano-seismic signals is difficult. The signal recorded is a convolution of the source signature, propagation and site effects,

and the instrument response, superimposed on noise which varies spatially and temporally. The source signature may be the result of complex interactions between a multiphase fluid and an unknown geometry of dikes and conduits. Further from the source, the signal-to-noise ratio decreases, and increased scattering and separation between P, S, and surface waves lengthen the signal, making it less clear. Inelastic attenuation may be greatest at shallow depths where poorly consolidated, highly fractured material results in higher frequencies being filtered out. Attenuation may also be significant at depths of 20 km or more, because of high heat flow and partial melt. Similar-looking signals at two different volcanoes, or even at the same volcano, might be caused by different physical mechanisms.

Classification is also subjective. Different seismic analysts may disagree on the classification of a particular event. Discrepancies in classification also appear in published literature: an LP in Lahr et al. (1994) looks similar to a hybrid in Luckett et al. (2007). Classification might artificially separate signals that lie along a continuum. The classification scheme may also evolve during volcanic unrest as a wider variety of signals are recorded. For these reasons it is helpful to reassess the event classes used before interpreting trends in their rates of occurrence or locations.

Terminology also differs. Some of the terms describe the frequency content of the signal and others imply a physical mechanism. VT earthquakes are also called high-frequency events and A-type events. LP earthquakes are also called B-type events. Collectively, long-period and hybrid events are called low-frequency (LF) events, because they are both dominated by low-frequency coda. Low-frequency seismicity includes low-frequency events and volcanic tremor.

Interpreting Volcanic Seismicity

A common pattern of volcanic seismicity is described by the Generic Volcanic Earthquake Swarm Model (McNutt 1996). Background seismicity varies from volcano to volcano and can only be established for a particular volcano through a long-established seismic monitoring

record. The first sign of unrest is often an increase in the rate of *VT earthquakes*, indicating rock fracture due to changes in stresses caused by rising magma. In some cases it may be the occurrence of *DLP earthquakes*.

When magma reaches shallower depths, volatiles exsolve, and *LP earthquakes* are recorded. Volatiles cause a reduction in the acoustic velocity and an increase in the impedance contrast of magma cavities, trapping energy and leading to longer, lower frequency signals. As the volatile content increases, more energy is trapped, and this may lead to LP earthquake events called *tornillos*. Narrowband *tremor* may be generated by continuous vesiculation as magma rises further or by the boiling of groundwater. If groundwater is heated rapidly, a phreatic explosion may result.

At any time in the sequence, magma rise may stall. This may be the end of the unrest, or it may recommence as more magma is injected at depth adding more heat and volatiles to the system. *Tremor* due to groundwater boiling may subside as the system dries out, leading to a period of relative quiescence. For basaltic systems, an eruption may then occur without further warning.

For andesitic to rhyolitic systems, *hybrid earthquake swarms* often occur due to repeated shear failure of viscous magma as it gets close to the surface. *Hybrid earthquake swarms* are frequently associated with growth of a lava dome. If the events merge into a continuous tremor, *harmonic tremor* and *gliding spectral lines* might be observed and are the strongest signs that an explosive eruption may be about to occur. During an explosive eruption, violent ground shaking typically manifests as a broadband tremor signal. *Explosion earthquakes* may be recorded, identifiable by the shockwave that travels through the air and is coupled back into the ground.

Explosive eruptions are short-lived, and hopefully before they occur, any vulnerable populations will have been evacuated. But extrusive eruptions may last for years or decades, with varying levels of activity, spawning rockfall, pyroclastic flows, and lahars. Monitoring these long-lived eruptions presents challenges of its own, since communities are often in close

proximity. Escalations in activity are often preceded by short sequences of VT earthquakes, low-frequency earthquakes, and tremor, perhaps suggesting a batch of new magma rising toward the surface. However, volcano-seismic data have also been used to track the rate, energy, and location of debris flows and estimate extrusion rates.

An important question to answer once an eruption is underway is when will it stop. The first sign may be the cessation of low-frequency seismicity, indicating that volatiles are no longer present in the system, which may be confirmed with gas flux measurements. More significant may be the occurrence of VT earthquakes at depth with fault-plane solutions consistent with magma withdrawal (Roman et al. 2006). Geodetic data may also indicate deflation of the volcanic edifice.

A more detailed discussion of the interpretation of volcano-seismic data is beyond the scope of this entry. There are several excellent summaries of our evolving understanding of volcano-seismology including Chouet and Matoza (2013), Chouet (2003), and McNutt (2000, 2002, 2005).

Observatories

Volcano observatories vary greatly in their level of sophistication. The simplest observatory may be a hut on the flank of a volcano, with a single seismograph recording on paper on a revolving drum. A modern observatory, however, will have a network of seismic stations with data telemetered to the observatory using a variety of communication systems such as FM radio, satellite, and cellular networks. The observatory may be on the flanks of the volcano, within a few miles of the active vent. Or it may be tens or hundreds of miles away, colocated with a university or a government agency. Many observatories monitor a single volcano, but some monitor several. The Alaska Volcano Observatory has operated seismic networks on as many as 32 volcanoes. The Japan Meterological Agency monitors 47 volcanoes with real-time seismic data. This poses unique challenges: how to monitor so many

volcanoes in parallel, how to interpret the data without being able to see the volcanoes, how to maintain so many networks, and how to manage all the data.

Participation in a volcanic crisis, is an opportunity to help society and also witness some of nature's most spectacular phenomena. Having the right team is crucial. The leader of a volcano observatory must be experienced in volcano monitoring and be committed to the job. The work can be stressful, the hours long, and hard decisions have to be made in the heat of the moment, e.g., regarding evacuations. Weekly meetings play a vital role at many observatories by pulling the team together, integrating data from many monitoring techniques, identifying technical problems, and prioritizing work. All observatory staff must conduct themselves professionally and have the authority to perform the roles assigned to them. Fostering good communications within the observatory and with the authorities and the public builds trust and enhances public safety.

The backbone of a volcano observatory is the seismic monitoring program, and the heart of the observatory is the operations room. It is from there that seismologists track live data streams, coordinate field teams and warn them of hazardous activity, and alert the authorities (and the public) to escalations in activity. When volcanoes are at background level, the operations room may be unoccupied. Periodic data checks, perhaps coupled with automatic alarm systems, may be enough to keep seismologists abreast of significant changes in seismicity. When unrest begins, periodic checks may become more frequent. At some point the operations room is activated and manned 24 h a day.

These different stages of volcanic activity, from background to unrest to impending eruption, greatly impact the level of seismic monitoring that can be done. At background level, seismologists may be able to conduct research. When a crisis is underway, it will be difficult to do anything more than interpret available real-time data using operational tools that are in place. Part of the job of an observatory seismologist is to engineer the infrastructure of the seismic

monitoring program and operations room to be able to respond effectively during a crisis. This infrastructure includes seismic networks and software systems used to monitor volcanic seismicity. These are the subjects of the next two sections.

Volcano-Seismic Networks

Around 550 volcanoes have erupted in historical times, and about 200 of these are seismically monitored. The best networks tend to be around volcanoes that pose a particularly high threat to large population centers, e.g., Etna and Vesuvius (Italy), Rainier and St. Helens (continental USA), and Kīlauea (Hawaii). Many volcanoes are monitored only by a single station, often as part of a regional network. Nevertheless, by tracking the number and cumulative energy of different types of seismic signals recorded each day on a single station, a volcanic eruption can be anticipated. It may then be possible to rapidly deploy sufficient additional stations to locate earthquakes and forecast where and when an eruption may take place.

The most commonly used seismometers for volcano monitoring are short-period sensors. These have a corner frequency of 0.5–2 Hz and come in single (vertical) and three component varieties. Short-period seismometers are usually deployed with analog telemetry. Both have a dynamic range (the ratio between the largest and smallest amplitudes that can be represented) of only a few thousand. This results in signal amplitudes often being "clipped," which greatly diminishes the monitoring and research value of the data. Three-component portable broadband seismometers, available since the late 1980s, use a force-feedback circuit and typically have a corner frequency of 30–120 s and a dynamic range of several million, allowing them to record signals on-scale even right up on the volcano summit. They have to be coupled with 24-bit digital telemetry systems, which have a similar dynamic range.

Analog-telemetered data are time stamped at the observatory by a time signal from a GPS clock, but modern field digitizers have an input

for a GPS clock and can time stamp data on site (which is more accurate). On-site data recording enables data to be retrieved or retransmitted later if there is a communication outage. Two-way telemetry allows data packets to be resent automatically if they are not received intact and allows troubleshooting and reconfiguring of stations from the observatory without the expense and delay of a site visit.

While a digital broadband network offers many advantages over an analog short-period network, the equipment is more expensive and power requirements are 2–3 times higher, meaning that additional solar panels and batteries are needed. Broadband sensors also require precise leveling. Many analog networks have been upgraded to digital telemetry to improve dynamic range, and most new networks use digital telemetry, but it is still common to find analog telemetry at volcano observatories.

Designing a volcano-seismic monitoring network is complicated and involves multiple tradeoffs. For the same budget a volcano observatory may be able to install a few digitally telemetred broadband stations or many more analog-telemetered short-period stations. The best choice depends on the goal.

For a volcano that is far removed from population centers, the goal might be to detect eruptions so that aviation authorities can be warned of hazardous ash clouds, requiring a minimal level of monitoring. For a densely populated region around a frequently active volcano, or one capable of devastating eruptions, higher quality monitoring is needed to provide as much lead time as possible. The magnitude detection threshold may vary from 2 for a regional network to below 0 for a dense volcano-seismic network. Moran et al. (2008) identified four levels of volcano-seismic monitoring and the number of short-period, broadband, strong motion, and infrasound sensors needed within different radii to make those levels of monitoring achievable. Their findings are summarized in Table 1.

A key consideration for locating earthquakes accurately is to minimize the azimuthal gap. A ring of stations, which encloses most of the volcanic-earthquake epicenters, is ideal with a

Seismic Monitoring of Volcanoes, Table 1 Recommended instrumentation to achieve different levels of volcano-seismic monitoring from Moran et al. (2008)

Level	Goal	Recommendation
1	**Minimal monitoring/eruption detection** Detect $M > 1.5$ earthquakes; crudely locate $M > 3$ earthquakes	Site a total of five seismic stations within 200 km, including two within 50 km of the volcanic center
2	**Limited monitoring/unrest detection** Detect $M > 1$ earthquakes; crudely locate $M > 2$ earthquakes; determine event type; detect energetic seismic tremor	Site a total of five seismic stations within 50 km, including two within 10 km of the volcanic center
3	**Basic real-time monitoring** Detect $M > 0.5$ earthquakes; accurately locate $M > 1$ earthquakes; determine event type; detect seismic tremor; on-scale recording of energetic seismicity on at least one station; detect very-long-period events Detect changes in travel time; detect broad-scale changes in seismic velocity Use fault-plane solutions and b-values to determine generalized stress fields near the volcanic center	Site six to eight seismic stations within 20 km of the volcanic center, including two or three stations with at least one three-component sensor and at least one broadband station, within 5 km
4	**Advanced real-time monitoring** Detect and accurately locate $M > 0$ earthquakes; determine event type; detect and crudely locate seismic tremor; on-scale recordings of energetic seismicity on multiple stations; detect and crudely locate very-long-period events and other very-low-frequency seismicity Determine detailed source properties of tornillos; construct 3-D velocity models (provided local seismicity is sufficient) Detect explosions and possible infrasonic precursors to explosions at restless and (or) frequently active volcanoes Detect detailed stress-field changes by calculating well-constrained fault-plane solutions and (or) moment tensors, mapping b-values at high spatial resolution, and detecting changes in S-wave-splitting directions over time	Site 12 to 20 seismic stations within 20 km of the volcanic center, including at least six broadband stations, as many as possible within 5 km; at least one strong-motion station; and at least two infrasonic stations (with at least two infrasonic sensors per station) at erupting, restless, and (or) frequently active remote volcanic centers

few stations closer to the volcanic center where they can help detect smaller seismic signals and be used to locate summit events with greater precision and improve depth resolution. For stratovolcanoes, all the earthquakes may be within 1–2 km of the volcanic center, whereas for calderas the seismicity might be diffused over an area 15–20 km in radius. It is useful also to have real-time data from at least one more distant station (which may be part of a regional network) to help constrain the depths of deeper events, and to discriminate more effectively between regional and local volcanic earthquakes. These steps make it possible to locate earthquake reasonably well relative to each other, but how well

they match the true locations of the earthquake will depend on the velocity model and to a lesser degree, on the algorithm used to locate the events. The sophistication of the velocity model may vary from a simple one-dimensional regional model with constant velocity layers to 3-D models determined from seismic tomography.

Site selection depends on the geology, topography, accessibility, and noise. It is preferable to install seismometers in solid bedrock, but volcanoes are typically comprised of layers of ash, flow deposits, and boulders. Repeaters may be required on ridges to rebroadcast signals or boost them if transmitting over tens of kilometers, but represent additional expense and

Seismic Monitoring of Volcanoes, Table 2 Some of the best volcano-seismic monitoring networks. The numbers in parentheses indicate number of borehole instruments. All of these networks surpass level 4 as defined in Table 1 of Moran et al. (2008)

Volcano	Stations within 5 km			Stations within 20 km		
	Broadband	Short period	Total	Broadband	Short period	Total
St. Helens (USA)	2	7	9	2	16 (4)	18 (4)
Soufrière Hills (Montserrat)	7 (1)	0	7 (1)	9 (1)	6 (4)	15 (5)
Piton de la Fournaise (Reunion)	14	6	20	16	17	33
Halemaumau, (Hawaii)	12	5	17	19	14	33
Stromboli (Italy)	18	0	18	18	0	18
Etna (Italy)	?	?	13	?	?	34
Vesuvius (Italy)	?	?	17	?	?	21
Erebus (Antartica)	6	6	12	6	6	12

potential points of failure. Sites that require helicopter access will be expensive to maintain, so often it is better to find sites that can be easily reached with four-wheel-drive vehicles or on foot. Sites should be far from traffic and other human noise and also away from any tall obstacles that may be vibrated by the wind, such as trees, cliffs, and radio towers. Burying sensors a few feet below the surface helps suppress high-frequency wind noise and low-frequency noise due to temperature and pressure variations.

Among the best volcano-seismic networks are those used to monitor Piton de la Fournaise (Reunion), Halemaumau (Hawaii), St. Helens (USA), and Soufrière Hills (Montserrat) volcanoes (Table 2). According to the criteria in Table 1, these are all level four networks. A map of the digital seismic network that has been operational in Montserrat since 2006 is shown in Fig. 1. This utilizes a mixture of Guralp CMG-40T broadband and Mark L4-C short-period seismometers, coupled with Guralp DM24 digitizers and FreeWave Spread Spectrum serial port radios. One of the broadband stations has a CMG-3 T seismometer in a 30-m borehole and less than 4-km from the dome. Four more sites feature a dilatometer and 2-Hz seismometer in 200-m boreholes, coupled with surface GPS. The Piton de la Fournaise network also uses CMG-40T and L4-C seismometers, coupled with Kinemetrics Q330 digitizers and a combination of wireless internet and FM radio telemetry.

Real-Time Monitoring

Volcanic activity may escalate suddenly, and the ability to rapidly identify anomalous volcano seismicity is critical. Real-time data visualization systems allow scientists to rapidly assess multiple parameters derived from seismic data such as hypocenters, magnitudes, event rates, spectral variations in tremor, etc. Alarm systems alert scientists to large events and the occurrence of high amplitude tremor and volcanic earthquake swarms. Data-rich websites and remote desktop connections allow scientists to respond rapidly to alarm: This is convenient and enhances safety. The US Geological Survey, through its Earthquake and Volcano Hazards Programs, has played a pivotal role in disseminating free, open-source software to aid in volcano-seismic monitoring. This section describes common techniques and software used to examine continuous seismic data and events in more detail.

Continuous Data

Digital Helicorder Plots
The most basic form of real-time data display is the helical drum recorder (often called "helicorders" or "drums"). A pen etches a trace on a smoked sheet (or draws an ink trace on a blank sheet) of paper wrapped around a cylindrical drum. Helical drum recorders have played a vital role in volcano-seismic monitoring, allowing rapid visualization of seismic amplitudes and event identification of signal types.

S

Seismic Monitoring of Volcanoes, Fig. 1 The Montserrat seismic network from April 2006 (Modified after Luckett et al. (2007)). Telemetry paths are shown as *lines*. The observatory is marked with MVO. Repeater sites are marked with a *gray circle*. Short-period stations are marked with a *triangle*. Broadband stations are marked with a *star*, except for the borehole broadband, MBLY, marked with a *square*. Not shown are the four 200-m borehole short-period stations, operated by the CALIPSO consortium since 2002
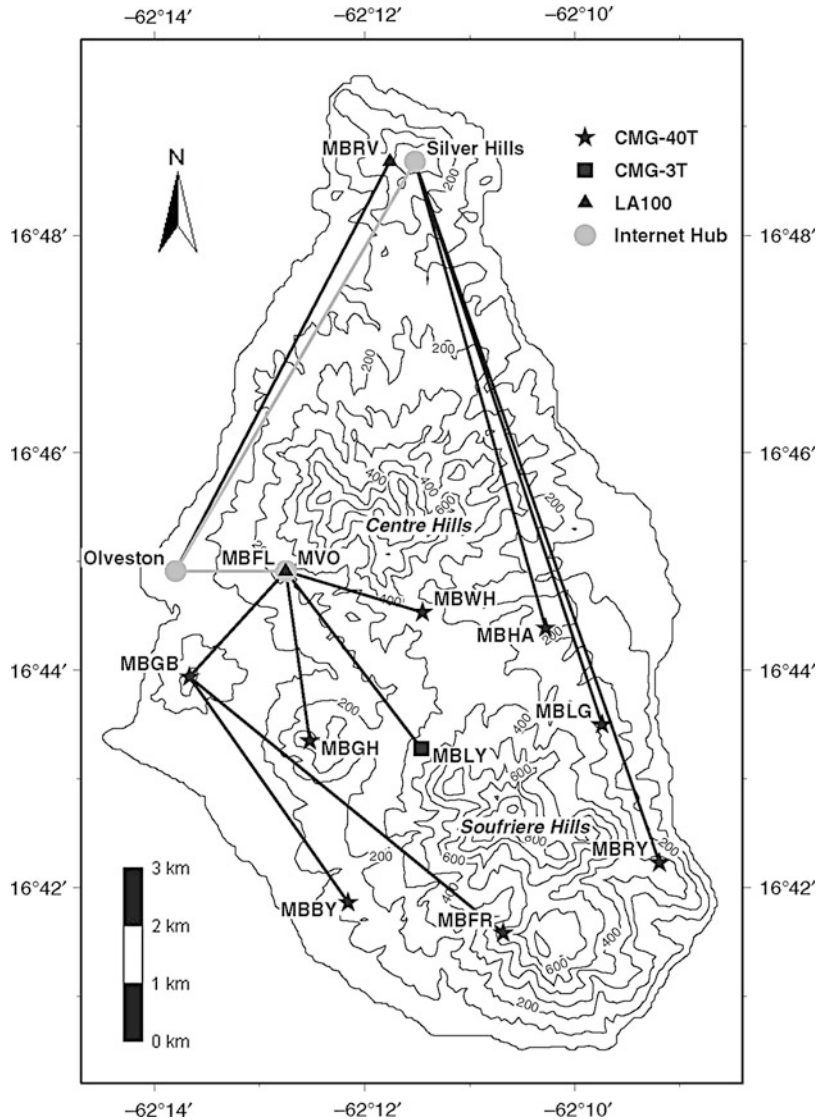
Helicorders were ubiquitous until the 1990s and are still used at many observatories today, but they have numerous drawbacks. They require considerable maintenance and provide limited dynamic range. Adjacent traces often overlap, making them hard to read. The data are not amenable to other forms of analysis, and the sheets require storage space. So there has long been a desire to replace helicorders with a software equivalent. Seismic Waveform Analysis and Real-time Monitor (SWARM) (http://volcanoes. usgs.gov/software/swarm/index.php) is an excellent solution that allows data to be plotted dynamically: The user can select the time range and the scale; the data can be filtered, and short segments of data can be highlighted and replotted as spectra or spectrograms. Many observatories have now replaced large collections of helicorders with multiple screens showing SWARM displays from different seismic channels. Figure 2 shows 48 h of seismicity at the Soufrière Hills Volcano in June 1997.
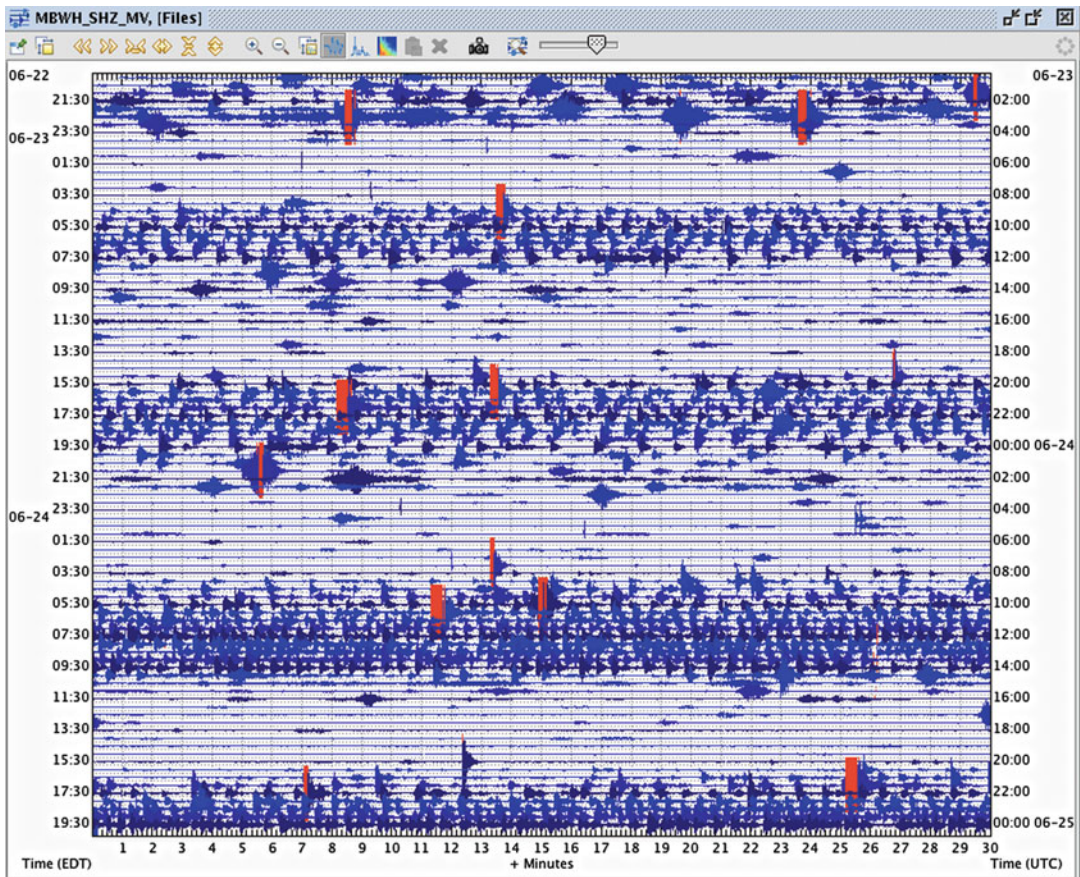
Seismic Monitoring of Volcanoes, Fig. 2 Screenshot of the software Swarm, showing a digital helicorder plot from station MBWH channel SHZ for 48 h from June 23, 1997. Remarkable cyclic hybrid swarms are visible. In June 25, a moderate dome collapse sent pyroclastic flows the northern flanks of the Soufrière Hills Volcano, claiming 19 lives

RSAM Plots

Continuous data are often downsampled to one sample per minute, which makes it easier to identify long-term trends in continuous seismic amplitude. The Real-time Seismic Amplitude Measurement system (Endo and Murray 1991) took the average amplitude of the seismic signal in each 1-min time window, and recorded these data into a file. RSAM data do not discriminate between different types of seismicity.

RSAM data show cyclic seismicity clearly (Fig. 3a). Changes in earthquake activity associated with dome-building episodes (Fig. 3b), weather, and instrumental difficulties are recognized as distinct patterns in RSAM datasets. RSAM data for dome-building episodes gradually develop into exponential increases that terminate just before the time of magma extrusion. Volcanic earthquakes and rockfall show up as isolated spikes on RSAM plots for seismic stations close to the edifice, but seldom for more distant stations. Weather-related noise shows up as low-level, long-term disturbances on all seismic stations, regardless of distance from the volcano. The RSAM system proved valuable in providing up-to-date information on seismic activity for three Mount St. Helens eruptive

Seismic Monitoring of Volcanoes, Fig. 3 (**a**) Plot of RSAM data with one sample per minute for station MBWH channel SHZ from June 22 to 29, 1997. The cyclic hybrid swarms mentioned in Fig. 2 are shown clearly. The spikes represent discrete events, e.g., pyroclastic flows. (**b**) RSAM plot showing the major dome collapse of the Soufrière Hills Volcano (Montserrat) which occurred on July 29, 2001
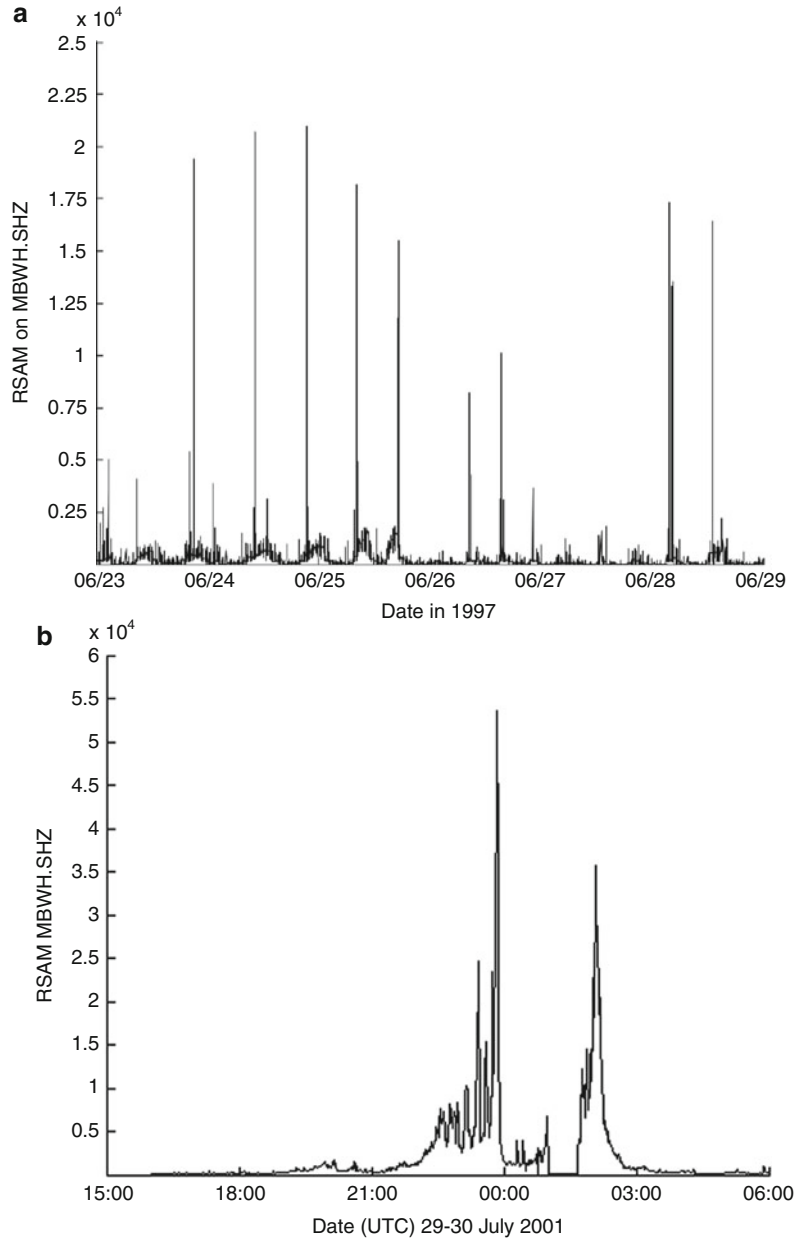
episodes from 1985 to 1986 and in numerous eruptions at other volcanoes since. Exponential increases in RSAM data commonly precede explosive eruptions. Inspired by RSAM, many other parameters have been computed on 1-min (or 10-min) timescale. For example, RSEM provides a relative measurement of energy on a 1-min (or 10-min) timescale.

Computing the median (rather than mean) amplitude in each time window provides a measurement of tremor amplitude less biased by events or spikes in the data - but is no longer strictly 'RSAM' data. Corrected for the instrument response and geometrical spreading, and then integrating the data, produces reduced displacement. Since 1996 the Alaska Volcano

Observatory has recorded an instrument corrected seismic amplitude and the peak frequency for each vertical-component seismic channel, on a 10-min timescale.

Spectrograms

Changes in tremor spectra may indicate different flow regimes or changes in source parameters such as geometry, sound speed, or ascent rate (Thompson et al. 2002). Harmonic tremor and gliding spectral lines are frequently followed by eruptions. Spectral monitoring can help differentiate between VT and low-frequency earthquakes, or tremor; wind noise; and electronic noise.

The Alaska Volcano Observatory (AVO) began experimenting with high-resolution spectrograms in 1996. For up to eight stations, a spectrogram is plotted underneath a normalized seismic trace. The time resolution is about 10 s and the frequency resolution 0.1 Hz. Each image file (Fig. 4) displays a 10-min seismic trace and spectrogram for up to eight seismic-data channels. These are ordered in increasing distance from the volcano. The data are instrument corrected and color coded, so that ground motions at one station can be easily compared with those at another station, even at a different volcano.

These web-based spectrograms have been a core AVO monitoring tool since 1998, and a seismologist reviews them twice a day. A convenient web interface (http://www.aeic.alaska.edu/spectrograms/mosaicMaker.php) shows mosaics of 10-min spectrogram image files (Fig. 5) and enables 12 h of data from more than 20 volcanoes to be reviewed in just 30–45 min. Similar web-based spectrograms are used for monitoring volcanic seismicity in the Cascades and Hawaii.

Tremor Alarms

Tremor is common precursor to eruptions, and strong tremor frequently accompanies eruptions. Observatory seismologists therefore need to be aware when tremor is being recorded. The earliest widely used volcano-seismic alarm system was part of the RSAM system (Endo and Murray

1991), and this is now incorporated in Earthworm (discussed below). The seismologist can choose which stations to monitor and define an amplitude threshold and a duration for each station. Both must be exceeded to trigger a station. The seismologist also must define the number of stations that must trigger simultaneously to declare an alarm. There is also a mechanism which reduces the false alarms due to regional or teleseismic earthquakes: A far away station can be chosen which prevents an alarm being declared when it exceeds an amplitude threshold. The tremor alarm is typically configured for a moderate amplitude and a duration of several minutes (the RSAM system also included an event alarm, which would typically be configured for high amplitude and duration of a few to tens of seconds).

The RSAM tremor alarm system was used to monitor the Soufrière Hills Volcano (Montserrat) from 1999 to 2003. Figure 6 shows the number of alarms issued over this period.

Events

Real-Time Seismic Event Catalogs

During a volcanic crisis, it is vital to have near-real-time information about the rates, sizes, and locations of seismic events, and for this a real-time event catalog is needed. Providing there are at least four stations, event detection software automates the process of capturing anomalous signals in large volumes of continuous data. Each channel of seismic data is detrended, filtered, and rectified, and a running short-term average (typically 1 to a few seconds) and long-term average (typically a few tens of seconds) are computed. Where the STA (short-term average) to LTA (long-term average) ratio exceeds a threshold, a candidate P or S arrival is declared. Association is the process whereby candidate arrivals' "picks" are grouped together based on similar arrival times and geographic location to declare an event. Location techniques typically use a 3-D grid search to minimize the difference between measured and theoretical travel times, given a particular velocity model. For volcanic earthquakes, it is most common to compute a

S
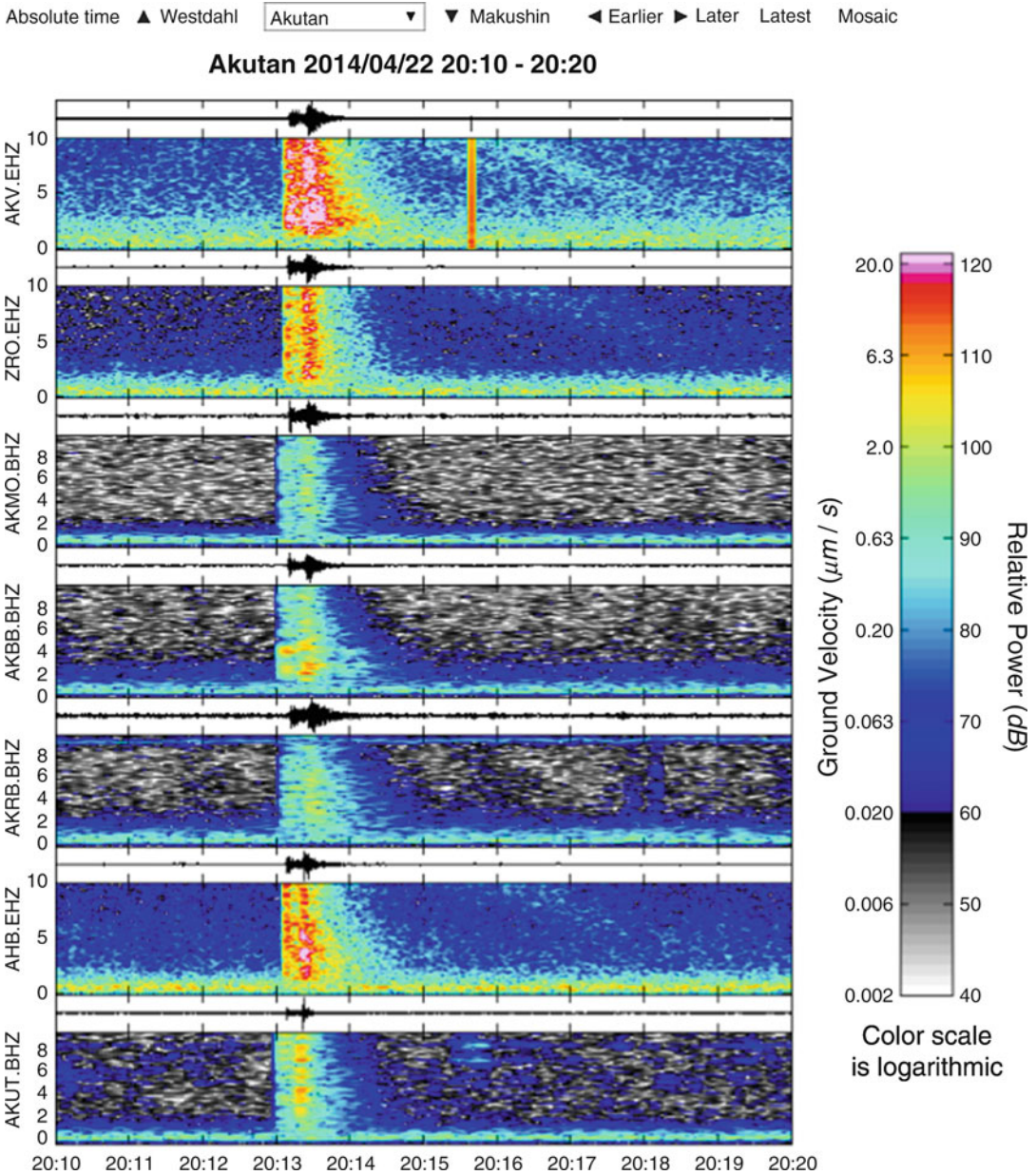
Seismic Monitoring of Volcanoes, Fig. 4 Seismic traces and corresponding spectrograms for the vertical components of seven seismic stations near Akutan Volcano, corresponding to 10 min of data on April 22, 2014. A regional earthquake is visible clearly around 20:13 UTC. The hottest colors shown (*pink*) suggest an S-wave ground velocity of 0.02 mm/s. Near-real-time spectrogram plots like this have been linked to the AVO internal website since 1996. The menu at the top of the screen allows easy navigation to other volcanoes or time periods

duration magnitude due to the prevalence of clipped signals from short-period analog telemetry. For on-scale recordings, local magnitude is often also computed. To visualize catalogs, it is helpful to view event rate and energy rate plots (Fig. 7) and hypocenter plots (Fig. 8).

There are some caveats with event catalogs. Discrete events are only one aspect of volcanic

Seismic Monitoring of Volcanoes, Fig. 5 A spectrogram mosaic, part of the same application shown in Fig. 4. Three hours of data are shown. Regional earthquakes are visible at 20:13 (Fig. 4) and 22:15. A calibration signal occurs on the top station (AKV) at 22:03. The apparent data dropouts are due to data latency

seismicity. Signals such as precursory earthquake swarms and tremor episodes, or those generated by explosive eruptions, dome collapses, pyroclastic density currents, and lahars, are of the greatest interest from a hazard perspective. Such signals generally evade standard earthquake-detection schemes based on comparing the short-term and long-term signal averages and are not systematically cataloged. The vast majority of volcano-seismic signals lack identifiable phases and so cannot be located with differential travel-time techniques. Earthquake catalogs therefore present only a narrow view of volcanic seismicity. Many volcano-seismic signals are emergent or long in duration, so STA/LTA detectors break down. Events may be masked by high background signals (e.g., tremor, wind, or electronic noise). Station outages also increase the

magnitude at which earthquakes become detectable. Times when an event catalog suggests there is little seismicity may actually be times of high seismicity. Real-time catalogs also suffer from poorly resolved locations and magnitudes – also a review by an analyst later corrects for these.

Swarm Alarms
Earthquake swarms, like tremor, are a common precursor to volcanic eruptions. A real-time event catalog can be used as the basis for detecting earthquake swarms.

Okmok Volcano (Alaska) erupted ash to a height of 15 km on July 12, 2008, after less than 5 hours of precursory seismicity which included an earthquake swarm. Concerned by this, the Alaska Volcano Observatory developed a swarm alarm system. The system identified the
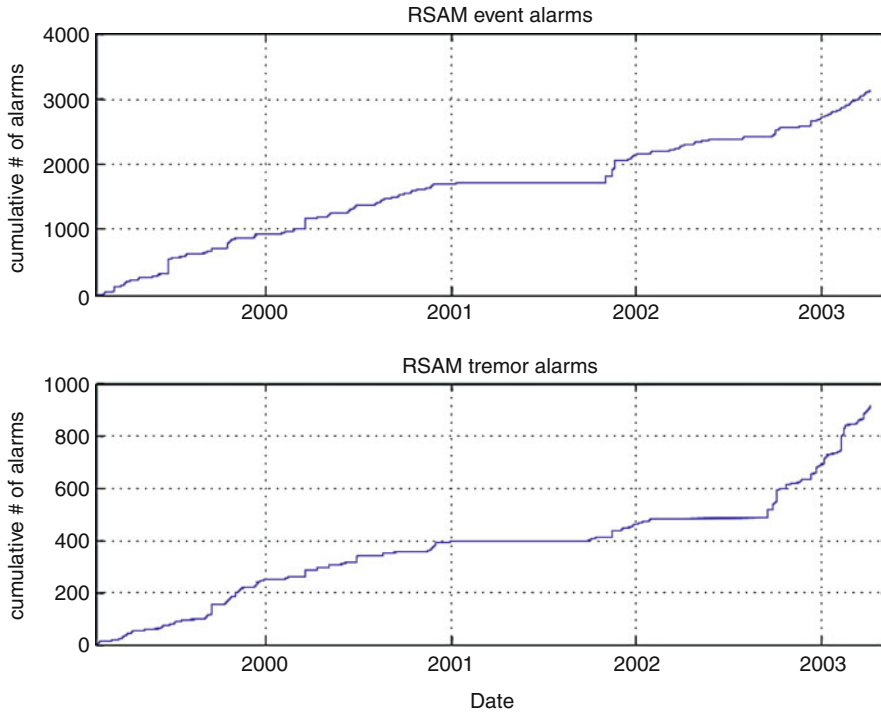
Seismic Monitoring of Volcanoes, Fig. 6 RSAM event alarms (*top panel*) and tremor alarms (*bottom panel*) issued at Montserrat Volcano Observatory between February 1999 and May 2003. In total there were about 4,000 alarms, an average of almost three alarms per day.

The alarm thresholds configured for each station were raised throughout 2000 and 2001 as activity increased. Event alarms typically corresponded to pyroclastic flows and regional earthquakes, and tremor alarms to dome collapses (i.e., a series of pyroclastic flows)

start and end of five swarms that occurred between February and April 2009, as well as significant escalations in the rate of earthquakes and energy release during those swarms (Thompson and West 2010).

Software

Earthworm (http://folkworm.ceri.memphis.edu/ew-doc/), is the most widely used system for earthquake and volcano-seismic monitoring today. It is an open-source, data acquisition, and earthquake detection framework developed by the US Geological Survey. It comprises of modules that communicate via messages on shared memory rings. It is used in the National Earthquake Information Center, tsunami warning centers, and regional seismic networks. Earthworm is in widespread use at volcano observatories

today, in conjunction with other tools that expand its real-time monitoring capabilities. Two reasons for this are that all US volcano observatories use them, and the USGS Volcano Disaster Assistance Program deployed them in many countries at the request of foreign governments.

Earthworm is not the only widely used seismic monitoring systems: Alternatives are SeisComP3 (https://www.seiscomp3.org/wiki/doc) and Antelope (http://www.brtt.com/software.html). Antelope is an excellent framework for developing new monitoring and research tools, but requires a commercial license except for US higher education institutions. SeisComP3 provides a rich set of graphical user interfaces for global and regional seismic monitoring and is particularly suited for tsunami early warning. However,
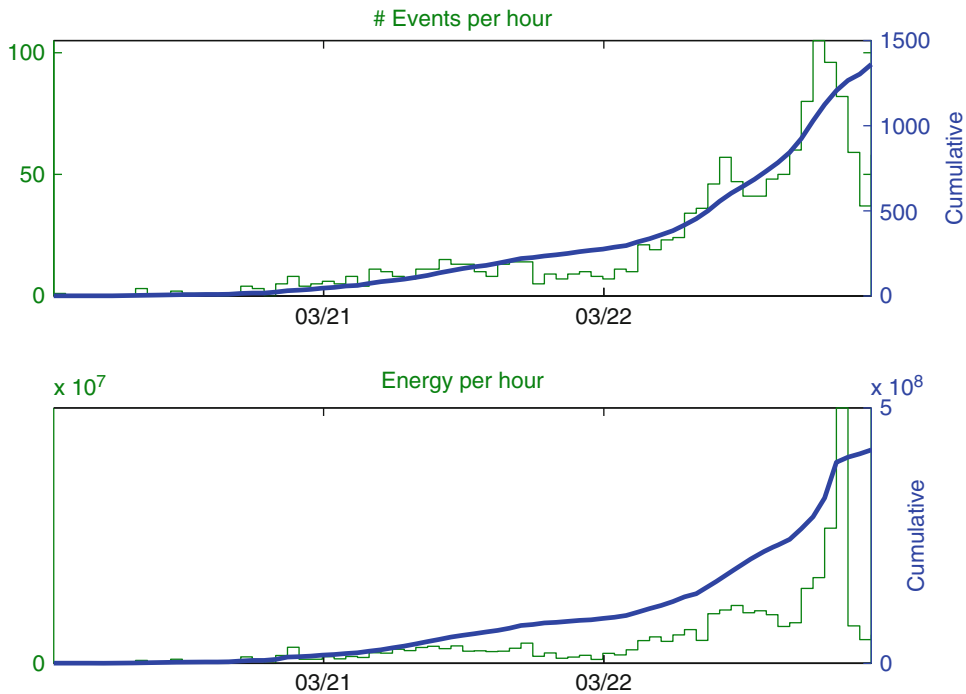
Seismic Monitoring of Volcanoes, Fig. 7 Two common ways of examining trends in detected events are (upper panel) event rate (also called 'counts'- the number of detected events per unit time) plots, and (lower panel) energy release rate plots (computed from magnitude data). The data shown are from Redoubt volcano, Alaska, from 2009/03/20 to 2009/03/23 and are produced using the GISMO toolbox. Energy units are arbitrary

neither currently has adaptations for volcano-seismic monitoring.

Earthworm, SeisComP3, and Antelope are capable of generating real-time earthquake catalogs within a few minutes of event occurrence, and they work well for regional earthquakes and volcano-tectonic earthquakes. However, other volcano-seismic signals lack clear P- and S-waves required to compute even an approximate location. Fortunately, Earthworm includes another mode of event detection called "subnet triggering," which determines only an approximate event time, and does not attempt to locate an event. This approach is employed by many volcano observatories. Custom code can be added to monitor these events and trigger other calculations. For example, at the Montserrat Volcano Observatory from 2000 to 2003, real-time measurements of the amplitude, energy, and frequency content of new events were made, and

used to regenerate plots on a private web site used by observatory scientists.

Earthworm has the ability to import data from a wide variety of field instrumentation and data servers and includes a variety of earthquake detection, location, and magnitude algorithms. There are modules which compute RSAM data and create daily helicorder and spectrogram plots. Tremor and swarm alarm systems are now included in Earthworm. Other modules can harvest messages about detected events, arrivals, locations, and magnitudes to create a real-time event catalog, serve up plots of event rates and Google Maps of epicenters, and serve data in QuakeML format.

SWARM can either draw data from a SeedLink server, an Earthworm wave server, or an FDSN (International Federation of Digital Seismograph Networks) web service. Another option is a Winston wave server, which emulates
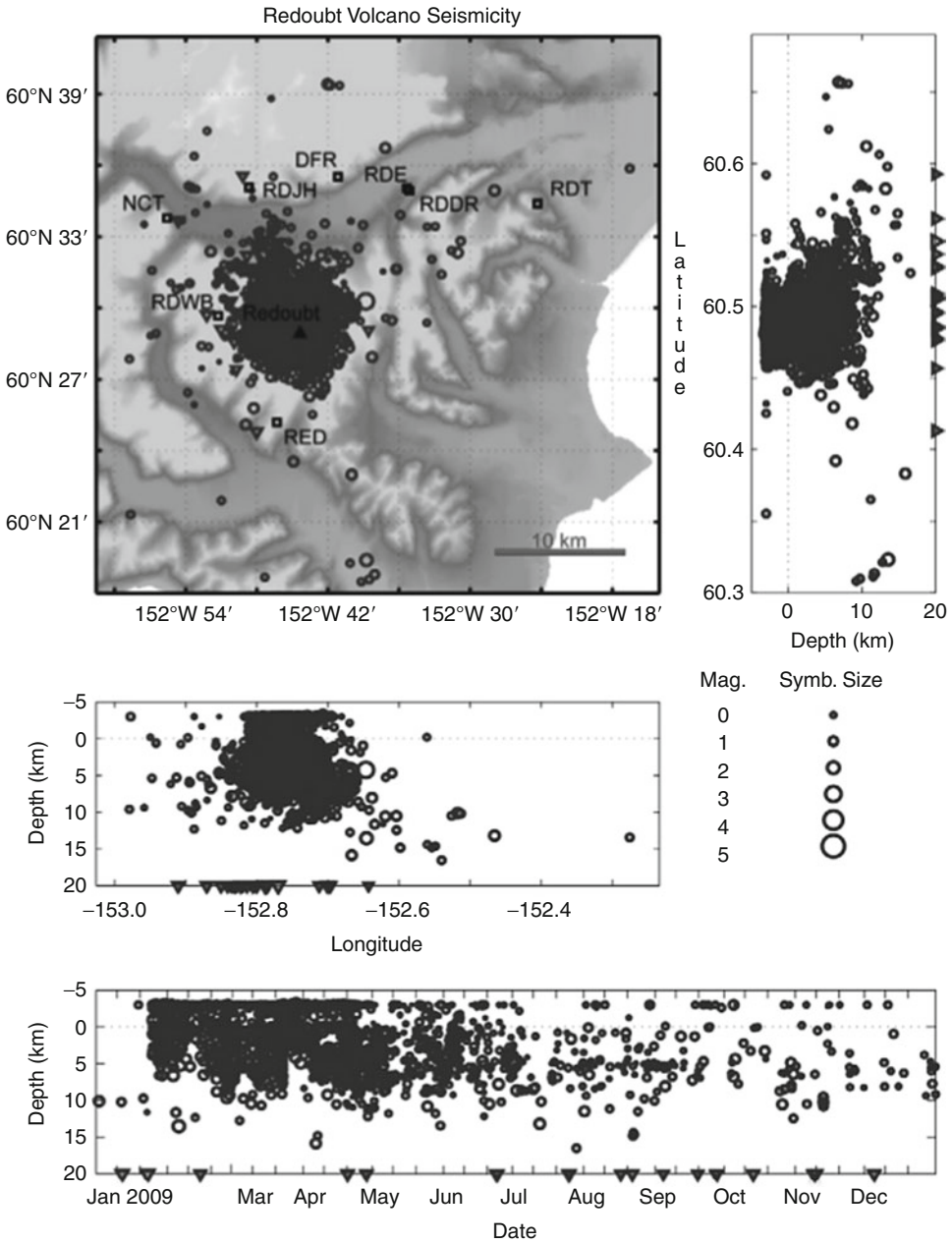
Seismic Monitoring of Volcanoes, Fig. 8 A common way of presenting 3-D hypocenter data is as a set of 2-D slices through the data. In this plot, generated by VolPlot, a program used for many years at the Alaska Volcano Observatory, earthquake hypocenters from Redoubt Volcano in 2009 are shown. The main panel shows epicenters in map view, and adjacent panels show depth versus latitude and depth versus longitude. These help scientists identify spatial relationships in hypocenter data. At the bottom, depth versus time is shown, which helps scientists recognize if earthquakes are getting closer to the surface. The size of the symbols indicates magnitude

an Earthworm wave server, but stores data in a MySQL database, providing rapid access and a deep data archive. Winston also provides a web interface that allows plotting of helicorder images and RSAM data (RSAM data are precomputed on import to Winston).

Data Analysis

Some analyses are not performed in real time, because it cannot be fully automated or just because the tools to do it well have not yet been developed; but they nevertheless provide timely information that can affect the way scientists interpret activity.

Analyst-Reviewed Event Catalogs
A real-time event catalog is valuable, but it cannot approach the quality of an analyst-reviewed catalog. A seismic analyst will periodically review the detected events, classify them, and delete false events. They might also pick P and S phases, run software to locate events, and compute magnitudes. Such a catalog is essential for research, but can be difficult to produce in a timely manner as volcanic unrest intensifies and decisions need to be made on timescales of minutes and hours, rather than days. During earthquake swarms (and major aftershock sequences), human analysts may be overwhelmed, increasing latency further just when it matters most.

While Antelope and SeisComP3 include an analyst review capability, Earthworm does not: So it is commonly paired with Seisan (http://seis. geus.net/software/seisan/seisan.pdf), which is an excellent free software package for processing and editing event catalogs. Seisan can estimate fault-plane solutions, moment tensors, and b-values too. Seisan includes some adaptations for volcano-seismic analysis: It supports event classes such as VT, LP, hybrid, and rockfall and can generate event-rate plots. Event waveforms are typically examined as spectra as well as time series. However, automated event classification is highly desirable because different real-time processing schemes could then be applied to

different event classes. For example, at the Montserrat Volcano Observatory from 2000 to 2003, rockfall signals were directed into an automated rockfall location system.

Automated techniques for classifying events have been tried at some observatories, but remain rare. Many researchers have had success with frequency-based analysis, artificial neural networks, and hidden Markov chains, but typically for only certain types of events recorded at a particular volcano on a particular station. Neither Earthworm, Antelope, or SeisComP3 currently include an auto-classification module.

Swarm Analysis
While event detection software captures many (perhaps most) volcanic earthquake signals, additional events can be found by using a match filter. This technique is particularly useful for tracking repeating, low-frequency earthquakes, since the emergent onsets of this type of event make them notoriously difficult to trigger using an STA/LTA detector. A known event waveform (or part of it) is cross-correlated against continuous data to find additional events in the same "family." Each family represents a source which is repeatedly activated in a volume that is small compared to the wavelength of the correlated signal. A number of volcanic processes are thought to be able to generate repeating earthquakes, including repeated resonances of fluid-filled cracks, a propagating crack tip that is driven by intruding magma, or even repeated shear failure within a body of magma. Additional analyses of the earthquake locations and the waveform frequency spectra are needed in order to distinguish between these various models. Double-difference relocation can be used to image the spatial evolution of the swarm more clearly.

A related technique is to cross-correlate waveforms from all events in a catalog against each other, to discover all event families. The MATLAB toolbox "GISMO" includes tools to automate this process. During unrest at Redoubt Volcano in 2009, this was used in near real time to monitor the evolution of swarms (Buurman et al. 2012) (Fig. 9).

S
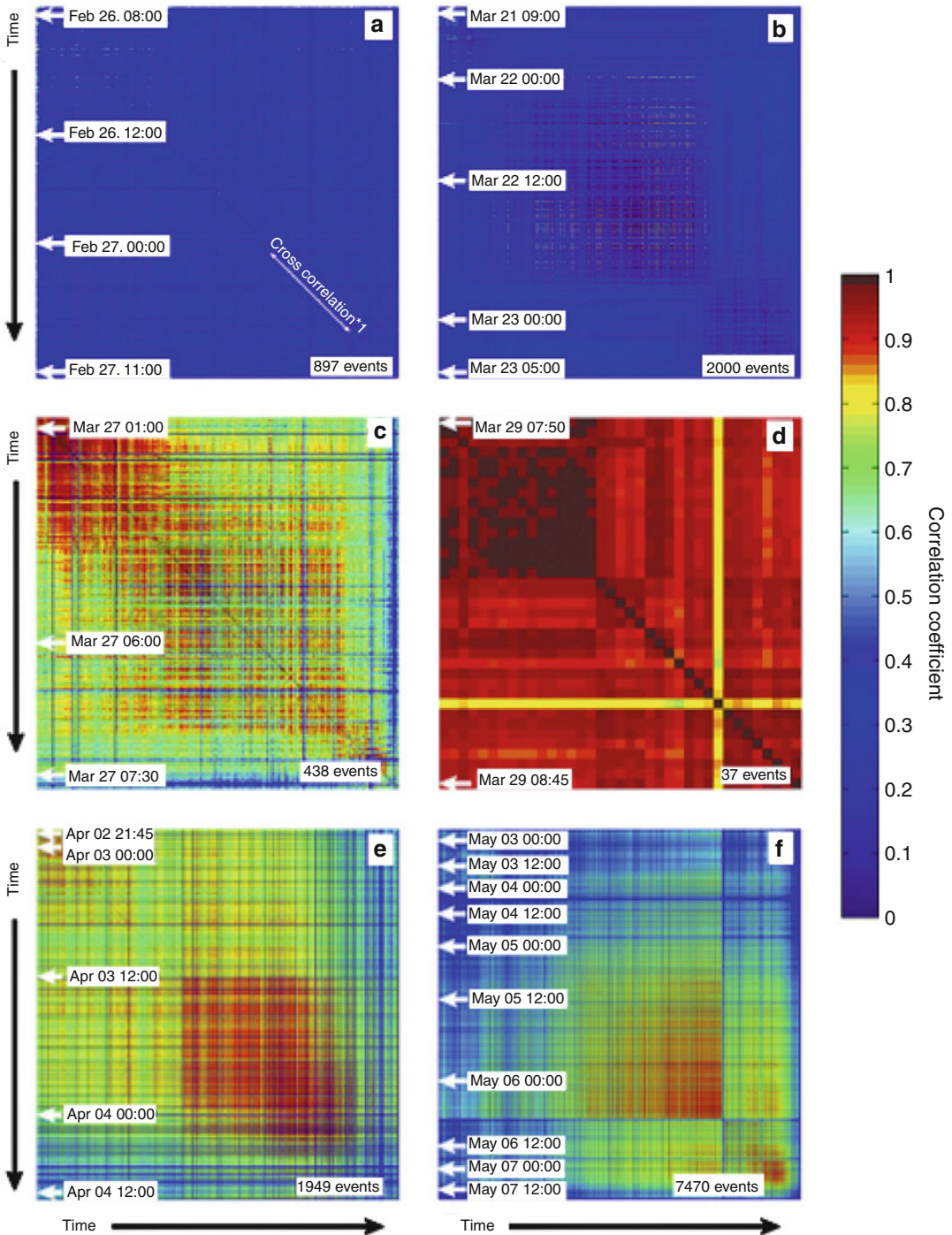
Seismic Monitoring of Volcanoes, Fig. 9 Cross-correlation plots for six swarms during the crisis at Redoubt Volcano in 2009 (reproduced from Buurman et al. (2012)). Each pixel represents of pair of waveforms that have been cross-correlated. Color represents the maximum cross-correlation coefficient. Time progresses from *top to bottom* and *left to right* in each panel. The *top-left* to *lower-right* diagonal represents the autocorrelation of each waveform. The swarms began on: (**a**) February 26, (**b**) March 21, (**c**) March 27, (**d**) March 29, (**e**) April 2, and (**f**) May 2

Locating Phaseless Seismic Signals

The tremor alarm systems mentioned previously are very simple and they lack the ability to locate tremor. Some observatories have exploited seismic data to track the location of tremor and map debris-flow trajectories, which could help mitigate hazards.

Phaseless signals, once identified (via an STA/LTA algorithm or visual inspection), can be approximately located by exploiting the manner in which seismic wave amplitudes decay with increasing distance from the source, due to geometrical spreading and attenuation (Jolly et al. 2002). Alternatively, they can be located by cross-correlating data to determine travel-time differences and then using traditional differential travel-time techniques to locate the source (assuming the wave type is known). The Waveform Envelope Clustering and Correlation (WECC) system (Wech and Creager 2008) reverses this approach. It continuously locates envelopes of continuous waveform data, and if the location errors are small, it declares an event. Designed to locate the episodic tremor associated with the slow slip, it is now being applied to volcanoes.

Detecting Lahars

Lahar signals may appear as high-frequency tremor signals on a volcano-seismic network, but may be difficult to detect and locate without designing a network for lahar monitoring. High-frequency surface waves attenuate rapidly with increasing distance, so the amplitude of high-frequency signals can be indicative of how close a flow is to a seismometer. More dilute flows are also richer in higher frequencies. An existing volcano-seismic network can be adapted for lahar monitoring by adding stations close to the flow channel at various points, enabling lahar signals to be more easily resolved and crudely located using the times of the peak amplitudes on each station, and also the amplitude distribution across the seismic network. Trip wires can be positioned across the valley at different points and at different heights, and the arrival time and location are known precisely when the wire is cut. Disadvantages are that trip wires only work once and a dilute flow may leave the wire intact. Sophisticated lahar-monitoring systems have been installed at many volcanoes including Rainier, Merapi, and Nevado del Ruiz.

Imaging Spatial Changes in Volcanic Seismicity

There are many techniques for imaging magma storage regions. A volume devoid of volcano-tectonic earthquakes may indicate a large magma storage body, incapable of shear failure. Seismic tomography can reveal volumes with low P- or S-wave propagation speeds. B-value mapping (see Sanchez et al., "▶ Frequency-Magnitude Distribution of Seismicity in Volcanic Regions," this volume) can reveal volumes incapable of sustaining large failures, which are often inferred as being associated with magma. In the region around Uturuncu Volcano (Bolivia), the presence of a sill at around 15–20-km depth has been inferred by an S-wave shadow zone. Roman et al. (2006) found that changes in the orientation of fault-plane solutions could be used (retrospectively) to differentiate between volcano-tectonic earthquakes related to the injection of new magma and posteruptive relaxation.

Ambient noise tomography is another technique, which can be used to detect velocity changes. In active source seismology, an impulsive source is generated at one location and recorded at another. The resulting seismogram is the Green's function of the path between the sites. If this were done repeatedly, seismic velocity changes along the travel path could be detected. At a volcano, a drop in velocity could be indicative of increased heat flow, faulting, fluid injection, or expansion of the volcanic edifice. MSNoise (http://www.msnoise.org/) is a package which has been developed to automate the computation of Green's functions between all station pairs on a daily basis, allowing velocity changes as small as 0.1 % to be identified. At Piton de la Fournaise Volcano, the locations of dike injections and eruptive fissures have been inferred.

S

Detecting Explosions

Cleveland Volcano (Alaska) has been erupting frequently in recent years, generating ash clouds that threaten aviation travelling between North America and Asia, and, without a functioning seismic station within a few tens of kilometers, many of these eruptions were previously missed or only recognized later in satellite images. De Angelis et al. (2012) describe a system which exploits data from a seismic network on Okmok Volcano, operated by the Alaska Volcano Observatory, to detect explosions at Cleveland Volcano (about 120 km away). An STA/LTA detector creates candidate arrivals. For those that fall within a 2-min sliding window, the differential travel times are computed by cross-correlating envelopes of the seismograms and then inverted for apparent slowness. If this is consistent with a sound speed of 340 +/−30 m/s in the direction of Cleveland, an alarm is sent by email. An alternative to the Cleveland explosion alarm mentioned above looks for a coherent signal across an infrasound array instead. The advantage is that infrasound sensors record explosion signals less ambiguously, and the technique could be used to detect explosions in other regions where installation of a local seismic network is unrealistic.

Technical Challenges

One of the major challenges for observatories and seismic networks is monitoring the state of health of a seismic network. There is only one chance to gather data, and in a crisis, data not collected in real-time have little value. Seismic and repeater stations need to be engineered to withstand weather variations which may include heavy rain, snow and ice accumulation, strong winds, and months of little solar energy. Spare equipment needs to be available so that stations can be fixed quickly when a seismic station component breaks. Vandalization by humans, damage by animals, and theft are other considerations. During unrest, ash build-up on solar panels can quickly cause power loss at a seismic station.

Problems will usually be apparent with tools volcano seismologists use every day. RSAM plots may reveal diurnal cycles in the data which may indicate undercharging of batteries or anthropogenic noise. Spectrograms and seismograms may show a flat signal (power loss), a one-sided signal (stuck seismometer), large offset (seismometer not level), white noise (loss of seismic signal), or spikes and dropouts (interference). These can cause havoc with automated detection and alarm systems, leading to corrupted event catalogs and false alarms, so one strategy is to identify and eliminate affected stations from automated processing schemes. SeisNetWatch (http://www.isti.com/products/seisnetwatch/) is used by many seismic network operators to monitor all ip-addressible nodes in a network, including field digitizers, to monitor data latency and other state-of-health parameters.

The volume of data flowing into an observatory can be overwhelming. A single channel of 24-bit, three-component data recorded at 100 Hz requires 620 MB of storage per day (uncompressed). Twenty channels of data, a typical amount for a small volcano observatory, are about 1.5 TB per year. Although this would fit on a single hard drive today, it is about 100 times the capacity of a hard drive available 20 years ago (the storage capacity of hard drives increases by a factor of about 10 every decade). Continuous digital volcano-seismic data before the mid 1990s are rare but, remarkably, the Pacific Northwest Seismic Network recently recovered continuous data saved to tape during the 1980 eruption of St. Helens Volcano.

To improve the reliability of data acquisition and processing systems, several measures can be taken. To protect against power failures, all mission-critical computers can be connected to uninterruptible power supplies and configured to automatically reboot and restart critical processes once power is restored. No one wants to arrive at an observatory on a Monday morning and finds that the alarm computer suffered a power outage on Friday night and has not been running since, especially if a significant event occurred. Another measure is to run all mission-critical computers in parallel, providing automatic failover or at

least have pre-configured spare computers ready to plug in whenever a primary computer failed. All potential points of failure can be monitored by a diagnostic alarm system. Finally, a daily checklist helps catch any other problems. All of these measures were taken at the Montserrat Volcano Observatory in 2000, leading to an improvement in the data capture rate and public safety.

Once captured, data need to be managed effectively to preserve them for further analysis and research. Given the capacity of modern hard drives, it is now possible to keep many years of data "online." Copying data to an off-site location is also becoming more common. Many observatories now transmit real-time data to the IRIS Data Management Center, which not only serves as a backup, but outsources the resources needed to disseminate the data more easily to the research community.

an expensive field visit or lengthy data outage. Free, community supported software is available for data acquisition, event detection and location (e.g. Earthworm), real-time data visualization (e.g. SWARM), event processing and catalog production (e.g. Seisan), instead of each observatory reinventing the wheel. Archiving data - including continuous waveform data, derived data such as RSAM, event catalogs and station metadata – is a crucial task that has become much easier thanks to the expanding capacity of hard drives, faster internet, and data management centers such as the IRIS DMC.

Real-time monitoring may rely on an operations room which is manned 24 h a day. Or it may rely on automated alarm systems that alert observatory staff to escalations in seismicity (and data outages) outside of normal office hours. Research and development is often driven by the need for better real-time monitoring tools – automated classification of volcano-seismic signals, for example. Ambient noise tomography and detection of multiplets are examples of techniques that might become routine at volcano observatories as computational power continues to increase. As automated monitoring becomes increasingly sophisticated, observatory seismologists will need to spend less time on troubleshooting and development and more time analyzing and interpreting volcanic-seismicity, and collaborating on research.

Summary

Volcano-seismologists forecast volcanic activity by analyzing the rates, energy release and spatial distribution of different types of characteristic seismic signals recorded in the vicinity of volcanoes. The most commonly identified signals are VT, LP or hybrid earthquakes. Explosive and effusive eruptions can cause rockfalls and pyroclastic flows, which also generate characteristic seismic signals. Volcanic earthquakes frequently occur in swarms. Swarms and tremor are both common precursors to escalations in volcanic activity. While there is still much that is not understood – the origin of LP earthquakes for example - no other technique provides such detailed information about the internal state of a volcano, or about debris flows as they occur, or provides such a detailed chronology of an eruption.

Volcano-seismic monitoring provides high-sample-rate data, 24 hours a day. Short-period seismometers and analog telemetry are gradually being phased out and replaced with modern broadband seismometers and field digitizers. Two-way telemetry allows observatory staff to diagnose and fix problems sometimes without

Cross-References

S

References

Barclay J, Johnstone J, Matthews A (2006) Meteorological monitoring of an active volcano: implications for eruption prediction. J Volcanol Geotherm Res 150(4):339–358. doi:10.1016/j.jvolgeores.2005.07.020

Benoit JP, McNutt SR (1996) Global volcanic earthquake swarm database and prelimary analysis of volcanic earthquake swarm duration. Annali Di Geofisica 39(2):221–229

Buurman H, West ME, Thompson G (2012) The seismicity of the 2009 redoubt eruption. J Volcanol Geotherm Res. doi:10.1016/j.jvolgeores.2012.04.024

Chouet B (2003) Volcano seismology. Pure Appl Geophys 160(3):739–788. doi:10.1007/PL00012556

Chouet BA, Matoza RS (2013) A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. J Volcanol Geotherm Res 252:108–175. doi:10.1016/j.jvolgeores.2012.11.013

De Angelis S, Fee D, Haney MM, Schneider DJ (2012) Detecting hidden volcanic explosions from Mt. Cleveland Volcano, Alaska with infrasound and ground-coupled airwaves. Geophys Res Lett 39:1–6. doi:10.1029/2012GL053635

Endo ET, Murray TL (1991) Volcanology a volcano monitoring and prediction tool. Bull Volcanol 53:533–545

Jolly AD, Thompson G, Norton G (2002) Locating pyroclastic flows on Soufriere Hills Volcano, Montserrat, West Indies, using amplitude signals from high dynamic range instruments. J Volcanol Geotherm Res 118(3–4):299–317. doi:10.1016/S0377-0273(02)00299-8

Lahr JC, Chouet BA, Stephens CD, Power JA, Page RA (1994) Earthquake classification, location, and error analysis in a volcanic environment: implications for the magmatic system of the 1989–1990 eruptions at redoubt volcano, Alaska. J Volcanol Geotherm Res 62(1–4):137–151. doi:10.1016/0377-0273(94)90031-0

Luckett R, Baptie B, Ottemoller L, Thompson G (2007) Seismic monitoring of the Soufriere Hills Volcano, Montserrat. Seismol Res Lett 78(2):192–200. doi:10.1785/gssrl.78.2.192

McNutt SR (1996) Seismic monitoring of volcanoes: A review of the state-of-the-art and recent trends. In: Scarpa R, Tilling R (eds) Monitoring and mitigation of volcano hazards, Chapter 3, Springer-Verlag, Berlin, pp 99–146

McNutt SR (2000) Volcanic seismicity. In: Houghton HB, McNutt SR, Rymer H, Stix J (eds) Chapter 63 of encyclopedia of volcanoes. Academic Press, San Diego CA, 1015-1033

McNutt SR (2002) Volcano seismology. In: Lee WHK, Kanamori H, Jennings PC (eds), Chapter 25 of international handbook of earthquake and engineering seismology, IASPEI, Palo Alto, CA, 81A: 383–406

McNutt SR (2005) Volcanic seismology. Annu Rev Earth Planet Sci 33(1):461–491. doi:10.1146/annurev.earth.33.092203.122459

Moran SC, Freymueller JT, LaHusen RG, McGee KA, Poland MP, Power JA, White RA (2008) Instrumentation recommendations for volcano monitoring at U.S. Volcanoes under the national volcano early warning system: scientific investigations report 2008 – 5114, p 47. Retrieved from http://pubs.usgs.gov/sir/2008/5114/

Powell TW, Neuberg JW (2003) Time dependent features in tremor spectra. J Volcanol Geotherm Res 128(1–3):177–185. doi:10.1016/S0377-0273(03)00253-1

Roman DC, Neuberg J, Luckett RR (2006) Assessing the likelihood of volcanic eruption through analysis of volcanotectonic earthquake fault–plane solutions. Earth Planet Sci Lett 248(1–2):244–252. doi:10.1016/j.epsl.2006.05.029

Thompson G, McNutt SR, Tytgat G (2002) Three distinct regimes of volcanic tremor associated with the eruption of Shishaldin Volcano, Alaska 1999. Bull Volcanol 64(8):535–547. doi:10.1007/s00445-002-0228-z

Thompson G, West ME (2010) Real-time detection of earthquake swarms at redoubt Volcano, 2009. Seismol Res Lett 81(3):505–513. doi:10.1785/gssrl.81.3.505

Wassermann J (2012) Volcano seismology. In: Peter B (ed) IASPEI new manual of seismological observatory practice 2 (NMSOP-2), second. Potsdam : Deutsches GeoForschungsZentrum GFZ, Potsdam, pp 1–77. doi:10.2312/GFZ.NMSOP-2_ch13

Wech AG, Creager KC (2008) Automated detection and location of Cascadia tremor. Geophys Res Lett 35(20), L20302. doi:10.1029/2008GL035458

Seismic Network and Data Quality

Peter H. Voss[1] and Lars Ottemöller[2]
[1]Geological Survey of Denmark and Greenland – GEUS, Copenhagen K, Denmark
[2]Department of Earth Science, University of Bergen, Bergen, Norway

Synonyms

Earthquake monitoring; Seismometer sensitivity; Instrument response; Ambient noise; Data latency

Introduction

Measuring of ground motion provides the most essential observations made from earthquakes and the basis for research within seismology and earthquake engineering. The measurements are done by seismic stations that are equipped with sophisticated instruments and are placed at specific selected locations. The good quality of these data is obviously important but can only be achieved if many different factors are considered. Often, the approach is adjusted to the scientific or monitoring needs and the available funding. This section deals with systems of seismic stations, so-called seismic networks and data quality-related issues.

Seismic Networks

Seismic networks are systems of individual seismic stations with the purpose to monitor ground motion. The size and configurations of seismic networks depend on the focus area and type of signals that are to be recorded. The data provided by seismic networks (seismograms; see Fig. 3) can be used for many different applications from

earthquake monitoring services to research of the Earth interior. Examples of seismic networks are global seismograph networks (Fig. 1) that are operated to primarily detect large earthquakes and other events globally, regional or national seismic networks to detect events within a country, microseismic arrays to detect induced earthquakes at a water reservoir or hydrocarbon field, and dense monitoring at volcanoes to measure the ongoing activity. Individual seismic stations are commonly part of several networks as sharing of data is common practice. Seismic networks can be either permanent or temporary (e.g., http://www.usarray.org/), i.e., deployed for a shorter limited time period often linked to very specific research projects. Normally, less effort is put into the installation of temporary networks.

Seismic Stations

The main elements of a seismic station are the vault and housing of the equipment, the power supply to the equipment, the sensor, the digitizer with a timing device and recorder, and communication. Figure 2 shows how these units can be connected at a seismic station, and Fig. 5 shows
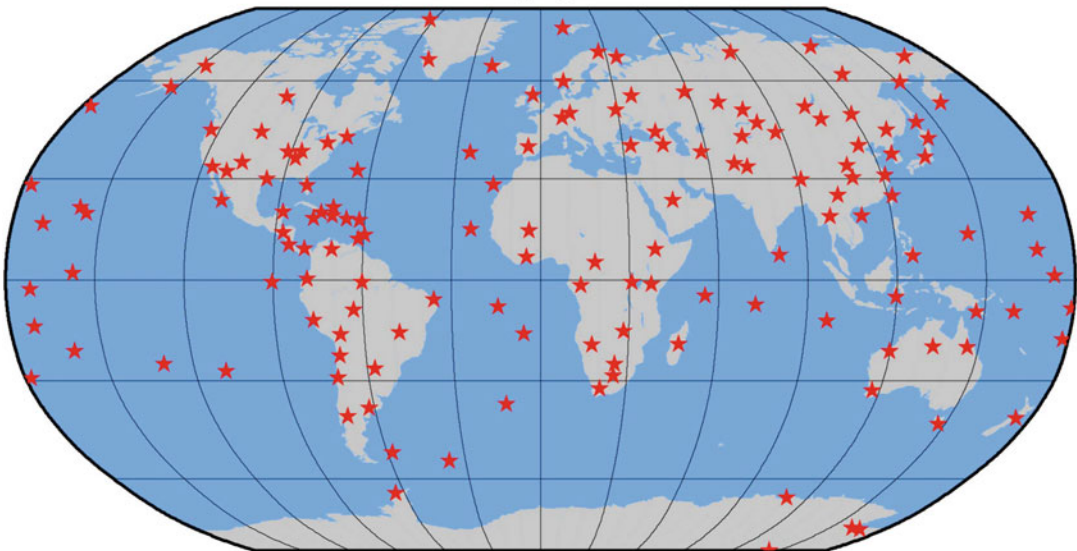
Seismic Network and Data Quality, Fig. 1 Map of the Global Seismograph Network (GSN)
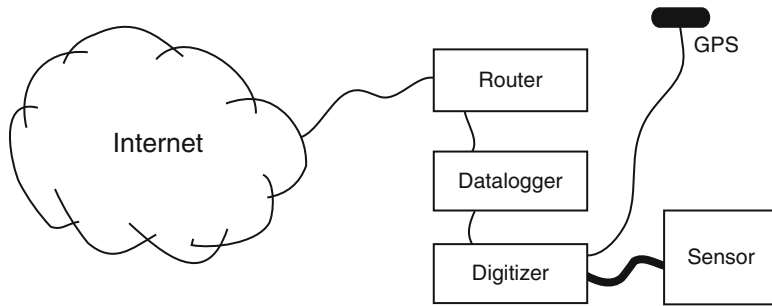
Seismic Network and Data Quality, Fig. 2 Equipment used at a typical seismic station. The analog signal from the sensor is converted to a computer-readable format in the digitizer that also time stamp the data with a GPS; a datalogger is storing and handling the transmission the data. A router and Internet connection are standard equipment for the data transmission. Some units will have integrated datalogger and digitizer or integrated digitizer and sensor. The units will naturally also need power supply

an example of a seismic station. All these elements will have an effect on the data quality and have to be considered carefully. When building a seismic station, there are normally financial constraints, and one has to find the optimum solution also considering the main monitoring purpose. The main choice for the sensor is between seismometer, also referred to as weak motion sensor, and accelerometer, also called strong motion sensor. The difference is that seismometers are more sensitive, particularly at lower frequencies, and the accelerometer will have a higher clip level than the seismometer. At some seismic stations, both a broadband seismometer and an accelerometer are installed to improve the dynamic range of the seismic station.

The main parameters describing the seismometer are the sensitivity, the natural frequency, and the damping. The response of the seismometer is given by its transfer function (impulse response) that has an amplitude and phase-delay part. Seismometers are available with different natural periods. Traditionally, long-period and short-period sensors were used, where the maximum gain would have been at 20 s and 1 s, respectively. Today, broadband sensors are the most common and typically have the high sensitivity between 120 s and 50 Hz. Seismometers are able to record the lowest possible expected ground noise levels within a given frequency band (limited by the instrument self-noise), and this is the most important technical specification of the sensor. Accelerometers remain on-scale even for the largest earthquakes and are the most common instrument for earthquake engineering applications.

The main considerations when choosing a digitizer are the dynamic range and the timing accuracy. An important setting on the digitizer is the sampling rate, which according to the Nyquist theorem allows to resolve frequencies less than half the sampling frequency. A normal sampling rate at a local seismic network is 100 Hz. Information of earthquake signals above 50 Hz is therefore not recorded, due to the Nyquist theorem. The sampling rate is thus set according to the expected signal frequencies but also restricted by limitations in data storage and communication. Communication to a seismic station can be achieved in many different ways (e.g., cell phone, satellite, the Internet, digital radios) and is an important element as this allows data to be available close to real time. Most seismic stations are able to send data in this way to one or several central recording systems, and this is what today defines a seismic network.

Before one selects instrumentation for a seismic station, it should be considered if the instrument will be able to properly record the earthquakes it is meant for. A way to verify these requirements is to display the instrument sensitivity and clip levels in frequency–amplitude diagrams, with the expected response of likely earthquakes. Figure 3
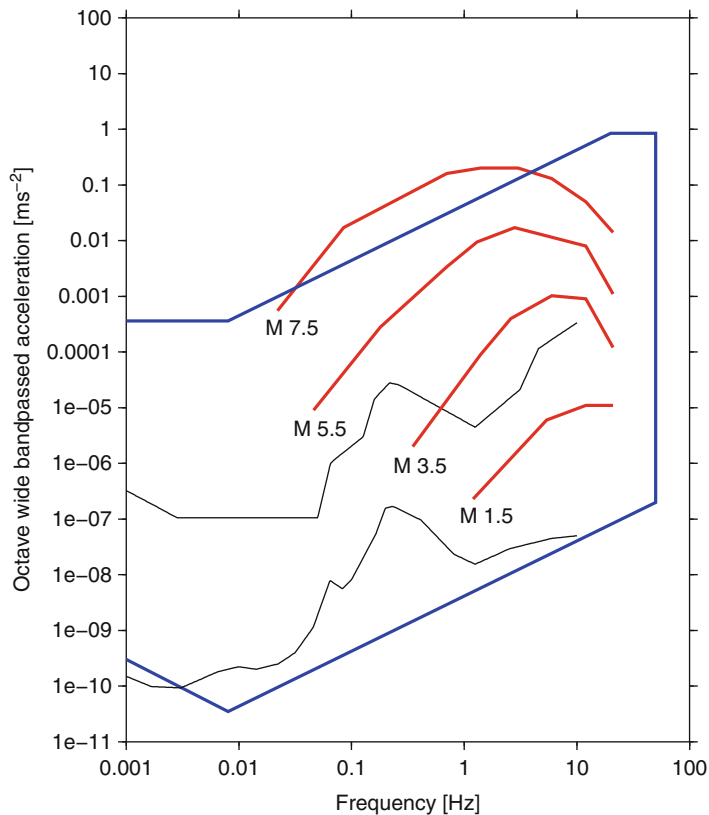
Seismic Network and Data Quality, Fig. 3 Frequency–amplitude plot for octave-wide band passes of ground-motion acceleration. The blue polygon represents the sensitivity limits of an STS-2 broadband seismometer, where the frequency limit at 50 Hz is given by the Nyquist frequency for a sample rate of 100 Hz. The lower acceleration amplitude limit in the band 0.001–50 Hz is the minimum sensitivity of the sensor, and the upper acceleration amplitude limit in the band 0.001–50 Hz is the clip level for the sensor. Sensor limits are scaled down to account for the band passing of the data. The red lines show the possible acceleration levels from different magnitude regional earthquakes that can be recorded at a distance of approximately 100 km. The two *black lines* show the USGS high- and low-noise models (Peterson 1993) (See also Clinton and Heaton 2002)

shows such a diagram for an STS-2 seismometer with regional earthquakes of magnitudes 1.5, 3.5, 5.5, and 7.5. Note that a magnitude 7.5 is above the clip level in this case, so another type of sensor should be considered for earthquakes of this strength. Figure 4 shows an example of a clipped seismic signal from a regional earthquake.

Data Exchange

The exchange of seismic data has been practiced since the installation of the very first stations as little could be done based on single stations. Prior to the exchange of digital data through the Internet, seismologists would mail the seismograms on paper or microfilm to each other. Today, it is still common practice (with some exceptions) to exchange and share data between the now enormous number of seismic stations. This sharing of data is essential for research projects as well as for seismic monitoring services as earthquakes often occur in between different networks and only a combination of the networks allows good determination of even the basic earthquake parameters. Data (including waveform data and derived data) can be exchanged in many different ways from email to data downloads from a server and receiving data in near real time. There are a
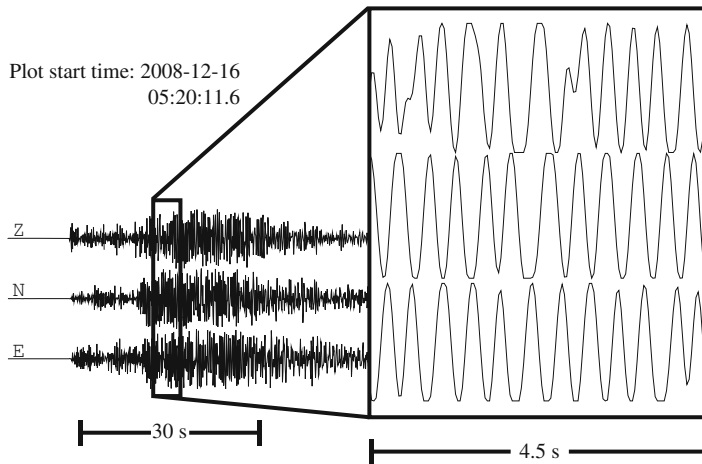
Seismic Network and Data Quality, Fig. 4 Seismogram showing a magnitude 4.2 earthquake recorded at a distance of 87 km on the GID station in Denmark. The channels represent individual sensors oriented in the vertical, north–south, and east–west directions, given by the Z, N, and E, respectively. The signal is clipped since the signal exceeded the digitizer clip level. The instrument is the short-period 4.5 Hz GeoSIG GBV-315 (Data are from the Geological Survey of Denmark and Greenland)

number of different public domain proprietary protocols for the near real-time exchange of data between the stations and data centers and also for the exchange between data centers. Most of these allow for data exchange in near real time and the recovery of missing data that may occur due to transmission errors. The public domain SeedLink protocol is one of the most used for real-time data transmission, especially in between data centers (Hanka et al. 2000). The SeedLink protocol transmits data in 512 byte packages, where station, channel, location, or network codes can be requested. The SeedLink protocol ensures that data are not lost in case of disconnected transmissions for shorter periods. This is also common in most commercial software solutions for data transmission in seismic networks. The largest data center receiving and providing seismic waveform data globally in a number of ways is the Incorporated Research Institutions for Seismology (IRIS) Data Management System (DMS) (see http://www.iris.edu/hq/programs/dms). It is also becoming possible to operate distributed data centers as, for example, done in Europe through the ORFEUS European Integrated Data Archive (EIDA) initiative (http://www.orfeus-eu.org/eida/eida.html). Exchange of parametric data, the data that contain information

on earthquake (e.g., phase travel times and amplitudes), is often carried out between neighboring networks. On a global scale, the main center for the parametric data is the International Seismological Center (ISC) (see http://www.isc.ac.uk).

Data Processing

The evaluation or processing of the data recorded by the seismic network normally takes place at the central recording site. Depending on the purpose of a seismic network, different automatic processing routines are applied to the data when it comes in. For most networks, the main focus is on detecting seismic events. This is done by checking the seismic signal and detecting changes from the background noise in the signal at individual stations. The information from a number of stations is combined to detect, locate, and estimate magnitude of the seismic events using the entire network. This is done for networks with different dimensions and monitoring purpose requiring different approaches. The outcome from the automatic processing is normally inspected before dissemination to the public in so-called earthquake early warning systems. This type of warning or information dissemination is

Seismic Network and Data Quality, Fig. 5 Example of a seismic station installed in southern Norway. The vault housing (with the seismometer and digitizer) is constructed from prefabricated concrete rings. (*left*).

Communication and power are available from a nearby hut (*right*). (Photo taken by Ole Meyer, University of Bergen)

an important responsibility of seismic monitoring systems. At a later stage, the data are further processed through manual interactive data processing. For both the automatic and manual data processing, there are a number of public domains as well as commercial software packages. Figure 6 shows phase readings done with the SEISAN software package (http://seisan.info/).

Metadata

The meta- or descriptive data are an essential part of the seismic data as without these the data is not complete. Data are identified by a so-called SCNL code (giving stations, component, network, and location) and the timing of the data. The station code represents the locality of the seismic station, the component code reflects the instrumentation, the seismic network code identifies the operator of the station, and the location code allows distinguishing between two instruments of the same type at one station. The metadata associated to the data streams will include information such as the location of the station (latitude, longitude, elevation), the orientation of the sensors (azimuth and dip), the sample rate, and the instrument response.

The main standard for seismic data is the Standard for the Exchange of Earthquake Data (SEED). One part of SEED refers to the waveform data, known as miniSEED; the other part of SEED, known as the dataless SEED, contains the metadata. The network codes for SEED are organized by the International Federation of Digital Seismograph Networks (FDSN, http://www.fdsn.org/). In addition to each component of the seismic data, the complete instrument description makes the data complete. This information makes it possible to convert the output recorded at the seismic station into true ground motion. Incorrect metadata is often the source to failed analysis of earthquake records.

Virtual Seismic Networks

Virtual seismic networks or virtual nets are the type of network where data from seismic stations in different seismic networks are collected and processed as a new "virtual" seismic network. The virtual seismic networks are often seen in regions in between two or more seismic networks where a joint processing of data from the region will improve the detection capabilities. An example of such virtual seismic network is the Greenland Ice Sheet Monitoring Network (http://www.glisn.info/), where data from nine different seismic networks are collected and processed jointly to improve the detection capability of glacial

S

Seismic Network and Data Quality,
Fig. 6 Seismogram displaying an earthquake recorded at a distance of about 50 km. The 9.5 s of data shows the three channels of the seismic station BLS5 in Norway. The channels represent individual sensors oriented in the vertical, north–south, and east–west directions, given by the Z, N, and E, respectively. The P and the S phase are seen at 19:22:36.5 and 19:22:42.5, respectively (The data is from the National Norwegian Seismic Network. Data are unfiltered)
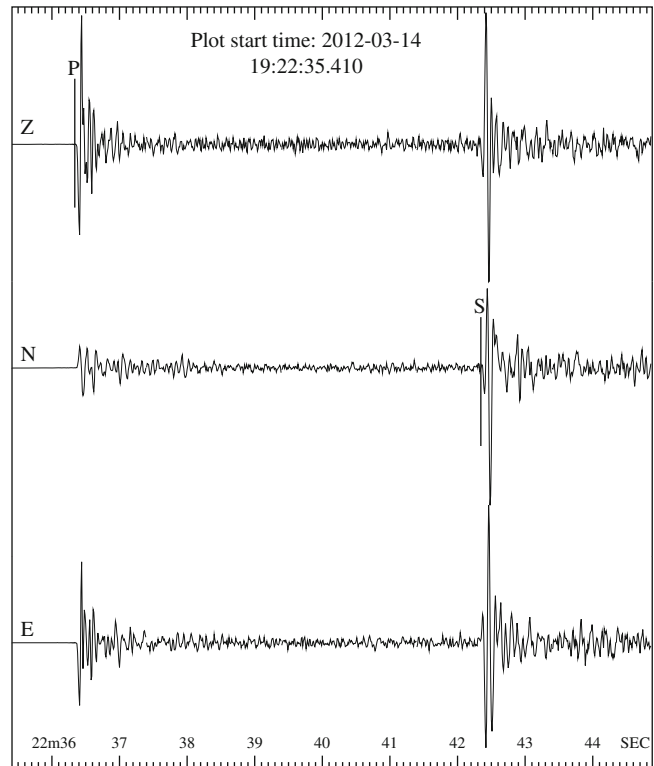
earthquakes in Greenland. Other virtual seismic networks collect data after quality or instrument type; see, e.g., the _REALTIME and _STS-1 virtual networks operated by IRIS (http://www.iris.edu/vnets). A precondition for such networks to operate is the use of common standards for real-time data transmission as described above.

Seismic Noise and Data Quality

With seismic waveform data, a number of parameters can be used to quantify data quality. The main parameters are ambient noise levels, timing accuracy, and data completeness. The noise levels at a seismic station depend on a number of factors including the true ambient ground noise, the vault construction, and the instrumentation. Increased noise levels at a seismic station result in a lower signal to noise ratio, which means that a seismic station is worse in detecting and observing seismic events.

When talking about "data quality," the question often is whether it is sufficient or insufficient. This depends on the application of the data. For example, a seismic station may perform well for local earthquake recording, but not provide sufficient data for global earthquake monitoring, if the used instruments are designed only to measure seismic signals above 5 Hz or if the ambient noise level is very high at the lower frequencies. If data quality is sufficient, it means that the collected data provides the information needed to answer the questions asked prior to the collection of the data. On the other hand, if data quality is insufficient, it means that the collected data cannot answer the questions at hand. Data quality is therefore a term used to describe if data is adequate for its intended use. The border between sufficient and insufficient data quality is not sharp since data are often collected to answer multiple questions or since the data quality might change over time during the collection of the data.

Improving data quality is first of all achieved by reducing the noise in the data, both for seismic
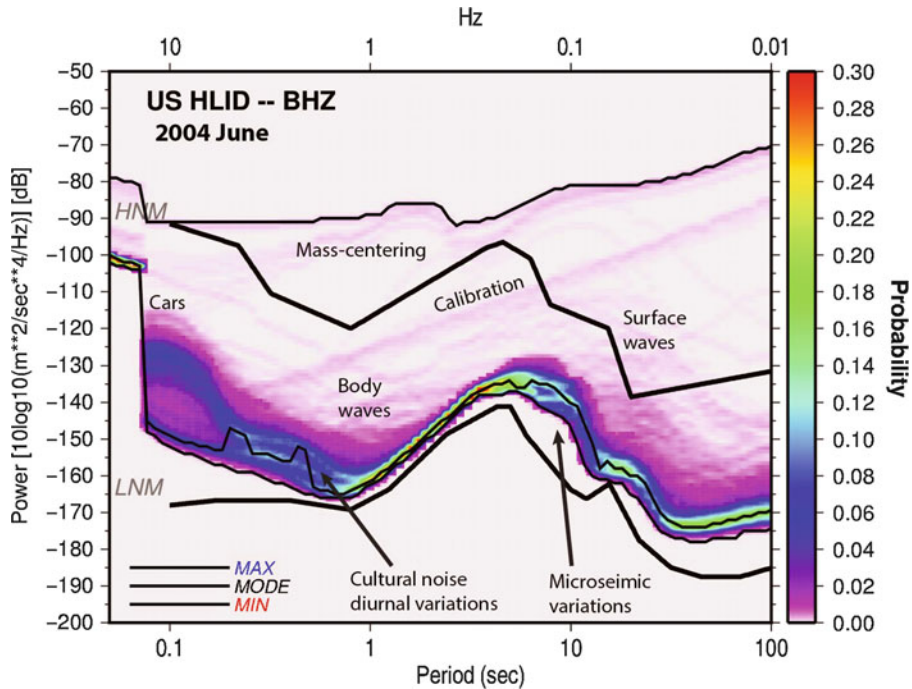
Seismic Network and Data Quality, Fig. 7 Ambient noise probability density functions (PDF) of data from the HLID station show how cars, calibrations, and other sources are clearly distinguished from the natural ambient noise (McNamara and Boaz 2005). Auto mass re-center and calibration signals are due to the station operator re-leveling the sensor and checking the amplitudes, respectively. Body wave and surface wave signals show the total power of all the earthquake records in the period of the PDF. LNM and HNM are the low- and high-noise models, respectively. Note the peak in the noise at about 5 s period; this is due to the ocean-generated noise (See also http://geohazards.cr.usgs.gov/staffweb/mcnamara/PDFweb/Noise_PDFs.html)

stations designed to record local or distant earthquakes. The ambient or microseismic noise has both natural and man-made sources. Typical man-made sources are roads, machinery, airports, railways, etc., and within some constraints it is often possible to stay away from these. The main natural source of seismic noise is ocean waves, and while stations can be deployed at some distance from the coastline, it will not be possible to get away from this noise completely. Wind and atmospheric changes also result in increased noise, and this can mostly be avoided by building better seismic vaults, by putting sensors in pressure-tight enclosures, and also by increasing the depth of installation. The seismic instrumentation, both seismometer and digitizer, also has self-noise, and this can be minimized when choosing the equipment but also during the installation process. The instruments should always be well grounded to avoid static electricity. Also ambient variations in the magnetic field should be avoided since it can limit the quality (Forbriger et al. 2010).

The most common way to present the microseismic noise at a station is to compute acceleration power density spectra and to display the probability of noise levels over a longer time period on probability density plots (McNamara and Buland 2004; McNamara and Boaz 2005). An example is shown in Fig. 7. This type of plot allows the identification of the different noise sources that act at different frequencies. However, one needs to be aware that the "real" ground noise is only resolved in a specific frequency range depending on the equipment.

The noise at frequencies greater than 2 Hz is related to cultural sources. The local maxima at around 4–8 s and 10–16 s are related to standing waves and waves breaking near the shore, respectively; see Fig. 7. Even more long-period noise is related to atmospheric changes. There can be strong differences at individual sites between day and night related to the changes in cultural activity but also between different seasons. For example, in northern Europe the noise caused by the ocean is much higher during the winter compared to the summer.

Borehole sensors are often used to improve data quality, since borehole sensors are not as affected by noise generated at the surface, as sensors placed at or near the surface. The effect is documented by Carter et al. (1991) stating that 945 m below the surface, noise levels between 15 and 40 Hz are reduced by 10 dB. For periods of between 30 s and 100 s, Hutt et al. (2002) report that the noise reduction is most prominent at the first few tens of meters. The disadvantage of borehole installations with respect to surface installations is that the instruments required are more expensive to buy and install and servicing the sensor is more difficult.

Network Maintenance and Quality Control

Seismic networks require a significant amount of maintenance work and attention to detail. Monitoring of the data quality provides an essential tool in the identification of problems with seismic stations. Being able to identify problems and respond quickly is key to successful operation of seismic networks. Recording of changes to instrumentation is also important.

Data quality can be looked at while the data is coming in, later to evaluate the performance of a network or at the end of a temporary deployment. Seismic networks make use of tools to monitor parameters for quality control. One such system is QUACK (Quality Analysis Control Kit) available at IRIS (http://www.iris.washington.edu/servlet/quackquery/welcome.do). QUACK provides information on several parameters such as availability,

data completeness, gaps (lack of data in a time interval), noise levels, and data latency (data arriving late). Most seismic networks record data continuously, and measuring data completeness is an efficient tool to evaluate the overall network performance. Checking for gaps is typically done to identify issues with the communication or equipment.

One parameter that can be checked in near real time is the data latency. An increase in latency will indicate unstable communication. If the latency increases continuously, it will indicate that the connection to the station is lost or worse the recorder, digitizer, or other units at the station are broken. Figure 8 shows an example on latency where the connection to the seismic station was down for a few days. Monitoring of latency will also allow evaluating the quality of communication during regular operation.

It is important that seismometers are installed with correct horizontal orientation so that the north and east components are oriented within the true geographic system. It is also possible to have different orientation, but this needs to be recorded with the station's metadata. The orientation of sensors can be computed retrospectively from the recordings of seismic events. For example, the "LDEO Seismology Project: Waveform Quality Center" measures misorientation by comparing computed to observed polarization of surface waves (http://www.ldeo.columbia.edu/%7Eekstrom/Projects/WQC.html). They also attempt to measure changes in gain of seismometers over time from recorded data. To record and possibly avoid changes in gain, regular calibration of the seismic equipment should be part of the network routine operation.

When plotting the noise spectra color scaled over time, one obtains spectrograms that are useful in evaluating the changes of noise. Failure of instrumentation will be seen easily as it will look like an apparent sudden change in noise levels. At the ORFEUS Data Center, PSD plots are generated in yearly intervals showing frequency bands around 0.01, 0.05, 0.5, and 2 Hz (http://www.orfeus-eu.org/data/data_quality.html). An example showing a 1-year PSD from the broadband station SFJD in Greenland is seen in Fig. 9.
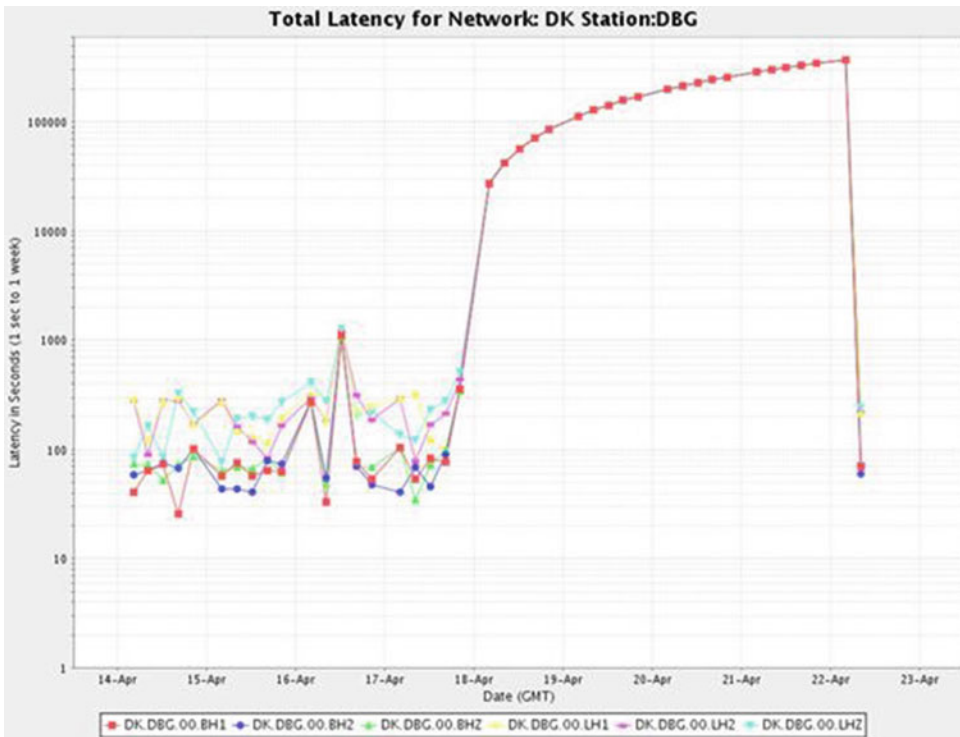
Seismic Network and Data Quality, Fig. 8 Latency of data streams from the DBG station in Greenland. The latency is the time between the last sample in the data record and GMT. The connection to the station was lost on April 18th and regained on the 22nd (This figure is generated at http://www.iris.edu/servlet/quackquery/)

Seismic Network and Data Quality, Fig. 9 Power spectral density (PSD) plot of data from the SFJD station in Greenland, covering the year 2011. The *black*, *red*, *green*, and *blue* lines show the frequencies 0.01 Hz, 0.05 Hz, 0.5 Hz, and 2.0 Hz, respectively (With permission from Dr. Sleeman, ORFEUS. This figure is obtained from http://www.orfeus-eu.org/)
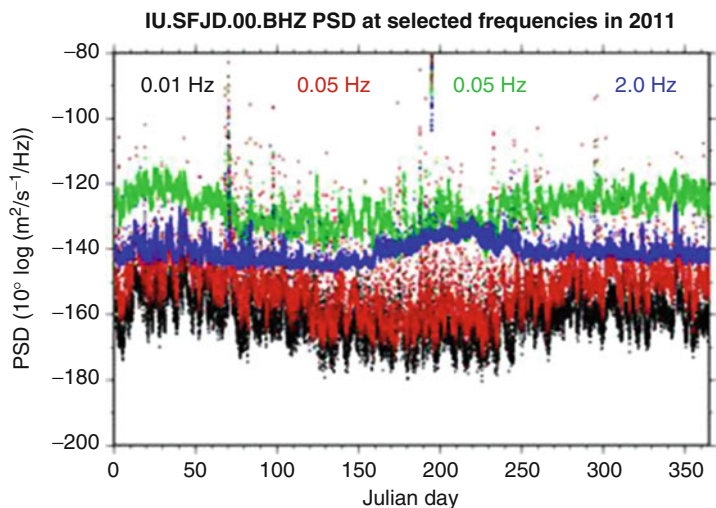
Site Selection and Vault Construction

The choice of site, selected to record seismic data, is essential to obtain data with a high quality. One can install instruments of the highest quality at sites where a high ambient noise level will make the data useless for the purpose they were intended for. Before deciding on a specific
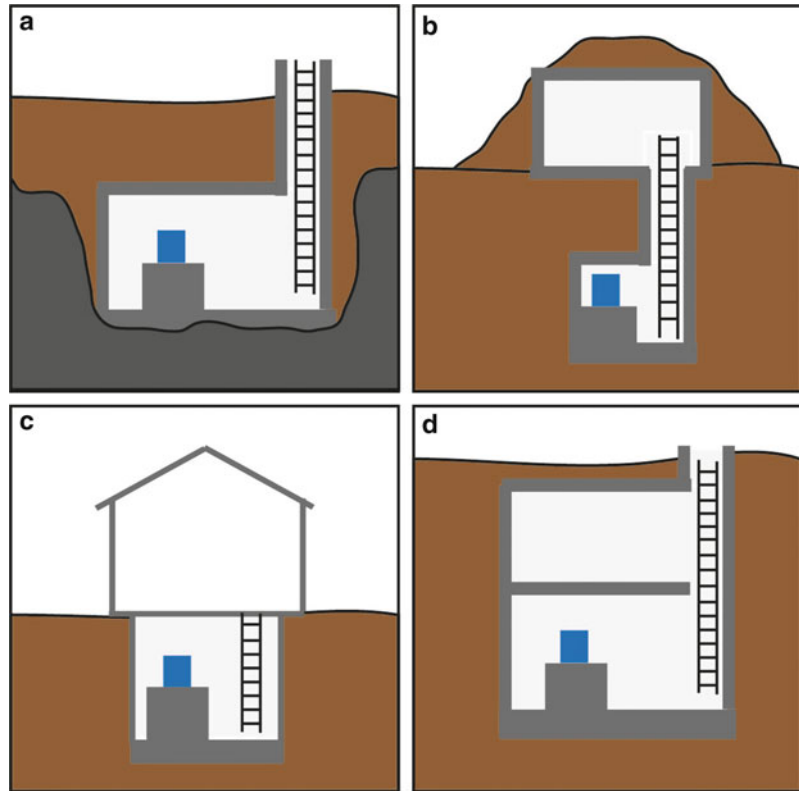
site, it is common practice to carry out a site survey where all relevant criteria are considered and noise measurements are done for a few days to identify potential noise sources. If the source of the seismic signal that one wishes to record is known, for example, a small area that is known to be seismically active, one should aim to get as close to the source as possible. One should also aim to have stations installed with the highest azimuthal coverage possible. It is of course important to stay away from the man-made sources of noise mentioned above. Site selection is often dependent on other factors than the ambient noise, such as access to power and communication as well as the physical access to the site. One might therefore not always be able to use the site with the lowest noise level, since it does not meet these requirements.

Constructing good vaults is also very important, and possibilities range from boreholes to relatively deep substantial vaults to rather shallow and simple installations (Fig. 10). The coupling of the sensor to bedrock is very important in all types of installation. If the budget is not limited, one would attempt to build a pier placed on bedrock several meters below the surface. Around the pier, a room would be constructed that would be isolated from the pier. It is essential to achieve rather stable temperatures and pressure around the sensor. Another room may be placed on top to house the remaining equipment. In a simpler setup, one can use a concrete structure with a smaller diameter that has the sensor at the bottom, a few meters below the surface. Care has to be taken to either drain water away or avoid water ingress completely. Many such designs have been made taking into account local conditions and available materials. The type of vault chosen should also be based on the specifications of the equipment that will be used.

Seismic Network and Data Quality, Fig. 11 Detection capability of the UK seismograph network. Unit is magnitude on the Richter scale. *Green* triangles are broadband stations, *red* triangles are short-period stations, and *blue* triangles are broadband stations operated by other agencies (http://www.earthquakes.bgs.ac.uk/)
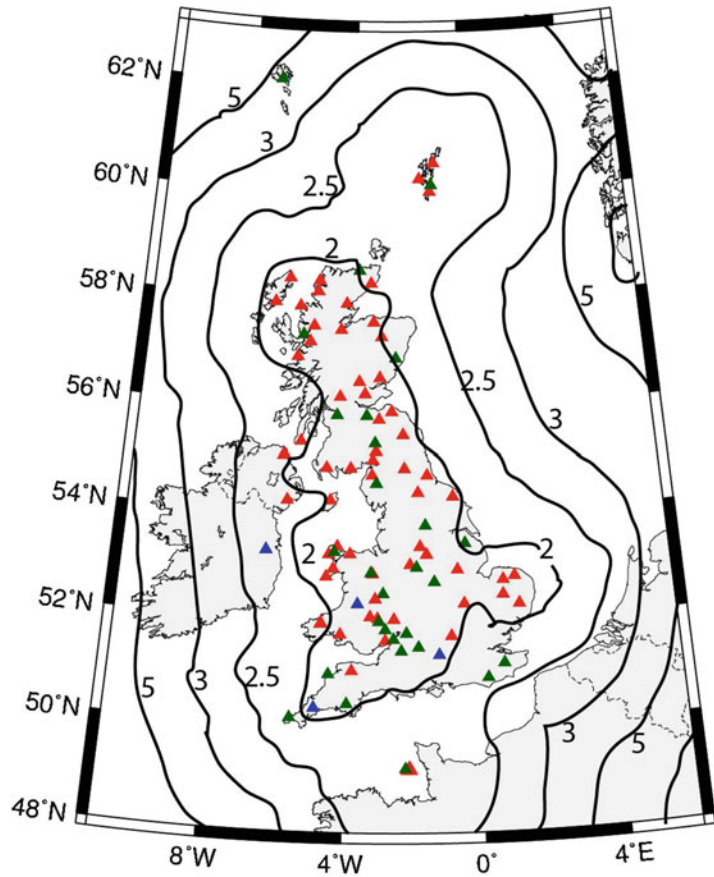
Detection Capability

If detection of earthquakes is the main purpose of a seismic network, knowledge on the detection capability of the network is essential when the network is designed and evaluated. The detection capability depends on the level of the background noise, the type of instrumentation, and the configuration of the network. The detection capability is mostly improved by lower background noise in the frequency range of interest, the distribution of stations, and their distance to the area of interest. An example of detection capability is shown in Fig. 11. Here it is assumed that the amplitude threshold for the detections of an earthquake must be twice the background noise level and that the earthquake is recorded on at least four stations (http://www.earthquakes.bgs.ac.uk/monitoring/detection_capability.html).

References

Carter JA, Barstow N, Pomeroy PW, Chael EP, Leahy PJ (1991) High-frequency seismic noise as a function of depth. Bull Seismol Soc Am 81:1101–1114

Clinton J, Heaton T (2002) Potential advantages of a strong-motion velocity meter as opposed to a strong motion accelerometer. Seismol Res Lett 73:332–342

Forbriger T, Widmer-Schnidrig R, Wielandt E, Mark Hayman M, Ackerley N (2010) Magnetic field background variations can limit the resolution of seismic broad-band sensors. Geophys J Int 183(1):303–312. doi:10.1111/j.1365-246X.2010.04719.x

Hanka W, Heinloo A, Jaeckel KH (2000) Networked seismographs: GEOFON real-time data distribution. ORFEUS Newsl 2(3):1–24

Havskov J, Alguacil G (2010) Instrumentation in earthquake seismology. Springer, Dordrecht, p 349

Havskov J, Ottemöller L (2010) Routine data processing in earthquake seismology with sample data, exercises and software. Springer, New York, p 380

Hutt CR, Bolton HF, Holcomb LG (2002) Seismograph networks. In: Lee WHK, Jennings P, Kisslinger C,

S

Kanamori H (eds) International handbook of earthquake & engineering seismology, part 1. Academic Press, p 1200

McNamara DE, Boaz RI (2005) Seismic noise analysis system, power spectral density probability density function: stand-alone software package, United States geological survey open file report, No. 2005–1438, p 30

McNamara DE, Buland RP (2004) Ambient noise levels in the continental United States. Bull Seismol Soc Am 94(4):1517–1527

Peterson J (1993) Observations and modeling of background seismic noise. Open file report 92–302, U.S. Geological Survey, Albuquerque

Scherbaum F (2001) Of poles and zeros. Fundamentals of digital seismology, vol 15, 2nd edn, Modern approaches in geophysics. Kluwer, Boston, p 269

Further Reading

The most complete coverage of these topics is probably given in the *New Manual of Seismological Observatory Practice* (http://http/nmsop.gfz-potsdam.de), giving an overview, the theoretical background, as well as examples and information sheets. Textbooks on seismic instrumentation are Havskov and Alguacil (2010) and Scherbaum (2001). A textbook on data processing for earthquake seismology is Havskov and Ottemöller (2010). Stein and Wysession (2003) and Lay and Wallace (1995) also have relevant chapters.

Seismic Noise

Peter Bormann
Formerly GFZ German Research Center for Geosciences, Potsdam, Germany

Synonyms

Ambient noise; Cultural noise; Digitization noise; Instrumental noise; Man-made noise; Microseismic noise; Microseisms; Self-noise; Signal-generated noise

Introduction

Seismic signals are usually transient waveforms radiated from a localized natural or man-made

---

Peter Bormann: deceased.

---

seismic source. They can be used to locate the source, to analyze source processes, and to study the structure of the medium of wave propagation. In contrast, the term "seismic noise" designates undesired components of ground motion that do not fit in our conceptual model of the signal under investigation. What we identify and treat as noise in seismic studies depends on the available data, on the aim of our study, and on the method of analysis. Accordingly, data treated as noise in one context may be considered as useful signals in other applications. For example, short-period noise (frequency >0.2 Hz) can be used for investigating the potential amplification of earthquake-generated ground motions due to local geological site conditions, also termed *microzonation* (e.g., Parolai 2012), and generally, both short- and long-period noise (frequency < 0.1 Hz) can also be used for tomographic studies of inhomogeneities in the Earth's crust (e.g., ▶ Noise-Based Seismic Imaging and Monitoring of Volcanoes).

According to Bormann and Wielandt (2013), disturbing noise in seismic records in a wider sense may comprise:

- Ambient vibrations due to natural sources (like ocean microseisms, wind, etc.)
- Man-made vibrations (from industry, traffic, etc.)
- Secondary signals resulting from Seismic Wave Propagation in inhomogeneous medium (due to *wave scattering* and therefore also termed *signal-generated noise*)
- Effects of gravity (like Newtonian attraction of atmosphere, horizontal accelerations due to surface tilt)
- Signals resulting from the sensitivity of ▶ Broadband Seismometers to ambient conditions (like temperature, air pressure, magnetic field, etc.)
- Signals due to technical imperfections or deterioration of the seismic ▶ Passive Seismometers (corrosion, leakage currents, defective semiconductors, etc.)
- Intrinsic ▶ Seismometer Self-Noise and Measuring Methods of the Seismic

Instrumentation (like Brownian noise, electronic and quantization noise)
- Artifacts from data processing

However, only the first four items are microseismic noise in a proper sense and elaborated in the following whereas the other ones relate to *seismometry* and *data processing* and are referred to only en passant.

The study of earthquakes or the imaging of the Earth with seismic wave arrivals requires ▶ Seismic Event Detection above background noise. Levels of natural ambient noise may vary by 60 dB (i.e., a factor of 1,000 in amplitude) depending on location, season, time of day, and weather conditions. This corresponds to differences in detection thresholds for seismic arrivals by about three magnitude units. Therefore, prior investigation of noise levels at potential recording sites is of utmost importance, as is environmental shielding and proper installation of seismic Sensors so as to reduce the influence of various sources of natural and man-made (cultural) noise (see sections "Causes and Basic Characteristics of Ambient Noise Observed on Land" to "Improving the Signal-to-Noise Ratio SNR by Data Processing" below; Trnkoczy et al. 2011; Forbriger 2012).

The essay will brief on:

- Different mathematical representations of noise and signals
- Origins of natural and man-made ambient noise
- Basic features of noise in the wide frequency range of interest for the recording of local micro-earthquakes to normal modes of the Earth (0.2 mHz $< f <$ 100 Hz)
- Influence of filter procedures and the sensor environment on the appearance of signals and noise in seismic records
- Task-dependent procedures for signal-to-noise ratio (SNR) improvement of earthquake and explosion records

More than 98 % of all seismological recordings are made on land, despite the growing interest in ▶ Ocean Bottom Seismometer (OSB) installations in order to improve the global ray path coverage for tomographic studies of the Earth and for regional studies of seismicity and Earth structure in the ocean environment. But OBS deployment, operation, and maintenance are rather expensive and logistically demanding, and the noise conditions on the ocean floor are usually worse than on land (Shinohara et al. 2011; Webb 2002). Therefore, this essay will deal, also because of strict space limitations, only with microseismic noise conditions on land. More details about such noise in both environments as well as on the specifics of seismic signals in different bandwidth ranges and related noise-dependent local and global detection thresholds can be found in Bormann and Wielandt (2013).

With respect to historical aspects of the recording and treatment of microseismic noise, it has to be considered that most of the early twentieth-century seismographs by Wiechert, Mainka, Galizyn, Bosch-Omori, Milne-Shaw, and others were mechanical medium-period broadband sensors of low magnification of the ground motion. Only the more sensitive ones with 100–500 times magnification were already able to record the ever-present ground motions with peak amplitudes around 6 s which constitute the dominating background noise for any seismic measurement with maximum amplitudes in the μ micrometer range (Fig. 1). They were termed microseisms. Such recordings were first reported by Algue in 1900. Wiechert (1904) wrote that these microseisms are caused by ocean waves on coasts. Later it was found that it is necessary to discriminate between: (a) the smaller primary ocean microseisms with periods between 10 and 20 s, typically around 14 ± 2 s and (b) the secondary or *double frequency microseism* which is related to the main noise peak around 6 ± 2 s (Figs. 1 and 2).

Other types of seismic as well as cultural and instrumental noise became more or less disturbing only when the sensitivity of electromagnetic seismographs, and later of electronically amplified seismographs, was raised so much that they were able to resolve even ground

S
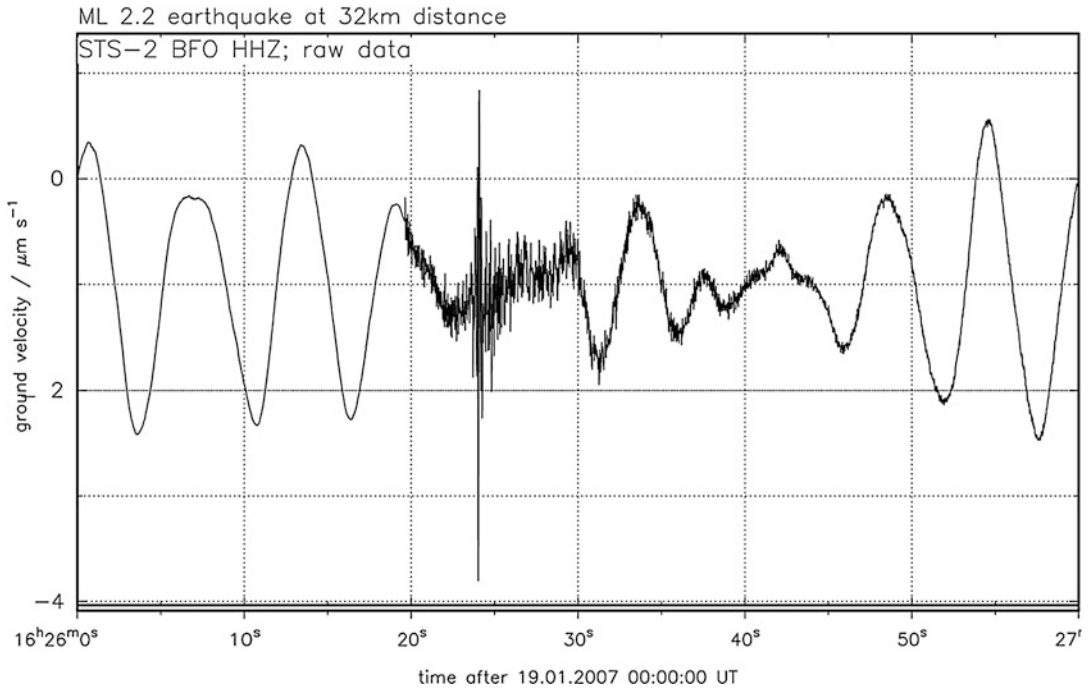
Seismic Noise, Fig. 1 STS-2 velocity broadband record of the 6 s secondary ocean microseisms at the Black Forest Observatory (BFO), Germany, superposed by high-frequency waves of a local earthquake of magnitude $M_L = 2.2$ that occurred 32 km away (Figure by courtesy of Thomas Forbriger for NMSOP Chapter 4 © IASPEI)

motion amplitudes in the size range of molecules and atoms (10^{-9} to 10^{-10} m).

Noise conditions at seismic recording sites vary globally in the wide range of periods of seismological interest between some upper and lower level, termed the New High Noise Model (NHNM) and the New Low Noise Model (NLNM) according to Peterson (1993). These noise models are expressed in dB of *power spectral density* (see section "Nature and Theoretical Representation of Signals and Noise") of acceleration related to 1 $(m/s^2)^2$/Hz (Fig. 2). However, the actual noise at any recording site corresponds nowhere in the whole range of periods to either the NHNM or the NLNM. Some sites and seismographs are more affected by long-period tilt-noise at a period above 30 s, and this differently in each record component, while others are mostly disturbed by high-frequency ($f > 1$Hz) man-made noise.

According to Fig. 2, variations in noise acceleration power are typically within about 30–80 dB, but they are much larger in velocity power (>130 dB) and displacement power (>250 dB). The latter is much more than the dynamic range of common modern digital seismographs with 24 bit *analog-to-digital converter* (ADC) and a dynamic range of 144 dB. In order to resolve, at good sites, at least the ambient noise with a few bits, even at the NLNM level, and still record without clipping also seismic signals from events in a reasonably large range of magnitudes and epicentral distance, current broadband seismographs used are *velocity sensors* such as the STS-1 and STS-2 ► Recording Seismic Signals. Their dynamic range is given in Fig. 3 with respect to the NLNM when measured in a constant relative frequency bandwidth of 1/3 octave and represented by "average peak" amplitudes (see section "Approximate Conversion of Power Densities into Recording Amplitudes"). In contrast, classical analog records on paper or film had only a dynamic range of

Seismic Noise,
Fig. 2 Envelope curves of acceleration noise power spectral density P_a (in units of dB relative to 1 $(m/s^2)^2$/Hz) as a function of noise period (according to Peterson 1993). They define the new global high (NHNM) and low noise models (NLNM) which are currently the accepted standard curves for generally expected limits of microseismic noise depending on period. No single station site, however, meets in the whole period range the NLNM or NHNM. However, some exceptional seasonal, diurnal, or local site noise conditions may exceed these limits. Also shown are for the NLNM the related curves calculated for the displacement and velocity power spectral density P_d and P_v in units of dB relative to 1 $(m/s)^2$/Hz and 1 m^2/Hz (Modified version of Fig. 2 from P. Bormann 1998 © Springer; with kind permission from Springer Science + Business Media)

ACCELERATION POWER SPECTRAL DENSITY [10 log ($m^2 \cdot s^{-4}$/Hz)]
VELOCITY POWER SPECTRAL DENSITY [10 log ($m^2 \cdot s^{-2}$/Hz)]
DISPLACEMENT POWER SPECTRAL DENSITY [10 log (m^2/Hz)]
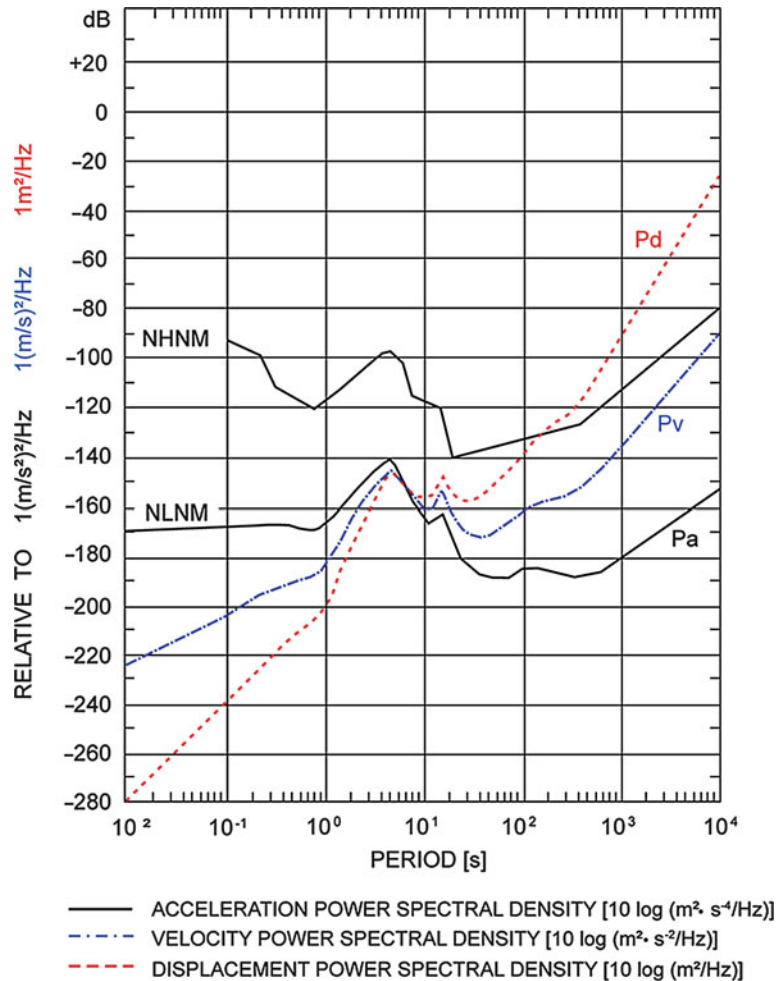
some 40 dB. They could cover only a very limited range of noise and record signal amplitudes between about 1 mm and 100 mm. Covering a range of 144 dB instead would correspond to covering an amplitude range between 1 mm and 16 km! However, what the actual usable dynamic range of a seismograph means has been discussed in Sect. 4.2.2 of Bormann and Wielandt (2013).

The reasons for the main types of natural and cultural microseismic noise observed on land and strategies for their avoidance or reduction in seismic records are discussed in the following sections after a brief discussion on the principal difference between seismic signals and noise, their mathematical representation, and the correct way of comparison.

Nature and Theoretical representation of Signals and Noise

Fourier Transformation of Continuous and Sampled Transient Signals

Seismograms contain noise and the desired signals. Sometimes the signal is completely masked by the noise. One of the main tasks in applied seismology is to ensure a good signal-to-noise ratio (SNR) both by selecting recording sites with low ambient noise and by data processing. The success of the latter largely depends on our understanding of the ways in which seismic signals and noise differ. The basic mathematical tool for this purpose is the harmonic or *Fourier analysis*, that is, the decomposition of a signal into sine waves. Depending on the type of signal
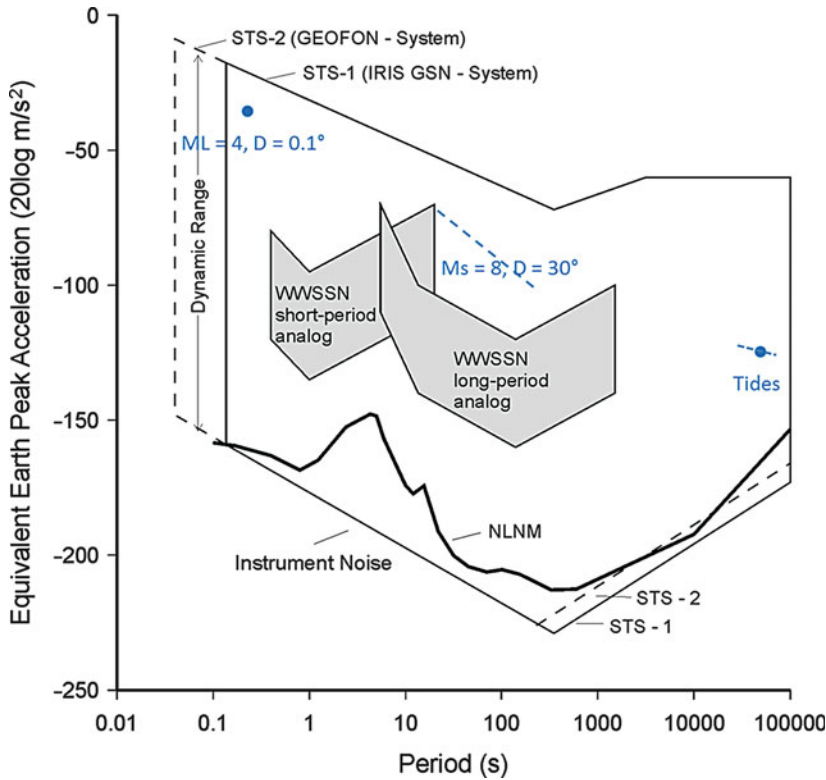
Seismic Noise, Fig. 3 A representation of the bandwidth and dynamic range of conventional analog World Wide Standard Seismograph Network (WWSSN) short- and long-period and digital *broadband seismographs* (STS1 and STS2 with good shielding - see Wielandt 2011 and Forbriger 2012) with respect to the equivalent Earth peak acceleration generated by solid Earth tides due to the attraction forces exerted by the Sun and the Moon, by a magnitude 8 earthquake at about 3,300 km distance or by a local earthquake of magnitude 4 at about 10 km distance (see ▶ Earthquake Magnitude Estimation). The depicted lower bound is determined by the *Instrumental Self-noise*

(also Sect. 4.3.2 in Bormann and Wielandt 2013). The scale is in decibels (dB) relative to $(1 \text{ m/s}^2)^2$. Noise is measured in a constant relative bandwidth of one third octave and represented by "average peak" amplitudes equal to 1.253 times the root-mean-square (RMS) amplitude. NLNM relates to the global New Low Noise Model according to Peterson (1993; see Fig. 2) The requirement to be fulfilled for presenting seismic signal, noise, and clip levels in one diagram has been outlined in Sect. 4.3.3 of Bormann and Wielandt (2013). Copy of Fig. 4.7 in Bormann and Wielandt (2013), compiled by using data of W. Hanka and J. M. Steim; © IASPEI

(transient or stationary, continuous or sampled) different mathematical formulations must be used which have been outlined in detail in Sect. 4.2 of Bormann and Wielandt (2013). The essence is that according to the *Fourier theorem*, any arbitrary *transient function* f(t) in the *time domain* can be represented by an equivalent function $F(\omega)$ in the *frequency domain*. The mathematical relationship between the two domains is defined by the *Fourier integral transformation*:

$$f(t) = (2\pi)^{-1} \int\limits_{-\infty}^{\infty} F(\omega) exp(i\omega t) d\omega \qquad (1)$$

$$F(\omega) = |F(\omega)| exp(i\phi(\omega)) = \int\limits_{-\infty}^{\infty} f(t) \exp(-i\omega t) dt$$

$$(2)$$

where ω (in units radian/s) $= 2\pi v$ is the circular frequency with $v =$ common frequency

(in units s^{-1}), $|F(\omega)|$ = *amplitude spectrum*, and $\phi(\omega)$ = *phase spectrum*. However, the Fourier integral transformation can only be applied to *transient signals* (i.e., signals that disappear or decay after some time so that they have a finite energy). For other signals, the integral may diverge. Moreover, the signals should have been radiated by localized finite sources thus being coherent signals with a defined phase relationship. In the case of digital data, the discrete transformation has to be applied. It also requires a sampled signal of finite duration. In contrast, signals which do not vary much within the interval of time over which earthquake signals are typically analyzed are mathematically treated as *stationary signals*, that is, signals whose statistical properties do not change with time. Such signals require a different mathematical treatment.

Normally, the detailed waveform of stationary signals is of no interest but rather subjected to efforts to remove the disturbing signal from the record. The exception is the above-mentioned application in *microzonation* studies (or *noise tomography*). However, it is of general interest to know how strong such stationary signals are in different parts of the frequency spectrum. This information is contained in a quantity named *power spectral density* (*PSD*).

Transient signals such as earthquake seismograms may have a complicated waveform which is determined by the earthquake source, the structure of the Earth, and the properties of the seismograph. The amplitudes and phases of the harmonic components of the signal are not random numbers but follow certain mathematical relationships. Such waveforms or spectra are sometimes called deterministic, in contrast to the "stochastic" or "random" waveform of noise. However, in a strict sense, even microseismic noise is deterministic as well. But the difference is that its sources and propagation paths are not known well enough to predict or interpret the waveform, and therefore it is assumed that the spectral amplitudes and phases are random numbers. Again, this is a mathematical simplification that may or may not be appropriate.

The concept of frequency filtering (e.g., high-, low-, or band-pass filtration) has its roots in analog signal processing and is therefore intimately connected to the *Fourier integral transformation*. Practical computer procedures of the *discrete Fourier transformation*, however, do not use integrals and are often only approximations to the corresponding analog procedures. Supplementary information on different aspects of the Fourier transformation can also be found in Sect. 5.2.3 of Wielandt (2011).

A general problem is that frequency filtering reduces the amplitudes of harmonic components of a signal in part of the frequency spectrum, usually in order to suppress noise. Examples are given in section "Frequency Filtering." But filtering may also change the waveform of the desired signal in the time domain, the more so the narrower bandwidth. Systematic differences between various magnitude scales as well as magnitude saturation effects are the consequence of filtering the input signals with different responses in different and often rather limited frequency ranges (see ▶ Earthquake Magnitude Estimation).

Spectral Analysis of Stationary Signals (Noise)

The analysis of stationary noise usually aims at assessing "how strong" the noise is or whether it will prevent the detection of a specific earthquake signal. Frequency filters may help to suppress noise in one part of the spectrum but still preserve signals in another part. Therefore it is necessary to consider the spectral distribution rather than the total strength of both types of signals. However, the formal measures of "strength" for transient and stationary signals are incompatible and cannot easily be compared. Mathematically speaking, transient signals have a finite energy and zero power (in the average over all times) while stationary signals have an infinite energy but a finite power. The same holds for the spectral densities. For a quantification of stationary noise, the concept of *power spectral density* (*PSD*) is therefore appropriate while transient signals are more adequately described by their *energy spectral density* (*ESD*).

The *Fourier integral transformation* as formulated in Eqs. 1 and 2 has the mathematical property (known as Rayleigh's or Parseval's theorem)

$$\int_{-\infty}^{\infty} f(t)^2 dt = \int_{-\infty}^{\infty} \frac{|F(\omega)|^2}{2\pi} d\omega = \int_{-\infty}^{\infty} |F(2\pi v)|^2 dv$$

(3)

The left side represents, in data processing, the total energy of a signal $f(t)$ which is not the physical energy but proportional to it. The right side represents the energy as an integral of $E = |F(2\pi v)|^2$ over all frequencies v. Thus, according to the above definition, the integrand is the *energy spectral density E* of the signal.

Note that negative frequencies must be included in the integration. For real signals, the ESD is a symmetric function, $E(-2\pi v) = E(2\pi v)$. Therefore, by multiplying by 2, the integration can be carried out over positive angular frequencies only. This convention is normally used in engineering and other practical applications; the ESD is then given as twice the "mathematical" value. The same applies to the PSD. It is important to clarify which convention is being used.

Power is energy per time, usually expressed as the average power over at least one cycle of an approximately periodic signal. The concept of power is mainly applied to stationary signals such as sine waves or seismic noise whose average power is constant or varies slowly. A Fourier transform in a strict sense is not defined for such signals, but the signal over any finite time interval of length T can be Fourier analyzed. When both sides of Eq. 3 are divided by T, the left side represents the average power, and $P = E/T = |F(2\pi v)|^2/T = |F(\omega)|^2/T$ on the right side is then the *power spectral density $P(\omega)$* for the interval T. But the PSD is not equivalent to a Fourier transform! The signal cannot be reconstructed from it with an inverse transformation because the phase information is lost. The smoothed and decimated PSD is however normally more useful because it contains all essential information in a concentrated form and eliminates arbitrary normalization factors.

The time derivative $\dot{f}(t)$ of a signal $f(t)$ has the Fourier transform $j\omega F(\omega)$. Thus, if the signal was originally measured as a displacement and the Fourier transform of the velocity is of interest then it is obtained by multiplying the original Fourier transform with $j\omega = j \cdot 2\pi v$ and the Fourier transform of the acceleration is then $-\omega^2 F(\omega)$. Going back from acceleration or velocity to displacement is achieved by dividing with the appropriate factor. The ESD and PSD are positive quadratic functions of the Fourier amplitudes. Therefore, differentiating the signal brings a factor ω^2, integration a factor ω^{-2}, etc. A PSD of velocity can thus easily be converted into that of displacement or acceleration.

Representing PSDs in Decibels

As shown in Fig. 2, it has become a standard to measure microseismic noise as a PSD of acceleration, that is, in $(m/s^2)^2$ per Hz or m^2s^{-3}. But "power" in this context is not the physical power in watts but simply the mean square amplitude of ground acceleration. As in acoustics, the power ratio $r = (a_2/a_1)^2$ between two signals of amplitude a_1 and a_2 is often expressed in decibels (dB). This is a logarithmic measure that expresses a ratio as a difference of logarithms. The difference in dB is $10 \log_{10}[(a_2/a_1)^2] = 20 \log_{10}[(a_2/a_1)]$. When expressing the power spectral density P_a of acceleration relative to the metric unit 1 $(m/s^2)^2$/Hz, it holds that

$$\mathbf{P_a}[dB] = 10 \log_{10}\left[P_a/1\left(m/s^2\right)^2/Hz\right] \quad (4)$$

In Fig. 2, however, a comparison has been made of the respective numerical values for physically incommensurable quantities such as acceleration power density with velocity and displacement power density. Therefore, Eqs. 2 and 3 and the following two text lines do not imply that the decibel units on both sides of the equal sign mean the same. They only look the same because the incommensurable reference levels were omitted. For this reason, three decibel scales for acceleration power, velocity power, and displacement

power have been given in Fig. 2. This clarification notwithstanding, it holds that numerically

$$\mathbf{P_v}[dB] = \mathbf{P_a}[dB] + 20 \log(\omega^{-1}) \qquad (5)$$

and

$$\mathbf{P_d}[dB] = \mathbf{P_a}[dB] + 40 \log_{10}(\omega^{-1})$$
$$= \mathbf{P_v}[dB] + 20 \log(\omega^{-1}) \qquad (6)$$

Consequently, for $\omega = 1$ (period $T = 2\pi/\omega = 6.28$ s) it holds that $\mathbf{P_a} = \mathbf{P_v} = \mathbf{P_d}$ (see Fig. 2) although only the numbers are equal! Also, $(\mathbf{P_d} - \mathbf{P_a}) = 2 \times (\mathbf{P_v} - \mathbf{P_a}) = $ constant for any given period, negative for $T < 2\pi$ s, and positive for $T > 2\pi$ s.

Cross-correlation and Coherence of Seismic Signals and Noise

The *cross-correlation* of two continuous stationary signals $f(t)$ and $g(t)$ at the time lag τ is defined as the expectation value (practically, the mean value) of the sliding dot product $f(t)\, g(t + \tau)$:

$$(f \star g)\ (\tau) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f*(t)\, g(t + \tau)dt, \qquad (7)$$

where $f*$ denotes the complex conjugate of f.

Properly normalized, the cross-correlation is a measure of the similarity between the two signals and is called their *coherence*. Like the power spectral density, it can be calculated via the Fourier transformation and specified as a function of frequency, in which case the name *coherence spectrum* is more adequate. Each signal is fully coherent with itself (*auto-correlation* with the coherence equaling 1). Independent random signals are incoherent (their coherence equals 0) although strictly this is only true in the limit of an infinite duration and after removing the mean. More interesting is the case that signals are partially coherent. Then, with some (frequency-dependent) fraction of the power, a common signal may be hidden in them, for example, a seismic signal masked by instrumental noise of the sensors.

Seismic signals depend not only on time but also on the location where they are recorded. It is

therefore meaningful to ask whether the signal recorded at one location is coherent with the same signal recorded at another location. This aspect of coherence is usually expressed as a "coherence length" over which the recorded signals have a specified coherence. The coherence length is only a fraction of a wavelength for omnidirectional noise (noise propagating in all directions) but can be much longer when noise originates in a limited source area and propagates essentially in one direction towards the station (Fig. 5). In a plane wave, the coherence length is theoretically infinite.

Comparing Spectra and Amplitudes of Transient and Stationary Signals

Bandwidth and Amplitude

Transient signals may be impulsive and oscillatory, and stationary signals either random signals (noise) or monochromatic sine waves. A typical seismogram belongs to the "oscillatory" category. The seismic signal near the focus of an earthquake should be impulsive. Ambient noise is often a superposition of all types, the "random" components being due to a large number of independent sources (traffic, wind, waves in the ocean), "transient" components being caused by identifiable sources such as passing cars, and nearly monochromatic components originating from machinery such as sawmills and pumps.

When the PSD of a stationary signal is constant over the bandwidth of a seismograph or a filter then, when reducing the filter bandwidth, the signal power is proportional to the bandwidth because it is the integral of a constant power density over the bandwidth. The rms amplitude of the signal, which is the square root of the power, must therefore be proportional to the square root of the bandwidth. This explains why a "spectral amplitude density" cannot be generally defined. Such a definition would imply that the amplitude is generally proportional to the bandwidth, which it is not (see Fig. 4.6 in Bormann and Wielandt 2013). It may nevertheless be so in special cases. Consequences of bandwidth differences of seismographs or filter

procedures applied for the evaluation of seismograms have been discussed in detail in Sect. 4.5 of Bormann and Wielandt (2013).

Approximate Conversion of Power Densities into Recording Amplitudes

According to Aki and Richards (1980), the *maximum* amplitude of a wavelet $f(t)$ near $t = 0$ can be *roughly approximated* by the product of what they call the "amplitude spectral density" $F(\omega)|$ (see Eq. 2) and the bandwidth of the wavelet, i.e.,

$$f(t)_{t=0} = |F(\omega)|2(v_u - v_1) \qquad (8)$$

with v_u and v_1 being the upper and lower corner frequencies of the band-passed signal. Such an approximation is only possible for signals whose energy is maximally concentrated in time. Likewise, if the power spectral density P of noise is defined as

$$\int\limits_{-\infty}^{\infty} f(t)^2 dt = \int\limits_{-\infty}^{\infty} \frac{|F(\omega)|^2}{2\pi} d\omega = \int\limits_{-\infty}^{\infty} |F(2\pi v)|^2 dv$$

$$(9)$$

then the *mean square amplitude* of noise in the time domain is

$$<f^2(t)> = 2P(v_u - v_1) \qquad (10)$$

This is of course a simplified version of the general rule that the power spectral density (PSD) must be integrated over the passband of a filter to obtain the power (or *mean square amplitude*) at the output of the filter. The square root of this power is then the *root mean square* (RMS) or effective amplitude

$$a_{RMS} = [2P \times (v_u - v_1)]^{1/2} \qquad (11)$$

However, if the power spectral density is defined according to the engineering approach only by its positive frequencies as $P_e = 2P$, then

$$a_{RMS} = [P_e \times (v_u - v_1)]^{1/2} \qquad (12)$$

Stationary signals must be characterized by their PSD, and *specifying noise by its RMS amplitudes is meaningless without definition of the bandwidth*. The values given by the NLNM and NHNM in Fig. 2 follow the engineering power ($\underline{P}_e$) convention, and the following related formulas as well.

Often it is necessary to represent the amplitude of noise (whether RMS or average-peak) in a *constant relative bandwidth RBW* over the whole frequency range. This means that the bandwidth **B** around each frequency $\boldsymbol{\nu}_0$ is a fixed fraction or multiple of $\boldsymbol{\nu}_0$. The lower and upper band limits, f_1 and f_u, are chosen so that f_0 is the geometric mean of the two (i.e., v_0 appears in the middle between v_1 and v_u on a logarithmic frequency scale). When the relative bandwidth is n octaves or m decades (n and m normally being fractions, not integers), then the following relationships hold:

$$v_1 = 2^{-n/2}v_0, \quad v_u = 2^{n/2}v_0, \boldsymbol{RBW}$$

$$= 2^{\boldsymbol{n}/2} - 2^{-\boldsymbol{n}/2} \text{ and } B$$

$$= \left(2^{n/2} - 2^{-n/2}\right)v_0 \qquad (13)$$

and (12) can then be written as

$$a_{RMS} == \left[P_e \times v_0 \times \left(2^{n/2} - 2^{-n/2}\right)\right]^{1/2}$$

$$= (P_e \times v_0 \times \boldsymbol{RBW})^{1/2} \qquad (14)$$

Octaves can be converted easily into decades and vice versa by using the relation

$$m = n \cdot \log_{10}(2) = n \cdot 0.3010 \qquad (15)$$

Typical response curves of short-period narrowband analog seismographs for recording transient teleseismic body waves (i.e., waves that have travelled more than 2000 km either as longitudinal waves, oscillating in the propagation direction, or transverse waves, oscillating perpendicular to the direction of propagation) typically have bandwidths between 1 to 2 octaves. Different authors have used an integration bandwidth of 1/2 to 1/3 octave (a standard bandwidth in acoustics) for computing RMS amplitudes from PSD. According
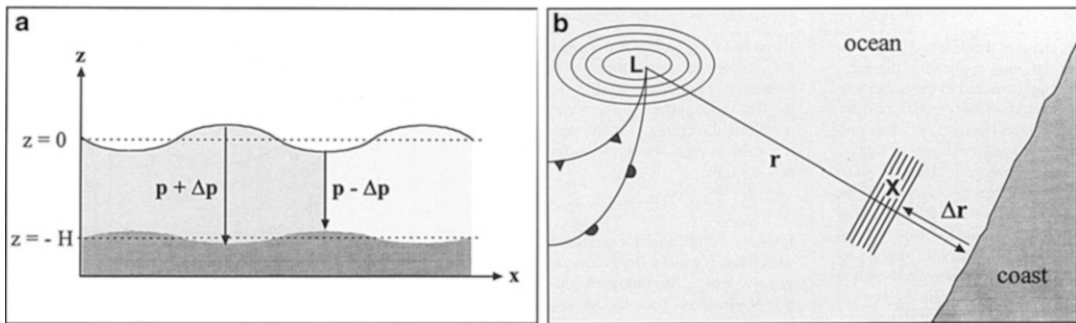
Seismic Noise, Fig. 4 Schemes for the generation of (**a**) primary and (**b**) secondary microseisms (for explanations see text). *L* cyclone low-pressure area, *X* area of interference where standing waves with half the period of ocean waves develop, *r* distance between L and the coast, Δr distance between the coast and *X*, *p* water pressure, Δp water pressure fluctuation proportional to wave amplitude (Copy of Fig. 12 from Friedrich et al. 1998 © Springer; with kind permission from Springer Science + Business Media)

to the above relationships, 1/3 octave RMS amplitudes will be only about 82 %, 70 %, and 50 % of the respective RMS amplitudes calculated for 1/2, 2/3, or 4/3 octave bandwidth, respectively.

There is a 95 % probability that the instantaneous amplitude of random noise with a Gaussian amplitude distribution will lie within a range of $\pm 2a_{RMS}$. Peterson (1993) showed that both broadband and long-period noise amplitudes follow closely a Gaussian probability distribution. This is true for largely natural ambient noise, often not, however, for short-period and broadband noise in urban and industrialized areas. There the noise is often dominated by transient or periodic signals (Groos and Ritter 2009 and section "Microseismic Noise in Urban Environments"). Yet, in case of a Gaussian probability distribution, the absolute peak amplitudes of the narrowband filtered signal envelopes should follow a Rayleigh distribution. In the case of an ideal Rayleigh distribution, the theoretical *average peak amplitudes* (APA) are 1.253 a_{RMS}. Therefore, RMS amplitudes in 1/6 decade bandwidth correspond approximately to average peak amplitudes in 1/3 octave bandwidth.

Causes and Basic Characteristics of Ambient Noise Observed on Land

Signals which have a similar waveform and polarization, so that they can interfere constructively, are termed coherent. This is usually the case for seismic signals generated and radiated by a common source process. The degree of coherence is defined by the ratio between the *cross-correlation* and the *autocorrelation* of the time series. The cross-correlation is squared and divided by both autocorrelations to render the result independent of the signal amplitudes. It may vary between 0 and 1. For noise, it shows distinct frequency dependence. Spatial coherence may be rather high for long-period ocean microseisms. Accordingly, the correlation radius, i.e., the longest distance between two seismographs for which the noise recorded in certain spectral ranges is still correlated, increases with the noise period. It may be several kilometers to tens of kilometers for $f < 1$ Hz but drops to just a few tens of meters or even less for $f > 50$ Hz. For random noise, it is usually not larger than a few wavelengths.

Primary and Secondary Microseisms

Primary ocean microseisms are generated only in shallow waters in coastal regions. Here the wave energy can be converted directly into seismic energy either through vertical pressure variations (Fig. 4a) or by the smashing surf on the shores. Therefore, primary ocean microseisms have the same period as the dominating water waves ($T \approx 10$ to 16 s). Haubrich et al. (1963) compared the spectra of microseisms and of swell at the beaches and could demonstrate a close

relationship between the two data sets. The relative stability in amplitude for continental sites around the globe suggests persistent, stable sources.

Contrary to the primary ocean microseisms, the secondary ocean microseisms could be explained by Longuet-Higgins (1950) and Hasselmann (1963) as being generated by the superposition of ocean waves of equal period traveling in opposite directions, thus generating standing gravity waves of half the period. These standing waves cause nonlinear pressure perturbations that propagate without attenuation to the ocean bottom and there couple into an elastic wave of much larger wavelength and higher phase velocity if the propagation directions are nearly opposite. The area of interference X (Fig. 4b) may be off-shore where the forward propagating waves, generated by a low-pressure area L, superpose with the waves traveling back after being reflected from the coast. But it may also be in the deep ocean when the waves, excited earlier on the front side of the low-pressure zone, interfere later with the waves generated on the back-side of the propagating cyclone. Thus, secondary ocean microseisms are mostly recorded with periods between 5 and 8 s and are the result of energetic ocean waves between the 10 and 16 s period. Horizontal and vertical noise amplitudes of marine microseisms are similar. The particle motion is dominatingly of Rayleigh-wave type, i.e., elliptical polarization of the particle motion in the vertical propagation plane. But some energy is present also as body waves, and heterogeneous Earth structure may even couple some energy into Love waves, i.e., surface waves that oscillate parallel to the Earth surface and perpendicular to the direction of propagation with amplitudes that decay exponentially with depth below the surface.

Medium-period ocean/sea microseisms experience low attenuation. They may therefore propagate hundreds of kilometers inland. Since they are generated in relatively localized source areas, they have, when looked at from afar, despite the inherent randomness of the source process, a rather well-developed coherent part, at least in the most energetic and prominent component.

This allows to locate the source areas and track their movement by means of seismic arrays (e.g., Cessaro, 1994; Friedrich et al. 1998, Fig. 5). This possibility has already been used decades ago by some countries, e.g., in India, for tracking approaching monsoons with seismic networks under the auspices of the Indian Meteorological Survey. While near-shore areas may be the source of both primary and secondary microseisms, the pelagic sources of secondary microseisms meander within the synoptic region of peak storm wave activity.

Note that the noise peak of secondary microseisms has a shorter period when generated in shallower inland seas or lakes (T ≈ 2–4 s) instead of in deep oceans because of generally shorter wavelengths and periods of the water waves in shallower and smaller water bodies. Also, off-shore interference patterns largely depend on coastal geometries and the latter may allow the development of internal resonance phenomena in bays, fjords, or channels (Fig. 6), which affect the detailed spectrum of microseisms. In fact, certain coastlines may be distinguished by unique "spectral fingerprints" of microseisms.

Long-Period Noise

Ground noise on land observed at frequencies between 0.2 and 50 mHz, i.e., lower than the microseism peak, is usually associated with atmospheric pressure fluctuations. At these frequencies, vertical component seismometers react to changes in gravity as the mass of the atmosphere above a site changes with atmospheric pressure. If seismometers are not fully sealed, they may also experience apparent accelerations due to buoyancy effects on the seismometer mass. Pressure fluctuations may also cause strong longperiod horizontal component noise by distorting the case of the seismometer and/or of the walls of the vault.

Long-period seismometers are also sensitive to temperature changes. Therefore, careful installations aim at sealing the sensor from both pressure and temperature fluctuations (Forbriger 2012). The vertical component noise levels between 0.2 and 1.7 mHz can be reduced by
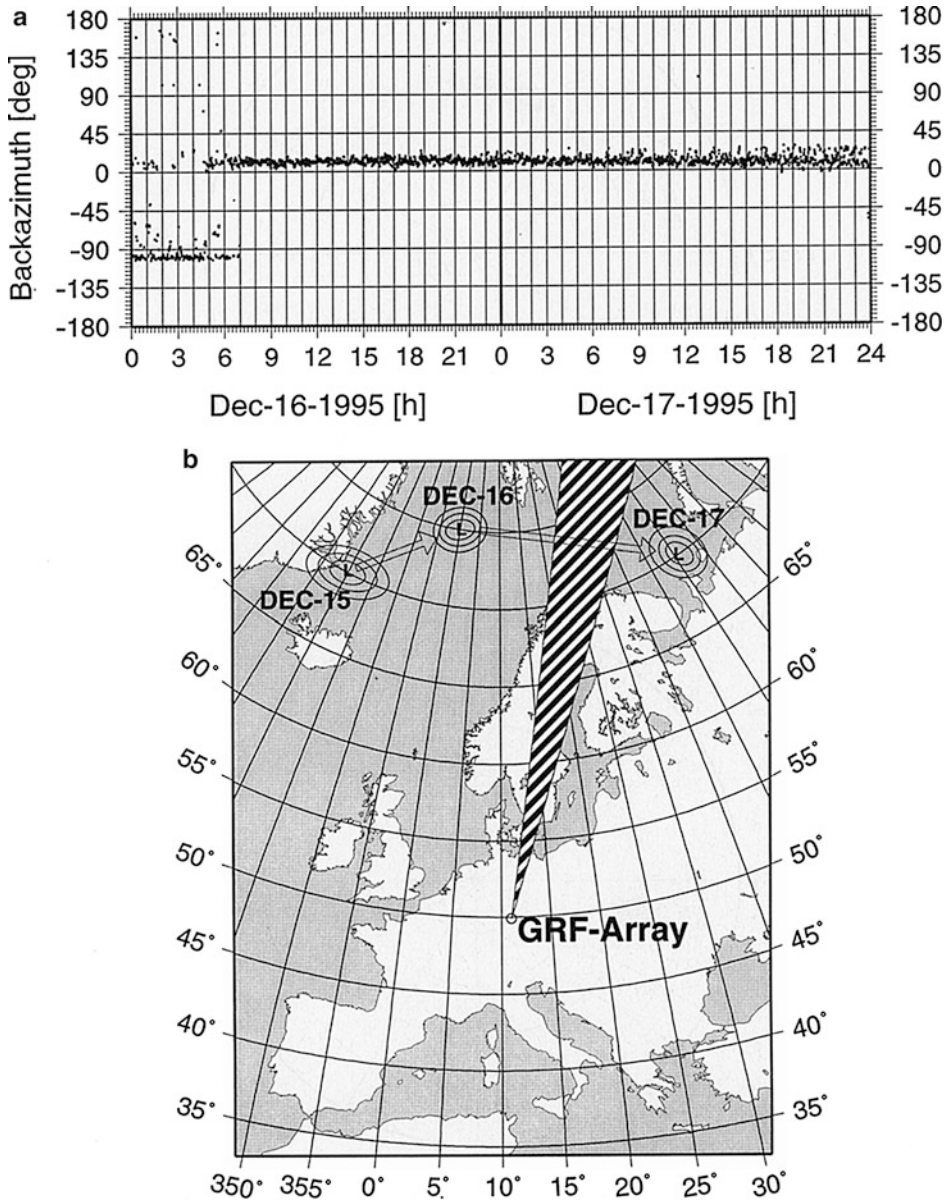
Seismic Noise, Fig. 5 An example of good spatial coherence of medium-period secondary ocean microseisms at larger distance from the source area which, in this case, allows rather reliable determination of the direction from the station towards the source (backazimut angle = angle measured clockwise from the north) by means of array analysis (see ▶ Seismometer Arrays). (**a**) Shows how the backazimuth determination changed from one day to the next, while (**b**) shows the location of the two storm areas and of the GRF seismic array. Observations by at least two arrays permit localization and tracking of the noise-generating low-pressure areas (Copy of Fig. 7 from Friedrich et al. 1998 © Springer; with kind permission from Springer Science + Business Media)

more than 10 dB by subtracting the scaled, locally recorded pressure signal from the vertical acceleration record, thereby improving the SNR for normal mode observations (e.g., Fig. 2.23 in Bormann et al. 2012 and Fig. 4.26 in Bormann and Wielandt 2013).

Turbulence in the atmospheric boundary layer produces pressure fluctuations at the Earth's
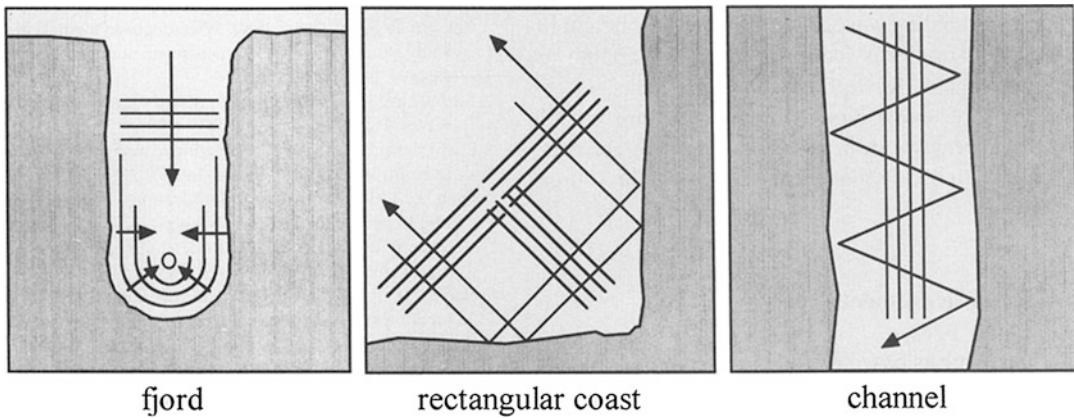
Seismic Noise, Fig. 6 Examples for coastline geometries that provide suitable interference conditions for the generation of secondary microseisms (Copy of Fig. 13 from Friedrich et al. 1998 © Springer; with kind permission from Springer Science+Business Media)

surface that cause significant deformation to depths of a few tens of meters below the surface. This noise source is a primary reason why broadband seismometers at permanent stations are often installed beneath the Earth's surface in shallow boreholes (see Sect. 7.4.5 and Fig. 7.59 in Trnkoczy et al. (2011)). For instruments on the surface, short-wavelength atmospheric pressure fluctuations produce a large tilt signal that makes the horizontal components particularly noisy. Tilt couples gravity into the horizontal components but much less into the vertical component (Chap. 5, Sect. 5.3.3 and Figs. 5.8 and 5.9 in Wielandt 2011). The ratio increases with the period and may reach a factor of up to 300 (about 50 dB). Horizontal component local tilt noise may also be caused by traffic and wind pressure.

At very long periods, deformation of the ground, of the seismometer vault, or of the building due to insolation gradients (uneven heating by sunshine) may become noticeable as well. Generally, recording sites poorly coupled to the ground or on poorly consolidated ground will tend to be noisier, both at long and short periods. Other reasons for increased long-period noise may be air circulation in the seismometer vault itself or underneath the sensor cover. Special care in seismometer installation and shielding is therefore required in order to reduce drifts and long-period environmental noise (Wielandt 2011; Trnkoczy et al. 2011; Forbriger 2012).

Short-Period Noise (0.5–50 Hz)

Ambient noise at frequencies higher than those associated with the microseisms peak may have natural causes such as wind (wind friction over rough terrain, trees, and other vegetation or built-up objects swinging or vibrating in the wind), rushing waters (waterfalls or rapids in rivers and creeks), etc. Wind-generated noise is broadband, ranging from about 0.5 Hz up to about 15–60 Hz. But the dominant sources of high-frequency noise are man-made (power plants, factories, rotating or hammering machinery, road and rail traffic, etc.; see Sects. 7.1 and 7.2 in Trnkoczy et al. 2011.). Wind-generated noise couples mostly into *surface-wave* modes; cultural noise, however, couples at least partly into *body waves* that can propagate also to greater depth. Cultural noise in records is in principle avoidable, although it is often impractical to place seismic station sites at sufficient distances away from cities or highways. Some rules of thumb suggest that it may be necessary to site short-period recording stations of high magnification as far as 25 km from power plants or rock-crushing machinery, 15 km from railways, 6 km from highways, and 1 km or more from smaller roads. And with respect to moving waters, it is recommended to place stations away from moving water between some 60 km for very large waterfalls and dams to 15 km for smaller rapids (Table 2.1 in Willmore 1979, http://www.seismo.
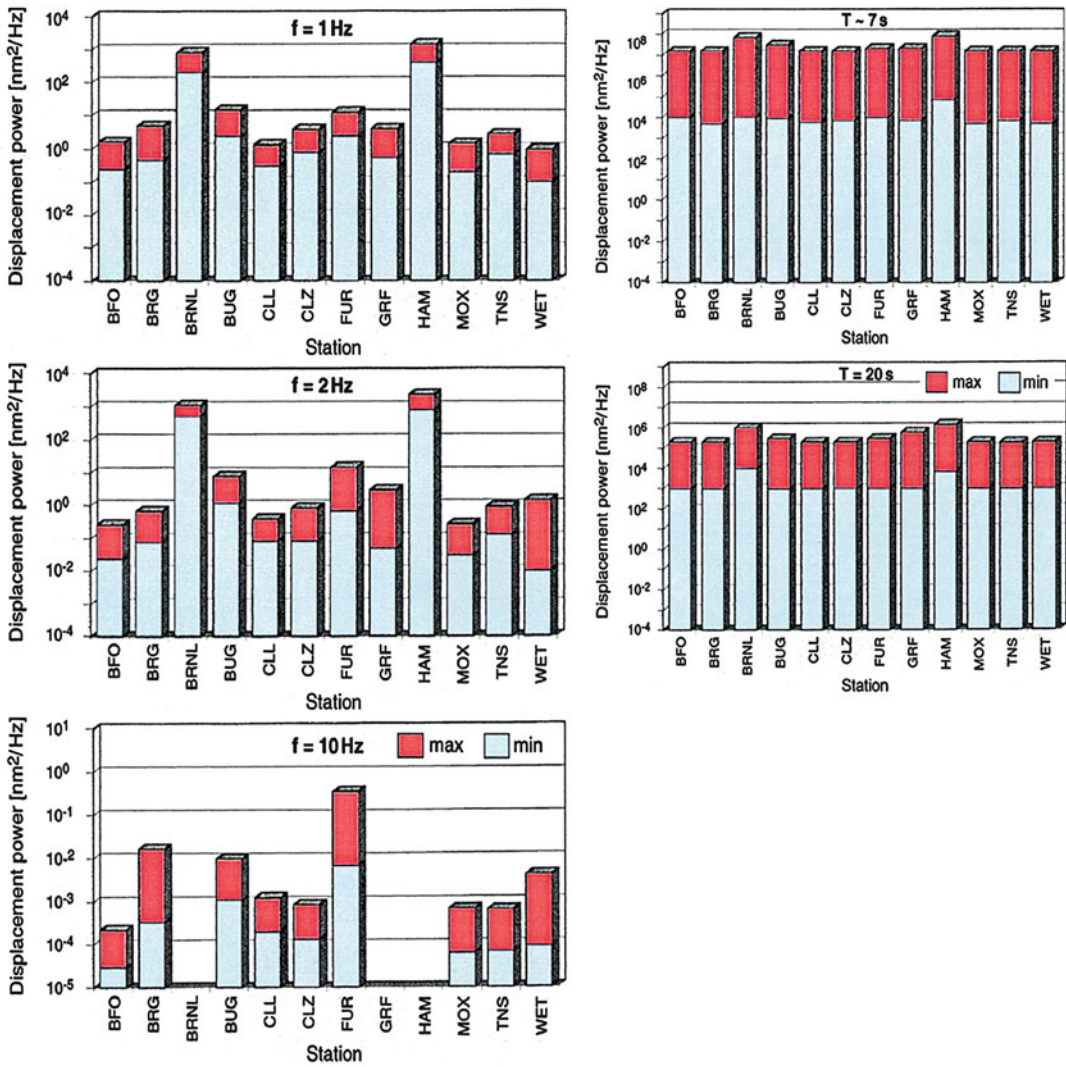
Seismic Noise, Fig. 7 Comparison of the minimum and maximum levels of short-period and long-period ambient noise power observed at stations of the German Regional Seismic Network (GRSN) (Color version of Fig. 9 from Bormann et al. 1997 © Springer; with kind permission from Springer Science + Business Media)

com/msop/msop79/sta/tab_2.1.gif or Table 7.5 in Trnkoczy et al. 2011). However, noise travels further across competent rock than through alluvial filled valleys. Therefore, "safe" distances may be from 1/2 to 2/3 of the values above, depending on propagation path.

Most of the short-period natural and man-made sources are distributed, stationary, or moving. Their contributions, coming from various directions, superpose to a rather complex, more or less stationary random noise field.

The particle motion of short-period noise is therefore more erratic than for long-period ocean noise and may vary more (by about 30 dB) than long-period noise (by about 10 dB) from station to station within a seismic network. Very pronounced is the seasonal variability of the secondary ocean microseisms by about 30 dB (Fig. 7).

Nevertheless, polarization analysis (the investigation of the different oscillation directions of the various types of seismic waves in space), averaged over moving time-windows, sometimes

reveals preferred azimuths of the main axis of horizontal particle motion, which hints at localized noise sources. Moreover, when the vertical component particle motion in three-component records is clearly developed and averaged, then this indicates the presence of fundamental Rayleigh-wave type polarization. A rather popular and cost-effective microzonation method is based on this assumption. It derives information about the fundamental resonant frequency of the soft-soil cover and estimates local site amplification of ground motion from the peak in the horizontal to vertical component spectral noise ratio (H/V method, e.g., Parolai 2012).

According to the above, every recording site has different noise characteristics depending on its distance from the ocean, industrial and settlement areas, major infrastructure facilities, the wind climate of the site, and the depth of burial of the sensor. At some locations, a seasonal cycle is evident due to variations in wind or water flow in nearby rivers. For example, according to Fyen (1990), spring runoff raises noise levels between 0.5 and 15 Hz at the NORESS array by up to 15 dB. In settlement and industrial areas, a diurnal cycle may dominate due to variable human activity during the day (Fig. 12).

High-frequency noise spectra on land during quiet intervals resemble each other at quieter sites. The rather featureless displacement power spectrum follows a power-law dependence in frequency, proportional to $\approx f^{-2}$ (Fig. 8). This spectral appearance changes drastically due to the strong impact of cultural noise, especially in the frequency band between 1 and 10 Hz (Figs. 13 and 14). Wind noise depends on the strength of the wind and the character of the site. Recorded noise power near 1 Hz in a surface vault may increase by 10–20 dB, as compared to records on calm days, when wind speed reaches about 5 m/s. Short-period wind noise becomes detectable in records at surface sites far from cultural sources at wind velocities above about 3 m/s and at 4 m/s for subsurface sensors as well (Fig. 9). The primary mechanism by which wind couples into highfrequency noise is probably by the direct action of the wind on trees, bushes, and other structures, although at lower frequency the
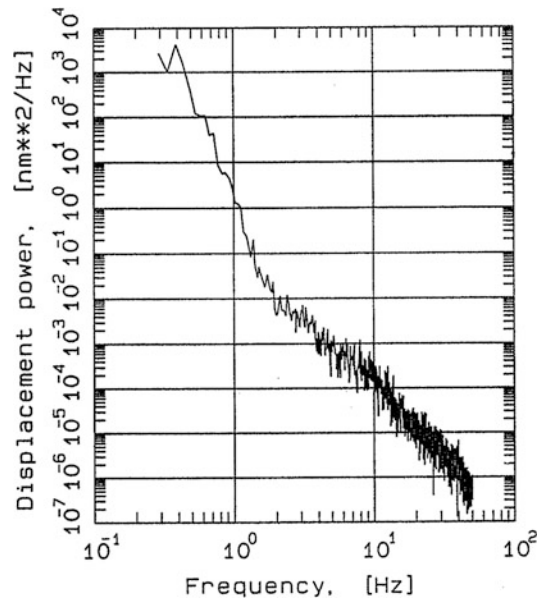
Seismic Noise, Fig. 8 Displacement power spectral density calculated from 6 moving windows, 50% overlapping, of short-period noise records, 4,096 samples long each, i.e., from a total record length of about 80 s at a rather quiet site in NW Iran (Copy of Fig. 4.6 in Bormann (2002) © IASPEI)

pressure fluctuations associated with wind can directly drive motions of the ground.

Generally, noise levels increase with higher wind speeds. However, there is apparently no linear relationship between wind speed and the amplitude of wind-generated noise. Wind noise increases dramatically at wind speeds greater than 3–4 m/s and may reach down to several hundred meters depth below the surface at wind speeds > 8 m/s (Young et al. 1996). But generally, the level and variability of wind noise is much higher at or near the surface and is reduced significantly with depth (Fig. 9). Moreover, at wind speeds below 3–4 m/s, the background noise may be omnidirectional with relatively large coherence length at frequencies below 15 Hz, but the coherence length is strongly reduced at higher wind speeds with increased air turbulence and local wind pressure fluctuations (Withers et al. 1996).

Comparisons of noise spectra from surface vaults and subsurface installations in mines (Fig. 9) or boreholes (Fig. 10) have repeatedly

Seismic Noise,
Fig. 9 Displacement
power noise spectra
measured at the surface
(*upper curves*) and at 420 m
below the surface in a
disbanded salt mine at
Morsleben, Germany
(*lower curves*), on a very
quiet day (*hatched lines*)
and on a day with light
wind on the surface (wind
speed about 4 m/s; *full
lines*) (Copy of Fig. 4.24 in
Bormann (2002) ©
IASPEI)
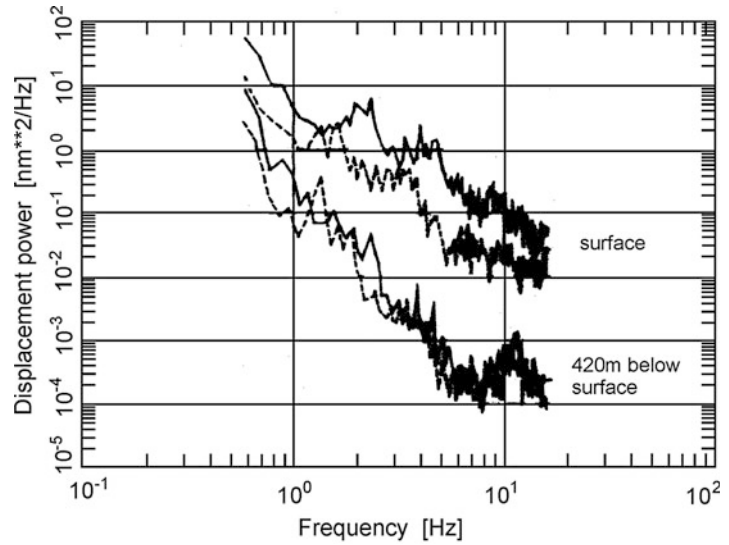
Seismic Noise,
Fig. 10 Velocity power
density spectra obtained
for noise records at the
surface (*top*) and at 300 m
depth in a borehole (*below*)
near Gorleben, Germany
(Figure by courtesy of
M. Henger; copy of
Fig. 4.23 in Bormann
(2002) © IASPEI)
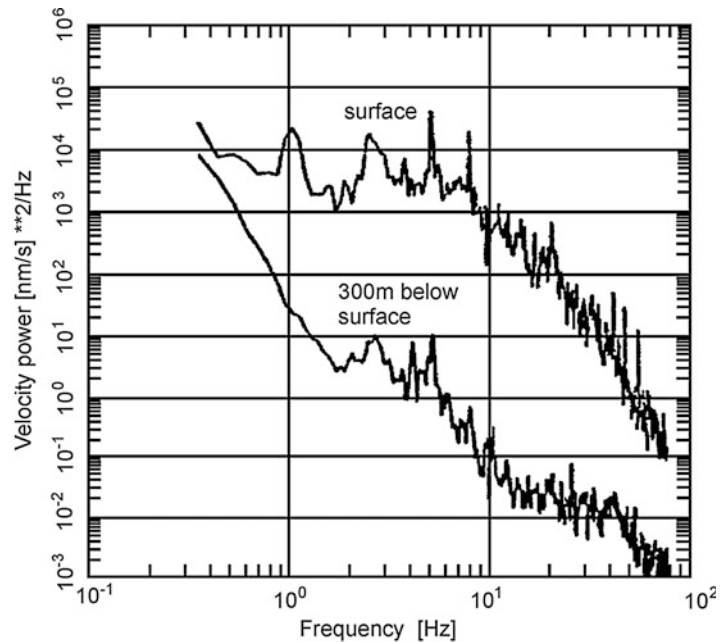
shown significant improvements in SNR for
high-frequency phases recorded at depths, some-
times even as shallow as a few meters. Large
reductions in noise level may occur already in
the first 5 m between the weathered rock of the
surface and the more competent rock in the bore-
hole. Note, however, that for periods longer than
2 s the noise reduction is only marginal, even for

sensor deployments at several hundred meters
depth.

In case of Fig. 10, the noise power at 300 m
depth was reduced, as compared to the surface, by
about 10 dB, at $f = 0.5$ Hz, 20 dB at 1 Hz and
35 dB at 10 Hz. The surface-wave nature of
ambient noise (including ocean noise) is the rea-
son for the exponential and frequency-dependent
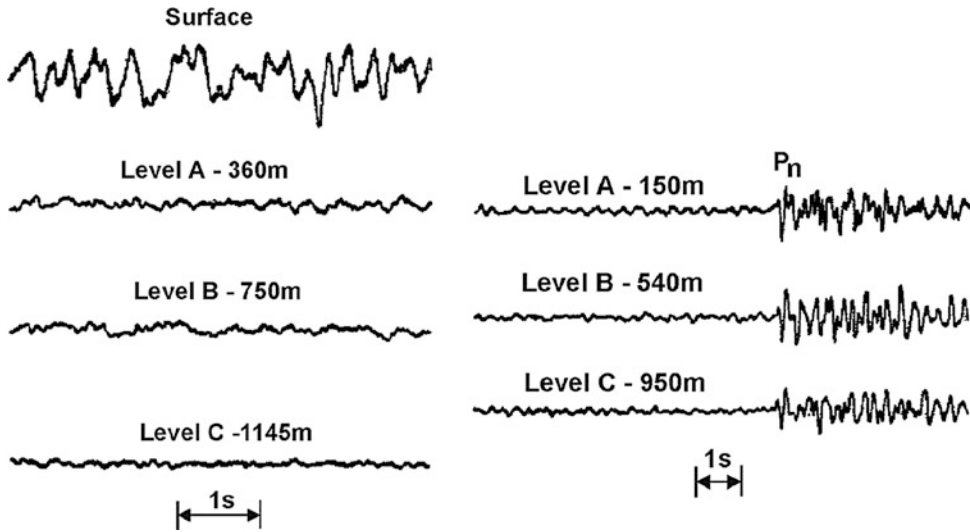
Seismic Noise, Fig. 11 Records of short-period noise (*left*) and signals (*right*) at the surface and at different depth levels of a *downhole seismometer array*. Note the significantly larger Pn amplitude closer to the surface. It is due to the surface amplification effect which may reach a factor of two for vertical incident waves but is less for incidence angles $>> 0°$ (Fig. copied from Bormann, 1966; modified from Broding et al. 1964)

decay of noise amplitudes with depth. This is not the case for body waves (Fig. 11). Since the penetration depth of surface waves increases with wavelength, high-frequency noise attenuates more rapidly with depth. But both noise reduction and signal behavior with depth depend also on local geological conditions (e.g., section "Installations in Subsurface Mines, Tunnels, and Boreholes"). Because of the surface-wave character of short- and medium-period noise, the horizontal propagation velocity of ambient microseismic noise depends on frequency. It is close to the shear-wave velocity in the uppermost crustal layers, which is about 2.5–3.5 km/s for outcropping hard rock and about 300–650 m/s for unconsolidated sedimentary cover. This is rather different from the apparent horizontal propagation velocity of P waves and all other steeply emerging teleseismic body-wave onsets.

Broadband spectral levels fall also with increasing borehole depth, but narrow spectral lines due to cultural sources and or waveguide trapping may show less depth dependence. This may result in very different noise decay and signal-to-noise ratio behavior with depth (see section "Installations in Subsurface Mines, Tunnels, and Boreholes").

Cultural noise may vary strongly, up to about 30 or even 50 dB at some higher frequencies between sites in different environments (Fig. 14) but also from day to day or with the time of the day at a given site in a busy urban environment (e.g., Fig. 12). Industrial and traffic noise levels are usually lowest on weekends but increase during holidays because of higher traffic. Sources such as power plants, transformer stations, etc. generate narrowband noise, producing energetic spectral lines at 50 or 60 Hz as well as related harmonics and subharmonics (e.g., at 30 Hz, 25 Hz, 12.5 Hz) depending on location (Fig. 13). The NORESS array data included lines at 6, 12, and 17 Hz due to reciprocating saws in a nearby sawmill, and records at Moxa (MOX) station in Germany have been spoiled over decades by a pronounced 2 Hz peak due to a heavy steam engine 16 km away driving a steel-press for rails. It may be possible to remove by frequency filtering such narrowband noise at frequencies that differ from the dominating signal frequency (see section "Frequency Filtering"). Other types of machinery, however, may also produce narrowband peaks in the noise spectra, but with time-variable frequencies as motor speed may vary. These noise sources are more difficult to remove.
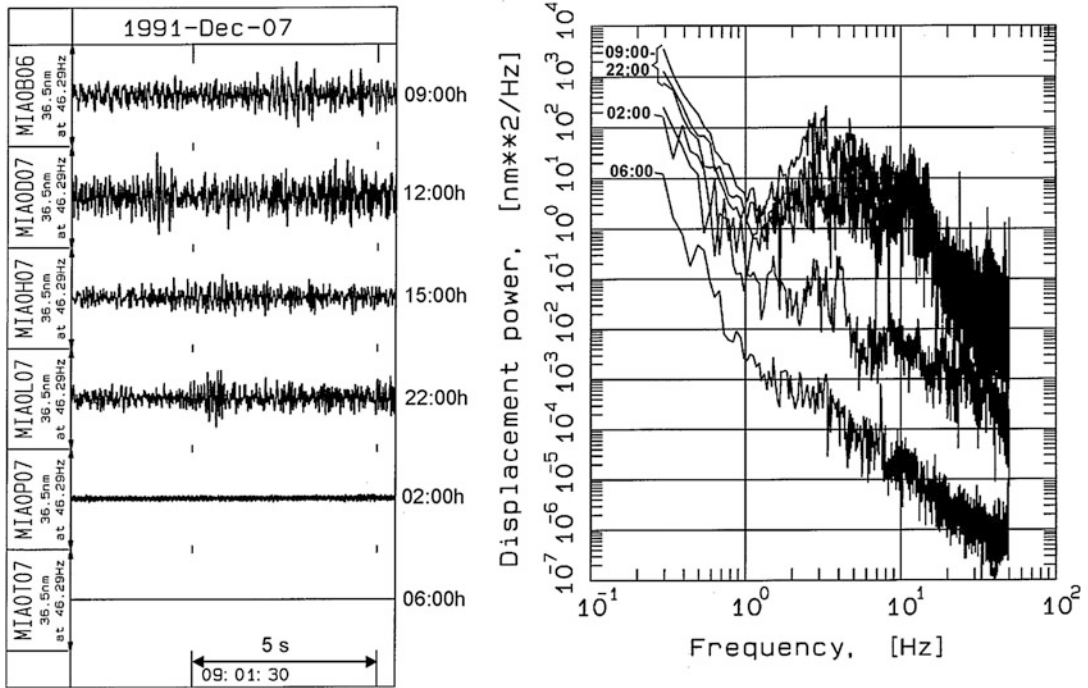
Seismic Noise, Fig. 12 Comparison of relatively quiet sections of vertical component noise records (*left*, without strong transients) and related power spectra (*right*) at a reference site in the district capital town of Miane in NW

Iran. The measurements were made at different times of the day (Copy of Fig. 7.15 from Bormann in Trnkoczy et al. (2002) © IASPEI)

Microseismic Noise in Urban Environments

In urban, densely populated, and industrialized environment both natural and cultural noise components may play a major role in a wide range of frequencies. Measuring and analyzing urban seismic noise (USN) requires broadband recordings and data analysis. Figure 14 compares the noise power density spectrum measured at the former seismological station in the city of Hamburg, Germany, with the spectrum measured on a hard rock site in a quiet Spa village some 70 km away, which is now the alternative station site in the GRSN.

From Fig. 14, the following conclusions can be drawn: While in a remote quiet urban setting the microseismic noise spectrum may decay at night almost monotonously towards higher frequencies, the cultural noise activity may increase the daytime PSD for frequencies above 1 Hz by some 10 dB. In contrast, in a very busy

industrialized city such as Hamburg, the microseismic noise level may be 20–40 dB higher for frequencies of about 0.5–30 Hz and may be reduced during night time by less than 10 dB only between 1 Hz and 10 Hz. In the metropolitan city of Berlin, the "cultural noise" patterns differ from that of HAM (see Fig. 7.31 in Trnkoczy et al. 2011) and may again be different in towns with more pronounced diurnal variability of human and industrial activities than in megacities, as in Fig. 12.

Groos and Ritter (2009) stated that the generally high variability of USN does not allow to characterize a sample time series of USN comprehensively by single measures such as the standard deviation of a noise series or the PSD at a given frequency. For the metropolitan city of Bucharest, Romania, they calculated therefore long-term spectrograms of up to 28 days duration from broadband seismic
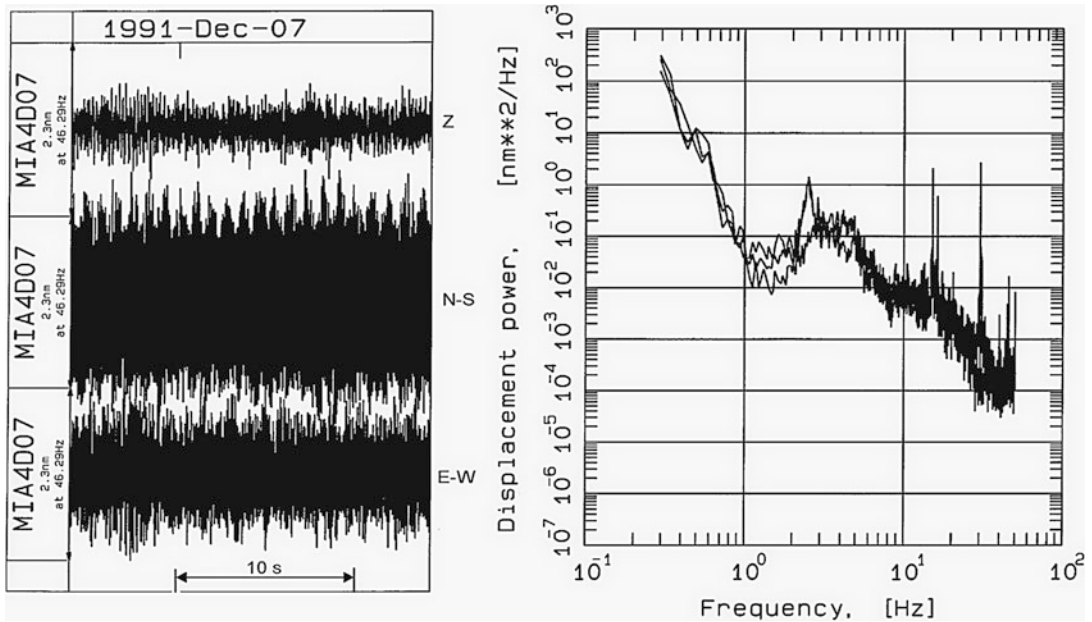
Seismic Noise, Fig. 13 Noise records and related power spectra near to a transformer house and power line. Note the monochromatic spectral lines around 13, 30, and 50–60 Hz, either induced by the AC current frequency and its lower harmonics and/or caused by the vibration of the transformer (Copy of Fig. 7.19 from Bormann in Trnkoczy et al. (2002) © IASPEI)

recordings. Spectrograms, or sonograms, are visual 2D representations of the spectrum of frequencies contained in a seismic, sound, or other signal as they vary with time or with some other variable. The aim was to identify the frequency-dependent behavior of the time-variable processes that contribute to USN. Based on the spectral analysis of the data in eight frequency ranges between 8 mHz and 45 Hz, they proposed a time-domain classification which allows to identify both Gaussian distributed noise time series as well as time series that are dominated by transient or periodic signals due to traffic, rotating machinery, etc. In the case of Bucharest, only some 40 % of the analyzed time series are characterized by Gaussian distributed random noise. The most common deviations are due to large-amplitude transient signals. Moreover, significant variations of the statistical noise properties during daytime were found in the wide frequency range between 0.04 and 45 Hz, pointing to a broadband human influence on USN. The authors recommend the use of such automatically derived broadband information about ambient noise for selecting suitable time windows for H/V noise studies (Parolai 2012) or ambient noise tomography.

Signal and SNR Variations due to Local Site Conditions

Compared to hard rock sites, both noise and signals may be amplified when recorded on soft soil cover. This signal amplification may partly or even fully outweigh the higher noise observed on such sites. Signal strength observed for a given event may vary strongly (up to a factor of about 10–30) within a given array or station network, even if its aperture is much smaller than the epicentral distance to the event (<10–20%), so that differences in backazimuth and amplitude-distance relationship are negligible (Figure 15 and Fig. 4.36 in Bormann and Wielandt 2013). Also, while one station of a network may record events rather weakly from a certain source area, the same station may do as well as other stations
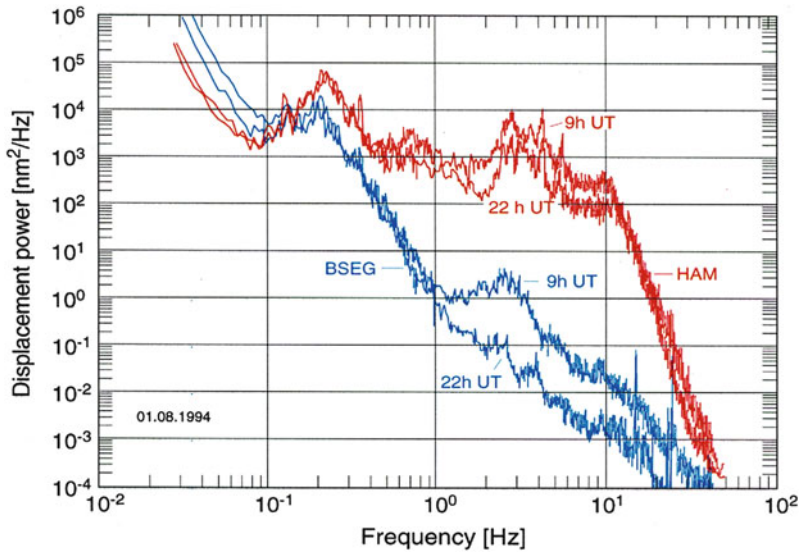
Seismic Noise, Fig. 14 Noise power spectra at HAM (*upper two curves*, *red*) and BSEG (*lower two curves*, *blue*) determined from vertical-component records at daytime and nighttime. Note that the noise reduction is almost negligible for periods longer than 2 s. The drastic reduction at higher frequencies has been achieved both by the reduced cultural noise sources at Bad Segeberg, as compared to the big city of Hamburg, and by installation of the broadband seismograph in an outcropping local hardrock salt-anhydrite diapir with very high contrast of its *acoustic impedance* (= product of wave velocity in the medium times density of the medium) against that of the surrounding alluvial softrock layers. This prevents much of the short-wavelength noise energy to penetrate into this local hard rock anomaly (Color version of Fig. 18 from Bormann et al. 1997 © Springer; with kind permission from Springer Science + Business Media)

(or even better) for events from another region, azimuth, or distance (e.g., station GWS in Fig. 16 left and right, respectively).

Figure 16 compares, for regional and teleseismic events, the short-period P-wave amplitude ratio (left) and SNR (right) of two stations of the German Regional Seismic Network (GRSN). In the same azimuth range, but at different epicentral distances, station BRG may record both > 3 times larger as well as > 3 times smaller amplitudes than station MOX. This corresponds to differences in station magnitude up to one unit! The same applies to the differences in SNR. Therefore, optimal site selection cannot be made only on the basis of noise measurements. Also, the signal conditions at possible alternative sites should be compared.

However, differences in local SNR conditions may become negligible in long-period recordings (Fig. 4.39 in Bormann and Wielandt 2013) and thus play a lesser role in site selection for broadband records. Thermal shielding and reduction of tilt noise then have the highest priority (see Forbriger 2012).

Installations in Subsurface Mines, Tunnels, and Boreholes

As shown in Figs 13–15, short-period microseismic noise is strongly reduced with the depth of sensor installations in mines and boreholes. Withers et al. (1996) found that for rather high frequencies monitoring, as in very local weak motion and industrial seismology measurements between 10 and 20 Hz, the SNR could be improved between 10 and 20 dB and for frequencies between 23 and 55 Hz as much as 20–40 dB by deploying a short-period sensor at only 43 m below the surface. However, depending on geological conditions and related peculiarities of the local noise field as well as the decreasing free-surface amplification effect of incident body waves, the SNR does not necessarily increase steadily with depth. A striking example have been short-period signal and noise measurements
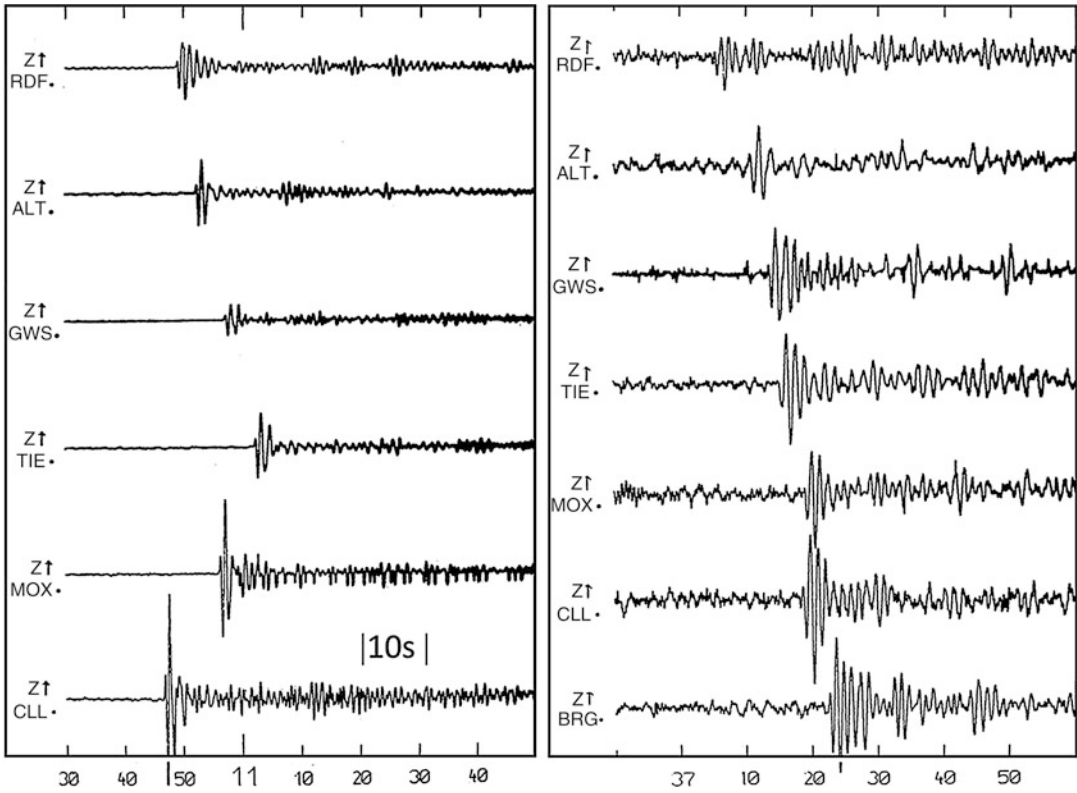
Seismic Noise, Fig. 15 Short-period records of underground nuclear explosions at the test sites of Semipalatinsk (*left*, $D \approx 41° \pm 1°$) and Nevada (*right*, $D \approx 81° \pm 1°$) at stations of the former East German seismic network. Note the differences in signal amplitudes both amongst the stations for a given event and for the same station, when comparing events in different backazimuth and distance (Copy of Fig. 4.35 in Bormann (2002) © IASPEI)
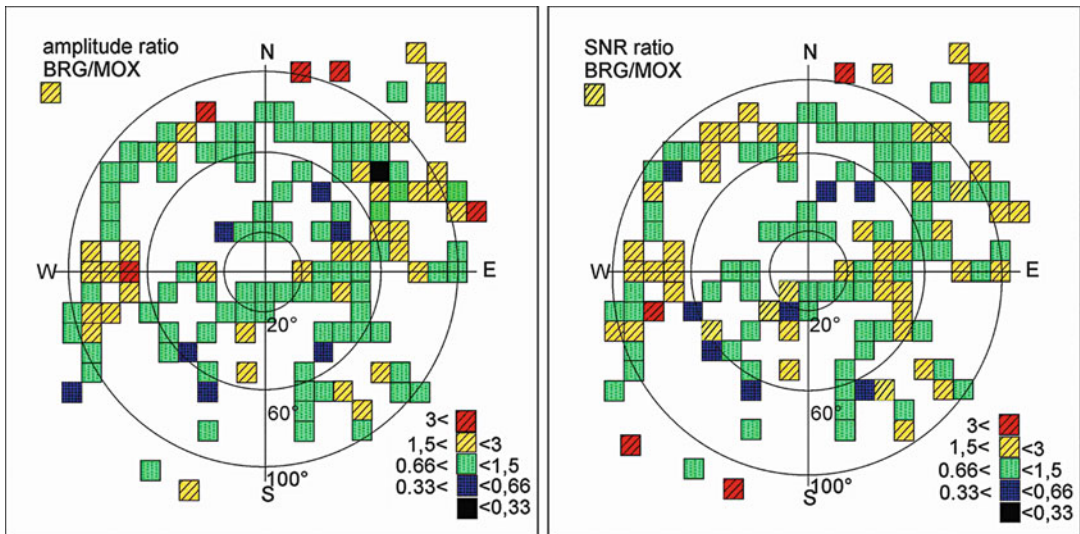
Seismic Noise, Fig. 16 Pattern of the relative short-period P-wave amplitudes (*left*) and of the related SNR (*right*) at station BRG normalized to those of station MOX (170 km apart) in a distance-azimuth polar diagram (Modified version of Fig. 7 from Bormann et al. 1992 © Elsevier; with permission from Elsevier)

in the early 1960s with 1 Hz borehole seismometers in two abounded deep oil drills in the USA (Douze, 1964). In the Texas hole, the average ratio of seismic signal in the borehole to that measured at the surface decreased to 1/10th at about 1,500 m depth and then increased again to ½ of the surface value at 3,000 m depth. In contrast, in an Oklahoma borehole, it dropped to about 1/3 at about 1,000 m depth and then remained roughly constant thereafter. Accordingly, despite the noise reduction with depth, there was no SNR improvement (on average) in the Texas borehole down to about 1,000 m depth, but then the SNR increased to a factor of about 15 at 3,000 m depth. Contrary to this, the SNR increased by a factor of 3 in the Oklahoma borehole within the first 800 m but then remained roughly constant (ranging between 1 and 5) up to 3,000 m depth (see Fig. 4.40 in Bormann and Wielandt 2013).

Nevertheless, generally a significant SNR improvement is to be expected within the first few hundred meters depth. This applies particularly to borehole installations of long-period and broadband sensors which benefit already at much shallower depth greatly from the very stable temperature conditions and strongly reduced atmospheric pressure fluctuations and related tilt noise, also in hard rock tunnels. In mines or boreholes, a depth of 100 m below surface is generally sufficient to achieve most of the practicable reduction by -20 to -30 dB of long-period noise with periods between 30 s and 1,000 s (see Fig. 7.59 in Sect. 7.4.5 of Trnkoczy et al. 2011). Since the cost of drilling and installation increases greatly with depth, no permanent seismic borehole installations deeper than 100–150 m have yet been made. In any event, the borehole should be drilled through the soil or cover of weathered rock and penetrate well into the compacted underlying rock formations. For more information about tunnel and long-period borehole installations, see Sects. 7.4.3 and 7.3.5 in Trnkoczy et al. 2011.

Thus, differences in the frequency spectrum, horizontal wave-propagation velocity, degree of coherence, and depth dependence between (short- and medium-period) microseismic noise and seismic waves allows to improve the signal-to-noise-ratio (SNR) by installing seismic sensors either at reasonable depth below the surface or by way of data processing.

Improving the Signal-to-Noise Ratio by Data Processing

Frequency Filtering

If the dominating frequency content of the seismic signal differs from that of the disturbing noise, then band-pass, high-pass, or low-pass filtering may significantly improve the SNR (Figs. 17 and 18).
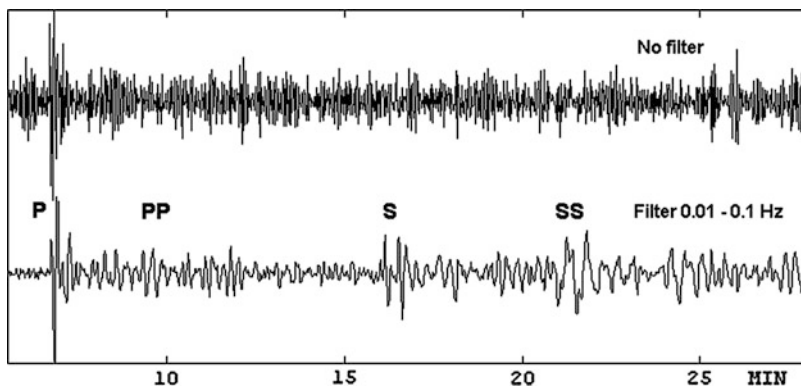
Seismic Noise, Fig. 17 *Top*: Velocity broadband record in the period range between 0.5 and 360 s of a teleseismic earthquake. Only the P-phase might be identified. *Bottom*: On the long-period band-pass filtered trace the signal-to-noise ratio is much improved and several later phases are clearly recognizable (Figure by courtesy of J. Havskov for NMSOP editions © IASPEI)
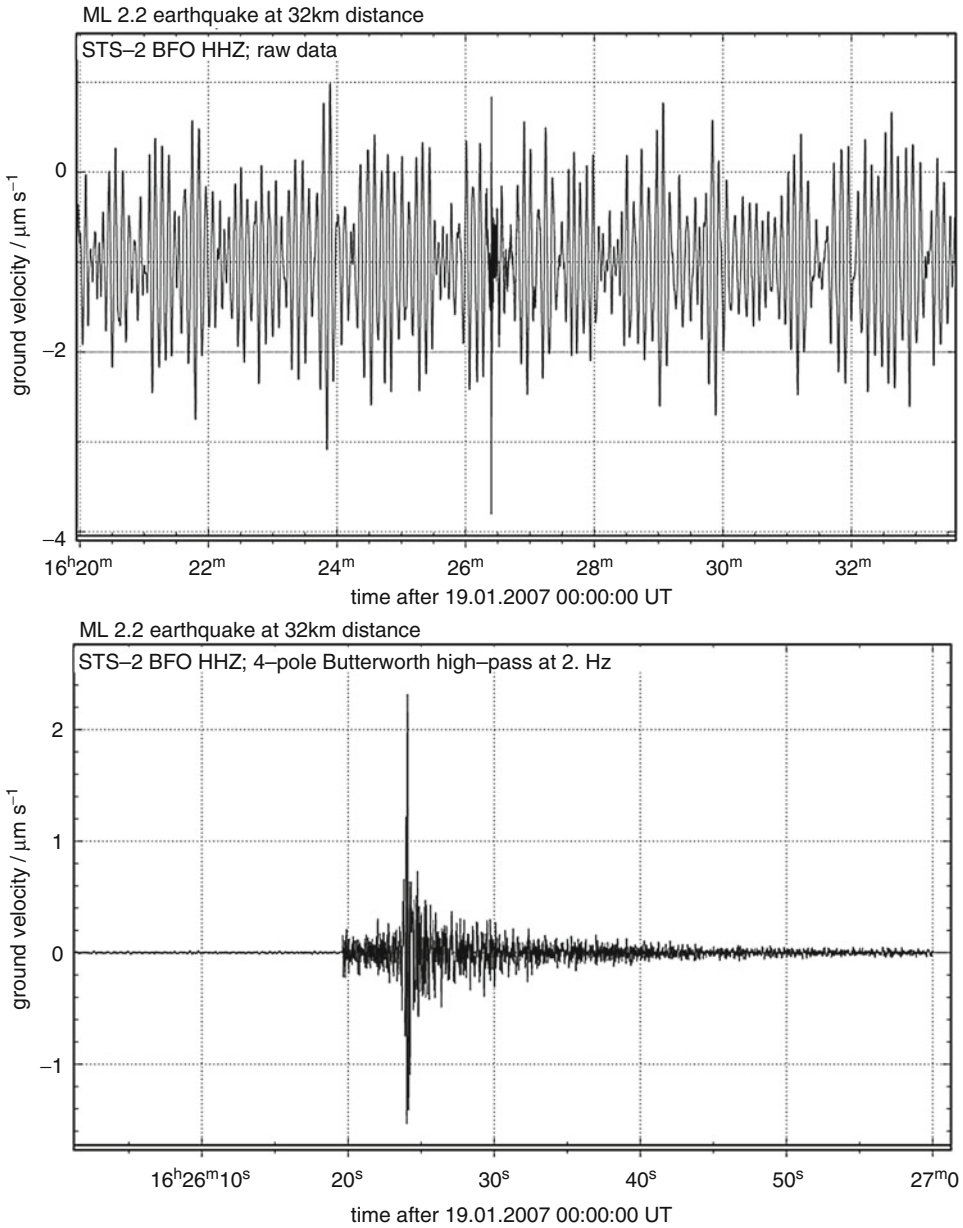
Seismic Noise, Fig. 18 Isolation of a local ML = 2.2 earthquake (*bottom trace*) that occurred 32 km away from the Black Forest Observatory (BFO), Germany, by 2-Hz high-pass filtering of the raw data of a velocity broadband STS-2 record (*top trace*), thus eliminating the dominating 6 s ocean microseisms (Figure by courtesy of Thomas Forbriger for NMSOP Chapter 4 © IASPEI)

Velocity Filtering and Beamforming

Often the dominant signal frequencies may coincide with that of strong noise. Then frequency filtering does not help to improve the SNR. On the other hand, the horizontal propagation velocity of noise, being dominatingly surface waves of Rayleigh-wave type, is much lower than that of P waves and also lower than that of teleseismic S waves with a steep angle of incidence. Then frequency-wavenumber (f-k)
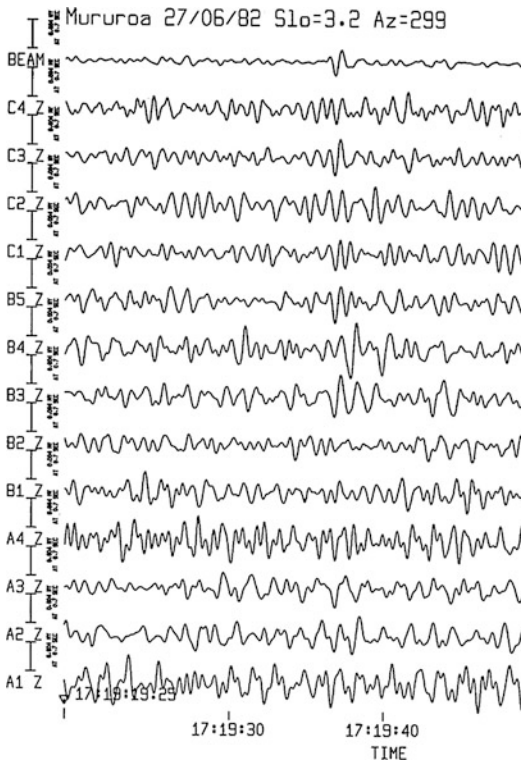
Seismic Noise, Fig. 19 Detection of a weak underground nuclear explosion in the 10 kt range at the Mururoa Atoll test site ($D = 145°$) by beam forming (*top trace*). No signal is recognizable in any of the 13 individual record traces from stations of the Gräfenberg array, Germany (*below*) (From Buttkus 1986 © Schweizerbart http://www.schweizerbart.de/)

filtering is suitable to improve the SNR in the beam trace of seismic array records (Schweitzer et al. 2011). Assuming that the noise within the array is random while the signal is coherent, even a simple direct summation of the n sensor outputs would already produce some modest SNR improvement. When the direction and velocity of travel of a signal through an array is known, it is possible to compensate for the differences in arrival time at the individual sensors and then sum up all the n record traces (beam forming). This increases the signal amplitude by a factor n while the random noise amplitudes increase in the beam trace only by $\sqrt{n}$, thus improving the SNR by $\sqrt{n}$. Figure 19 compares the (normalized) individual records of 13 stations of the Gräfenberg array, Germany, with the beam

trace. A weak underground nuclear explosion at a distance of $143.6°$, not recognizable in any of the single traces, is very evident in the beam trace.

Noise Prediction Error, Polarization, and Other Multi-parameter Methods of Filtering

In near real time, it is possible to use a moving time-window to determine the characteristics of a given noise field by means of cross- and autocorrelation of array sensor outputs. This allows the prediction of the expected random noise in a subsequent time interval. Subtracting the predicted noise time series from the actual record results in a much reduced noise level. Weak seismic signals, originally buried in the noise but not predicted by the noise "forecast" of the prediction-error filter (NPEF), may then stand out clearly. NPEFs have several advantages as compared to frequency filtering:

- No assumptions on the frequency spectrum of noise are required since actual noise properties are determined by the correlation of array sensor outputs.
- While frequency differences between signal and noise are lost in narrowband filtering, they are largely preserved in the case of the NPEF. This may aid signal identification and onset-time picking.
- Signal first-motion polarity is preserved in the NPEF whereas it is no longer certain after narrow-band or zero-phase band-pass filtering.

Three-component recordings allow to reconstruct the ground particle motion and to determine its polarization which may differ between different types of seismic waves and noise. One great advantage of polarization filtering is that it is independent of differences in the frequency and velocity spectrum of signal and noise and thus can be applied in concert with other procedures for SNR improvement.

Early examples of NPEF and polarization filtering have been given in Bormann and Wielandt (2013). For a briefing on modern three-component broadband procedures of SNR

improvement, signal detection, phase identification, and event location that have been implemented in procedures for fully automatic near real-time multi-parameter signal analysis, see Sect. 11.9 in Bormann et al. (2014). These procedures apply cross-correlation methods, use differences in the complexity and polarization properties between signal and noise as well as a compressed signal representation by a sequence of n-dimensional feature vectors which increase the robustness of the procedures. Some of these versions also adapt the signal detector to temporal/seasonal variations in the nature and power of ambient noise.

Summary

The essay defines the term of "seismic noise" and explains why coherent seismic signals and incoherent random noise are not commensurate, thus requiring different mathematical treatment and analysis. Also described are the different causes and ways of analysis of ambient microseismic noise as recorded at stations on land in the very wide frequency range between some mHz and 100 Hz. The article sketches some essential procedures of noise reduction by way of proper sensor shielding and installation, also at depth, as well as feature and problem adapted filtering aimed at improved signal-to-noise ratio and thus signal detection and identification. Many references for further readings are given, amongst them to relevant IASPEI Manual articles accessible via http://nmsop.gfz-potsdam.de.

Cross-References

▶ Downhole Seismometers
▶ Noise-Based Seismic Imaging and Monitoring of Volcanoes
▶ Ocean-Bottom Seismometer
▶ Seismic Event Detection
▶ Seismic Instrument Response, Correction for
▶ Seismic Network and Data Quality
▶ Seismometer Self-Noise and Measuring Methods
▶ Seismometer Arrays
▶ Sensors, Calibration of
▶ Site Response: 1-D Time Domain Analyses
▶ Site Response: Comparison Between Theory and Observation
▶ Stochastic Ground Motion Simulation

References

Aki K, Richards PG (1980) Quantitative seismology – theory and methods, vol II. Freeman, San Francisco

Bormann P (1998) Conversion and comparability of data presentations on seismic background noise. J Seismol 2(1):37–45

Bormann P (2002) Chapter 4: Seismic signals and noise. In: Bormann P (ed) IASPEI new manual of seismological observatory practice (NMSOP), vol 1. GeoForschungsZentrum, Potsdam, 34 pp

Bormann P, Wielandt E (2013) Chapter 4: Seismic signals and noise. In: Bormann P (ed) New manual of seismological observatory practice (NMSOP-2) (2012). IASPEI, GFZ German Research Centre for Geosciences, Potsdam, 63 pp. doi:10.2312/GFZ. NMSOP-2_CH4. http://nmsop.gfz-potsdam.de

Bormann P, Wylegalla K, Strauch W, Baumbach M (1992) Potsdam seismological station network: processing facilities, noise conditions, detection threshold and localization accuracy. Phys Earth Planet Int 69:311–321

Bormann P, Wylegalla K, Klinge K (1997) Analysis of broadband seismic noise at the German Regional Seismic Network and search for improved alternative station sites. J Seismol 1:357–381

Bormann P, Engdahl ER, Kind R (2012) Chapter 2: Seismic waves and Earth models. In: Bormann P (ed) New manual of seismological observatory practice (NMSOP-2) (2012). IASPEI, GFZ German Research Centre for Geosciences, Potsdam, 105 pp. doi:10.2312/GFZ.NMSOP-2_CH2. http://nmsop.gfz-potsdam.de

Bormann P, Klinge K, Wendt S (2014) Chapter 11: Data analysis and seismogram interpretation. In: Bormann P (ed) New manual of seismological observatory practice (NMSOP-2) (2012). IASPEI, GFZ German Research Centre for Geosciences, Potsdam, doi:10.2312/GFZ.NMSOP-2_CH11. http://nmsop.gfz-potsdam.de

Broding RA, Bentley-Llewellyn NJ, Hearn DP (1964) A study of three dimensional seismic detection system. Geophysics 29(2):221–249

Buttkus B (ed) (1986) Ten years of the Gräfenberg array: defining the frontiers of broadband seismology. Geologisches Jahrbuch E-35, Hannover, 135 pp

Cessaro RK (1994) Sources of primary and secondary microseisms. Bull Seismol Soc Am 84(1):142–148

Douze EJ (1964) Signal and noise in deep wells. Geophysics 29(5):721–732

Forbriger Th (2012) IS 5.4: Recommendations for seismometer deployment and shielding. In: Bormann P (ed) New manual of seismological observatory practice (NMSOP-2) (2012). IASPEI, GFZ German Research Centre for Geosciences, Potsdam, 10 pp. doi:10.2312/GFZ.NMSOP-2_IS_5.4. http://nmsop.gfz-potsdam.de

Friedrich A, Klinge K, Krüger F (1998) Ocean-generated microseismic noise located with the Gräfenberg array. J Seismol 2:47–64

Fyen J (1990) Diurnal and seasonal variations in the microseismic noise observed at the NORESS array. Phys Earth Planet In 63:252–268

Groos JC, Ritter JRR (2009) Time domain classification and quantification of seismic noise in an urban environment. Geophys J Int 179:1213–1231

Hasselmann KA (1963) A statistical analysis of the generation of microseisms. Rev Geophys 1:177–209

Haubrich RA, Munk WH, Snodgrass FE (1963) Comparative spectra of microseisms and swell. Bull Seismol Soc Am 53:27–37

Longuet-Higgins MS (1950) A theory for the generation of microseisms. Phil Trans R Soc Lond A 243:1–35

Parolai St (2012) Chapter 14: Investigation of site response in urban areas by using earthquake data and seismic noise. In: Bormann P (ed) New manual of seismological observatory practice (NMSOP-2) (2012). IASPEI, GFZ German Research Centre for Geosciences, Potsdam, 38 pp. doi:10.2312/GFZ.NMSOP-2_CH14. http://nmsop.gfz-potsdam.de

Peterson J (1993) Observations and modeling of seismic background noise. U.S. Geological Survey Open-File Report, 93–322, 95 pp

Schweitzer J, Fyen J, Mykkeltveit S, Gibbons SJ, Pirli M, Kühn D, Kværna T (2011) Chapter 9: Seismic arrays. In: Bormann P (ed) (2012) New manual of seismological observatory practice (NMSOP-2) (2012). IASPEI, GFZ German Research Centre for Geosciences, Potsdam, 80 pp. http://nmsop.gfz-potsdam.de

Shinohara M, Suyehiro K, Shiobara H (2011) Chapter 7, section 7.5: Marine seismic observation. In: Bormann P (ed) (2012) pp 112–135. http://nmsop.gfz-potsdam.de

Trnkoczy A, Bormann P, Hanka W, Holcomb LG, Nigbor RL (2002) Chapter 7: Site selection, preparation and installation of seismic stations. In: Bormann P (ed) IASPEI New Manual of Seismological Observatory Practice (NMSOP), vol 1. GeoForschungsZentrum Potsdam, Potsdam, 108 pp

Trnkoczy A, Bormann P, Hanka W, Holcomb LG, Nigbor RL, Shinohara M, Suyehiro K, Shiobara H (2011) Chapter 7: Site selection, preparation and installation of seismic stations. In: Bormann P (ed) New manual of seismological observatory practice (NMSOP-2) (2012). IASPEI, GFZ German Research Centre for Geosciences, Potsdam. doi:10.2312/GFZ.NMSOP-2_CH7. http://nmsop.gfz-potsdam.de

Webb SC (2002) Seismic noise on land and on the sea floor. In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) International handbook of earthquake and engineering seismology, Part A. Academic, Amsterdam, pp 305–318

Wiechert E (1904) Verhandlungen der zweiten internationalen Seismologischen Konferenz. Beitr Geophys, Ergänzungsband II, pp 41–43

Wielandt E (2011) Chapter 5: Seismic sensors and their calibration. In: Bormann P (ed) New manual of seismological observatory practice (NMSOP-2) (2012). IASPEI, GFZ German Research Centre for Geosciences, Potsdam, 51 pp. doi:10.2312/GFZ.NMSOP-2_CH5. http://nmsop.gfz-potsdam.de

Withers MM, Aster RC, Young CJ, Chael EP (1996) High-frequency analysis of seismic background noise as a function of wind speed and shallow depth. Bull Seismol Soc Am 86(5):1507–1515

Young CJ, Chael EP, Withers MW, Aster RC (1996) A comparison of the high- frequency (>1Hz) surface and subsurface noise environment at three sites in the United States. Bull Seismol Soc Am 86(5):1516–1528

Seismic Reliability Assessment, Alternative Methods for

Hossein Ebrahimian and Raffaele De Risi
Department of Structures for Engineering and Architecture, University of Naples Federico II, Naples, Italy

Synonyms

Bayesian method; Fragility curve; Limit state; Mean annual frequency of exceedance; Monte Carlo simulation; PEER performance-based approach; Reliability assessment; Second-moment reliability methods; Simulation-based reliability; System reliability

A Perspective into the Seismic Reliability Assessment

The treatment of structural behavior under seismic actions can be considered as one of the most challenging aspects of structural performance assessment and design. It encompasses numerous

sources of uncertainty associated with the seismic action, structural response, and structural capacities. The various parameters involved in the seismic performance assessment problem can be generically designated in terms of seismic *demand* (*D*) and *capacity* (*C*) of the structure. The seismic demand and capacity can be characterized as a functional (i.e., function of function) of a set of *uncertain parameters* (also denoted as "random variables"). Hence, probabilistic methods are necessary in order to take into account the uncertainties in demand and capacity within the performance assessment or design.

Traditionally, seismic-design codes have addressed the uncertainties by allowing some degree of conservatism in evaluating demand and capacity at the level of structural components; nevertheless, the link to the overall performance of the structure remains unclear. Strictly speaking, the reliability of the structure to withstand future events remains more-or-less unknown to the designer while employing various established code-based approaches. Therefore, in the past decades, several research efforts have been carried out and substantial progress has been made towards the inclusion of various sources of uncertainty into structural performance assessment and design frameworks.

This entry introduces the main sources of uncertainty in demand and capacity, as well as their treatments through alternative reliability methods. In this regard, three general families of methods are briefly summarized in this entry: the second-moment reliability methods, the simulation-based reliability approaches, and finally the probabilistic performance-based methodology developed within the joint efforts of Pacific Earthquake Engineering Research Center (PEER). The closing section provides a brief discussion about future challenges in seismic reliability assessment.

An Introduction to Seismic Reliability Assessment

Seismic structural reliability is concerned with the probabilistic performance evaluation of structures under seismic actions. This issue is addressed by means of defining a prescribed set of *limit states*, which describe specific stages of structural behavior and its consequences (e.g., states associated with consequences in terms of costs, loss of lives, impact on the environment, or serviceability; see Bozorgnia and Bertero 2004). Reaching a limit state denotes that an unacceptable behavior for the structure occurs. This condition is generically referred to as *failure*. The limit state exceedance is verified by comparing a measurable quantity, associated with the structural response, denoted as the *demand D* (also referred to as the engineering demand parameter (EDP), e.g., maximum inter-story drift ratio, maximum chord rotation, maximum component shear, etc.), with the corresponding limit value, denoted as the *capacity C* (e.g., yielding rotation, ultimate rotation, shear strength, etc.). As a result, the inequality $D \leq C$, also called *safety margin formulation*, defines acceptable performance respect to the current limit state in mathematical terms (Pinto et al. 2007). This formulation is generally expressed as *limit state function* (or *performance function*) $G = C\text{-}D$. This type of formulation is also employed in the *load and resistance factor design* (LRFD) format (Galambos et al. 1982), which is familiar to engineers as the safety checking equation. As an alternative to the safety margin formulation, the *safety factor formulation* is introduced as $G = \ln(C/D)$ (DNV 1992). For both the *safety margin formulation* and the *safety factor formulation*, failure is defined as $G \leq 0$.

As mentioned before, the evaluation of the structural performance in terms of demand and capacity may encompass several sources of uncertainty. In such context, the performance with respect to a given limit state is evaluated in terms of the probability that the limit state function will be less than zero (*probability of failure P_F*) in a specified reference time period. Identifying $\mathbf{X} = [X_1, X_2, \ldots, X_n]$ to be the group of *n* uncertain parameters involved in the reliability problem, the probability of failure may be determined by the following integral:

$$P_F = P[G(\mathbf{X}) \leq 0] = \int_{G(\mathbf{X}) \leq 0} f_{\mathbf{X}}(\mathbf{X}) \mathrm{d}\mathbf{x} \quad (1)$$

where $f_{\mathbf{X}}$ is the joint probability distribution for uncertain parameters $\mathbf{X}$. This means that one can evaluate the probability of failure by integrating the joint density function over the failure domain defined as $G(\mathbf{X}) \leq 0$.

Conventionally, the *generalized reliability* (or *safety*) *index* β associated with a given limit state function is defined by one-to-one mapping relationship with failure probability as follow:

$$\beta = -\Phi^{-1}(P_F) \quad (2)$$

where $\Phi(\cdot)$ is the standard normal Cumulative Distribution Function (CDF). Equation 2, however, does not mean to imply that uncertain parameters are jointly normal. The inverse of the standard normal CDF simply provides a convenient one-to-one mapping between the computed probability of failure and a reliability index.

The evaluation of the integral in Eq. 1 can be computationally difficult; some examples are as follows: $f_{\mathbf{X}}$ is often not well-defined because of the incompleteness of the statistical information available; $G(\mathbf{X})$ may have a nonlinear form; the computation of the multifold integral can be very difficult if the number of uncertain parameters is high. Various methods have been proposed for solving the integral form in Eq. 1. These approaches range from the classical moment methods for structural reliability (e.g., first-order second-moment reliability method) to the simulation-based approaches (i.e., Monte Carlo family of methods), and also the PEER approach, which is quite different compared to the other two techniques. In this entry, alternative methods for estimating the probability of failure are described.

Types of Uncertainty

There are alternative ways for classifying the sources of uncertainty. Arguably, such classification depends to some extent on the reliability method adopted. Some classical reliability methods provide the following classification of uncertainties in seismic structural engineering: (1) *inherent randomness* due to the intrinsic variability in material mechanical properties as well as environmental actions (i.e., loads or seismic excitation); (2) *statistical uncertainties* due to the lack of knowledge associated with the estimation of the parameters of probability distributions from observational samples of limited size; and finally (3) *model uncertainties* due to the imperfections of the mathematical models used to describe complex physical phenomena (i.e., model describing load or element capacity) (see also Der Kiureghian 1996; Ditlevsen and Madsen 1996). Some other reliability methods (mostly simulation based) classify the sources of uncertainty into two general categories: those related to the representation of the seismic action and those related to the structural modeling (see, e.g., Papadimitriou et al. 2001). Another useful classification is based on the definition of the *aleatory* and *epistemic* uncertainties (Cornell et al. 2002; Der Kiureghian and Ditlevsen 2009). These two categories are also known as *inherent* randomness (i.e., uncertainty impossible to mitigate) and *statistical* uncertainties (i.e., uncertainties that can be mitigated with reasonable effort), respectively (McGuire 2004). On the other hand, in the Bayesian probabilistic framework, the various sources of uncertainty are treated in a unified manner. In subsequent sections, the uncertainties in the seismic reliability problems are divided (if required) into those related to the representation of the seismic action, also referred to as the record-to-record (RTR) variability, and other sources of uncertainty. As mentioned above, the set of uncertain parameters involved in the seismic performance assessment or design problem are denoted hereafter by the vector of parameters $\mathbf{X}$.

Structural Component and Structural System Reliability

A *structural component* is a basic structural element whose performance is defined by a single limit state function $G(\mathbf{X}) = 0$ (Fig. 1a). In the hyperspace of the designated uncertain parameters $\mathbf{X}$, the limit state function represents the

S
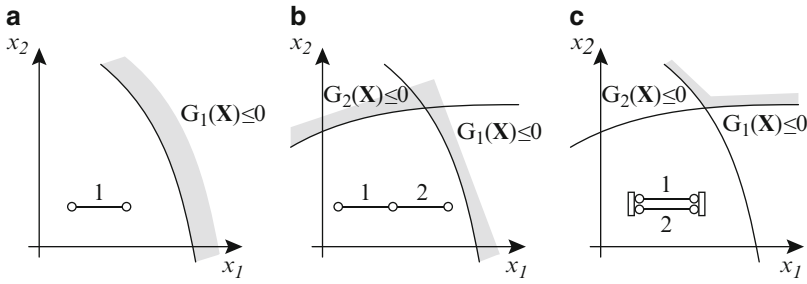
Seismic Reliability Assessment, Alternative Methods for, Fig. 1 Reliability problems for the special case of only two uncertain parameters x_1 and x_2: (**a**) Component, (**b**) series system, (**c**) parallel system. The *gray* shaded area represents the side of the failure domain

boundary (i.e., *limit state surface*) between safe and failure zones (the latter is also called the *failure domain*), as shown in Fig. 1a for the special case of only two uncertain variables (Ditlevsen and Madsen 1996).

A *structural system* is an assemblage of components, whose performance is described by individual component limit state function. Therefore, in a structural system reliability problem, the failure domain may be described as the union and/or intersection of several limit state surfaces. In particular, a *series system* (also known as the *weakest-link*) reliability problem is defined as the union of the failure domains (Fig. 1b), and a *parallel system* reliability problem is defined as the intersection of the failure domains (Fig. 1c). In analytic terms, the probability of failure for series and parallel systems with n components is defined, respectively, as

$$P_{F,\text{series}} = P\left[\bigcup_{i=1}^{n}\left(G_i(\mathbf{X}) \leq 0\right)\right] \quad P_{F,\text{parallel}}$$
$$= P\left[\bigcap_{i=1}^{n}\left(G_i(\mathbf{X}) \leq 0\right)\right] \qquad (3)$$

A cut-set is made as a combination of series and parallel systems generating a *generalized system* reliability problem. A cut-set (Fig. 2) is any set of components whose joint failure represents a failure of the system (Ditlevsen and Madsen 1996; Au and Beck 2003b; Jalayer et al. 2007a). Analytically, the probability of failure for a cut-set system is evaluated as
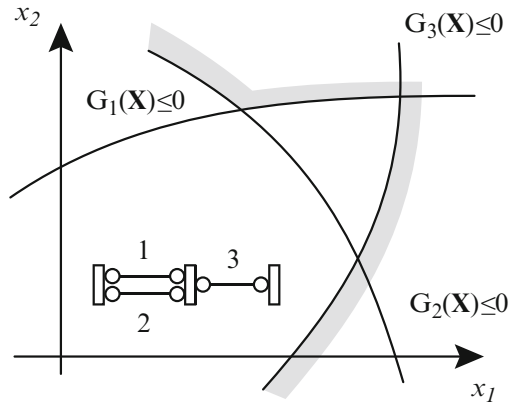
Seismic Reliability Assessment, Alternative Methods for, Fig. 2 Reliability problems for the special case of only two uncertain parameters x_1 and x_2: cut-set system. The *gray* area represents the side of the failure domain

$$P_{F,\text{cut-set}} = P\left[\bigcup_{j=1}^{N}\bigcap_{i=1}^{n_j}\left(G_i(\mathbf{X}) \leq 0\right)\right] \qquad (4)$$

where N defines the associated number of the subsystem that is connected in series to the other subsystems and n_j is the number of components connected in parallel belonging to the jth subsystem.

Seismic Reliability Assessment, a Time Variant Nonlinear Problem

If the limit state function reveals time-dependent properties, then the reliability problem can be referred to as *time variant*. Hence, given the stochastic nature of the seismic excitation, the seismic reliability assessment can be classified as a

time-variant problem. Therefore, the failure event constitutes the outcrossing of the time-variant limit state surface. This represents the so-called *first-passage* (or *first-excursion*) problem (e.g., Au and Beck 2001b). Moreover, in many practical cases, the limit state function is nonlinear. Therefore, the seismic reliability problems can be classified, in general, as time-variant nonlinear problems.

The Second-Moment Reliability Methods

In general, the nth moment of a mono-variate distribution (i.e., a distribution with a single scalar random variable) is defined as the expected value of the variable raised to the power of n. Using moments of finite order for describing the uncertainty in given parameter is a well-known problem in statistics referred to as the *moment methods*. A comprehensive review of the *moment methods* in the structural reliability analysis is presented in Zhao and Ono (2001). The second-moment reliability methods can be dated back to the 1960s (some historical notes can be found in Ditlevsen and Madsen 1996). Since then, this family of methods has been refined and extended significantly. At present, they represent one of the most important and widespread approaches for seismic reliability evaluation of structures. As reflected by their title, the second-moment reliability methods employ the first two moments of the vector of uncertain parameters $\mathbf{X}$ (i.e., expected value $\mathbb{E}[\mathbf{X}]$ and covariance $\mathbf{\Sigma_X}$).

Linear (First-Order) Limit State Function and Its First Second Moments

The simplest mathematical form for a performance function is the linear (first-order) polynomial:

$$G(\mathbf{X}) = a_0 + a_1 X_1 + \ldots + a_n X_n = a_0 + \boldsymbol{a}^T \mathbf{X} \tag{5}$$

where the coefficients a_0 and $\boldsymbol{a} = [a_1, \ldots, a_n]$ are constants. It is possible to write the two central

moments of $G(\mathbf{X})$ in terms of the first two (central) moments of the vector of uncertain parameters X:

$$\mu_G = \mathbb{E}[G(\mathbf{X})] = a_0 + \boldsymbol{a}^T \mathbb{E}[\mathbf{X}] = a_0 + \boldsymbol{a}^T \boldsymbol{\mu_X}$$

$$\sigma_G^2 = \mathbb{E}\left[(G(\mathbf{X}) - \mu_G)^2 \right]$$

$$= \boldsymbol{a}^T \mathbb{E}\left[(\mathbf{X} - \boldsymbol{\mu_X})(\mathbf{X} - \boldsymbol{\mu_X})^T \right] \boldsymbol{a} = a^T \boldsymbol{\Sigma_X} a \tag{6}$$

One of the simplest forms of reliability measure was proposed by Cornell (1969), in which the second-moment reliability index is defined as the ratio between the mean value and the standard deviation of the performance function G (i.e., the so-called "Cornell index"):

$$\beta = \frac{\mu_G}{\sigma_G} \tag{7}$$

Assuming that the unacceptable performance is reached when $G \leq 0$, it is possible to establish a one-to-one mapping between the Cornell index and the probability of failure, as already shown in Eq. 2:

$$P_F = P[G(\mathbf{X}) \leq 0]$$

$$= P\left[\frac{G(\mathbf{X}) - \mu_G}{\sigma_G} \leq \frac{0 - \mu_G}{\sigma_G} \right]$$

$$= \Phi\left(-\frac{\mu_G}{\sigma_G} \right) = \Phi(-\beta) \tag{8}$$

This mapping is exact if the variables are normal. Otherwise, Eq. 8 returns an approximate value of the failure probability.

First-Order Second-Moment (FOSM) Method in Component Reliability

In the case of a nonlinear limit state surface, the reliability index fails to be constant under different but mechanically equivalent formulations of the performance function. The issue of a lack of invariance was first recognized by Ditlevsen (1973) and later resolved by Hasofer and Lind (1974). In order to overcome this issue, they proposed the transformation of the uncertain
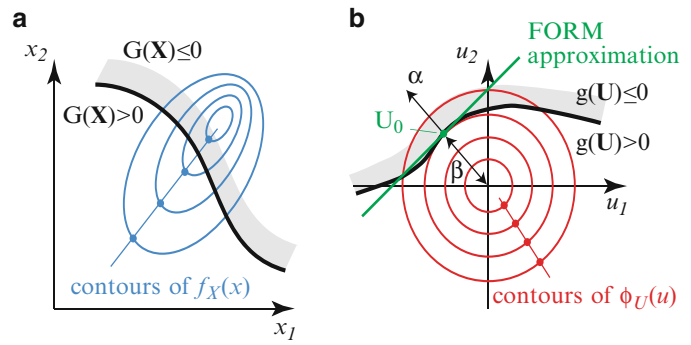
parameters $\mathbf{X}$ into a set of uncorrelated standard normal deviates (variables) $\mathbf{U}$ (i.e., $\boldsymbol{\mu}_{\mathbf{U}} = \mathbf{0}$, $\boldsymbol{\Sigma}_{\mathbf{U}} = \mathbf{I}$). To this aim, the basic random variables, $\mathbf{X}$, can be transformed into a set of uncorrelated standard variates $\mathbf{U}$ through the linear transformation:

$$\mathbf{U} = \mathbf{L_X}^{-1}\mathbf{D_X}^{-1}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{x}}) \quad \text{or} \quad \mathbf{X} = \mathbf{D_X}\mathbf{L_X}\mathbf{U} + \boldsymbol{\mu}_{\mathbf{X}} \tag{9}$$

where $\mathbf{D_X}$ is diagonal matrix of the standard deviations of the basic random variables and $\mathbf{L_X}$ is the lower triangular matrix obtained as the Cholesky decomposition of the correlation coefficient matrix $\mathbf{R_X}$ (i.e., $\mathbf{R_X} = \mathbf{L_X}\mathbf{L_X}^T$). It is then possible to define the joint-normal PDF $\phi_{\mathbf{U}}(\boldsymbol{u})$ and $g(\mathbf{U}) = G[\mathbf{X}(\mathbf{U})]$ as the joint-normal distribution of the uncertain parameters and the limit state function, respectively, transformed in the space of the standard uncorrelated variables (Fig. 3). By considering a linear performance surface (see Eq. 5), $g(\mathbf{U})$ can be expressed as

$$\begin{aligned} g(\mathbf{U}) = G[\mathbf{X}(\mathbf{U})] &= a_0 + \mathbf{a}^T(\mathbf{D_X}\mathbf{L_X}\mathbf{U} + \boldsymbol{\mu}_{\mathbf{X}}) \\ &= a_0 + \mathbf{a}^T\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{a}^T\mathbf{D_X}\mathbf{L_X}\mathbf{U} = b_0 + \mathbf{b}^T\mathbf{U} \end{aligned} \tag{10}$$

It can be easily demonstrated that the first and second moments of this performance function are $\mu_g = b_0$ and $\sigma^2_g = \mathbf{b}^T\mathbf{b}$. Substituting these two terms in Eq. 7, the reliability index in the standard space can be obtained as follows:

$$\beta = \frac{\mu_g}{\sigma_g} = \frac{b_0}{\sqrt{\mathbf{b}^T\mathbf{b}}} \tag{11}$$

The reliability index in Eq. 11 can be interpreted as the distance from the origin to the hyperplane (in standard normal space) defined by $b_0 + \mathbf{b}^T\mathbf{U} = 0$. In other words, the reliability index for a linear limit state function is numerically equivalent to the distance from the origin to the limit state hyperplane in the space of the standard normal variables. Using this geometric interpretation, Hasofer and Lind (1974) defined the reliability index as the minimum distance between the origin of standard normal space and the function $g(\mathbf{U})$ (Fig. 4a).

The point $\mathbf{U}_0$ belonging to the limit state surface that has the minimum distance from the origin of standard normal space is called *design point* (or *performance point*). The latter is the most likely failure point. Strictly speaking, $\mathbf{U}_0$ is the vector of uncertain parameters that has the maximum contribution towards the failure probability.

As also mentioned before, the limit state function is generally nonlinear. Therefore, the function $g(\mathbf{U})$ usually cannot be characterized by a first-order polynomial. An approximate way to solve this problem is to replace the nonlinear function with a first-order Taylor series expansion. In other words, the performance surface, in the neighborhood of the design point $\mathbf{U}_0$, is approximated by the tangent hyperplane at $\mathbf{U}_0$ (Fig. 4b). In analytical terms, the approximate limit state function $g_F(\mathbf{U})$ becomes

$$g_F(\mathbf{U}) = g(\mathbf{U}_0) + \nabla g^T(\mathbf{U}_0) \cdot (\mathbf{U} - \mathbf{U}_0) \tag{12}$$

where $\nabla g(\mathbf{U}_0)$ is the gradient of g evaluated at the performance point and the subscript F represents
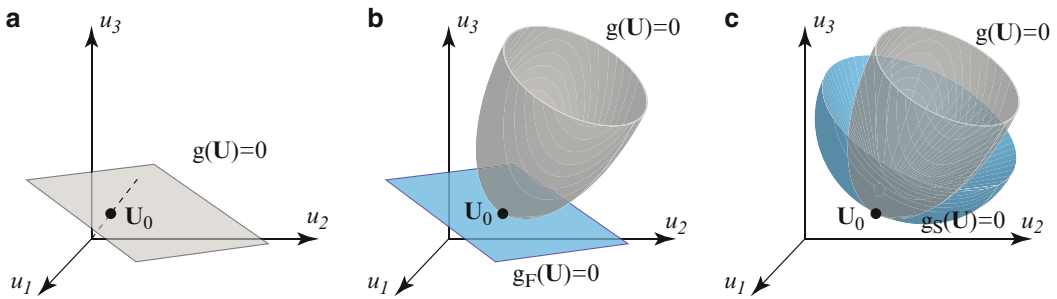
Seismic Reliability Assessment, Alternative Methods for, Fig. 4 Different types of limit state surface and their approximations: (**a**) linear, (**b**) nonlinear with linear approximation, $g_F(U) = 0$ at performance point, (**c**) nonlinear surface approximated with a paraboloid, $g_S(U) = 0$

the first-order approximation to g. In particular, the first two moments of $g(U)$ can be calculated as follows (by using the relationships in Eq. 6):

$$
\begin{aligned}
\mu_g &= g(U_0) + \nabla g^T(U_0)(\boldsymbol{\mu}_U - U_0) \\
&= -\nabla g^T(U_0)(\boldsymbol{\mu}_U - U_0)\sigma_g^2 \\
&= \nabla g^T(U_0)\boldsymbol{\Sigma}_U\nabla g(U_0) \\
&= \nabla g^T(U_0)\nabla g(U_0)
\end{aligned} \tag{13}
$$

where $g(U_0)$ is equal to zero since U_0 is on the limit state function, and the mean is equal to zero ($\boldsymbol{\mu}_U = 0$) and the covariance is equal to one ($\boldsymbol{\Sigma}_U = I$) by definition. For a nonlinear performance function, the first-order reliability index can be defined as equal to the distance of the linearized failure surface (at the performance point U_0) to the origin of the standardized uncorrelated variables (Fig. 3b). With reference to Eq. 12, it is possible to calculate the first two moments of the linearized limit state function (according to Eq. 13) and obtain the following reliability index (in literature also referred to as first-order second-moment reliability index β_{FOSM} or Hasofer-Lind reliability index β_{HL}):

$$
\beta = \frac{\mu_g}{\sigma_g} = \frac{-\nabla g^T(U_0) \cdot U_0}{\sqrt{\nabla g(U_0)^T \nabla g(U_0)}} = \boldsymbol{\alpha}^T \cdot U_0 \quad (14)
$$

where $\boldsymbol{\alpha}$ is the unit normal (directed towards the failure region) on the hyperplane that is tangent to the limit state surface at the performance

point U_0 (Fig. 3b). The components of $\boldsymbol{\alpha}$ are called the sensitivity factors of the reliability index β with respect to uncertain parameters U. The smaller is α_i, the smaller is the contribution of the related uncertain parameter U_i to the failure probability assessment.

Clearly, the determination of β requires the finding of the design point U_0. Therefore, one has to solve the following constrained optimization problem:

$$
\begin{cases}
\beta = \min|U| \\
g(U) = 0
\end{cases} \tag{15}
$$

There are many algorithms that are able to solve this kind of problems (see, e.g., Liu and Der Kiureghian 1991). The most common algorithm used in structural reliability is the one developed by Rackwitz and Flessler (1978). A comprehensive description of the algorithm and relevant examples can be found in Pinto et al. (2007), while a first review of the efficiency of the available algorithms was presented in Liu and Der Kiureghian (1991). Once β is determined, the probability of failure can be approximated by Eq. 8.

Extension to System Reliability

It is noteworthy that in order to compute a system reliability problem, one must be able to compute (or approximate) the probability of a union of N events (series system) or the probability of an intersection of N events (parallel systems). The probability of failure for these two cases

S

can be approximated as (Hohenbichler and Rackwitz 1983)

$$P_{F,\text{series}} \approx 1 - \Phi_N(\boldsymbol{\beta}, \mathbf{R}) \quad P_{F,\text{parallel}} \approx \Phi_N(-\boldsymbol{\beta}, \mathbf{R}) \tag{16}$$

where Φ_N is the multi-normal CDF, $\boldsymbol{\beta}$ is the N-vector of the reliability indexes associated with the structural components, and $\mathbf{R}$ is the $N \cdot N$ correlation matrix, in which the correlation coefficients, $\rho_{ij} = \alpha_i^T \alpha_j$, quantify the correlation between failure modes i and j. In general, the evaluation of $\Phi_N(\boldsymbol{\beta}, \mathbf{R})$ and $\Phi_N(-\boldsymbol{\beta}, \mathbf{R})$ in Eq. 16 is not trivial; however, an efficient Monte Carlo simulation algorithm has been developed for this purpose (Ambartzumian et al. 1998). The previous results can be combined to estimate the probability of failure of a general system (Ditlevsen and Madsen 1996).

FORM and SORM Method in Component Reliability

For the reliability index presented so far, only the two moments of each uncertain parameter have been considered. However, in many applications, information beyond the second moments is available. The methods that make use of that additional information are called *full distribution reliability methods*. If the probability distribution is available for some or all the basic variables, it would be sensible to incorporate it in the reliability analysis. If the joint probability distribution for the vector of uncertain parameters in the original space is not normal and the variables are not independent, the multifold integral presented in Eq. 1 becomes very hard to solve. In fact, several efficient methods for approximating the integral take advantage of the special properties of the standard normal distribution of uncorrelated variables. Therefore, the transformation of independent/dependent non-normal uncertain parameters to independent joint-normal uncertain parameters is quite important.

Historically, the transformation approach was developed as an extension of the FOSM and is

therefore often called *advanced* or *extended* FOSM, since the limit state function is approximated with a first-order approximation, while the probability distributions are no longer approximated only with the first two central moments. In particular, if the limit state function is approximated with its first-order Taylor expansion, these methods are known as FORM (*First-Order Reliability Methods*), and the analytical relations to calculate the reliability index are similar to those reported from Eqs. 12 to 15. If the limit state function is approximated with its second-order Taylor expansion, these methods are known as SORM (*Second-Order Reliability Methods*). In both cases, the Taylor expansion is done at the *design point* $\mathbf{U}_0$.

A SORM analysis reduces the error associated with the non-flatness of the limit state function at the design point in a FORM approximation. In particular, this approach fits a second-order surface (usually a paraboloid) at the performance point as shown in Fig. 4c. A parabolic approximation to the limit state surface at the design point is

$$g_S(\mathbf{U}) = g(\mathbf{U}_0) + \nabla g^T(\mathbf{U}_0) \cdot (\mathbf{U} - \mathbf{U}_0)$$
$$+ \frac{1}{2}(\mathbf{U} - \mathbf{U}_0)^T \boldsymbol{H}(\mathbf{U}_0)(\mathbf{U} - \mathbf{U}_0) \tag{17}$$

where $\boldsymbol{H}(\mathbf{U}_0)$ is the Hessian matrix of g evaluated at $\mathbf{U}_0$ and the subscript S denotes the approximation to the second order of g. In this case, it is more difficult to express the first two moments of the limit state function (for more details, see Pinto et al. 2007); instead, the probability of failure is approximated in terms of β and the $n-1$ principal curvatures κ_i of the paraboloid in the standardized uncorrelated n-space defined by vector $\mathbf{U}$. The exact expression of this probability is given in terms of single-fold integral defined by Tvedt (1990) which can be solved numerically in lack of closed-form solution. A simpler approximation was derived on the basis of asymptotic analysis by Breitung (1984):

$$P_F \approx \Phi(-\beta) \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 + \beta \kappa_i}} \tag{18}$$

where β is calculated as the minimum distance between the origin of the standard normal space and the paraboloid.

Reliability Methods in Design Codes: The Safety Factors in LRFD

Until the 1960s, most design codes did not provide an explicit implementation of reliability concepts. The recommended methods mainly involved stress checks in critical sections, which aimed to verify that the unit stresses under the action of working, or service, loads do not exceed predesignated allowable values (i.e., values prescribed by the building code to provide a factor of safety against attainment of some limiting stress, such as the minimum specified yield stress or the stress at which the buckling occurs); this design philosophy is also known as the working stress design. The issue of uncertainty in the various parameters was implicitly addressed by employing safety factors much larger than unity on the material strength. During the 1960s and 1970s, reliability analyses were addressed more explicitly by the codes, giving rise to reliability-based design. However, because professional design engineers found the application of the reliability-based methods somewhat difficult to access, a *semi-probabilistic* code-based approach known as the *load and resistance factor design* (LRFD) format was proposed (Galambos et al. 1982):

$$R/\gamma_R \geq \gamma_S \cdot S \quad \text{or} \quad R_d \geq S_d \qquad (19)$$

where R_d and S_d are the design resistance and load, respectively; R and S are the nominal resistance and nominal action (due to loads); and γ_R and γ_S are the associated resistance and load safety factors. The codes aim to link these coefficients with the target (admissible) reliability and with the amount of uncertainty in each uncertain parameter. The two safety factors in Eq. 19 have the general form

$$\gamma_R = \frac{\mu_R}{R} \left(1 + \alpha'_R \beta \delta_R\right) \text{ or } \gamma_S = \frac{\mu_S}{S} \left(1 + \alpha'_S \beta \delta_S\right) \tag{20}$$

where μ_R and μ_S denote the mean resistance and load; α'_R and α'_S are the sensitivity coefficients (function of the sensitivity factors presented in Eq. 14) for the resistance and load corresponding to R and S, respectively; β is the FOSM reliability index associated to the target reliability (obtained from the inverse of the standard normal cumulative distribution); and finally, δ_R and δ_S are the coefficient of variation (i.e., the standard deviation divided by the mean) of the resistance and load. For example, assuming that the nominal values are equal to the mean values, it can be observed that the safety factor is only a function of α, β, and δ.

Clearly, the safety factors depend on the chosen limit states. For the ultimate limit state, for instance, most of design codes provide safety factors that are calibrated with respect to nominal values. These nominal values are chosen as the value with a predefined probability of being exceeded; typically a lower (usually 5 %) percentile is selected for the resistance and an upper (usually 95 %) percentile is selected for the load effect. The European standard (CEN 2002) and ASCE standard (ASCE 2010) both define the acceptable values of the reliability index. In particular, the European standard defines minimum values that are established as a function of the consequences related to the reaching of a given limit state (with β varying from 3.3 to 4.3 with respect to 50 years of service period), whereas the ASCE standard defines the acceptable reliability index (which is related to a specified failure probability) as a function of the structural damage and occupancy category of the construction (with β varying from 2.5 to 4.5 for 50 years of service period). This approach, however, allows for reliability checking at the member level rather than at the system level. As mentioned before, establishing the link between component reliability and system reliability is not a trivial problem.

Alternative Methods Based on Time-Invariant Reliability Methods

When an equivalent linear static procedure or a response spectrum analysis is performed, the

S

reliability analysis can be carried out in a straight-forward manner using the time-invariant reliability methods. However, in the case of dynamic nonlinear analyses, the problem becomes much more complicated. In fact, in the latter case, it is generally difficult to express the limit state of interest directly in terms of time-invariant basic load and resistance uncertain parameters. Therefore, the combination of the different source of uncertainties becomes an intricate problem to tackle. Nowadays, it is possible to combine the reliability methods and the finite element analysis to overcome this problem (Der Kiureghian 1996). Currently, useful specialized software is available to address this issue (Der Kiureghian et al. 2006). Alternatively, a few *hybrid* approaches have been recently proposed in order to address the time variance in reliability problem.

These methods combine the RTR variability (taken into account by alternative methods like the PEER approach, which is discussed subsequently in this entry) with other source of uncertainties (treated by current FORM/SORM reliability methods). Typical hybrid approaches consist of the confidence interval approach (Cornell et al. 2002; Ellingwood and Kinali 2009) and the mean estimate approach (Cornell et al. 2002). Other relevant examples can be found in Ibarra and Krawinkler (2005), Haselton and Deierlein (2007), Dolšek (2009), Liel et al. (2009), Vamvatsikos and Fragiadakis (2010), and Celarec and Dolšek (2013). In these studies, many approximations are presented; for instance, if the limit state function does not have a defined functional form, the required gradients of the linearized limit state function can be obtained through perturbation of individual uncertain parameters in a series of sensitivity analyses.

Simulation-Based Reliability Methods

The general integral form in Eq. 1 can also be estimated by employing various simulation techniques. These techniques are especially efficient when the limit state function is non-differentiable or has several design points that make significant contributions to the failure probability (Faber 2012; Pinto et al. 2007). Although a large variety of simulation techniques may be found in the literature, they generally have their origin in the so-called *Monte Carlo* (MC) method (Rubinstein 1981; Fishman 1996). The general principles of standard MC simulation technique are outlined in the following section. To have a further insight into the basis of simulation techniques, the probability integral in Eq. 1 is rewritten as follows:

$$P_F = \int_{\Omega_{\mathbf{X}}} I_F(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \mathbb{E}[I_F(\mathbf{X})] \qquad (21)$$

where $\Omega_{\mathbf{x}}$ denotes the entire sample space of $\mathbf{X}$ and $I_F(\mathbf{x})$ is an indicator function that equals one when $G(\mathbf{x}) \leq 0$ and zero otherwise. The integral in Eq. 21 is the expected value of the indicator function denoted as $\mathbb{E}[I_F(\mathbf{x})]$. In other words, P_F is the expected value of $I_F(\mathbf{x})$ based on the joint PDF $f_{\mathbf{X}}(\mathbf{x})$. It is noteworthy that throughout section "Simulation-Based Reliability Methods," $p(\mathbf{x})$ might be used in place of $f_{\mathbf{X}}(\mathbf{x})$ interchangeably. This is done in order to homogenize and to simplify the probability notations within different equations. Hence, the term $p(\mathbf{x})$ is used in cases where the conditioning on a preposition is defined, the use of $f_{\mathbf{X}}$ might create further confusions, or in the robust reliability expressions defined subsequently in section "Robust Reliability Assessment Using Simulation-Based Approaches." However, the marginal PDFs are expressed generally by $f_{\mathbf{X}}$. By employing the MC method, various realizations of the uncertain parameters $\mathbf{x}$ (i.e., $\mathbf{x}_k$, $k = 1$: N_S) can be independently sampled so that the multidimensional integral in Eq. 21 is estimated as follows:

$$P_F = \mathbb{E}[I_F(\mathbf{X})] \cong \frac{1}{N_S} \sum_{k=1}^{N_S} I_F(\mathbf{x}_k) = \frac{N_F}{N_S} \qquad (22)$$

where N_S is the total number of simulations and N_F is the number of realizations for which $I_F(\mathbf{x}_k) = 1$ (i.e., the realizations leading to failure).

It can be shown that Eq. 22 is an unbiased estimator for P_F since

$$\mathbb{E}(\mathbb{E}[I_F(\mathbf{X})]) = \frac{1}{N_S}\sum_{k=1}^{N_S}\mathbb{E}[I_F(\mathbf{x}_k)] = \frac{N_S P_F}{N_S} = P_F \tag{23}$$

The variance of the estimator can be calculated by considering that $I_F(\mathbf{x}_k)$'s are independently distributed Bernoulli variables with variance equal to $P_F(1-P_F)$:

$$\begin{aligned}\mathbb{VAR}(\mathbb{E}[I_F(\mathbf{X})]) &= \mathbb{VAR}\left(\frac{1}{N_S}\sum_{k=1}^{N_S}I_F(\mathbf{x}_k)\right)\\ &= \frac{N_S P_F(1-P_F)}{N_S{}^2}\\ &= \frac{P_F(1-P_F)}{N_S}\end{aligned} \tag{24}$$

This equation reveals that the uncertainty in the estimator decreases with increasing N_S. Therefore, in order to accurately quantify the likelihood of rare events (with small P_F), which is a characteristic feature in seismic assessment, N_S should be increased. Besides the integral form in Eq. 21 for the structural reliability formulation, many problems in probabilistic inference require the calculation of multidimensional integrals or summations over very large outcome spaces. In the general form, they can be expressed as

$$R = \int_{\Omega_{\mathbf{x}}} h(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})\mathrm{d}\mathbf{x} = \mathbb{E}[h(\mathbf{X})] \tag{25}$$

where R is a performance measure of interest and $h(\mathbf{x})$ is a response quantity of interest. The general idea of MC integration is used to approximate the expectation. Hence, a set of N_S samples can be drawn independently from distribution $p(\mathbf{x})$, and the expectations can be approximated by a finite sum:

$$R = \mathbb{E}[h(\mathbf{X})] \cong \frac{1}{N_S}\sum_{k=1}^{N_S}h(\mathbf{x}_k) \tag{26}$$

In this procedure, the analytic integration is replaced with summation over a suitably large set of samples. Generally, the approximation can be made as accurate as needed by increasing N_S.

Fundamentals of MC Method

The MC family of methods employs various sampling techniques in order to generate realizations of the prescribed uncertain parameters. The standard MC simulation uses the random sampling approach in order to generate a large amount of realizations of uncertain parameters. The probability of failure can then be estimated by directly implementing Eq. 22 regardless of the complexity of the problem. Checking whether the structure has failed for each sample usually requires a structural analysis. To simulate $k = 1:N_S$ outcomes of the joint density function (corresponding to Eqs. 21–22 or Eqs. 25–26) in standard MC simulation, a random number, denoted as $z_{j,\,k}$, between 0 and 1 is generated for each of the components in $\mathbf{x}_k$ (i.e., $x_{j,\,k}$, $j = 1:n$). Assuming that X_j's are independent, the random numbers $z_{j,\,k}$ are mapped to $x_{j,\,k}$ by

$$x_{j,\,k} = F_{X_j}^{-1}\left(z_{j,\,k}\right) \tag{27}$$

where F_{Xj} is the marginal Cumulative Distribution Function (CDF) associated with the uncertain parameter X_j. Generally, the standard MC simulation is not justifiable computationally for adequate estimation of very small failure probabilities as a very large number of simulations are required.

The Latin hypercube sampling technique (LHS; Helton and Davis 2003) can be employed in order to reduce the number of simulations, N_S, in addition to achieving an acceptable level of accuracy for the statistical characteristics of response. The LHS is a special type of MC simulation that uses the stratification of the theoretical CDFs of uncertain parameters. Stratification divides the CDF curve into N_S equal intervals on the probability scale (i.e., 0.0 to 1.0). A sample is then randomly drawn from each interval or "stratification" of the input CDFs based on the

technique of "sampling without replacement." According to this technique, once a sample is taken from a designated stratification, this stratification is not sampled again.

As a result, the samples based on LHS reflect more accurately the distribution of values in the input probability distribution. LHS offers great benefits in terms of increased sampling efficiency and faster runtimes (due to fewer N_S that is required). Its roles in different aspects of reliability engineering have been described by, e.g., Novák et al. (1998) and Olsson et al. (2003). Stein (1987) has shown that the LHS reduces the variance of the response function compared to the crude Monte Carlo method.

Various methodologies were proposed for sampling based on LHS method; however, the most efficient strategy to perform sample selection, which deals also with samples of the tails of the PDF, is the sampling of interval mean values denoted as LHS mean (Huntington and Lyrintzis 1998). In this context, consider that the range of probability distribution F_{Xj} between 0.0 and 1.0 is divided into N_S equidistant intervals each with equal probability of $1/N_S$. The expression $\xi_{j,\,k}$ is the value of the uncertain parameter associated with the kth stratification with the probability k/N_S where $k = 1:N_S$:

$$\xi_{j,\,k} = F_{X_j}^{-1}(k/N_S) \tag{28}$$

The sampled data $x_{j,\,k}$ within the interval $[\xi_{j,\,k-1}, \xi_{j,\,k}]$ can be derived as

$$
\begin{aligned}
x_{j,k} &= \frac{\displaystyle\int_{\xi_{j,\,k-1}}^{\xi_{j,\,k}} x f_{X_j}(x)\,dx}{\displaystyle\int_{\xi_{j,\,k-1}}^{\xi_{j,\,k}} f_{X_j}(x)\,dx} \\
&= N_S \int_{\xi_{j,\,k-1}}^{\xi_{j,\,k}} x f_{X_j}(x)\,dx, \; j = 1, \ldots, n, \\
&\quad k = 1, \ldots, N_S
\end{aligned}
\tag{29}
$$

where f_{Xj} is the PDF of the uncertain parameter X_j. In spite of high efficiency of LHS technique, there are generally two issues concerning statistical correlation (Vořechovský and Novák 2009): (1) diminishing undesired and spurious correlation between uncertain parameters generated during sampling procedure, particularly in the case of a very small number of simulations, and (2) introducing the prescribed statistical correlations between pairs of uncertain parameters defined by the target correlation. Hence, in order to impose a prescribed correlation into the sampling scheme, an optimization problem for minimizing the difference between the target correlation and the actual correlation (estimated from samples) should be solved. For this purpose, a stochastic optimization approach called simulated annealing (SA) has been recently proposed (Vořechovský and Novák 2009; Golafshani et al. 2011) for application of LHS together with the SA methodology in structural reliability assessment).

The main drawback of the standard MC simulation methods is that it is not computationally efficient for problems dealing with small probabilities of failure (i.e., $P_F \leq 10^{-3}$). The main reason is due to the fact that the number of samples and hence the number of structural analyses, which are required to achieve a given accuracy, is inversely proportional to P_F. It requires information from rare samples that lead to failure, and on average, many samples are required before one such failure sample occurs. For instance, in case of achieving the P_F in order of 10^{-4} (which is typical in seismic reliability assessment) with a coefficient of variation around 10 %, nearly 10^6 simulations are required. Essentially, for rare events, the chance of generating samples from $f_{\mathbf{X}}(\mathbf{x})$ for which $I_F(\mathbf{x})$ equals one is extremely small (i.e., success rate is slow); therefore, many samples are needed to fulfill this requirement.

To overcome this obstacle, different techniques have been proposed in order to reduce the number of simulations. One of the most commonly applied techniques is namely the *importance sampling* method, which is described in the next section.

Importance Sampling Simulation Technique
The underlying idea of the importance sampling simulation (Rubinstein 1981; Shinozuka 1983;

Schuëller and Stix 1987; Au and Beck 1999) is to carry out Monte Carlo simulation with samples having a higher rate of falling in the failure region since these samples contribute only to the evaluation of P_F. In this class of methods, the integral is written in the following form:

$$P_F = \int\limits_{\Omega_{\mathbf{X}}} \mathrm{I}_F(\mathbf{x}) \frac{p(\mathbf{x})}{p_{\mathrm{IS}}(\mathbf{x})} p_{\mathrm{IS}}(\mathbf{x}) d\mathbf{x} = \mathbb{E}\left[\mathrm{I}_F(\mathbf{X}) \frac{p(\mathbf{X})}{p_{\mathrm{IS}}(\mathbf{X})} \right]$$

(30)

where $P_{\mathrm{IS}}(\mathbf{x})$ is denoted the importance sampling density function to increase the probability of sampling from the failure region. Applying Monte Carlo simulation to the integral, P_F is then estimated by

$$P_F \cong \frac{1}{N_S} \sum_{k=1}^{N_S} \mathrm{I}_F(\mathbf{x}_k) \frac{p(\mathbf{x}_k)}{p_{\mathrm{IS}}(\mathbf{x}_k)}$$

(31)

The efficiency of this method relies on a proper choice of the importance sampling distribution, which inevitably requires some knowledge about failure zone. Many past studies successfully applied importance sampling to time-invariant reliability problems where the number of uncertain parameters is not too large (see, e.g., Schuëller and Stix 1987; Papadimitriou et al. 1997; Der Kiureghian and Dakessian 1998; Au and Beck 1999).

For instance, Schuëller and Stix (1987) proposed the selection of an importance sampling density function as an N-dimensional joint-normal PDF with uncorrelated components, where the mean value equals the design point as obtained from FORM analysis and standard deviation values corresponding to the standard deviations of the components of $\mathbf{X}$. Although design points are a good notion to characterize the important region in $\mathbf{X}$ and hence importance sampling densities centered on them are often a good choice, they are not a necessary ingredient for forming a good importance sampling density. Hence, adaptive or iterative schemes and in particular the kernel density estimators, which do not involve the notion of design points, have attracted

considerable attention (see Au and Beck (1999) for a complete discussion and suggestion of an efficient kernel method).

Au and Beck (2001b) have also developed a very efficient importance sampling method for the first-excursion problem for linear dynamical systems under Gaussian stochastic excitation. However, for time-dependent problems, which are often characterized by a large number of uncertain parameters with complexity arising from their dynamic nature, the application of importance sampling is much more difficult (Au and Beck 2003a).

Hence, Au and Beck (2001a) proposed a *subset simulation* technique that is capable of efficiently computing the small failure probabilities encountered in engineering reliability analysis of general dynamical systems.

Calculation of Failure Probability Using the Subset Simulation

Subset simulation is an efficient and adaptive stochastic simulation algorithm for estimating small failure probabilities in high dimensions (Au and Beck 2001a, 2003b; Zuev et al. 2012). The underlying idea is to express the small failure probability as a product of larger probabilities conditional on some intermediate events. This allows converting the simulation of a rare event into a sequence of simulations of more frequent events (see also Jalayer and Beck 2008; Jalayer et al. 2010).

In this methodology, the failure region is modeled as the last element in a sequence of embedded failure regions (i.e., decreasing nested sequence of failure regions) $F = F_m \subset \ldots \subset F_2 \subset F_1$. F_1 is the first element in the failure sequence (i.e., largest failure region), and $F = F_m$ is the target failure region and the last element in the failure sequence. Therefore, by considering the product rule in probability, the failure probability can be derived as follows:

$$P_F = P(F_m) = P(F_m|F_{m-1})P(F_{m-1})$$
$$= P(F_1) \prod_{i=1}^{m-1} P(F_{i+1}|F_i)$$

(32)

According to Eq. 32, the rough idea of subset simulation is to estimate the failure probability as a product of a sequence of conditional probabilities denoted as $\{P(F_{i+1}|F_i): i = 1:m–1\}$ and $P(F_1)$. Hence, although P_F is small, the conditional probabilities can be made sufficiently large by appropriately adopting m.

Various simulation techniques can be used for sampling from PDFs conditional on a failure region F_i. Subset simulation proceeds in advance by simulating N_S samples $\mathbf{x}^{(0)}$ from $p(\mathbf{x})$ by MC simulation (note that superscript "0" denotes that the samples correspond to the first level):

$$P(F_1) = \int_{\Omega_{\mathbf{x}}} I_{F_1}(\mathbf{x})p(\mathbf{x})d\mathbf{x} \cong \frac{1}{N_S}\sum_{k=1}^{N_S} I_{F_1}\left(\mathbf{x}_k^{(0)}\right)$$

$$(33)$$

From these samples, one can readily obtain new samples that are distributed as $p(\mathbf{x}|F_1)$ (i.e., these samples lie in F_1). These samples can be used to estimate $P(F_2|F_1)$ by sample averaging similar to what has already done in Eq. 33. These samples provide seeds for simulating more samples according to $p(\mathbf{x}|F_2)$ and hence for estimating $P(F_3|F_2)$. Repeating this process, the conditional probabilities of the higher-conditional levels can be computed until the failure region of interest F_m is attained. In the ith conditional level, where $1 \leq i \leq m-1$, let $\mathbf{x}^{(i)}$ be the samples with distribution $p(\mathbf{x}|F_i)$; the conditional probability can be expressed as

$$P(F_{i+1}|F_i)\int_{\Omega_{\mathbf{x}}} I_{F_{i+1}}(\mathbf{x})p(\mathbf{x}|F_i)d\mathbf{x} \cong \frac{1}{N_S}\sum_{k=1}^{N_S} I_{F_{i+1}}\left(\mathbf{x}_k^{(i)}\right)$$

$$(34)$$

The subset simulation is shown to be especially efficient for modeling rare failure events (i.e., when the probability of failure is very small). The statistical properties of each conditional failure probability as well as the P_F are derived in Au and Beck (2001a, 2003b).

Considering that the target failure region can be stated as $F = \{\mathbf{x}: D(\mathbf{x}) > C(\mathbf{x})\}$, the choice

of the sequence of embedded failure regions, F_i, is a key issue to be addressed. It can significantly affect the efficiency of the subset simulation procedure. If F is introduced with a single parameter, the sequence of intermediate failure regions F_i can be generated by adaptively varying that parameter. The (scalar) demand-to-capacity ratio, which is suitable within the structural reliability assessment, is defined as (Au and Beck 2003b; Jalayer et al. 2007a):

$$Y(\mathbf{x}) = \max_{j=1}^{N_{\text{mech}}} \min_{i=1}^{n_j} \frac{D_{i,j}(\mathbf{x})}{C_{i,j}(\mathbf{x})}$$

$$(35)$$

where N_{mech} is the number of potential failure mechanisms and n_j is the number of components in the jth mechanism (cf. Eq. 4). This generalized performance variable can be interpreted as the component demand-to-capacity ratio that brings the system closer to the desired limit state. It is interesting to note that the performance variable in Eq. 35 takes into account a range of potential failure mechanisms (*cut-sets*) as noted in section "Structural Component and Structural System Reliability." As a result, the target failure region can be rewritten as $F = \{\mathbf{x}: Y(\mathbf{x}) > 1\}$, and the sequence of embedded intermediate failure regions can be generated as $F_i = \{\mathbf{x}: Y(\mathbf{x}) > y_i\}$ where $0 < y_1 < \cdots < y_m = 1$ are the intermediate thresholds. Au and Beck (2003b) proposed an adaptive methodology for choosing y_i values so that the estimated conditional probabilities are equal to a fixed value. They found that a fixed value equal to 0.1 yields good efficiency.

This section attempts to briefly introduce the main concepts of the subset simulation as an advanced stochastic simulation method for estimation of small probabilities corresponding to rare failure events. However, a detailed introductory description of methodology can be found in the chapter "▶ Subset Simulation Method for Rare Event Estimation: An Introduction" of the general section "Reliability and Robustness" of the Encyclopedia of Earthquake Engineering.

Robust Reliability Assessment Using Simulation-Based Approaches

The characterization of uncertainties can be treated in two levels: (1) prior probability distributions for the uncertain parameters based on available information and/or (qualitative) professional judgment and (2) the results of in situ tests, various types of inspections, pseudo-dynamic health-monitoring tests, and the consideration of the corresponding measurement errors that can be used to update the prior distributions of uncertain parameters. The treatment of the latter issue is tackled by the Bayesian updating procedure (e.g., Box and Tiao 1973; Jaynes 2003).

In the *Bayesian* approach, the probability is always conditional on the amount of information available. In this probabilistic framework, probability represents the degree of belief in a certain outcome based on the amount of information. The Bayesian approach was not assessed objectively by many statisticians until late last century because of the absence of a strong rationale behind the theory (see Jaynes 2003; Zuev et al. 2012 for a brief summary of development corresponding to the Bayesian approach). However, the seminal work of the physicist Cox (1961) expounded by the physicist Jaynes (2003) has significantly enhanced the Bayesian probability theory as a convenient mathematical language for inference and uncertainty quantification. The Bayesian approach usually leads to high-dimensional integrals that often can be evaluated neither analytically nor numerically by straightforward quadrature. Nevertheless, the development of new algorithms as well as increasing computing power has led to an explosive growth of using Bayesian approach in different branches of science.

In the *robust* (updated) reliability approach, the plausibility of all the possible structural models conditional on the amount of information available is taken into account in a Bayesian framework (for more details, see Jaynes 2003; Papadimitriou et al. 2001; Beck and Au 2002; Jalayer et al. 2010; Zuev et al. 2012). The robust failure probability then can be calculated (in comparison to Eq. 21) by the following integral:

$$P_F = p(F|D, M) = \int_{\Omega_{\mathbf{x}}} p(F|\mathbf{x}, D, M)p(\mathbf{x}|D, M)\mathrm{d}\mathbf{x}$$

$$= \int_{\Omega_{\mathbf{x}}} I_F(\mathbf{x})p(\mathbf{x}|D, M)\mathrm{d}\mathbf{x}$$

$$(36)$$

where D denotes some test data and M is the set of possible structural models used to specify (both the structural and the probabilistic) modeling assumptions in the analysis; $p(\mathbf{x}|D, M)$ is the posterior PDF of the model parameters, $\mathbf{x}$, based on the observed data and assumed structural model; and $P(F|\mathbf{x}, D, M)$ is the failure probability given the model parameters defined by $\mathbf{x}$, which can be reduced to a deterministic index function, I_F. This index is a function of the model parameters which equals one if failure occurs and zero otherwise. The conditioning on M is included here to stress the fact that all probabilities involved in model updating are always conditional on the choice of the modeling assumptions.

The Bayesian framework is used herein to provide a rigorous method for updating the plausibility of each of the models in representing the structure that is quantified by the probability distribution over the vector of model parameters or the posterior PDF, $p(\mathbf{x}|D, M)$:

$$p(\mathbf{x}|D, M) = p_D(\mathbf{x}) = c^{-1}p(D|\mathbf{x}, M)p(\mathbf{x}|M)$$

$$(37)$$

where $p(\mathbf{x}|M)$ is the prior probability distribution for $\mathbf{x}$ specified by M, which reflects the relative plausibility of each model before utilizing the data D, $p(D|\mathbf{x}, M)$ is the (updated) probability distribution (also known as the likelihood function) for observed data D based on a model specified by the model parameters $\mathbf{x}$, and c^{-1} is a normalizing constant. Thus, the robust failure probability in Eq. 36 can be shown as

$$P_F = c^{-1} \int_{\Omega_{\mathbf{x}}} I_F(\mathbf{x}, M)p(D|\mathbf{x}, M)p(\mathbf{x}|M)\mathrm{d}\mathbf{x} \quad (38)$$

Similarly, the general performance measure of Eq. 25 can also be rewritten based on the robust reliability concept as follows:

$$R_D = \int_{\Omega_{\mathbf{x}}} h(\mathbf{x})p(\mathbf{x}|D,M)\mathrm{d}\mathbf{x} = \int_{\Omega_{\mathbf{x}}} h(\mathbf{x})p_D(\mathbf{x})\mathrm{d}\mathbf{x}$$

$$(39)$$

Substituting Eq. 37 into 39, R_D can also be expressed as

$$R_D = c^{-1} \int_{\Omega_{\mathbf{x}}} h(\mathbf{x})p(D|\mathbf{x},M)p(\mathbf{x}|M)\mathrm{d}\mathbf{x} \quad (40)$$

Equations 36 and 38 (similarly, Eqs. 39 and 40) suggest two ways of evaluating the robust reliability by simulation. Equations 36 and 39 suggest estimating the P_F or R_D as the average of I_F or h over samples simulated from p_D, while Eqs. 38 and 40 indicate that the integral and the normalizing constant should be estimated, individually and then combined.

Particular difficulties are encountered in the evaluation of P_F or R_D based on the expressions in Eq. 36 or 39, respectively. The updated probability density function p_D is known only up to a multiplicative constant (see Eq. 37), since the normalizing constant c^{-1} is usually given by a high-dimensional integral that is also difficult to evaluate (Beck and Au 2002; Jalayer et al. 2010). Hence, the application of MC simulation or importance sampling is not generally feasible in this case, because these methods cannot simulate independent samples from p_D.

Additionally, when using importance sampling (see Eq. 30), it is necessary to choose a sampling density that is concentrated in the aforementioned small zone; otherwise similar problems, as in the MC simulation, will arise. However, this task is not trivial to accomplish since information about the zone, where $p(D|\mathbf{x}, M)$ is concentrated, is not directly available (see Beck and Au 2002 for more details). Strictly speaking, the posterior PDF p_D occupies a much smaller volume than that of the prior PDF $p(\mathbf{x}|M)$, so samples in its high probability region cannot

be generated efficiently by sampling from the prior PDF using direct MC method. Of course, if the posterior PDF p_D is evaluated first from Eq. 37, then there would be no problem in using the standard MC from Eq. 36. However, in this case, the factor c^{-1} needs to be calculated.

According to the drawbacks of the current MC methods as well as the importance sampling techniques in solving robust reliability problems, Markov chain Monte Carlo (MCMC) simulation techniques, in particular, the Metropolis-Hastings (MH) algorithm (Metropolis et al. 1953; Hastings 1970), are discussed briefly in the next section. It offers a feasible and powerful way for simulating samples according to an arbitrary distribution (when the target PDF is known only up to a scaling constant), at the expense of introducing dependence among the samples.

Generating Samples in Robust Reliability Based on Adaptive MCMC Simulation Technique

The goal of MCMC is to design a Markov chain such that the stationary distribution of the chain is exactly the distribution that we are interested in sampling from (i.e., the target distribution). The idea is to use specially designed methods for setting up the transition function, such that no matter how each chain is initialized, it will converge to the target distribution. The MH algorithm is normally used to generate samples according to an unscaled PDF when the target PDF is known only up to a scaling constant, i.e., there is no need to know the normalizing constant (see Eq. 33) in advance. The fact that this procedure allows us to sample from un-normalized distributions is one of its major attractions especially in case of Bayesian model updating. The MH algorithm was originally developed by Metropolis and his coworkers (Metropolis et al. 1953) in statistical physics and later generalized by Hastings (1970) in Bayesian statistics (see Fishman 1996 for comprehensive discussions on the MCMC methods). However, its potential use for solving reliability problems in structural engineering has been only recently demonstrated (see Au and Beck 1999, 2001a, 2003b; Beck and Au 2002).

The MH algorithm can be used to generate samples according to the target updated PDF p_D. Using the Markov chain samples generated from the MH procedure, P_F or R_D is estimated as the average over the samples, which is the same approach as the usual MC method, except that the samples are simulated from a Markov chain instead of being independent and identically distributed. In order to reduce the initial effect of the choice of the transient probability distributions (e.g., non-normalized target PDF) on the estimate, the first few samples are often not included in the sample averaging.

In the MH procedure, a proposal PDF is chosen to generate a candidate point that is conditional on the previous state of the sampler. The next step is to either accept the proposal or reject it. However, as noted previously, the updated posterior PDF p_D is concentrated in a small zone, and direct adaptation using a proposal PDF, which varies with a vastly different length scale from that of the target PDF, will not be effective (see Beck and Au 2002 for more details). Therefore, a sequence of intermediate proposal PDFs that vary gradually between the prior PDF and the target PDF p_D are introduced (Au and Beck 1999; Beck and Au 2002). According to the proposed methodology, however, the proposal PDF is chosen as the kernel sampling density constructed using the Markov chain samples from the previous simulation level (for the first simulation level, the prior PDF is used as the proposal PDF). Thus, the adaptation is done from one simulation level to the next. This methodology is especially useful in the subset simulation described previously in section "Calculation of Failure Probability Using the Subset Simulation." Accordingly, a modified MH algorithm was developed by Au and Beck (2003b) that can obtain samples from a posterior PDF conditioned on a failure region F_i.

PEER Performance-Based Approach

In recent years, an effective foundation for the development of probabilistic Performance-Based Earthquake Engineering (PBEE) for the design and assessment of building structures has been developed by the Pacific Earthquake Engineering Research Center (PEER; Cornell and Krawinkler 2000; Moehle and Deierlein 2004). There are several stages to this process: (1) calculation of ground motion hazard by quantifying uncertainty in ground motion with a probabilistic model for a parameter related to ground motion and known as the intensity measure (IM); (2) estimation of the uncertainty in structural response D (e.g., force and deformation engineering parameters), for each IM level; (3) estimation of the uncertainty in damage measure DM (i.e., physical states of damage) given response level D; and (4) estimation of the uncertainty in the decision variable DV as the resulting consequences (e.g., financial losses, fatalities, and business interruption) given damage measure.

Each stage of the process is performed and executed (more or less) independently and then linked back together, as expressed in the following integral form:

$$\lambda(DV) = \iiint G[DV|DM] \cdot \left| \frac{dG[DM|D]}{dDM} \right|$$
$$\cdot \left| \frac{dG[D|IM]}{dD} \right| \cdot \left| \frac{d\lambda(IM)}{dIM} \right| \cdot dIM$$
$$\cdot dD \cdot dDM \qquad (41)$$

where $G[Y|X]$ denotes generically the conditional Complementary Cumulative Distribution Function (CCDF) of Y given a certain value of X and $\lambda(Y)$ denotes the mean annual exceedance rate (mean annual frequency) of Y. As it can be observed from Eq. 41, the PEER framework enjoys a modular structure and benefits from the (hypothetical) conditional independence between the main parameters (i.e., the conditional independence of $DV|DM$ from D and IM and other ground motion parameters, $DM|D$ from IM and other ground motion parameters, and $D|IM$ from other ground motion parameters, such as, but not limited to, magnitude and distance).

This (hypothetic) Markovian independence between various intermediate parameters is one of the main factors that distinguish the PEER approach from the two other approaches

discussed in this entry, namely, the classic and the simulation-based reliability methods. In these approaches, the probability of failure (which can generically represent a decision variable DV exceeding a certain threshold value) is calculated directly and without considering intermediate parameters such DM, D, and IM. Another important factor that distinguishes the PEER framework equation (Eq. 41) from the other two approaches, discussed herein, is that Eq. 41 yields a rate of exceedance and not a probability. This is while the classic and the simulation-based reliability methods lead to a probability of exceedance. In other words, the rate of exceedance reported in Eq. 41 needs to be translated into probability by assuming an underlying probability model (see Der Kiureghian 2005).

In this section, the second stage of the PEER integral expression dedicated to estimating the conditional probability of exceeding the engineering demand parameter D given IM, denoted as G $[D|IM]$, is discussed in detail. The estimation of the mean annual frequency of exceeding a specific IM values, denoted as $\lambda(IM)$, is the main focus of the chapter "▶ Probabilistic Seismic Hazard Models." Finally, an outlook into the whole framework is provided in the chapter entitled "▶ Performance-Based Design Procedure for Structures with Magneto-Rheological Dampers".

Taking into account other sources of uncertainty imposes an overall increase in the dispersion for the failure probability (Jalayer 2003; Jalayer and Cornell 2003), while an eventual shift in the median failure probability is not envisioned in this method (see the "hybrid" methods discussed in section "Alternative Methods Based on Time-Invariant Reliability Methods").

In the PEER PBEE methodology, the evaluation of structural performance is defined by probabilistic performance objectives, which can be expressed in terms of the mean annual frequency (MAF) of exceeding a specified limit state (also referred to as the limit state frequency) denoted as λ_{LS}. Therefore, it is worth noting that the usual outcome of this framework is an estimate of the rate of exceeding a designated limit state threshold rather than a probability.

The main advantage of this methodology is that it evaluates the limit state frequency, λ_{LS}, by decoupling the ground motion hazard and nonlinear dynamic analyses via a site- and structure-specific intermediate variable known as the ground motion intensity measure, IM (see also Shome et al. 1998; Shome and Cornell 1999; Cornell and Krawinkler 2000; Luco 2002; Jalayer 2003; Jalayer and Cornell 2009). The ground motion IM serves as link between seismic hazard analysis, typically provided by seismologists, and structural analysis conducted by engineers.

This IM approach is appealing because it allows the first two analysis stages in the PEER PBEE methodology to be performed (almost) independently. The benefit of this approach is that the number of analyses needed can be substantially reduced because most of the uncertainties are concentrated in ground motion hazard in terms of the MAF of exceeding a certain level of IM, λ_{IM}, which can directly be estimated from probabilistic seismic hazard analysis (PSHA; Cornell 1968; McGuire 2004) for the area where the structure is located.

The conditional probability of D given IM can be predicted through alternative nonlinear dynamic analysis procedures. Herein, three distinct nonlinear dynamic procedures are described: (1) the so-called *cloud analysis* method, in which the structure is subjected to a set of (as-recorded) ground motions (Cornell et al. 2002; Jalayer and Cornell 2003; Elefante et al. 2010), and subsequently, the distribution of D given IM (D and IM are denoted as *cloud data*) is directly obtained by performing a logarithmic linear regression on the cloud data; (2) the incremental dynamic analysis (Vamvatsikos and Cornell 2004) where a selected suite of ground motions are incrementally scaled to different levels of the IM; and (3) multiple-stripe analysis (MSA; Jalayer 2003; Jalayer and Cornell 2009), where analysis is performed at a specified set of IM levels known as *stripes*, each of which may be performed for a different suite of ground motions (i.e., structural dynamic analyses for multiple stripes of IM through selected set of records for each IM level). It is noteworthy that although IDA uses the same set of ground motions for all

IM levels, the MSA approach allows the re-selection of ground motions at each *IM* level to be consistent with the seismicity of the designated site (based on the disaggregation of seismic hazard; see, e.g., Baker and Cornell 2005). Thus, the distribution of *D* given *IM* is estimated for multiple *IM* levels.

The structural fragility curve for a range of *IM* values is a useful probabilistic structural response quantity that can directly be estimated from the aforementioned nonlinear dynamic analysis procedures. The structural fragility for a given limit state, *LS*, can be defined as the conditional probability of exceeding the limit state capacity for a given level of *IM* (Ellingwood 2001; Jalayer 2003; Baker and Cornell 2005; Jalayer and Cornell 2009; Baker 2014). If the considered limit state corresponds with the collapse condition, the associated fragility curve is known as the collapse capacity. Subsequently, this result can be combined with a ground motion hazard, i.e., λ_{IM}, to compute the MAF of exceeding any given limit state (Shome and Cornell 1999; Jalayer 2003; Jalayer and Cornell 2003; Ibarra and Krawinkler 2005; Haselton and Deierlein 2007).

Mean Annual Frequency of Exceeding a Limit State

The MAF of exceeding a specified limit state, denoted as λ_{LS}, can be obtained from two approaches: (a) the *IM*-based approach (also known as the fragility/hazard format) and (b) the *D*-based (i.e., demand-based) approach, which in the literature is also called the engineering demand parameter (*EDP*)-based approach (Cornell and Krawinkler 2000). The two approaches are expressed by the two following equations, respectively:

$$\lambda_{LS} = \int_{im} P\left[D > C_{LS}|IM = im\right] |\mathrm{d}\lambda_{IM}(im)| \quad (42)$$

$$\lambda_{LS} = \int_{im} P\left[IM_{cap} < IM|IM = im\right] |\mathrm{d}\lambda_{IM}(im)| \quad (43)$$

where C_{LS} is the capacity associated with the given *LS*, IM_{cap} is an uncertain parameter
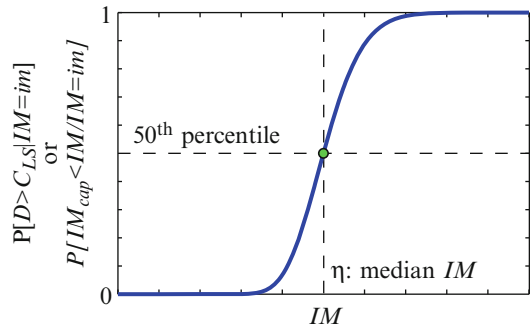
Seismic Reliability Assessment, Alternative Methods for, Fig. 5 Typical fragility curve

representing the capacity of structure in terms of *IM* values (distribution of *IM* values) that result in a *D* level equal to C_{LS}, and the probability terms denote the structural fragilities for the desired limit state. The typical shape of the fragility curve is presented in Fig. 5.

The fragility curve can be estimated by cloud analysis, MSA, or IDA.

Cloud Analysis Approach

The probability term presented in Eq. 42 can be derived by assuming *D* to be a lognormal variable. This is a common assumption that has been confirmed as reasonable in many past studies (e.g., Porter et al. 2007; Jalayer and Cornell 2009):

$$P[D > C_{LS}|IM - im] = 1 - \Phi\left(\frac{\ln C_{LS} - \ln\eta_{D|IM}(im)}{\beta_{D|IM}(im)}\right) \quad (44)$$

where Φ is the standardized Gaussian CDF and η and β are conditional median and standard deviation (dispersion) of the natural logarithm of *D* given *IM*. In order to estimate the statistical properties of the cloud data, i.e., conditional median and dispersion, conventional linear regression (using least squares method) is used in the natural logarithmic scale, as shown in Fig. 6. Thus,

Seismic Reliability Assessment, Alternative Methods for, Fig. 6 Typical logarithmic linear regression on cloud data
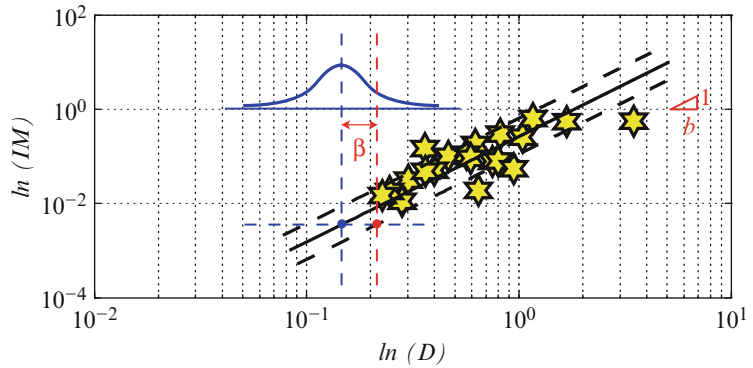
$$\ln\eta_{D|IM}(im) = \ln a + b\ln(im),$$

$$\beta_{D|M} = \sqrt{\frac{\sum_{j=1}^{N_{cloud}}\left(\ln\frac{C_{LS}}{a\cdot im_j^b}\right)}{N_{cloud} - 2}} \qquad (45)$$

where a and b are the coefficients of the logarithmic linear regression and N_{cloud} is the number of the cloud data.

Rare events may exert a very large demand D (i.e., even cause instability in the nonlinear dynamic analysis) that can affect the general trend in the linear regression of the cloud data. In such cases, a logistic regression (Agresti 2002) can generally be used. In this methodology, binary values of 0 (for non-collapse data) and 1 (for collapse cases) are assigned for each IM, and the logistic regression allows the construction of the associated fragility curve based on these binary data (Baker and Cornell 2005; Elefante et al. 2010).

IDA Approach

In a great number of research works (see, e.g., Jalayer 2003; Vamvatsikos and Cornell 2004; Baker and Cornell 2005; Ibarra and Krawinkler 2005; Zareian and Krawinkler 2007; Haselton and Deierlein 2007; Jalayer and Cornell 2009; Baker 2014), the collapse capacity is estimated by repeatedly scaling a ground motion using the IDA procedure until the ground motion causes collapse of the structure; hence, each ground motion can be associated with a single IM value
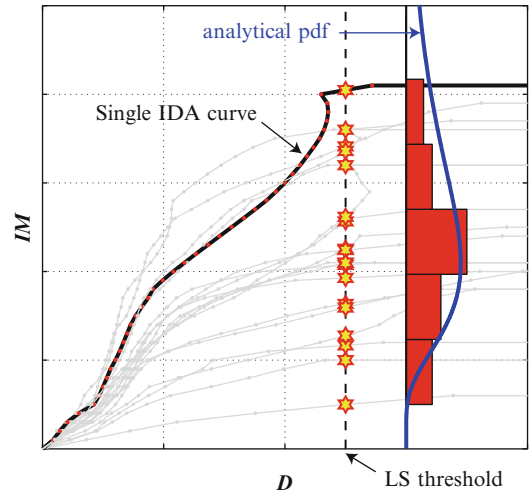
Seismic Reliability Assessment, Alternative Methods for, Fig. 7 Typical IDA curves and fitting of the empirical distribution of the IM related to the overcoming of limit state

associated with collapse. By repeating this process for a set of ground motions, one can obtain a set of IM values corresponding to the onset of collapse. Subsequently, the probability of collapse for a specific IM equal to im can be estimated as the fraction of records for which collapse occurs at a level lower than the predefined value, im. A lognormal distribution is often fitted to the empirical distribution of IM levels that cause the structure to collapse. In Fig. 7, the black dashed line represents the LS threshold, and the red stars illustrate the

IM_{cap} values that cause the onset of the limit state. On the same diagram, the histogram of the IM_{cap} data and the related empirical distribution are shown as well.

Multiple-Stripe Analysis Approach

In cases where different records are used at each *IM* level, the IDA method (section "IDA Approach") cannot be used to estimate the probability of collapse. A substitute for the conventional IDA approach should be utilized in this particular case. Therefore, instead of an *IM* value associated with the onset of collapse for each record (i.e., the IDA approach), the fraction of records that cause collapse at each IM level can be taken into account. The latter can be obtained from individual stripes by calculating the number of accelerograms for which the limit state has been exceeded with respect to the total number of records (Baker and Cornell 2005; Ebrahimian 2012; Baker 2014).

As a result, the empirical distribution of the collapse capacity for a set of *IM* values can be obtained. Once the ratio of the number of collapse cases to the total number of ground motions is calculated for each IM level, a fragility curve can be generated through the logistic regression. The observed fraction of collapse and the estimated

fragility function are shown schematically in Fig. 8 (Ebrahimian 2012).

Demand and Capacity Factor Design (DCFD) Format (IM-Based Version)

The *IM*-based closed-form solution for the annual frequency of exceeding a desired limit state, λ_{LS}, has been derived by Jalayer (2003) (see also Jalayer and Cornell 2003) under a set of simplifying assumptions. It is used as a performance-oriented design procedure proposed in the SAC/FEMA Steel Project (Cornell et al. 2002) as

$$\lambda_{LS} = \lambda_{IM}\left(\eta_{IM_{cap}}\right)\exp\left(\frac{1}{2}k^2\beta_{IM_{cap}}^2\right) \qquad (46)$$

where η_{IMcap} and β_{IMcap} are the parameters of the estimated lognormal distribution assuming that the IM_{cap} capacity is a lognormal variable and k reflects the steepness of the hazard curve, λ_{IM}, in the vicinity of the point η_{IMcap} (the hazard curve is approximated in the region of interest by the power-law relationship as $\lambda_{IM}(\eta_{IMcap}) = k_o(\eta_{IMcap})^{-k}$). The first term in Eq. 46 shows the first-order approximation to the limit state probability, and the exponential expression is a magnifying factor that accounts for the
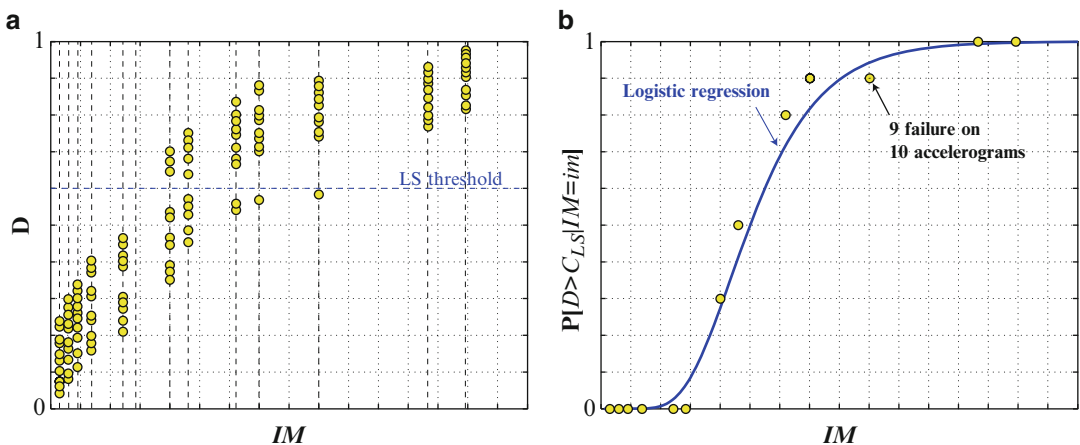
Seismic Reliability Assessment, Alternative Methods for, Fig. 8 (**a**) Typical MSA output and (**b**) fragility curve obtained through the logistic regression applied on the MSA results

sensitivity of the limit state probability to the randomness in the IM_{cap}.

The closed-from expression for the limit state frequency can represent an alternative interpretation for the seismic design or assessment considering an IM-based design criterion. A certain design criterion is to check whether the MAF of exceeding a certain limit state (limit state frequency) is less than or equal to an allowable annual frequency, λ_o. Thus,

$$\lambda_{LS} = \lambda_{IM}\left(\eta_{IM_{cap}}\right)\exp\left(\frac{1}{2}k^2\beta_{IM_{cap}}^2\right) \leq \lambda_o \quad (47)$$

After some simple rearrangements with the objective of allocating the parameters corresponding to the capacity to one side, Eq. 47 takes the following form (Jalayer 2003; Jalayer and Cornell 2003):

$$im^{\lambda_o} \leq \eta_{IM_{cap}}\exp\left(-\frac{1}{2}k^2\beta_{IM_{cap}}^2\right) \quad (48)$$

where $im^{\lambda o}$ is the indicator IM for a hazard level equal to the allowable annual frequency, λ_o. The right-hand side of this expression represents the *factored capacity* (expressed in terms of the adopted IM). Therefore, this expression illustrates a design criterion based on the fragility/hazard format in terms of the IM-based factored capacity being greater than or equal to the IM for a given allowable annual rate.

Summary and Future Challenges

A summary of general methodologies used in the seismic reliability assessment has been presented. In particular, the historical evolution of each method that allows for accounting uncertainties in demand and capacity is briefly illustrated. It is noted that considerable progress has been made in the development of these methodologies, and great efforts have been performed to implement them with the aim of loss reduction and risk mitigation. Considering that the seismic reliability of structures is a specific time-dependent problem with many sources of

uncertainty, applying any of the aforementioned methods is not a trivial task.

Arguably, the first-order and second-order reliability methods cannot be adopted in a straightforward manner when solving time-dependent nonlinear reliability problems. On the other hand, the PEER approach works best in cases where the primary source of uncertainty is reflected in the RTR variability (Der Kiureghian 2005). Therefore, in the recent years, various hybrid methods (section "Alternative Methods Based on Time-Invariant Reliability Methods") have been proposed (see, e.g., Liel et al. 2009). Other approaches include those developed by Vamvatsikos and Fragiadakis (2010) and Celarec and Dolšek (2013) where the PEER approach is mixed with other reliability methods (FOSM and Latin hypercube sampling). As a result, the treatment of various sources of uncertainty in both aforementioned reliability methods deserves further research and investigations.

It is noteworthy that both the modern construction codes and the professional community of engineers around the world are becoming more inclined towards the application of nonlinear dynamic analyses. This raises the professional interests to the PEER approach, since it provides a more accessible way of taking into account and propagating the RTR variability in design and assessment problems. Moreover, this approach leads directly to the estimation of the economic losses, which are viewed as essential risk metrics in decision making.

In order to take into account various sources of uncertainty (e.g., modeling, RTR, etc.) in the PEER approach while maintaining a low number of structural analyses, Jalayer et al. (2013) proposed a method inspired from the concept of robust reliability discussed in section "Calculation of Failure Probability Using the Subset Simulation." Conditioned on a prescribed analytical fragility model, they suggest using the standard MC simulation (with few samples) in order to create various structural model realizations based on the modeling uncertainties present in the problem. A suite of recorded ground motions employed to represent the RTR variability is then applied to each of the above-mentioned structural

models in order to calculate a desired structural response. In other words, the number of structural response realizations is equal to the number of ground motion records present in the suite of records. These structural response realizations are then used as data in order to both update the joint probability distribution for the parameters of the adopted analytic fragility model and also to directly obtain the robust fragility.

In the simulation-based approach, consideration of RTR variability requires the adoption of a suitable stochastic ground motion model (see the chapter "▶ Stochastic Ground Motion Simulation") conditional on seismic source parameters (see Au and Beck 2003b; Jalayer and Beck 2008). Accordingly, the consideration of RTR variability together with other sources of uncertainties will provide a complete probabilistic treatment of the structural response (see, e.g., Jalayer et al. 2007b where RTR variability and modeling uncertainty are considered together in the seismic reliability analysis of RC frames based on the subset simulation). One of the main challenges in applying this approach is to ensure that the stochastic ground motion model provides a realistic description of the characteristics of the ground motions expected to happen at the building site (see, e.g., Atkinson and Silva 2000; Rezaeian and Der Kiureghian 2010). Moreover, there are difficulties in the application of advanced simulation-based algorithms (e.g., subset simulation) so that these techniques cannot be performed in a straightforward and practical manner. However, the fast and incredible increase in computational power is going to render the simulation-based methods ever more accessible for resolving high-dimensional reliability problems.

In comparison to the PEER approach, the advanced simulation-based methods have the advantage of being able to directly estimate losses without passing through multiple stages (e.g., hazard, demand, damage, losses). Undoubtedly, the application of advanced simulation-based techniques is another major concern in the future of reliability methods due to the usually high dimension of the vector of uncertain parameters.

The Bayesian network methodologies have been developed during the past 25 years mostly in the field of artificial intelligence (Russell and Norvig 2003). This method has recently been applied in order to solve high-dimensional structural reliability problems such as infrastructure seismic risk assessment for spatially distributed systems (see, e.g., Grêt-Regamey and Straub 2006; Straub and Der Kiureghian 2010). The Bayesian networks rely on graphical visualization of the uncertain parameters and their correlation structure. They are arguably going to become efficient means of resolving systemic reliability problems due to their propensity for being automatized and also due to the fact that they lend themselves quite well to the implementation of advanced simulation-based methods.

References

Agresti A (2002) Categorical data analysis. Wiley, New York, 710 pp

Ambartzumian R, Der Kiureghian A, Ohanian V, Sukiasian H (1998) Multinormal probability by sequential conditioned importance sampling: theory and application. Probab Eng Mech 13(4):299–308

ASCE American Society of Civil Engineers (2010) ASCE/SEI 7–10, Minimum design loads for buildings and other structures, New-York

Atkinson G, Silva W (2000) Stochastic modelling of California ground motions. Bull Seismic Soc Am 90(2):255–274

Au SK, Beck JL (1999) A new adaptive importance sampling scheme for reliability calculations. Struct Saf 21(2):135–158

Au SK, Beck JL (2001a) Estimation of small failure probabilities in high dimensions by subset simulation. Probab Eng Mech 16(4):263–277

Au SK, Beck JL (2001b) First excursion probabilities for linear systems by very efficient importance sampling. Probab Eng Mech 16(3):193–207

Au SK, Beck JL (2003a) Important sampling in high dimensions. Struct Saf 25(2):139–163

Au SK, Beck JL (2003b) Subset simulation and its application to seismic risk based on dynamic analysis. J Eng Mech 129(8):901–917

Baker JW (2014) Efficient analytical fragility function fitting using dynamic structural analysis. Earthq Spectra doi: http://dx.doi.org/10.1193/021113EQS025M

Baker JW, Cornell CA (2005) Vector-valued ground motion intensity measures for probabilistic seismic demand analysis. Report No. 150, John A. Blume Earthquake Engineering Center, Stanford, 321 pp

S

Beck JL, Au SK (2002) Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. J Eng Mech 128(4):380–391

Box GEP, Tiao GC (1973) Bayesian inference in statistical analysis. Wiley, New York

Bozorgnia Y, Bertero VV (2004) Earthquake engineering: from engineering seismology to performance-based engineering. CRC Press, ISBN 0-8493-1349-9

Breitung K (1984) Asymptotic approximations for multinormal integrals. J Eng Mech 110(3):357–366

Celarec D, Dolšek M (2013) The impact of modelling uncertainties on the seismic performance assessment of reinforced concrete frame buildings. Eng Struct 52:340–354

CEN (2002) European Committee for Standardisation, Eurocode-Basis of Structural Design, EN 1990:2002

Cornell CA (1968) Engineering seismic risk analysis. Bull Seismol Soc Am 58(5):1583–1606

Cornell CA (1969) A Probability-based structural code. ACI J 66(12):974–985

Cornell CA, Krawinkler H (2000) Progress and challenges in seismic performance assessment. PEER Cent News 3(2):1–3

Cornell CA, Jalayer F, Hamburger RO, Foutch DA (2002) Probabilistic basis for 2000 SAC Federal Emergency Management Agency steel moment frame guidelines. J Struct Eng 128(4):526–533

Cox RT (1961) The algebra of probable inference. Johns Hopkins Press, Baltimore

Der Kiureghian A (1996) Structural reliability methods for seismic safety assessment: a review. Eng Struct 18(6):412–424

Der Kiureghian A (2005) Non-ergodicity and PEER's framework formula. Earthq Eng Struct Dyn 34(13):1643–1652

Der Kiureghian A, Dakessian T (1998) Multiple design points in first and second-order reliability. Struct Saf 20(1):37–49

Der Kiureghian A, Ditlevsen O (2009) Aleatory or epistemic? Does it matter? Struct Saf 31(2):105–112

Der Kiureghian A, Haukaas T, Fujimura K (2006) Structural reliability software at the University of California, Berkeley. Struct Saf 28(1):44–67

Det Norske Veritas (1992) Structural reliability analysis of marine structures, classification notes no 30.6

Ditlevsen O (1973) Structural reliability and the invariance problem. Research report no 22, Solid Mechanics Division, University of Waterloo, Waterloo

Ditlevsen O, Madsen HO (1996) Structural reliability methods. Wiley, Chichester

Dolšek M (2009) Incremental dynamic analysis with consideration of modeling uncertainties. Earthq Eng Struct Dyn 38(6):805–825

Ebrahimian H (2012) Assessment of existing offshore platforms in Persian Gulf region for life extension or operative conditions. Ph.D. dissertation, Department of Civil and Environmental Engineering, Sharif University of Technology, Tehran, 363 pp

Elefante L, Jalayer F, Iervolino I, Manfredi G (2010) Disaggregation-based response weighting scheme for seismic risk assessment of structures. Soil Dyn Earthq Eng 30(12):1513–1527

Ellingwood BR (2001) Earthquake risk assessment of building structures. Reliab Eng Syst Saf 74(3):251–262

Ellingwood BR, Kinali K (2009) Quantifying and communicating uncertainty in seismic risk assessment. Struct Saf 31(2):179–187

Faber MH (2012) Statistics and probability theory: in pursuit of engineering decision support, vol 18. Springer, New York

Fishman GS (1996) Monte Carlo: concepts, algorithms, and applications. Springer, New York

Galambos TV, Ellingwood B, MacGregor JG, Cornell CA (1982) Probability based load criteria: assessment of current design practice. J Struct Div 108(5):959–977

Golafshani AA, Ebrahimian H, Bagheri V, Holmas T (2011) Assessment of offshore platforms under extreme waves by probabilistic incremental wave analysis. J Constr Steel Res 67(5):759–769

Grêt-Regamey A, Straub D (2006) Spatially explicit avalanche risk assessment linking Bayesian networks to a GIS. Nat Hazards Earth Syst Sci 6(6):911–926

Haselton CB, Deierlein GG (2007) Assessing Seismic Collapse Safety of Modern Reinforced Concrete Moment-Frame Buildings, PEER Report 2007/08, Pacific Engineering Research Center. University of California, Berkeley

Hasofer AM, Lind NC (1974) Exact and invariant second-moment code format. J Eng Mech Div 100(1):111–121

Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57(1):97–109

Helton JC, Davis FJ (2003) Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab Eng Syst Saf 81(1):23–69

Hohenbichler M, Rackwitz R (1983) First-order concepts in system reliability. Struct Saf 1(3):177–188

Huntington DE, Lyrintzis CS (1998) Improvements to and limitations of Latin hypercube sampling. Probab Eng Mech 13(4):245–253

Ibarra LF, Krawinkler H (2005) Global collapse of frame structures under seismic excitations. John A. Blume Earthquake Engineering Center, Report no 152, Department of Civil and Environmental Engineering, Stanford University, Stanford, 324 pp

Jalayer F (2003) Direct probabilistic seismic analysis: implementing non-linear dynamic assessments. Ph.-D. dissertation, Department of Civil and Environmental Engineering, Stanford University, Stanford, 244 pp

Jalayer F, Beck JL (2008) Effects of two alternative representations of ground-motion uncertainty on probabilistic seismic demand assessment of structures. Earthq Eng Struct Dyn 37(1):61–79

Jalayer F, Cornell CA (2003) A technical framework for probability-based demand and capacity factor design (DCFD) seismic formats. Pacific Earthquake Engineering Research Center (PEER) 2003/08

Jalayer F, Cornell CA (2009) Alternative non-linear demand estimation methods for probability-based seismic assessments. Earthq Eng Struct Dyn 38(8):951–972

Jalayer F, Franchin P, Pinto PE (2007a) A scalar damage measure for seismic reliability analysis of RC frames. Earthq Eng Struct Dyn 36(13):2059–2079

Jalayer F, Franchin P, Pinto PE (2007b) Considering structural modelling uncertainty in seismic reliability analysis of RC frames: use of advanced simulation methods. In: Proceedings of the conference: COMPDYN 2007, paper ID: 1218, Crete, Greece, 13–16 June 2007

Jalayer F, Iervolino I, Manfredi G (2010) Structural modeling uncertainties and their influence on seismic assessment of existing RC structures. Struct Saf 32(3):220–228

Jalayer F, De Risi R, Elefante L, Manfredi G (2013) Robust fragility assessment using Bayesian parameter estimation. Paper presented at the Vienna Congress on Recent Advances in Earthquake Engineering and Structural Dynamics 2013 (VEESD 2013), Vienna, 28–30 Aug

Jaynes ET (2003) Probability theory: the logic of science. Cambridge University Press, New York

Liel AB, Haselton CB, Deierlein GG, Baker JW (2009) Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. Struct Saf 31(2):197–211

Liu P-L, Der Kiureghian A (1991) Optimization algorithms for structural reliability. Struct Saf 9(3):161–177

Luco N (2002) Probabilistic seismic demand analysis, SMRF connection fractures, and near-source effects. Probabilistic seismic demand analysis, SMRF connection fractures, and near-source effects. Ph.D. dissertation, Department of Civil and Environmental Engineering, Stanford University, Stanford, 260 pp

McGuire RK (2004) Seismic hazard and risk analysis. Earthquake Engineering Research Institute, Berkeley, 240 pp

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21:1087–1092

Moehle J, Deierlein GG (2004) A framework methodology for performance-based earthquake engineering. In: Proceedings of the 13th world conference on earthquake engineering, Vancouver, 1–6 Aug 2004, 13 pp

Novák D, Teplý B, Keršner Z (1998) The role of Latin Hypercube Sampling method in reliability engineering. In: Shiraishi N, Shinozuka M, Wen YK (eds) The 7th international conference on structural safety and reliability, ICOSSAR 97, vol 2. Balkema, Rotterdam (Kyoto, Japan)

Olsson A, Sandberg G, Dahlblom O (2003) On Latin hypercube sampling for structural reliability analysis. Struct Saf 25(1):47–68

Papadimitriou C, Beck JL, Katafygiotis LS (1997) Asymptotic expansions for reliability and moments of uncertain systems. J Eng Mech 123(12):1219–1229

Papadimitriou C, Beck JL, Katafygiotis LS (2001) Updating robust reliability using structural test data. Probab Eng Mech 16(2):103–113

Pinto PE, Giannini R, Franchin P (2007) Seismic reliability of structures. IUSS Press, Pavia

Porter K, Kennedy R, Bachman R (2007) Creating fragility functions for performance-based earthquake engineering. Earthq Spectra 23(2):471–489

Rackwitz R, Flessler B (1978) Structural reliability under combined random load sequences. Comput Struct 9(5):489–494

Rezaeian S, Der Kiureghian A (2010) Simulation of synthetic ground motions for specified earthquake and site characteristics. Earthq Eng Struct Dyn 39(10):1155–1180

Rubinstein RY (1981) Simulation and the Monte Carlo method. Wiley, New York

Russell SJ, Norvig P (2003) Artificial intelligence: a modern approach, 2nd edn. Prentice-Hall, Upper Saddle River

Schuëller GI, Stix R (1987) A critical appraisal of methods to determine failure probabilities. Struct Saf 4(4):293–309

Shinozuka M (1983) Basic analysis of structural safety. J Struct Eng 109(3):721–740

Shome N, Cornell CA (1999) Probabilistic seismic demand analysis of nonlinear structures. Report no RMS-35, Department of Civil and Environmental Engineering, Stanford University, Stanford, 320 pp

Shome N, Cornell CA, Bazzurro P, Carballo JE (1998) Earthquakes, records, and nonlinear responses. Earthq Spectra 14(3):469–500

Stein M (1987) Large sample properties of simulations using Latin hypercube sampling. Technometrics 29(2):143–151

Straub D, Der Kiureghian A (2010) Bayesian network enhanced with structural reliability methods: methodology. J Eng Mech 136(10):1248–1258

Tvedt L (1990) Distribution of quadratic forms in normal space-application to structural reliability. J Eng Mech 116(6):1183–1197

Vamvatsikos D, Cornell CA (2004) Applied incremental dynamic analysis. Earthq Spectra 20(2):523–553

Vamvatsikos D, Fragiadakis M (2010) Incremental dynamic analysis for estimating seismic performance sensitivity and uncertainty. Earthq Eng Struct Dyn 39(2):141–163

Vořechovský M, Novák D (2009) Correlation control in small-sample Monte Carlo type simulations I: a simulated annealing approach. Probab Eng Mech 24(3):452–462

Zareian F, Krawinkler H (2007) Assessment of probability of collapse and design for collapse safety. Earthq Eng Struct Dyn 36(13):1901–1914

Zhao Y-G, Ono T (2001) Moment methods for structural reliability. Struct Saf 23(1):47–75

Zuev KM, Beck JL, Au S-K, Katafygiotis LS (2012) Bayesian post-processor and other enhancements of Subset Simulation for estimating failure probabilities in high dimensions. Comput Struct 92:283–296

S

Seismic Resilience

Domenico Asprone and Gaetano Manfredi
Department of Structures for Engineering and
Architecture, University of Naples "Federico II",
Naples, Italy

Synonyms

City resilience; Community resilience; Disaster
resilience; Ecosystem resilience; Engineering
resilience; Urban resilience

Introduction

Resilience has several meanings in academic dis-
course. It is derived from the Latin term *resilire*,
which means "to bounce back." It is used in
multiple scientific contexts to identify the capa-
bility to recover, absorb shocks, and restore equi-
librium after a perturbation. First, the concept of
resilience was introduced in the nineteenth cen-
tury in physics to indicate the ability of materials
to withstand impulsive loads without suffering
damages. Then, resilience was also used in med-
icine (Pfeiffer 1929) and psychology (Werner
1971; Garmezy 1973).

Recently, resilience is triggering increasing
interest in other scientific contexts, referred to
communities, urban systems, and built environ-
ment, as the capability to recover from natural
and human-induced disasters. The advent of the
concept of resilience in this context is the result of
an increasing need for a response to new and
intense threats to modern societies. Increasing
interdependence and complexity of contempo-
rary cities along with more severe events induced
by climate change is making modern societies
ask for prevention, preparation, reduced impacts
and damages, and rapid recovery; that is resil-
ience. This urgent need is pushing scientific
community to discuss about the best approach to
resilience against disasters and, first of all, to
define disaster resilience (see "▶ Earthquake
Disaster Recovery: Leadership and Governance";

"▶ Earthquakes and Their Socio-economic
Consequences").

Contemporary cities can be interpreted as
complex systems, composed of dynamic relation-
ships between physical environment, i.e., infra-
structural systems (e.g., utility and transportation
networks) and more in general all lifelines, natu-
ral environment and social environment,
consisting of communities and their internal rela-
tionships. Hence, according to a general defini-
tion, cities can be considered resilient if able to
cope with extreme events without suffering dev-
astating losses and damages to their physical
systems or reduced quality of life for the inhabi-
tants (Godschalk 2003). However, a comprehen-
sive definition is still not available, given the
complexity in defining the properties of urban
systems and the response of cities to extreme
events.

What are the real operations taking place in
urban systems? What about the dynamic equilib-
rium at the basis of the urban system operations?
What is meant by limited damages and preserva-
tion of functionality for urban systems after
extreme events? Does the optimal response of
urban systems to extreme events, i.e., the "resil-
ient" response, depend on the type of extreme
event? These are just some of the questions that
make the concept of resilience exploding with
different and multidisciplinary meanings.

Defining Resilience

The extreme events that urban systems can be
subjected into can be divided into four categories
(O'Brien et al. 2006), and each of them may need
different approaches to define a resilient
response:

– Natural events, such as earthquakes, tsunamis,
  floods, etc.
– Technological events, or human-induced
  events, such as accidents on transport net-
  works, industrial accidents, terrorist attacks,
  etc. In developed countries, these events are
  extensively considered in risk management
  strategies, aimed at public awareness of risks,

risk mitigation, and building capacity to withstand and recover from extreme events.

– Humanitarian emergencies, i.e., man-made or natural events, hitting very vulnerable and weak populations, such as droughts, famines, epidemics, wars, etc.

– Events induced by climate changes, i.e., events that may still be counted as natural events (floods, landslides, etc.) or as humanitarian emergencies (drought, heat waves, etc.) but that are induced by recent climate changes and can affect unprepared populations with unexpected intensity. In addition, for the climate-induced events, the most effective mitigation techniques are indirect, undertaken on a global scale, through the reduction of greenhouse gases.

Hence, resilience of urban systems against different events varies, and different strategies can be implemented to build resilient cities, in terms of risk mitigation actions, shock preparedness, and recovery capability from damages.

Given the multiplicity of points of view to disaster resilience, numerous definitions of resilience are available in literature; an excellent review of these definitions is presented by Zhou and coworkers (Zhou et al. 2010), who elaborated the most of the following list:

• Holling (1973, 1986), Holling et al. (1995)

Resilience is defined as the amount of disturbance that can be sustained by a system before a change in system control or structure occurs. It could be measured by the magnitude of disturbance the system can tolerate, still persisting in its pre-disturbance state

• Timmerman (1981)

Resilience is the ability of human communities to withstand external shocks or perturbations to their infrastructure and to recover from such perturbations

• Pimm (1984)

Resilience is the speed with which a system returns to its original state following a perturbation

• Pimm (1984), Holling et al. (1995), Gunderson et al. (1997)

Resilience of an ecological system relates to the functioning of the system, rather than the stability of its component populations, or even the ability to maintain a steady ecological state

• Wildavsky (1991)

Resilience is the capacity to cope with unanticipated dangers after they have become manifest, learning to bounce back

• Dovers and Handmer (1992)

Re-active and pro-active resilience of society can be distinguished, based on the major difference between ecosystems (that react to disturbances) and societies (that can plan in advance, due to human capacity for anticipation and learning)

• Adger (1997, 2000)

Social resilience could be measured through proxies of institutional change and economic structure, property rights, access to resources, and demographic change

• Horne and Orr (1998)

Resilience is a fundamental quality of individuals, groups and organizations, and systems as a whole to respond productively to significant change that disrupts the expected pattern of events without engaging in an extended period of regressive behavior

• Mallak (1998)

Resilience is the ability of an individual or organization to expeditiously design and implement positive adaptive behaviors matched to the immediate situation, while enduring minimal stress

• Miletti (1999)

Local resiliency with regard to disasters means that a locale is able to withstand an extreme natural event without suffering devastating losses, damage, diminished productivity or quality of life and without a large amount of assistance from outside the community

• Comfort (1999)

The capacity to adapt existing resources and skills to new systems and operating conditions

• Miletti (1999), Geis (2000), Chen et al. (2008)

In the context of disaster management, resilience is used to describe the ability to resist or adapt to stress from hazards, and the ability to recover quickly

- Adger (2000), Kimhi and Shamai (2004)

  Social resilience is understood as having three properties: resistance, recovery and creativity, in which (1) resistance relates to a social entity's efforts to withstand a disturbance and its consequences, and can be understood in terms of the degree of disruption that can be accommodated without social entity undergoing long-term change; (2) Recovery relates to an entity's ability to pull through the disturbance, and can be understood in terms of the time taken for an entity to recover from a disruption. (3) Creativity is represented by a gain in resilience achieved as part of the recovery process, and it can be attained by adapting to new circumstances and learning from the disturbance experience

- Carpenter et al. (2001)

  The Resilience Alliance consistently refers to social-ecological systems (SES) and defines their resilience by considering three distinct dimensions: (1) the amount of disturbance a system can absorb and still remain within the same state or domain of attraction; (2) the degree to which the system is capable of self-organization; (3) the degree to which the system can build and increase the capacity for learning and adaptation

- Paton et al. (2000)

  Resilience describes an active process of self-righting, learned resourcefulness and growth—the ability to function psychologically at a level far greater than expected given the individual's capabilities and previous experiences

- UN/ISDR (2009)

  The ability of a system, community or society exposed to hazards to resist, absorb, accommodate to and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions. Resilience means the ability to "resile from" or "spring back from" a shock. The resilience of a community in respect to potential hazard events is determined by the degree to which the community has the necessary resources and is capable of organizing itself both prior to and during times of need.

- Bruneau et al. (2003)

  An analysis of seismic resilience at four levels: (1) technical, how physical systems perform when subjected to earthquake forces; (2) organizational, the ability to respond to emergencies and carry out critical functions; (3) social, the capacity to reduce the negative social consequences of loss of critical services; and (4) economic, the capacity to reduce both direct and indirect economic losses. Resilience has four dimensions: (1) robustness, strength to withstand a given level of stress without loss of function; (2) redundancy, the extent to which elements and systems are substitutable; (3) resourcefulness, the capacity to identify problems, establish priorities and mobilize resources; and (4) rapidity, the capacity to meet priorities and achieve goals in a timely manner. A resilient system has: (1) reduced probability of failures; (2) reduced consequences from failures; and (3) reduced time to recovery

- Kendra and Wachtendorf (2003)

  The ability to respond to singular or unique events

- Cardona (2003)

  The capacity of the damaged ecosystem or community to absorb negative impacts and recover from these

- Pelling (2003)

  The ability of an actor to cope with or adapt to hazard stress

- Rockstrom (2003)

  Strategies of social resilience building include manageable strategies, such as institutional development, land reform, land tenure, diversification, marketing, human capacity building, and unmanageable ones, such as relief food, cereal banks, social networks, virtual water imports

- Rose (2004, 2007)

  Resilience includes inherent resilience (ability under normal circumstances) and adaptive resilience (ability in crisis situations due to ingenuity or extra effort)

- Aguirre (2006)

  A resilient social entity absorbs, responds and recovers from the shock; and improvises and innovates in response to disturbances

- Maguire and Hagan (2007)

  In broad terms, social resilience is the capacity of a social entity (e.g., a group or community) to bounce back or respond positively to adversity

- Kang et al. (2007)

  Resilience is the ability of the system to recover once hazard has occurred and can be measured by the duration of an unsatisfactory condition.

- Asprone and Manfredi (2014)

  An extreme event and the resulting changes, moving urban systems to new dynamic equilibrium, represent a phase in urban life cycle; resilience represents the sustainability of this phase, from the economic, social and environmental point of view, for all the present and future actors, directly and indirectly involved in the recovery process.

Resilience of Ecosystems and Engineering Resilience

To apply the concept of resilience to complex systems, such as cities, two approaches can be followed: the resilience of ecosystems (a) and the engineering resilience (b). In the first, proposed and developed by Holling (1973, 1986, 2001), resilience can be defined as the ability of a system in dynamic equilibrium, subject to external shocks, to move back to a dynamic equilibrium state. On the contrary, engineering resilience, developed by Pimm and other authors (Pimm 1984; Bruneau et al. 2003) can be defined as the ability of a system to absorb an external shock and quickly return to the initial state.

Apparently the first definition may be more complete and suitable for urban systems; in fact, moving from the fact that a complex system in dynamic equilibrium (as the urban system, which consists of physical and social subsystems linked by a dynamic network of relationships) can present different equilibrium states (i.e., can "work") in various configurations, it can be concluded that a positive response to a negative external shock can be also represented by a new equilibrium state, different from the previous one. For example, looking at the terrorist attack on the World Trade Center in New York on 11 September 2001, it can be said that the city of New York had a "resilient" response. New York quickly recovered from the social and economic effects induced by the event, even if a new equilibrium was reached in a different configuration of the physical system, i.e., without rebuilding the World Trade Center towers and relocating the activities that took place there, that is, re-thinking that space (i.e., "ground zero" site).

Furthermore, the social value of the towers, representing a crucial symbol for the collective identity of the city of New York, has been preserved by reconfiguring the city in a different dynamic equilibrium. The towers' values still exist and their physical absence was recovered from the social and cultural point of view. Nevertheless, engineering resilience is also extremely meaningful. In fact, one could argue that a complex and dynamic system, as the city, is always able to reach a state of equilibrium after a shock, because the ability of cities to adapt to changes is extremely high. But the new post-event dynamic equilibrium could be "worse" than the previous one. In this case only with an engineering resilience approach a "negative" response can be appreciated; for example, quality and performance indicators of the urban system can be used for this scope. Hence, aiming at synthesizing the different approaches, it can be concluded that the urban system is resilient if, after a shock, it can reach a dynamic equilibrium state, even if different from the previous one, but the urban system is really resilient if, at the same time, certain indicators of quality and performance of the system return to pre-shock values. This concept was also introduced by Dalziell and McManus (2004), which affirmed the need to introduce metrics of resilience. Hence, the questions the scientific community is currently dealing with refer to the indicator system that should be used. Are the indicators currently in use to assess city sustainability suitable for the scope? Which quality and performance indices can describe the "effectiveness" of the response to external shocks?

A further approach to the concept of resilience of cities leads to the definition of the social resilience. It moves from the centrality of communities in urban systems, i.e., the predominance of the social over the physical system; according to this approach, social resilience is defined as the ability of communities to deal with external shocks, managing the changes induced on infrastructures, on the external environment, and on the economic and social systems (Adger 1997, 2000). Social resilience, according to Adger, can be measured by three characteristics: resistance

S

to external shocks (a), the ability to recover from external shocks (b), and creativity (c), that is, the ability to adapt to new circumstances (Adger 2000; Kimhi and Shamai 2004). With a similar approach, Lorenz (2010) affirms that social resilience consists of the adaptive capacity (a), that is, the ability to change in order to withstand the external shocks; the coping capacity (b), that is, the ability to preserve and give continuity to the system of relationships, given the external shocks; and the participative capacity (c), that is, the self-organization ability, aimed at coping with external shocks. Hence, the approach to social resilience affirms the centrality of communities, able to manage the other physical elements and determine resilient urban systems.

In all the approaches so far analyzed, however, resilience is perceived as the ability of the city to have a "positive" response, when exposed to an external shock, as an extreme event. The main issue is the need to define a "positive" response: the ability to return to the previous equilibrium configuration or even a different reconfigured equilibrium state. And, in a complex and dynamic system, what is an equilibrium state? Furthermore, is resilience to be reached separately both in the social and the physical systems, or does social system resilience entail physical system resilience? Thus, is the physical system resilience condensed in community resilience? Hence, is community, representing the only decision maker for urban management, the only master of a city's destiny, the key to a resilient city?

Quantifying Resilience

The increasing interest in resilience requires methodological frameworks to measure and assess it. Measuring disaster resilience would help understand and improve resilience of urban systems against risks and implement the most effective strategies to "bounce back" from disasters. Aimed at this goal, different studies have been developed, proposing measurement frameworks of disaster resilience and other properties related to resilience.

The most important methodologies available in literature can be divided into two categories: (a) the physical resilience approach and (b) the social-economic resilience approach. In the former, attention is focused on performances of physical systems, e.g., single structures, urban lifelines, and transportation systems. In this case, resilience is measured as the capability of the physical components and systems to recover their functionality. Mainly, these methods are developed and proposed within engineering community. In the latter, attention is focused on social systems, and resilience is measured as the capability of communities to recover a good life quality level. Mainly, these methods are proposed in social sciences community.

Physical Resilience

Bruneau et al. (2003) define resilience as characterized by four main properties: robustness, rapidity, redundancy, and resourcefulness (4 R's), to be managed and computed as proxies of resilience. *Robustness* is related to the "strength, or the ability of elements, systems, and other units of analysis to withstand a given level of stress or demand without suffering degradation or loss of function." *Rapidity* is "the capacity to meet priorities and achieve goals in a timely manner in order to contain losses and avoid future disruption." *Redundancy* refers to the availability of substitutable elements or systems in the aftermath of a disruption, and *resourcefulness* is the capacity to mobilize material and human resources. Within this approach, different methods have been proposed, whose final scope is to compute resilience as the ability to cope with degradation in system performance $Q(t)$, over time. Numerically, resilience R is often computed as the area underneath function $Q(t)$, divided by the time to restore the pre-event performance (Fig. 1):

$$R = \frac{\int_{t_0}^{t_1} Q(t)dt}{t_1 - t_0} \qquad (1)$$
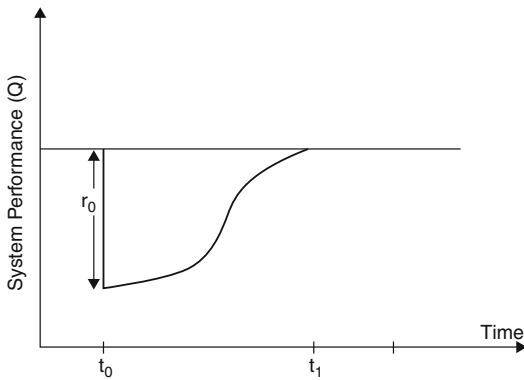
Seismic Resilience, Fig. 1 Physical resilience

Being t_0 the time of the event and t_1 the time of the total recovery of the pre-event performance. This approach has been applied to buildings (Bruneau and Reinhorn 2007), bridges (Decò et al. 2013), road networks (Arcidiacono et al. 2012), and urban infrastructure systems (Ouyang and Dueñas-Osorio 2012; Franchin and Cavalieri 2013), using different performance functions $Q(t)$.

Further works are increasingly conducted on this topic and made available in literature, focusing on different urban systems and different performance functions $Q(t)$; however, the most of the works in recent literature share the theoretical scheme in Eq. 1 to compute resilience. Recently, Cavallaro et al. (2014) and Franchin and Cavalieri (2013) applied this approach to social-physical graphs (merging physical networks and social components) using as performance $Q(t)$ the efficiency of the network in the "social" nodes, aiming at measuring the capability of the physical systems to serve their end users.

Social-Economic Resilience

Studies aimed at computing resilience from a social perspective focus on economic, demographic, and institutional variables, in time and space. In example, economic growth and the distribution of income among people are fundamental aspects of resilience (Adger 2000) and are often used to compute resilience. Attitude to mobility and migration or amount of young people is also related to resilience (Ruitenbeek 1996; Adger 2000). Social memory of past changes and impacts (Olick and Robbins 1998) also relates to the capacity of communities to adapt and cope with disasters, that is, resilience. Hence, different authors refer to this kind of variables to estimate community resilience, in terms of preparedness and copying capacity to disasters. Specific indicators have been also developed, moving from social-economic variable. This is the case of the Disaster Deficit Index (DDI), proposed by Cardona et al. (2008), measuring country resilience against disasters from a macroeconomic perspective:

$$DDI = \frac{L_R}{R_E} \tag{2}$$

being L_R the maximum expected direct economic impact of possible disasters and R_E the available internal and external resources that can be made available to face disasters.

A recent attempt to integrate physical and social-economic perspectives of resilience has been done with the PEOPLES Resilience Framework (Renschler et al. 2010), linking different resilience dimensions (technical, organizational, societal, and economic) and resilience properties (robustness, redundancy, resourcefulness, and rapidity) as proposed by Bruneau et al. (2003).

Summary

An increasing research interest in scientific community is focusing on disaster resilience of urban environment, from different scientific perspectives, including those from risk engineering and social sciences communities. A comprehensive approach to resilience is still not available, but the need for a general methodology to manage and measure urban resilience is urged due to the

increasing interdependence and complexity of contemporary cities along with more severe events induced by climate change.

This entry deals with the most recent approaches to urban disaster resilience. It moves from the description of a large variety of definitions of disaster resilience, as proposed by researchers of different disciplines. The differences between engineering resilience and resilience of ecosystems are discussed. Thus, different approaches to compute and measure resilience are described, focusing on physical resilience, i.e., resilience of infrastructures and lifelines, and social-economic resilience, i.e., resilience of social systems and communities.

Cross-References

▶ Earthquake Disaster Recovery: Leadership and Governance
▶ Earthquakes and Their Socio-economic Consequences

References

Adger WN (1997) Sustainability and social resilience in coastal resource use. CSERGE working paper series, Centre for Social and Economic Research on the Global Environment, University of East Anglia, Norwich and University College London

Adger WN (2000) Social and ecological resilience: are they related? Prog Hum Geogr 24(3):347–364

Aguirre B (2006) On the concept of resilience. Disaster Research Center, University of Delaware, Delaware

Arcidiacono V, Cimellaro GP, Infuso A, Reinhorn AM, Road network's disaster resilience assessment methodology (2012) Bridge maintenance, safety, management, resilience and sustainability. In: Proceedings of the sixth international conference on bridge maintenance, safety and management, Como (Italy), pp 1912–1919

Asprone D, Manfredi G (2014) Linking disaster resilience and urban sustainability: a global approach for future cities. Disasters J 39(1):s96–s111

Bruneau M, Reinhorn AM (2007) Exploring the concept of seismic resilience for acute care facilities. Earthq Spectra 23(1):41–62

Bruneau M, Chang S, Eguchi R, Lee G, O'Rourke T, Reinhorn A, Shinozuka M, Tierney K, Wallace W, von Winterfeldt D (2003) A framework to quantitatively assess and enhance seismic resilience of communities. Earthq Spectra 19:733–752

Cardona OD (2003) The notions of disaster risk: conceptual framework for integrated management. Information and indicators program for disaster risk management. Inter-American Development Bank, Manizales

Cardona OD, Ordaz MG, Marulanda MC, Barbat AH (2008) Estimation of probabilistic seismic losses and the public economic resilience – an approach for a macroeconomic impact evaluation. J Earthq Eng 12(S2):60–70

Carpenter S, Walker B, Anderies JM, Abel N (2001) From metaphor to measurement: resilience of what to what? Ecosystems (N Y, Print) 4(8):765–781

Cavallaro M, Asprone D, Latora V, Manfredi G, Nicosia V (2014) City ecosystem resilience analysis in case of disasters. Comput-Aided Civ Infrastruct Eng 29(8):608–625

Chen SC, Ferng JW, Wang YT, Wu TY, Wang JJ (2008) Assessment of disaster resilience capacity of hillslope communities with high risk for geological hazards. Eng Geol 98(3–4):86–101

Comfort L (1999) Shared risk: complex systems in seismic response. Pergamon, New York

Dalziell EP, McManus ST (2004) Resilience, vulnerability, and adaptive capacity: implications for system performance. Stoos, Switzerland: 1st International Forum for Engineering Decision Making (IFED), 17 pp

Decò A, Bocchini P, Frangopol DM (2013) A probabilistic approach for the prediction of seismic resilience of bridges. Earthq Eng Struct Dyn 42(10):1469–1487

Dovers SR, Handmer JW (1992) Uncertainty, sustainability and change. Glob Environ Chang 2(4):262–276

Franchin P, Cavalieri F (2013) A framework for physical simulation of critical infrastructures, accounting for interdependencies and uncertainty. In: 11th international conference on structural safety & reliability ICOSSAR, Taylor & Francis, New York

Garmezy N (1973) Competence and adaptation in adult schizophrenic patients and children at risk. In: Dean SR (ed) Schizophrenia: the first ten Dean award lectures. MSS In- formation Corporation, New York

Geis DE (2000) By design: the disaster resistant and quality-of-life community. Nat Hazard Rev 1(3):106–120

Godschalk D (2003) Urban hazard mitigation: creating resilient cities. Nat Hazard Rev 4:136–143

Gunderson LH, Holling CS, Pritchard L, Peterson GD (1997) Resilience in ecosystems, institutions and societies. Discussion paper 95. Beijer International Institute of Ecological Economics, Stockholm

Holling CS (1973) Resilience and stability of ecological systems. Annu Rev Ecol Syst 4:1–23

Holling CS (1986) The resilience of terrestrial ecosystems: local surprise and global change. In: Clark WC, Munn RE (eds) Sustainable development of the

biosphere. Cambridge University Press, Cambridge, pp 292–317

Holling CS (2001) Understanding the complexity of economic, ecological, and social system. Ecosystems (N Y, Print) 4:390–405

Holling CS, Schindler DW, Walker BW, Roughgarden J (1995) Biodiversity in the functioning of ecosystems: an ecological synthesis. In: Perrings C, Ma̋ler KG, Folke C, Holling CS, Jansson BO (eds) Biodiversity loss: economic and ecological issues. Cambridge University Press, Cambridge, pp 44–83

Horne JF, Orr JE (1998) Assessing behaviors that create resilient organizations. Employ Relat Today 24(4):29–39

Kang B, Lee SJ, Kang DH, Kim YO (2007) A flood risk projection for Yongdam dam against future climate change. J Hydro Environ Res 1(2):118–125

Kendra MJ, Wachtendorf T (2003) Elements of resilience after the world trade center disaster: reconstructing New York city's emergency operation center. Disasters 27(1):37–53

Kimhi S, Shamai M (2004) Community resilience and the impact of stress: adult response to Israel's withdrawal from Lebanon. J Community Psychol 32(4):439–451

Lorenz DF (2010) The diversity of resilience: contributions from a social science perspective. Nat Hazards 2010:1–18

Maguire B, Hagan P (2007) Disasters and communities: understanding social resilience. Aust J Emerg Manag 22(2):16–20

Mallak L (1998) Resilience in the healthcare industry. In: Paper presented at the seventh annual engineering research conference, Banff, 9–10 May

Miletti D (1999) Disasters by design: a reassessment of natural hazards in the United States. Joseph Henry Press, Washington, DC

O'Brien G, O'Keefe P, Rose J, Wisner B (2006) Climate change and disaster management. Disasters 30(1):64–80

Olick JK, Robbins J (1998) Social memory studies: from "collective memory" to historical sociology of mnemonic practices. Annu Rev Sociol 24:105–140

Ouyang M, Dueñas-Osorio L (2012) Time-dependent resilience assessment and improvement of urban infrastructure systems. Chaos 22(3):art no 033122

Paton D, Smith L, Violanti J (2000) Disasters response: risk, vulnerabilities and resilience. Disaster Prev Manag 9(3):173–179

Pelling M (2003) The vulnerability of cities: natural disasters and social resilience. Earthscan, London

Pfeiffer K (1929) Untersuchungen über die Resilienz der durch die Prothesen beanspruchten Gewebe und ihre Bedeutung für die Okklusion der Prothesen. Berichthaus, Zürich

Pimm SL (1984) The complexity and stability of ecosystems. Nature 307:321–326

Renschler C, Frazier A, Arendt L, Cimellaro GP, Reinhorn AM, Bruneau M (2010) Framework for defining and measuring resilience at the community scale: the PEOPLES resilience framework. MCEER technical report –MCEER-10-006, pp 91, University at Buffalo (SUNY), The State University of New York, Buffalo

Rockstrom J (2003) Resilience building and water demand management for drought mitigation. Phys Chem Earth 28:869–877

Rose A (2004) Defining and measuring economic resilience to disasters. Disaster Prev Manag 13:307–314

Rose A (2007) Economic resilience to natural and man-made disasters: multidisciplinary origins and contextual dimensions. Environ Hazards 7:383–398

Ruitenbeek HJ (1996) Distribution of ecological entitlements: implications for economic security and population movement. Ecol Econ 17:49–64

Timmerman P (1981) Vulnerability, resilience and the collapse of society: a review of models and possible climatic applications. Institute for Environmental Studies, University of Toronto, Toronto

United Nations International Strategy for Disaster Reduction (UNISDR) (2009) Global assessment report on disaster risk reduction: risk and poverty in a changing climate. UNISDR, Geneva

Werner E (1971) The children of Kauai. A longitudinal study from the prenatal period to age ten. University of Hawaii Press, Honolulu

Wildavsky A (1991) Searching for safety. Transaction, New Brunswick

Zhou HJ, Wang JA, Wan JH et al (2010) Resilience to natural hazards: a geographic perspective. Nat Hazards 53(1):21–41

Seismic Response Prediction of Degrading Structures

Ching Hang Ng[1], Nopdanai Ajavakom[2] and Fai Ma[3]
[1]US Nuclear Regulatory Commission, Office of Nuclear Reactor Regulation, Washington, DC, USA
[2]Department of Mechanical Engineering, Chulalongkorn University, Bangkok, Thailand
[3]Department of Mechanical Engineering, University of California, Berkeley, Berkeley, CA, USA

S

Synonyms

Degrading structures; Hysteresis; Nonlinear response; System identification

Introduction

All structures degrade when acted upon by cyclic forces such as those associated with earthquakes, high winds, and sea waves. Development of a practical model of degrading structures that would match experimental observations is an important task. Furthermore, identification and prediction of deterioration is a problem of considerable practical significance. Under cyclic excitation, degradation manifests itself in the evolution of the associated hysteresis loops. Theoretical research in internal friction in the last few decades has noticeably increased the conceptual understanding of hysteresis (Kojic and Bathe 2005; Krasnoselskii and Pokrovskii 1989). Practical issues related to internal friction, however, have not been adequately addressed. The lack of a practical theory of hysteretic evolution is at present a major barrier to successful design of structures against performance deterioration.

In the past 30 years, cyclic performance testing of structural joints and subassemblies around the world has generated a substantial amount of experimental data on load–displacement traces. In the same period, the generalization of the Bouc–Wen differential model of hysteresis permits curve fitting of practically any hysteretic trace with a suitable choice of its 13 control parameters (Bouc 1967; Wen 1976). Using system identification techniques, it appears highly feasible to utilize the generalized differential model of hysteresis and the extensive database of experimental hysteretic traces to deduce a working model for degrading structures. A fundamental objective of this research project is to do just that.

Many questions naturally spring to mind. What are the advantages of the differential model over other empirical models of hysteresis? How can an empirical model of hysteresis be used for prediction of nonlinear system response? What are the factors that would affect the precision of prediction? These questions will be answered along the way. Two principal tasks in connection with hysteretic evolution will be addressed. First, a robust identification algorithm will be used to generate hysteretic models of a deteriorating structure from its experimental load–displacement traces. Second, it will be shown that a hysteretic model obtained by system identification can be used to predict the future performance of the same deteriorating structure. The organization of this entry is as follows. In section "Differential Model of Hysteresis," the smoothly varying Bouc–Wen differential model of hysteresis in both its classical (non-degrading) and generalized forms is described. A robust identification algorithm is constructed in section "System Identification" to generate hysteretic models of a deteriorating structure from its experimental load–displacement traces. This algorithm is based upon the generalized Bouc–Wen model and differential evolution, streamlined through global sensitivity analysis. The model thus generated can account for degradation and pinching effects, which are prominent features of real-life structural deformation. To obtain experimental data for model validation, cyclic performance tests of simple joints and subassemblies are reported in section "Prediction of Performance." Using an experimental load–displacement trace, a working hysteretic model is identified. It will be shown that the hysteretic model obtained by identification may be used for predicting the nonlinear response of the same structure when driven by other cyclic loads. Finally, the requirements for accurate prediction of system response will be discussed. The terms "Bouc–Wen model" and "differential model" will be used interchangeably throughout the entry.

Differential Model of Hysteresis

When a structure is subjected to severe cyclic loading, the hysteresis loops associated with the structural response are memory dependent. That means the evolution of hysteresis loops depends not only on the instantaneous deformation but also on the history of deformation. This memory nature of degradation makes modeling and analysis challenging. The generalized Bouc–Wen model is one of the widely used empirical models capable of modeling the memory effects.

Suppose the equation of motion of a multi-degree-of-freedom system can be decoupled and, along the direction of the generalized coordinate x, the system is governed by

$$m\ddot{x} + c\dot{x} + r(x,z) = f(t) \qquad (1)$$

where m and c are, respectively, the mass and damping coefficients, z is an imaginary hysteretic displacement, and $r(x, z)$ is the total restoring force. It is assumed that the excitation $f(t)$ is cyclic. In the development of differential model, the restoring force $r(x, z)$ is separated into an elastic (linear) component and a hysteretic (nonlinear) component by

$$r(x,z) = \alpha kx + (1 - \alpha)kz \qquad (2)$$

where k is the stiffness coefficient and $0 \leq \alpha \leq 1$ is a weighting parameter. Obviously, the restoring force is purely hysteretic if $\alpha = 0$; it is purely elastic if $\alpha = 1$. A diagrammatic representation of the system is shown in Fig. 1.

Hysteresis loops may be generated if the hysteretic displacement z and the total displacement x are connected by the nonlinear differential equation (Bouc 1967; Wen 1976)

$$\dot{z} = A\dot{x} - \beta|\dot{x}||z|^{n-1}z - \gamma\dot{x}|z|^n \qquad (3)$$

There are five unspecified loop parameters A, α, β, γ, and n in Eqs. 2 and 3, which together represent the classical Bouc–Wen model. The parameters A, β, and γ are basic hysteresis shape parameters. The parameter n controls the sharpness of yield. Over the years, the original Bouc–Wen model has been extended, and new parameters have been added to fit hysteretic shapes arising from deteriorating systems. The result is a contemporary model with thirteen control parameters given by

$$\dot{z} = h(z)\left\{\frac{A\dot{x} - v\left(\beta|\dot{x}||z|^{n-1}z + \gamma\dot{x}|z|^n\right)}{\eta}\right\} \qquad (4)$$

In the above expression, v and η are degradation shape functions (Baber and Wen 1981), and $h(z)$
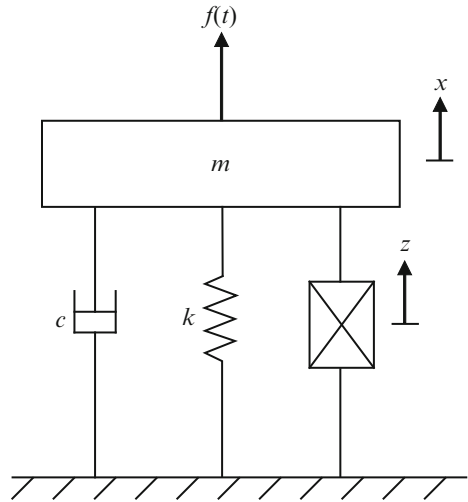
Seismic Response Prediction of Degrading Structures, Fig. 1 Schematic diagram of a hysteretic degrading system

is a pinching shape function. In general, degradation depends on the response duration and severity. A convenient measure of such combined effect is the energy

$$E(t) = \int_0^t (1 - \alpha)kz\dot{x}\,dt \qquad (5)$$

dissipated through hysteresis from initial time $t = 0$ to present time t. Since

$$\varepsilon(t) = \int_0^t z\,\dot{x}\,dt \qquad (6)$$

is proportional to $E(t)$, it may also be used as a measure of response duration and severity. Both degradation shape functions v and η are assumed to depend linearly on ε as the system evolves:

$$v(\varepsilon) = 1 + \delta_v\,\varepsilon \qquad (7)$$

$$\eta(\varepsilon) = 1 + \delta_\eta\varepsilon \qquad (8)$$

Two unspecified degradation parameters δ_v and δ_η are thus introduced. Under cyclic excitation, the pinching of hysteresis loops is often observed.

For example, pinching may be associated with slippage of longitudinal reinforcement in reinforced concrete or with X-braced steel frames driven by high-shear loads. Baber and Noori (1986) introduced a slip-lock element behaving quite similarly to a hardening nonlinear spring, with the special characteristics that the "slip" zone stiffness is nearly zero, while the "lock" zone stiffness is infinite. The "slip" zone stiffness is initially zero and increases continuously as the system degrades with time. Thus, the pinching shape function $h(z)$ takes the form (Foliente 1995)

$$h(z) = 1 - \zeta_1 \; e^{-[z \; \mathrm{sgn}(\dot{x}) - q \; z_u]^2/\zeta_2^2} \qquad (9)$$

where sgn is the signum function and z_u is the ultimate value of z given by

$$z_u = \left(\frac{A}{v(\beta + \gamma)}\right)^{1/n} \qquad (10)$$

The two functions $\zeta_1(\varepsilon)$ and $\zeta_2(\varepsilon)$ control the progress of pinching and are written as

$$\zeta_1(\varepsilon) = \zeta_s \left[1 - e^{(-p \; \varepsilon)}\right] \qquad (11)$$

$$\zeta_2(\varepsilon) = \left(\psi + \delta_\psi \; \varepsilon\right)(\lambda + \zeta_1) \qquad (12)$$

Six pinching parameters ζ_s, q, p, ψ, δ_ψ, and λ are therefore present. Altogether there are thirteen loop parameters of hysteresis: A, α, β, γ, n, δ_v, δ_η, ζ_s, q, p, ψ, δ_ψ, and λ. This generalized model of hysteresis possesses all the important features observed in real structures, which include strength degradation, stiffness degradations, and pinching of the successive hysteresis loops.

Compared with other parametric models, a differential model of hysteresis has many advantages (Foliente 1995; Song and Der Kiureghian 2006). The primary one is its ability to generate a large variety of realistic hysteresis loops. Another advantage is the coupling of the equation of motion (1) to either Eqs. 3 or 4 to form an overall differential system. This greatly facilitates any theoretical and numerical manipulations.

Nonetheless, it should be noted that the differential model is one-dimensional and cannot account for interactions between loadings in different directions. There are generalizations in which the differential model is augmented to a multi-degree linear system (Roberts and Spanos 1990). In such generalizations, however, it is also assumed that the system could be decoupled. Since no new physics is obtained, these generalizations would not be pursued herein. Finally, it must be emphasized that the extended Bouc–Wen model of hysteresis is an empirical model. As such, it is not derivable from the fundamental postulates of mechanics and the exact physical meanings of its thirteen parameters are not fully understood. The probable role played by each parameter is summarized in Table 1. Also contained in Table 1 are the sensitivity rankings (Ma et al. 2004) of the control parameters. As will be explained in the next section, in system identification, it is required to estimate only a subset of the 13 unspecified parameters.

System Identification

In analysis, the response of a system is sought if the system model and excitation are known. This is sometimes termed a forward problem. In this interpretation, the inverse problem of finding a system model given the excitation and response is called system identification. System identification in the time domain involves the determination of unspecified parameters of an assumed system model. This can always be formulated as an optimization problem. In the present context, suppose the differential model of hysteresis is adopted and a set of measured excitation–response data from cyclic performance tests of an inelastic structure is given. How can the loop parameters of hysteresis be estimated from the measured data? For each choice of the parameters, the response of the degrading structure subjected to the given excitation can be obtained by numerical simulation. The calculated response data can then be compared to the measured data to see if there are large errors. Obviously, the assumed loop parameters

Seismic Response Prediction of Degrading Structures, Table 1 Parameters of the generalized differential model of hysteresis

Parameter	Description	Local sensitivity ranking (highest = 1)	Global sensitivity ranking (highest = 1)
α	Ratio of linear to nonlinear response	1	2
A	Basic hysteresis shape control	Not varied	Not varied
β	Basic hysteresis shape control	5	4
γ	Basic hysteresis shape control	6	5
n	Sharpness of yield	8	7
δ_v	Strength degradation	12	9
δ_η	Stiffness degradation	4	8
ζ_s	Measure of total slip	2	1
q	Pinching initiation	9	6
p	Pinching slope	3	10
ψ	Pinching magnitude	7	3
δ_ψ	Pinching rate	11	12
λ	Pinching severity	10	11

provide a good fit if the errors are small. Thus, this amounts to estimating the 13 control parameters of the differential model of hysteresis when a load–displacement trace is given. The optimization problem can be stated as the determination of the parameter vector

$$\mathbf{p} = \left(A, \alpha, \beta, \gamma, n, \delta_v, \delta_\eta, \zeta_s, q, p, \psi, \delta_\psi, \lambda \right) \quad (13)$$

such that the objective function

$$g(\mathbf{p}) = \frac{1}{N} \sum_{j=1}^{N} \left[x(t_j) - \hat{x}(t_j | \mathbf{p}) \right]^2 \quad (14)$$

is minimized. In the above expression, t_j is a sequence of time instants and $x(t_j)$ is the given system displacement at t_j, where $j = 1, 2, \cdots, N$. On the other hand, $\hat{x}(t_j | \mathbf{p})$ is the system displacement at t_j calculated from Eqs. 1 and 4 when the parameter vector is equal to $\mathbf{p}$. The chosen objective function is simply the mean-square error in the displacement. Minimization of the objective function is subjected to the constraint that all parameters in $\mathbf{p}$ with the exception of γ are positive (Ma et al. 2004).

Early studies of parametric identification of hysteresis typically employed the non-degrading classical differential model containing only five parameters (Kyprianou et al. 2001; Ni

et al. 1998). In the few studies that involved degradation (Furukawa and Yagawa 1997; Sues et al. 1988; Zhang et al. 2002), either a restricted Bouc–Wen model containing less than thirteen parameters or a restricted identification algorithm was used. For example, Zhang et al. (2002) used gradient-based local-search algorithms to estimate some of the hysteretic control parameters for degrading structures. However, the generalized differential model of hysteresis is highly nonlinear and any gradient-based method tends to be trapped near local minima and therefore fails to converge. A robust and efficient identification algorithm is needed to estimate all thirteen unspecified parameters of the differential model of hysteresis (Ma et al. 2006).

Reduction of Parameters

In order to streamline the identification of the control parameters of differential hysteresis, the generalized Bouc–Wen model has been reexamined (Ma et al. 2004). Two significant issues have been uncovered. First, it was discovered that the unspecified parameters of the differential model are functionally dependent. One of the 13 control parameters can be eliminated through suitable transformations in the parameter space. The number of unspecified parameters can thus be reduced from 13 to 12 without any loss of

generality. Elimination of a parameter will appreciably accelerate the convergence of any identification algorithm. As explained by Ma et al. (2004), a convenient way to eliminate one parameter is to map the parameter A into 1. Henceforth, $A = 1$ will be assumed, and there remain twelve loop parameters in the differential model of hysteresis.

The second issue uncovered is the existence of insensitive parameters in the generalized Bouc–Wen differential model: the variations of three or four control parameters in the generalized model do not appreciably alter the computed hysteretic evolution of the system. Through local and global sensitivity analyses (Iman and Helton 1991; Ma et al. 2004), the twelve remaining parameters are ranked in order of decreasing sensitivity, as shown in Table 1. The method of global analysis employed is a probabilistic method recently expounded by Sobol (1993). It can account for the mutual interactions of the twelve parameters. On an overall basis, δ_v, δ_ψ, and λ are the least sensitive parameters. In the identification of hysteresis, these parameters are likely to slow down numerical convergence. One extreme measure to deal with this problem is to set the insensitive parameters to constants in the beginning; at any rate they need not be estimated with high precision. It is decided in this investigation that a two-stage procedure will be adopted to streamline the identification process. In the first stage, the three least sensitive parameters δ_v, δ_ψ, and λ are fixed, and a crude value of the nine-parameter vector

$$\mathbf{p}_1 = \left(\alpha, \beta, \gamma, n, \delta_\eta, \zeta_s, q, p, \psi \right) \qquad (15)$$

is first estimated by minimization of the objective function in Eq. (14). This crude value of $\mathbf{p}_1$, together with the fixed values of δ_v, δ_ψ, and λ, is then used as seeds in a second-stage identification to estimate the optimal value of the twelve-parameter vector

$$\mathbf{p}_2 = \left(\alpha, \beta, \gamma, n, \delta_v, \delta_\eta, \zeta_s, q, p, \psi, \delta_\psi, \lambda \right) \qquad (16)$$

In the identification of hysteresis, this two-stage procedure economizes on both computer core memory and computing time.

Nonlinear Optimization Algorithms

The goal of an optimization problem is to find a vector $\mathbf{p}^*$ in the search space S so that certain quality criterion is satisfied, namely, the error norm $g(\mathbf{p})$ in Eq. 14 is minimized. The vector $\mathbf{p}^*$ is a solution to the minimization problem if $g(\mathbf{p}^*)$ is a global minimum in S. For the constrained nonlinear optimization problem associated with the identification of differential hysteresis, the error surface defined by the objective function and constraints can exhibit many local minima or can even be multimodal. For this reason, different solution techniques will have dramatically different performance. The primary consideration in evaluating an optimization algorithm may be convergence speed or the minimum error achieved. Secondary consideration may be consistency, robustness, computational efficiency, or tracking capabilities.

Solution techniques are classified into local or population search. Basic local-search methods rely on iterative descent toward a minimum in which the direction of descent depends either on the gradient $\nabla g(\mathbf{p})$ or the Hessian matrix $\nabla^2 g(\mathbf{p})$ of the objective function. There are drawbacks of most gradient-based techniques, which include difficulties in calculating the gradient and frequent trapping of the iterates near a local minimum (Bertsekas 1999). In view of that, Zhang et al. (2002) used a non-gradient method based upon the Nelder and Mead simplex method to minimize the objective function $g(\mathbf{p})$ in Eq. 14. The simplex method explores the search space S either by reflecting, contracting, or expanding away from the worst vertex or shrinking toward the best vertex. An appropriate sequence of such movements converges to the nearest local minimum. This downhill simplex method requires only function evaluations and not derivatives. However, even this method does not have desirable convergence properties.

Population-based methods, also known as evolutionary computations, search the entire solution space S by maintaining a group of candidate vectors. Evolutionary techniques are inspired by natural evolution and adaptation with the essence of survival of the fittest. During the iterative process, new candidate vectors are

generated from existing ones by means of variation and selection such that the set of candidates would move toward increasingly favorable regions of the search space. Variation is achieved by first combining several candidate vectors to form a new one and then performing a random modification of the components of the resulting vector. Selection is introduced by comparing the fitness of the candidate vectors; the smaller the objective function $g(\mathbf{p})$, the fitter the vector $\mathbf{p}$. The fitter vectors will be selected to survive and become members of a new generation. Popular evolutionary techniques include genetic algorithms (Goldberg 1989), evolution strategies (Schwefel 1995), and differential evolution (Kyprianou et al. 2001; Lampinen and Storn 2004). Instead of performing sequential search as in the von Neumann architecture, evolutionary techniques employ massively parallel schemes with many computational elements connected by links of various weights. As a consequence, these techniques tend to be relatively robust because the risk of being trapped near a local minimum is significantly reduced. The trade-off is that evolutionary techniques are relatively slow compared with gradient methods. Genetic algorithms require conversion of floating-point values into bit strings and are less efficient for problems in which the parameters are real valued. For most floating-point problems, evolution strategies are more complicated and computationally involved than differential evolution. Differential evolution emerges as the best evolutionary method for use in the identification of hysteretic parameters of the generalized Bouc–Wen model.

Differential Evolution

Differential evolution is a relatively new and efficient approach for minimizing real-valued nonlinear and non-differentiable continuous-space functions (Price and Storn 1997). Let K be the dimension of the parameter vector $\mathbf{p}$ in Eq. 13. Suppose each generation in differential evolution consists of a population of P vectors. New vectors are generated in successive generations through certain mechanisms such that the vectors are getting closer and closer to the optimal value.

Basic mechanisms of differential evolution can be described as follows:

Initialization. In the beginning, the population can be initialized with random values chosen within boundaries specified by constraints if there is no prior knowledge about the system. Alternatively, a uniform probability distribution of the initial parameters may be assumed.

Mutation. Suppose all P vectors in generation G have been constructed. Denote these vectors by $\mathbf{v}_{i,G}$, where $i = 1, 2, \ldots, P$ and every vector has dimension K. For each population member $\mathbf{v}_{i,G}$, a mutated vector $\mathbf{w}$ is generated by adding the weighted difference between two population members to a third member such that

$$\mathbf{w} = \mathbf{v}_{r_3, G} + F\left(\mathbf{v}_{r_1, G} - \mathbf{v}_{r_2, G}\right) \qquad (17)$$

In the above expression, $0 < r_1, r_2, r_3 \leq P$ are three different integers chosen randomly, and $0 < F \leq 1$ is a preselected real-valued scaling factor. A common variation of mutation is to use the fittest vector $\mathbf{v}_{\text{best},G}$ of the current generation in place of $\mathbf{v}_{r_3, G}$ in Eq. 17.

Crossover. In order to increase the diversity of the population, the operation of crossover between $\mathbf{v}_{i,G}$ and $\mathbf{w}$ is performed. First, independent uniform random numbers $z_1, z_2, \ldots, z_K$ are generated. With a preselected crossover constant $0 \leq CR \leq 1$, a trial vector $\mathbf{u} = [u_1, u_2, \cdots, u_K]^T$ is constructed by means of

$$u_j = \begin{cases} w_j & \text{if } z_j < CR \\ (\mathbf{v}_{i,G})_j & \text{otherwise} \end{cases} \qquad (18)$$

The constant CR determines whether each component of the vector $\mathbf{u}$ will originate from the mutated vector $\mathbf{w}$ or from the current vector $\mathbf{v}_{i,G}$. In essence, CR controls the probability that a crossover would occur for each component of the trial vector $\mathbf{u}$.

Selection. The selection process is performed by a one-to-one competition in which the fitness of a vector is determined by its objective value, which is calculated by Eq. 14 with a fourth-order Runge–Kutta method for numerical integration. The trial vector $\mathbf{u}$ is then compared to the

S

(a) pick a current vector
(b) obtain the difference of two randomly chosen vectors
(c) multiply the difference by F
(d) add a third vector
(e) perform crossover
(f) decide to use $\mathbf{u}$ or not for next generation
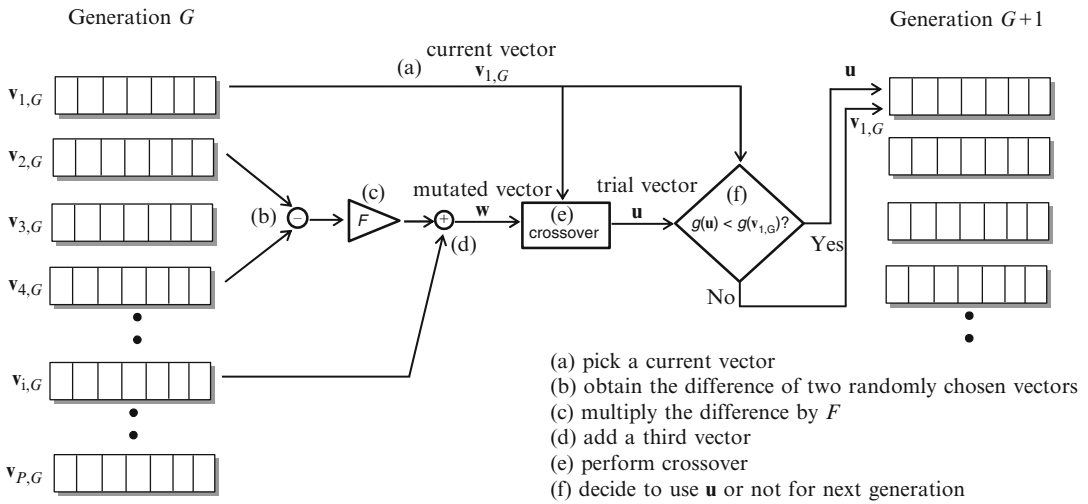
Seismic Response Prediction of Degrading Structures, Fig. 2 Basic mechanisms of differential evolution

vector $\mathbf{v}_{i,G}$. If the vector $\mathbf{u}$ yields a smaller objective value, it replaces $\mathbf{v}_{i,G}$ as $\mathbf{v}_{i,G+1}$ in generation $G + 1$. Otherwise, the original vector $\mathbf{v}_{i,G}$ will be kept. The process of mutation, crossover, and selection of the differential evolution algorithm is summarized in Fig. 2. The process is repeated P times to construct all vectors of generation $G + 1$. As evolution progresses, new generations of vector members become fitter and fitter as the objective values decrease.

In the differential evolution algorithm, three control parameters P, F, and CR must be preselected. Lampinen and Storn (2004) indicated that convergence of differential evolution is not particularly dependent upon the choice of these control parameters. It suffices to choose F and CR as multiples of 0.1 and P as a multiple of 10. Finally, a convergence criterion is needed to determine if the evolution process has converged. A common procedure is to set a desired objective value or to establish a threshold based on the relative fitness between the best and the worst vectors in each generation. Another method is to require a minimum percentage of improvement in fitness in order to continue the iteration process. One may also limit the number of generations or the computing time or both. Each convergence criterion has its advantages and drawbacks. Thus, a combination of convergence criteria is often used.

Based upon the exposition in this section, a robust algorithm using the latest theory of differential evolution is constructed for the identification of differential hysteresis. As explained before, a two-stage procedure is adopted whereby a crude value of the vector $\mathbf{p}_1$ in Eq. 15 is first obtained before the optimal value of $\mathbf{p}_2$ in Eq. 16 is computed. In simulations reported in this entry, the population size P is set to 100, while both the scaling factor F and the crossover constant CR are set to 0.5. The computing time is dependent on the amount of input data used for identification. The runtime can be reduced appreciably if some of the insensitive parameters are fixed and parallel computing is utilized.

Prediction of Performance

With a robust and relatively efficient algorithm devised in section "System Identification," a differential model of hysteresis of any degrading system may be identified from a given load–displacement trace. Suppose a hysteretic model is generated with a given load–displacement trace. It will be shown that, under fairly broad conditions, the model predicts the response of the same structure when driven by other cyclic loads. This can be carried out experimentally by building a number of identically configured structural joints

and subjecting them to different cyclic loads. One set of load–displacement trace is chosen to generate a hysteretic model of the joint by system identification. Subsequently, this identified model is used to compute the displacements of the joint under other cyclic loads. The computed displacements will be compared with experimental measurements. A reasonable match will validate the identified model and will demonstrate its prediction capability. In order to reduce the front-end costs of this investigation, only wood joints are used in the experiments. It must be emphasized that steel and concrete structures may be used as easily. Although a limited set of experimental data will be presented in this entry, many different structural joints have been built and tested to support any observations made herein.

Test Setup

Cyclic tests of degrading structures are performed in the Forest Products Laboratory located at the Richmond Field Station of the University of California, Berkeley. Twenty identical specimens of a simple *T*-connection consisting of two wood members joined by plywood gusset plates, as shown in Fig. 3, are constructed. The two wood members are made out of laminated veneer lumber (LVL), which is a highly predictable engineered wood product with relatively small material variability. Each LVL member is a uniform beam with a square cross section of 3.5×3.5 in. and an elastic modulus of 2.0×10^6 psi. The LVL lumber is sawn to consistent sizes and is virtually free from warping and splitting. The two LVL members are connected on both sides by plywood gusset plates of 0.5 in. thickness. Six 2-in. metal screws are arranged in three rows to fasten the plates to the vertical beam. The two plates are also connected securely to the horizontal beam with four metal screws and two retrofit bolts. Cyclic loads are applied parallel to the lower beam at a height 36 in. above the lower beam. The arrangement is such that the *T*-connection will fail through detachment of vertical LVL member from the gusset plates.
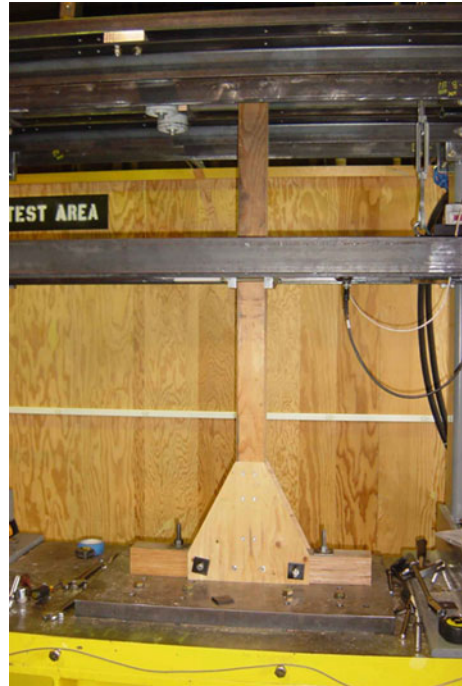
Seismic Response Prediction of Degrading Structures, Fig. 3 Reaction frame test system with a specimen

Common seismic testing protocols are employed to drive the *T*-connections. It is claimed that such protocols would simulate damage accumulation and degradation of the structural joint under real-life loads. Three of the cyclic excitations used in the experiments are shown in Figs. 4, 5, and 6. Excitation *A* is proposed by the Consortium of Universities for Research in Earthquake Engineering (CUREE), developed exclusively for wood testing (Krawinkler et al. 2001). The CUREE protocol is designed to model ordinary ground motion typical of most far-field locations. It is characterized by a pattern of a primary cycle followed by a number of trailing cycles that are 75 % of the previous primary cycle. Each primary cycle, which has higher amplitude, initiates a new level of damage to the structure. Damage then accumulates and changes the structural response in the trailing cycles. In addition, the number of reversed cycles is limited in order to induce failure modes resembling those in actual structures.
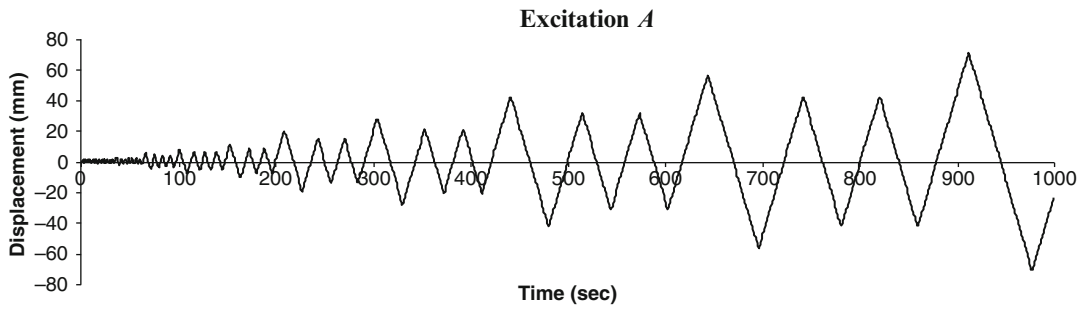
Excitation *A*

Seismic Response Prediction of Degrading Structures, Fig. 4 Excitation of CUREE
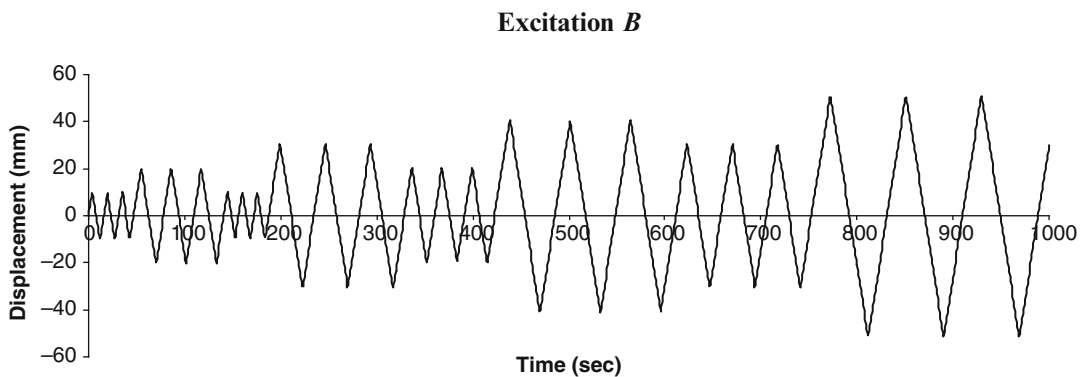
Excitation *B*

Seismic Response Prediction of Degrading Structures, Fig. 5 Excitation of Forintek Canada Corporation

Excitation *C*– Reference
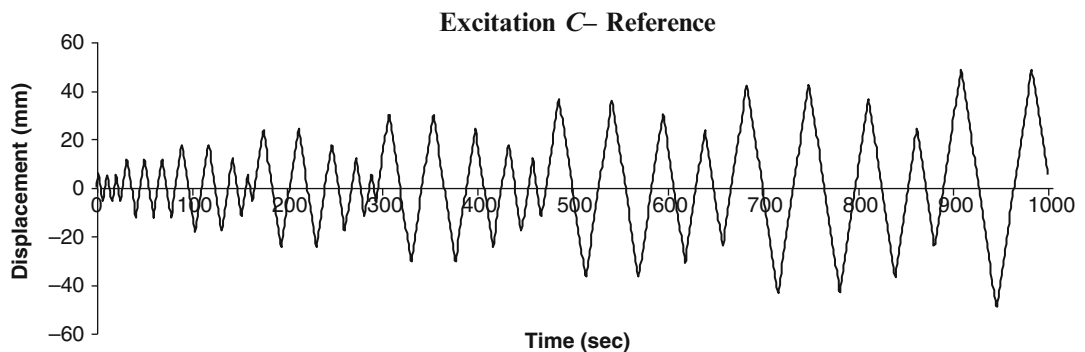
Seismic Response Prediction of Degrading Structures, Fig. 6 Excitation of Forest Products Laboratory

Excitation *B* is devised by Forintek Canada Corporation (FCC), a national wood products research institute in Canada. Compared with excitation *A*, equal-amplitude primary and trailing cycles are used in this excitation. The FCC protocol consists of a sequence of cyclic groups, each composed of three identical cycles. The amplitude of each group is taken as a percentage of the nominal yield slip and is increased stepwise, with inter-spaced decreasing cycles until the specimen fails. Excitation *C* was developed at the Forest Products Laboratory for the

Seismic Response Prediction of Degrading Structures,
Fig. 7 Hysteresis loops generated by excitation C in 500 s
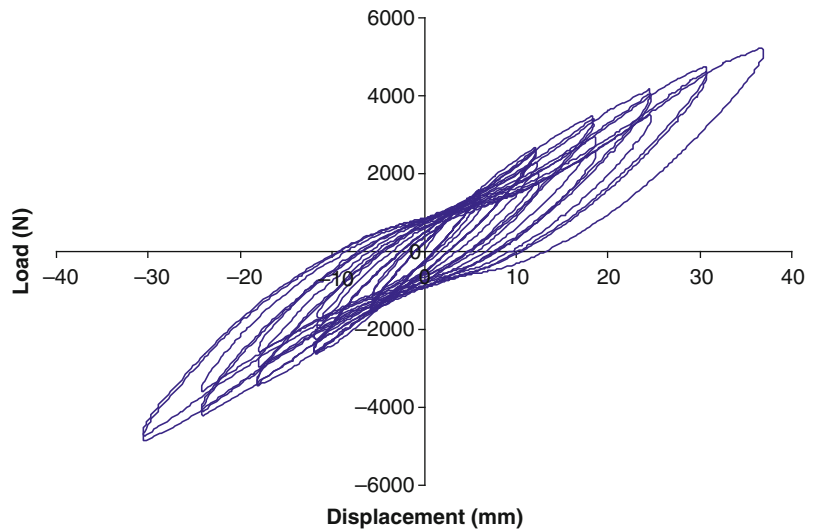
experiments reported herein. This protocol features trailing cycles that are decreasing in amplitude, which are different from the equal-amplitude trailing cycles in the CUREE and FCC excitations. In addition, excitation C does not have the series of low-amplitude cycles featured in the starting interval of excitation A.

Model Identification and Validation

Each specimen of the identically configured T-connections is driven by a cyclic excitation and 1,000 s of load–displacement data are recorded. Using any of these load–displacement traces, a model of hysteretic evolution can be identified. Two issues will now be examined. (1) Can a hysteretic model identified with a given cyclic load predict the future structural response if the *same* cyclic load continues beyond the duration used for identification? (2) Can a hysteretic model identified with a given cyclic load predict the structural response due to a *different* cyclic load?

In Fig. 7, the load–displacement trace associated with excitation C for the first 500 s is shown. As reflected in the hysteresis trace, the gusset joint is degrading in both strength and stiffness. There is also some pinching of the hysteretic loops, which is probably caused by slipping of the connecting screws. The load–displacement trace of Fig. 7 is used as input to the identification

algorithm devised in the last section. Recall that $A = 1$, and there remain twelve hysteretic parameters to be estimated of the gusset-plate connection. Once these twelve parameters are identified, a hysteretic model of the T-connection is obtained. This hysteretic model can now be used to compute the displacement history of the gusset joint associated with excitation C. As illustrated in Fig. 8, the computed displacement history closely matches the experimental measurements. This validates the identified hysteretic model over the first 500 s.

A different hysteretic model of the T-connection is generated when the duration of input data changes. In Fig. 9, the load–displacement trace associated with excitation C for $0 < t < 1,000$ s is shown. Compared with Fig. 7, pinching is more prominent after the first 500 s. Thus, the twelve control parameters estimated with input data of Fig. 9 are different from those associated with Fig. 7. The new hysteretic model must be validated again for a duration of 1,000 s because the dominant characteristics of degradation are different between the two intervals $0 < t < 500$ s and $0 < t < 1,000$ s. A hysteretic model identified with input data over $0 < t < 500$ s will not produce a displacement history closely matching the experimental data over $500 < t < 1,000$ s. In short, system identification cannot produce
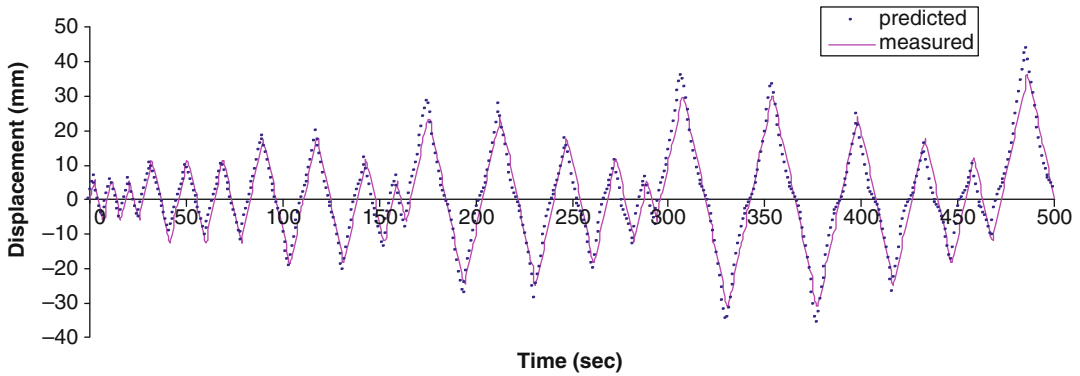
Seismic Response Prediction of Degrading Structures, Fig. 8 Comparison of computed and measured displacements under excitation C

Seismic Response Prediction of Degrading Structures,
Fig. 9 Hysteresis loops generated by excitation C in 1,000 s
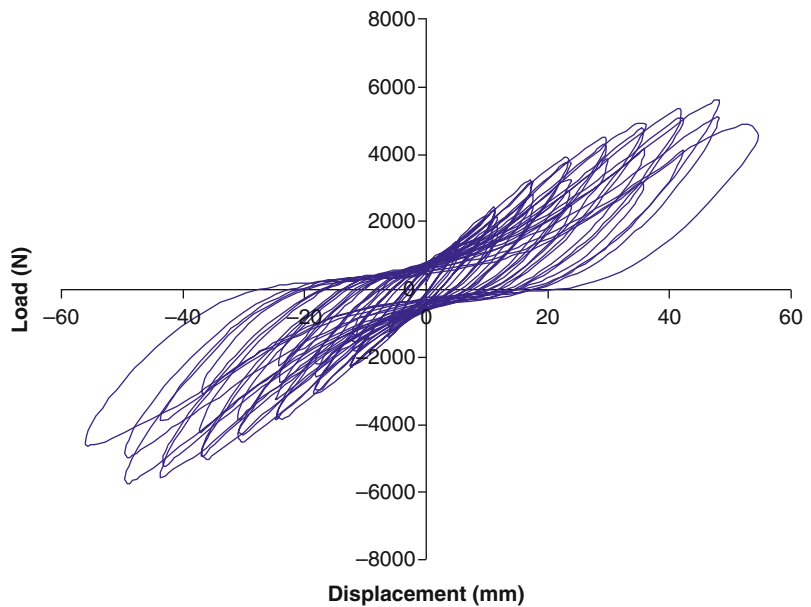
features and trends not already contained in the input data. If the duration of input used for identification already contains all features of hysteretic evolution, then a hysteretic model identified with a given excitation can accurately predict the future response when the *same* excitation continues beyond the duration used for identification. This basically answers the first question brought up at the beginning of this section.

In identifying a hysteretic model of the *T*-connection, the experimental load–displacement trace associated with excitation C has been used directly as input for identification. It can be argued that such input data are corrupted by Gaussian noise, and a more accurate model can be identified if noise is suppressed with a filter. To this end, the input was first taken through a median filter and then a least-squares low-pass filter. Such a filtering process is known to be effective in smoothing out random fluctuations caused by Gaussian noise (Papoulis and Pillai 2002; Zhang et al. 2002). However, it has been found that a better match than what is shown in Fig. 8 is not achieved with the noise-filtered data. It appears that differential evolution is not sensitive to a moderate level of
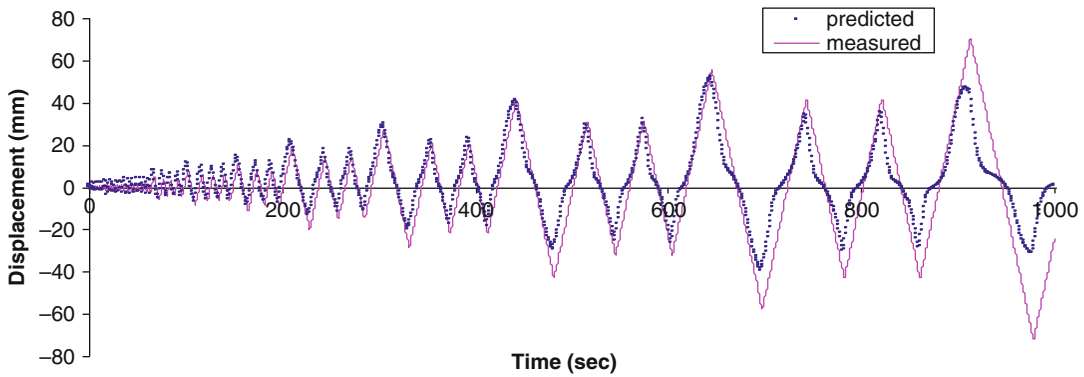
Seismic Response Prediction of Degrading Structures, Fig. 10 Predicted and measured displacements under excitation *A* in 1,000 s
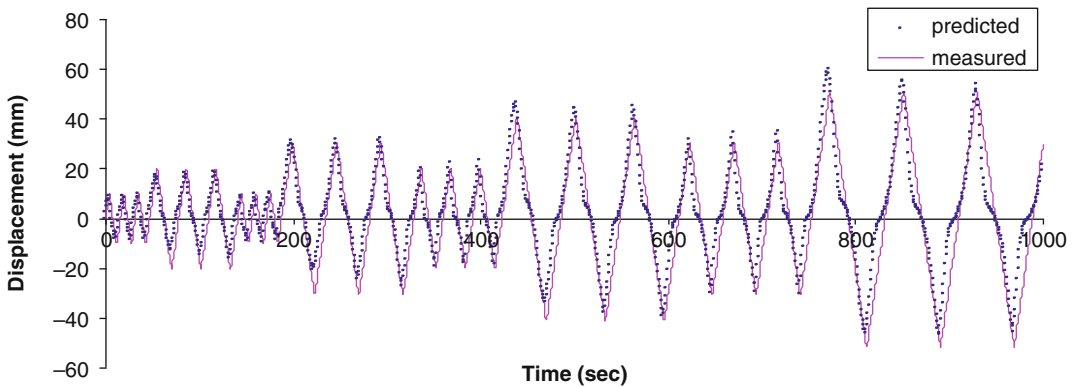
Seismic Response Prediction of Degrading Structures, Fig. 11 Predicted and measured displacements under excitation *B* in 1,000 s

Gaussian noise. Subsequently, noise-filtering is not used on any of the load–displacement traces.

Can a hysteretic model obtained by identification using a given excitation predict the structural response due to a *different* cyclic load? This is the second question brought up earlier. Using 1,000 s of the load–displacement trace associated with excitation *C*, a hysteretic model of the *T*-connection has been identified. Keep in mind that other cyclic loads do not play any part in the identification of hysteresis. The identified model is then used to compute the displacement of the gusset joint under excitation *A* of Fig. 4. However, load–displacement data associated with excitation *A* have already been collected experimentally using another specimen of the identically configured gusset joints. Thus, the

computed (or predicted) displacement can be compared with the measured displacement. As shown in Fig. 10, there is reasonably good match. A similar observation can be made with other cyclic loads. As illustrated in Fig. 11, there is also reasonably good match between the computed (or predicted) and experimental data when excitation *B* of Fig. 5 is used. This demonstrates that a hysteretic model identified with a given excitation may be used for response prediction under different excitations. However, the accuracy of prediction tends to diminish as time increases.

Requirements for Accurate Prediction
As demonstrated in previous sections, a hysteretic model obtained by identification can be used

Seismic Response Prediction of Degrading Structures,
Fig. 12 Failure configuration of *T*-connection under excitation *C* over 1,000 s
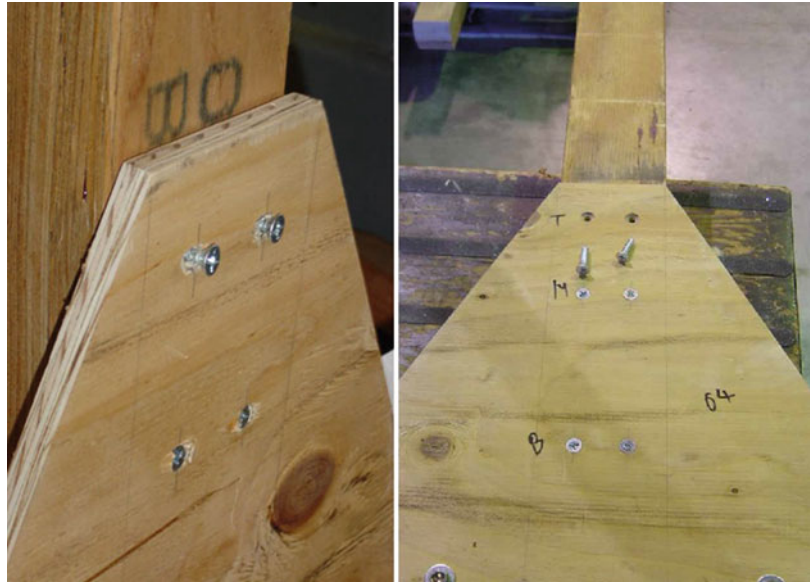
to predict future performance of a system. Large errors, however, can arise in many cases. For example, the prediction of response associated with excitation *A* is no longer satisfactory after 900 s, as shown in Fig. 10. Perhaps this is not surprising. As time progresses, damage accumulates and the *T*-connection begins to fail. There is, however, a divergence in the failure configuration due to different excitations.

As shown in Fig. 12, a typical failure configuration under excitation *C* over 1,000 s is that the top two screws in the vertical member are appreciably loosened due to wood crushing. Disassembly of the gusset joint reveals that both screws are broken into two pieces under excitation *C*. If 1,000 s of the load–displacement trace associated with excitation *C* is used for identification, a hysteretic model is generated that anticipates the breaking of the top two screws. Under excitation *A*, however, it has been observed that all four screws in the top two rows are broken over 1,000 s, as shown in Fig. 13. If a model identified with excitation *C* is used for response prediction when the structure is driven by excitation *A*, it fails to predict that all four screws in the top two rows would be broken over 1,000 s. This explains the discrepancy between the predicted and measured displacements under excitation

Seismic Response Prediction of Degrading Structures, Fig. 13 Failure configuration of *T*-connection under excitation *A* over 1,000 s

A in Fig. 10. If the failure mechanism can be controlled, the accuracy in response prediction will be increased.

The observations made in experimenting with the elementary *T*-connection can be extended to complex assemblies or structures. A complex structure consists of many different elementary connections. Since each elementary connection may have several failure configurations, the totality of failure configurations is very large for a complex structure. If structural degradation can be controlled so as to result in a failure configuration identical to what is contained in the input data used for identification, then nonlinear degrading response of a complex structure can be predicted. Otherwise, prediction is only reliable up to a point at which the input data and the existing complex structure progress into different failure configurations. A study of the influence of failure configurations on the reliability of response prediction will be worthwhile in a subsequent course of investigation.

Summary

A basic objective of this entry is to advance the methodology for predicting the performance of nonlinear deteriorating structures. Using the generalized Bouc–Wen differential model, it is shown how system identification can be utilized to predict the response of a degrading structure well beyond its linear range. Important observations reported in the paper are summarized in the following statements.

1. A robust identification algorithm based upon the generalized Bouc–Wen model and the theory of differential evolution can be used to generate practical models of hysteresis of degrading structures. Differential evolution is not sensitive to a moderate level of input noise.
2. If the duration of input used for identification already contains all features of hysteretic evolution, then a hysteretic model identified with a given cyclic load can accurately predict the future structural response when the same cyclic load continues beyond the duration used for identification.
3. A hysteretic model identified with a given cyclic load may be used to predict the nonlinear response under different cyclic loads. As damage accumulates, the precision of predicted response decreases, and the reliability of prediction ultimately depends on the degree of similarity between failure configurations under different cyclic loads.

In the absence of a fundamental theory of degradation, the response prediction of degrading structures is indeed a challenging task. System identification of hysteretic evolution provides a brute-force procedure for prediction that has the potential of allowing a closer representation of reality. The research reported herein addresses various aspects of an identification methodology for predicting the performance of real-life deteriorating structures well beyond their linear ranges. Among other things, it is hoped that this entry would point to directions along which further research efforts should be made. This entry could also contribute to the development of a general formulation of the mechanisms of degradation in the long term.

Cross-References

▶ Nonlinear System Identification: Particle-Based Methods
▶ Online Response Estimation in Structural Dynamics
▶ System and Damage Identification of Civil Structures
▶ Vibration-Based Damage Identification: The Z24 Bridge Benchmark

References

Baber TT, Noori MN (1986) Modeling general hysteresis behavior and random vibration application. ASME J Vib Acoust Stress Reliab Des 108:411–420

Baber TT, Wen YK (1981) Random vibration of hysteretic degrading systems. ASCE J Eng Mech 107(6):1069–1087

S

Bertsekas DP (1999) Nonlinear programming, 2nd edn. Athena Scientific, Belmont

Bouc R (1967) Forced vibration of mechanical systems with hysteresis. In: Proceedings of 4th conference on nonlinear oscillations, Prague, Czechoslovakia, p 315

Foliente GC (1995) Hysteretic modeling of wood joints and structural systems. ASCE J Struct Eng 121:1013–1022

Furukawa T, Yagawa G (1997) Inelastic constitutive parameter identification using an evolutionary algorithm with continuous individuals. Int J Numer Methods Eng 40:1071–1090

Goldberg DE (1989) Genetic algorithms in search, optimization & machine learning. Addison-Wesley, Reading

Iman RL, Helton JC (1991) The repeatability of uncertainty and sensitivity analyses for complex probabilistic risk assessments. Risk Anal 11:591–606

Kojic M, Bathe KJ (2005) Inelastic analysis of solids and structures. Springer, Berlin

Krasnoselskii MA, Pokrovskii AV (1989) Systems with hysteresis. Springer, Berlin

Krawinkler H, Parisi F, Ibarra L, Ayoub A, Medina R (2001) Development of a testing protocol for woodframe structures. CUREE-Caltech woodframe project report W-02, Consortium of Universities for Research in Earthquake Engineering, Richmond

Kyprianou A, Worden K, Panet M (2001) Identification of hysteretic systems using the differential evolution algorithm. J Sound Vib 248:289–314

Lampinen J, Storn R (2004) Differential evolution. In: Onwubolu GC, Babu BV (eds) New optimization techniques in engineering. Springer, Berlin, pp 125–166

Ma F, Zhang H, Bockstedte A, Foliente GC, Paevere P (2004) Parameter analysis of the differential model of hysteresis. ASME J App Mech 71:342–349

Ma F, Ng CH, Ajavakom N (2006) On system identification and response prediction of degrading structures. Struct Control Health Monit 13(1):347–364

Ni YQ, Ko JM, Wong CW (1998) Identification of nonlinear hysteretic isolators from periodic vibration tests. J Sound Vib 217:737–756

Papoulis A, Pillai SU (2002) Probability, random variables, and stochastic processes, 4th edn. McGraw-Hill, New York

Price K, Storn R (1997) Differential evolution. Dr Dobb's J 22:18–24

Roberts JB, Spanos PD (1990) Random vibration and statistical linearization. Wiley, New York

Schwefel HP (1995) Evolution and optimum seeking. Wiley, New York

Sobol IM (1993) Sensitivity estimates for nonlinear mathematical models. Math Model Comput Exp 1:407–414

Song J, Der Kiureghian A (2006) Generalized Bouc-Wen model for highly asymmetric hysteresis. ASCE J Eng Mech 132(6):610–618

Sues RH, Mau ST, Wen YK (1988) System identification of degrading hysteretic restoring forces. ASCE J Eng Mech 114:833–846

Wen YK (1976) Method for random vibration of hysteretic systems. ASCE J Eng Mech 102(2):249–263

Zhang H, Foliente GC, Yang Y, Ma F (2002) Parameter identification of inelastic structures under dynamic loads. Earthq Eng Struct Dyn 31(5):1113–1130

Seismic Risk Assessment, Cascading Effects

Paolo Gasparini and
Alexander Garcia-Aristizabal
Center for the Analysis and Monitoring of
Environmental Risk (AMRA), Naples, Italy

Synonyms

Domino effects; Short-term hazard and risk assessment; Time-dependent risk; Triggering effects

Introduction

Risk is conveniently assessed as a function of the probability that a certain event will occur and of the extent of the damage caused to humans, environment, and objects. Conceptually, it is the result of the operation

$$
\begin{aligned}
\text{Risk} = \text{Hazard} \times \text{Damage} = \text{Hazard} \\
\times \text{Vunerability} \times \text{Value at Risk}
\end{aligned}
\tag{1}
$$

where the *Hazard* can be understood as the probability that a given adverse event of a given intensity (hereinafter referred to as *intensity measure*, IM) will occur in a certain area within a defined time interval; the *Vulnerability*, a concept interpreted and applied in various ways, here is considered in its physical dimension as the probable damage to an element at risk given a level of intensity (IM) of the hazard; and the *Value at Risk* term is considered as a measure of the total potential loss due to an adverse event in a given area.

Note that both *Hazard* and *Vulnerability* are generally defined in probabilistic terms, while *Value at Risk* has some defined units of measure (monetary cost, number of human lives, etc.).

Seismic risk assessment methods have shown an important development over the last 40 years reaching an outstanding level of acceptance by the scientific community as a valid framework for sound risk mitigation and decision-making. In general the seismic risk assessment uses seismic hazard information combined with fragilities of structures and facilities in order to estimate the probabilities of damages and to measure expectancies of losses using a predefined metric of loss.

Different past disasters have highlighted the fact that natural or man-made events can trigger other events, leading to a significant increase of fatalities and damages. As a consequence, there is a growing demand from risk managers to implement multi-type hazard and risk assessments as well as assessing cascading effects. Consequently, an increasing number of scientists studying natural hazards and risk assessment are turning their interest towards the multi-hazard risk and multi-risk assessment practice. The transition from the single- to the multi-risk assessment represents a process in which both the complexity and the data requirements are significantly increased. This transition implies a shift from the "hazard-centered" perspective that characterizes the single-risk assessment to a "territorial-centered" one.

The cascading events are one of the fundamental concepts of the multi-risk assessment. It is worth noting that a different terminology is used in the practice of risk evaluation when addressing the problem of chains of events. In particular, the term "domino effect" is generally used in the field of industrial accidents, whereas the term "cascading effects" is mainly used in ecology and in natural hazards in the context of the multi-risk assessment. The objective of this entry is to discuss the role of cascading effects triggered by earthquakes and its importance for a holistic view of the seismic risk assessment process.

Multi-risk as a Framework for the Assessment of Cascading Effects

The transition from single- to multi-risk assessment represents a process in which both the complexity and the data requirements are significantly increased (Fig. 1a). This transition implies a shift from a "hazard-centered" perspective that characterizes the single-risk assessment, to a "territorial-centered" one. In fact, from a multi-risk perspective, the first element to be defined is the target area of interest, which is in general terms the piece of territory composed of elements at risk that are vulnerable in different ways to various sources of hazard.

The evolution from the single-risk to the multi-risk analysis consists of three different and complementary elements: the single-risk assessment (SR), the multi-hazard risk assessment (MHR), and the multi-risk assessment (MR).

Figure 1 represents the transition from the SR to the MR assessment (Fig. 1a). The hazard-centered perspective of the SR is represented in Fig. 1b; in this case, the first element of the analysis is generally the hazard source followed by the definition of the impact area and the assessment of the potential effects. The change towards a territorial perspective is represented in Fig. 1c. In the MHR, different independent hazard sources affecting a given common area of interest are considered; in this case, the first element of the analysis is the definition of the piece of territory which is the target area for the analysis; the target area includes the elements at risk exposed to adverse events. Note that this change in perspective is common to both the MHR and the MR assessments. Once the target area has been identified, the next element of the analysis is the identification of the potential hazard sources that may cause harm to the set of exposed elements in the target area.

Figure 1d illustrates how each element in this transition is a subset in the higher levels of the sequence. In fact, a set of SR analyses compose an MHR analysis, and when the MHR assessment is complemented with the analysis of interactions

S
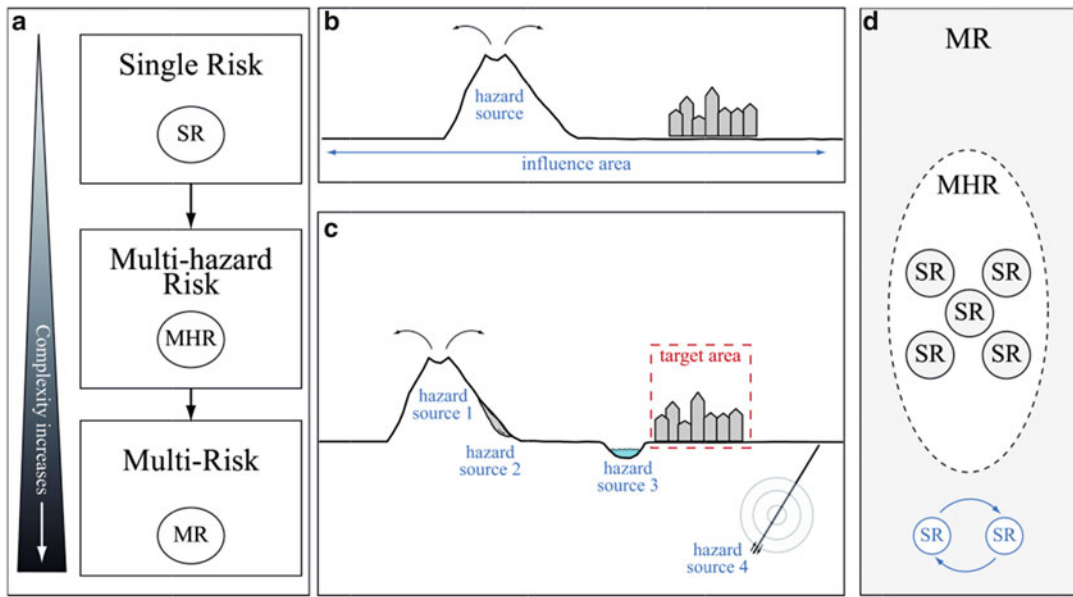
Seismic Risk Assessment, Cascading Effects, Fig. 1 Representation of the transition from the single- to the multi-risk assessment

among events, then the MR level is reached. This sequence is of course valid if the SR and MHR analyses are done coherently in a multi-risk framework, following some basic (and often non-simple) elements of harmonization.

The multi-risk assessment (MR) is then a generalization of the MHR assessment in the sense that it integrates also the possible interactions or cascading effects (Fig. 1c, d); therefore, the main objective of the MR assessment level is to assess the effects of possible interactions or cascading effects among the different hazardous events. Considering long- and short-term assessments, Marzocchi et al. (2012) identify the following main steps of the MR assessment procedure:

1. Definition of the space-time window for the risk assessment and the metric for evaluating the risks
2. Identification of the risks impending on the selected area
3. Identification of selected hazard scenarios covering all possible intensities and relevant hazard interactions
4. Probabilistic assessment of each scenario

5. Vulnerability and exposure assessment for each scenario, taking into account the vulnerability of combined hazards
6. Loss estimation and multi-risk assessment

In this framework, a set of scenarios considering adverse events from different sources are identified along with the potential interactions. Each chain of adverse events is defined in a series/parallel sequence of happenings and is generally represented in an "event tree." Each branch of the event tree is quantified by a probabilistic analysis considering the sequence of the events and the vulnerability and the exposed values of the specified targets. This concept is illustrated in Fig. 2, where the possible kinds of interactions are identified using discontinuous red lines. For example, a sequence defined for independent hazards interacting at the vulnerability (or damage) level is represented by the red discontinuous line in the sequence at the left side of Fig. 2. Likewise, an example of interaction at the hazard level is presented with source 3 triggering source 2 (see the path defined with the red discontinuous line at the right in Fig. 2). Note that
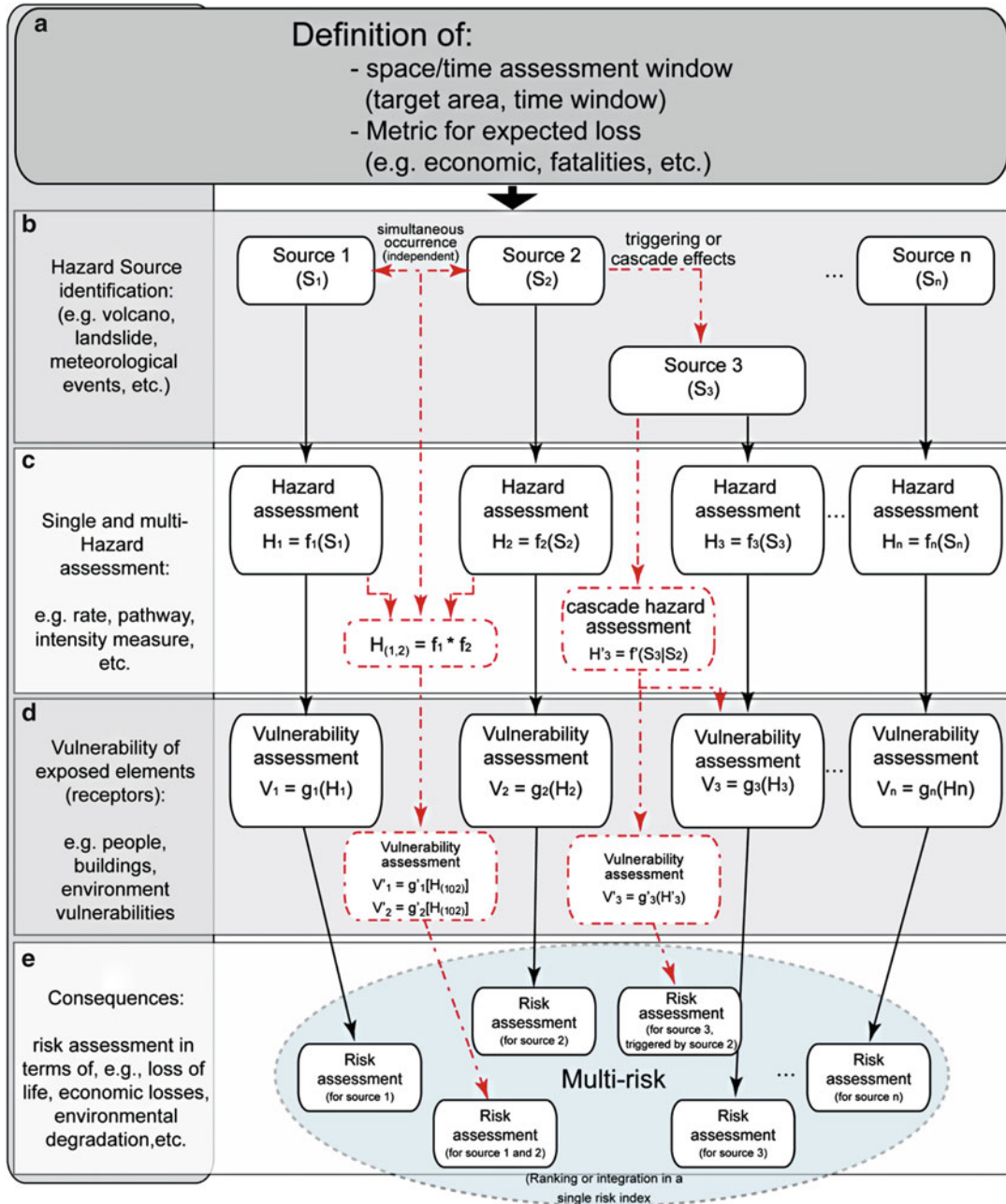
Seismic Risk Assessment, Cascading Effects, Fig. 2 Multi-risk framework considering possible interactions at the hazard and the vulnerability levels (Reproduced from Marzocchi et al. 2012, with permission from Springer®)

in this case, the interaction at the vulnerability level may or may not be present. Details about the kinds of interactions are more deeply presented in the following sections of this entry.

Within this context of interactions and triggering mechanisms, earthquake-triggered events have drawn much attention in the last decades because of the severe damages that the triggered

events may cause. Sometimes, the losses caused by the secondary events may exceed the losses caused by the initial, triggering earthquake itself. For example, in a number of earthquake-triggered landslides, the losses caused by the triggered mass movements have been comparable (if not higher) to the losses caused by the direct impact of ground shaking. An outstanding case is the Huascaran rock-debris avalanche triggered by the M7.7 Peru earthquake in 1970 (e.g., Plafker et al. 1971), in which the impact of the avalanche counts for more than 50 % of the total casualties.

The kind of events that can be triggered and the possible interactions that can be analyzed vary in a wide spectrum of events including natural (as other earthquakes, landslides, tsunamis, volcanic unrest and eruptions, etc.) and nonnatural (as Natech) events. A non-exhaustive description of known examples is presented in the section "Examples of Cascading Effects Caused by Earthquakes."

Cascading Effects: An Overview

From the risk assessment point of view, the cascades of events occurring after an earthquake may represent an important source of loss amplifications. In this context, we consider as "*loss amplification*" all the losses not caused directly by the earthquake (as the triggering event) but caused by other secondary events occurring as a consequence of the initiating triggering event. Therefore, the loss amplification can be a factor resulting from comparing the total losses caused by the chain of events triggered by an earthquake with the direct losses caused by the initial earthquake itself.

The cascading effects result as a consequence of interactions generated by cause/effect relationships among different phenomena. The nature of these interactions may be described by a wide set of phenomenological relationships, and it makes difficult to set a generalized procedure for the quantification of cascading effects. For this reason and to simplify the setting of the problem for quantitative purposes, a generic taxonomy of the

possible kinds of interactions can be defined. Following the framework presented in Marzocchi et al. (2012), Garcia-Aristizabal et al. (2013), and Garcia-Aristizabal et al. (forthcoming-2015), we can consider the following two major sets of interactions: (1) interactions at the hazard level and (2) interactions at the vulnerability level. This generic taxonomy of possible interactions is qualitatively described in the following paragraphs, and a more quantitative definition is presented in the last section of this entry.

Interactions at the Hazard Level

In this case the interaction problem is understood as the assessment of possible chains of adverse events in which the occurrence of a given initial "triggering" event entails a modification of the probability of occurrence of a secondary event. Even if this typology of problem can be assessed on a long-term basis, its utility can be highlighted in short-term problems. It can be seen that the typology of interactions that can be grouped under this name are in fact phenomena in which the physical process of interest is a pure triggering mechanism in which a first occurring event produces a perturbation that, when acting over a given system, may bring it to unstable conditions forcing it to find a new equilibrium. Reaching this new equilibrium may imply the occurrence of an event that in this case may be said as *triggered* by the initial triggering one.

The link between the intensity of the triggering event (e.g., the ground shaking caused by an earthquake) and the intensity of the triggered event (e.g., a volume of mass moving in a slope) is governed by complex physical mechanisms that are intrinsically related to the specific triggering and triggered events. This fact and the ubiquitous random effects that may affect these processes make of probabilistic approaches the most promising way for the quantitative characterization of such interactions. Therefore, assessing interactions at the hazard level in probabilistic terms implies to quantify the conditional probabilities of occurrence of the triggered event, given the occurrence of the triggering event (with given intensity).

Interactions at the Vulnerability Level

In this case the problem fundamentally intends to assess the effects that the simultaneous occurrence of two or more events (not necessarily linked) may have on the response of a given typology of exposed elements. In this case, the process of interest is the additive effect of different events acting on the same elements. More specifically, this kind of interaction is referred to the case in which the occurrence of one event (the first one occurring in time) may alter the response of the exposed elements against another event (that may be of the same kind as the former but also a different kind of hazard). In this case, it is assumed that two or more events act on a system (e.g., a structure) and that the additive or cumulated effect produces a change in the response of the system with respect to the conditions before the occurrence of the former event. In practice, it implies that two or more hazardous events act simultaneously or in a short-time window (in general short enough that the system cannot be repaired) on the same structure.

The cases that can be grouped under this kind of interaction are of different nature; in general, the physical processes of interest are those related with the response of the system (in this case, an exposed element) to the loads caused by different events taking in account their additive or cumulated effects. Note that the different events may be of the same nature (i.e., same kind of hazards, as, e.g., two earthquakes shaking the same structures in a short-time window) or coming from different kinds of phenomena (e.g., the shaking caused by an earthquake as the first event, followed by loads caused by river flooding or strong wind). Note that the different events causing additive loads can be themselves the result of a common triggering event or may be independent events, and this relationship is important for the quantitative analyses.

As in the previous case, the quantitative assessment of this kind of interactions relies on probabilistic analyses. It may be understood as a multidimensional development of fragility functions for the different typologies of exposed elements. In general, it involves the determination of the conditional probabilities of having certain damage state given the simultaneous or cumulative solicitation caused by different combinations of intensities of the different events considered. Note that this information can be represented in different ways, according with the nature of the acting events. For instance, if the acting events are not causing cumulated damages in the same elements (e.g., seismic loads from one side and the loading caused by snow accumulation or flooding water on the other), then the conditional probabilities define a multidimensional fragility surface. Conversely, if the different acting events produce cumulated damages in the same elements (i.e., when the first event pre-damages a given element, as, e.g., two earthquakes acting one after the other in a short time interval), the conditional probabilities are generally referred as *time-dependent* or *damage-dependent* fragilities (or vulnerabilities).

Examples of Cascading Effects Caused by Earthquakes

Many examples of cascading effects caused by earthquakes can be cited as, for example, the well-known cases of tsunamis triggered by earthquakes or earthquake-triggered landslides. As described in the previous section, there may be different typologies of interactions, and the physical processes involved in these interactions are extremely case dependent. In this section, we describe some outstanding cases of interactions (mainly at the hazard level) that highlight the potential loss amplifications that these triggering effects may produce.

Earthquake-Earthquake Triggering

Earthquake mainshocks cause aftershocks to occur, which in turn activate their own subsequent aftershock sequences, resulting in a cascade of triggering that may increase the damage caused to the exposed elements with respect to the effects directly produced by the mainshock alone. For this reason, contrary to the normal practice in seismic hazard and risk assessment, the cascade of seismic events triggered during a

seismic sequence must be of interest, at least in the short term.

A cascade of earthquakes triggering earthquakes causes the seismicity to develop in complex patterns (see, e.g., Marsan and Lengliné 2008). Different physical mechanisms have been proposed to study mainly the direct triggering problem; these include static stress changes (i.e., the difference in the stress field just before an earthquake to shortly after the seismic waves have decayed, see, e.g., Hill et al. 2002; Manga and Brodsky 2006), quasi-static stress changes (e.g., associated with slow viscous relaxation of the lower crust and upper mantle beneath the epicenter of a large earthquake, see, e.g., Hill et al. 2002), dynamic stress changes (i.e., induced by the seismic waves from a large earthquake, see, e.g., Hill et al. 2002; Manga and Brodsky 2006), and fluid flow (e.g., migrating pore fluids and pore pressure changes as a mechanism controlling the timing of aftershocks, see, e.g., Harris 1998 and references therein). The distinction between static and dynamic stress transfer mechanisms breaks down at distances less than a fault length from a given earthquake. That region, the *near field*, is generally the spatial domain of aftershock occurrences (Hill et al. 2002).

The evident complexity of the triggering problem explains the objective difficulty to efficiently discriminate between direct and indirect triggering in the observed data, limiting the possibility of implementing pure deterministic approaches for the analysis of the events in a sequence. Therefore, the main efforts, as found in literature, attempt to identify and separate earthquakes between mainshocks and aftershocks, with the most sophisticated methods employing stochastic approaches. Our discussion in this section is restricted to the description of stochastic modeling of aftershock sequences.

The statistical properties of the aftershock sequences have long been of interest to study the earthquake generation processes. It has resulted in the development of a number of point process models (a type of random process for which any realization consists of a set of isolated points in time or space) to describe the standard activity of earthquake sequences.

Considering in particular the *trigger models* (Vere-Jones and Davies 1966), it is assumed that the conditional probability of a shock occurring at a time t after a given triggering mainshock is proportional to a *decay function* $\lambda(t)$. In general, it considers that the triggering events occur by following a Poisson process with a constant rate, while the number of triggered events (i.e., the group of shocks initiated by a single triggering event) decays by following a prescribed function. Regarding the nature of $\lambda(t)$, Vere-Jones and Davies (1966) considered in detail two functions:

- An exponential decay, $\lambda(t) = pe^{-pt}$, $(p > 0)$
- An inverse power-law decay, $\lambda(t) = pc^p / (c + t)^{p+1}$, $(c > 0)$

(where the model parameters arise from the Omori formula). Vere-Jones and Davies (1966) suggested that the inverse power-law type of decay function gives a better fit than the negative exponential. Considering a restricted form of the trigger model, Ogata (1988) proposed the epidemic-type aftershock sequences (ETAS) model. ETAS is a stochastic point process in which each earthquake has some magnitude-dependent ability to trigger its own Omori law-type aftershocks. The total occurrence rate can be described, in time, as the superposition of a background uncorrelated seismicity μ_0 and the events triggered by another earthquake:

$$\lambda(t) = \mu_0 + \sum_{i:t_i < t} \lambda_i(t) \qquad (2)$$

where $\lambda_i(t)$ is the rate of aftershocks induced by an event occurred at time t_i with magnitude M_i, defined as (Ogata 1988):

$$\lambda_i(t) = \frac{k}{(c + t - t_i)^p} e^{\alpha(M_i - M_c)} \qquad (3)$$

for $t > t_i$. The parameter k measures the productivity of the aftershock activity; α defines the relation between triggering capability and magnitude M_i of a triggering event; c measures the incompleteness of the catalog in the earliest part of each cluster; p controls the temporal decay of

Seismic Risk Assessment, Cascading Effects, Fig. 3 Percentage of landslide events in Italy classified by the triggering mechanism (Source of the data: CNR-GNDCI – database of areas affected by landslides and floods in Italy)

Other causes, 5.5%

Ground water variation, 1.4%

Erosion, 11.9%

Anthropic effects, 5.1%

Earthquakes, 2.6%

Precipitation and infiltration; 73.5%
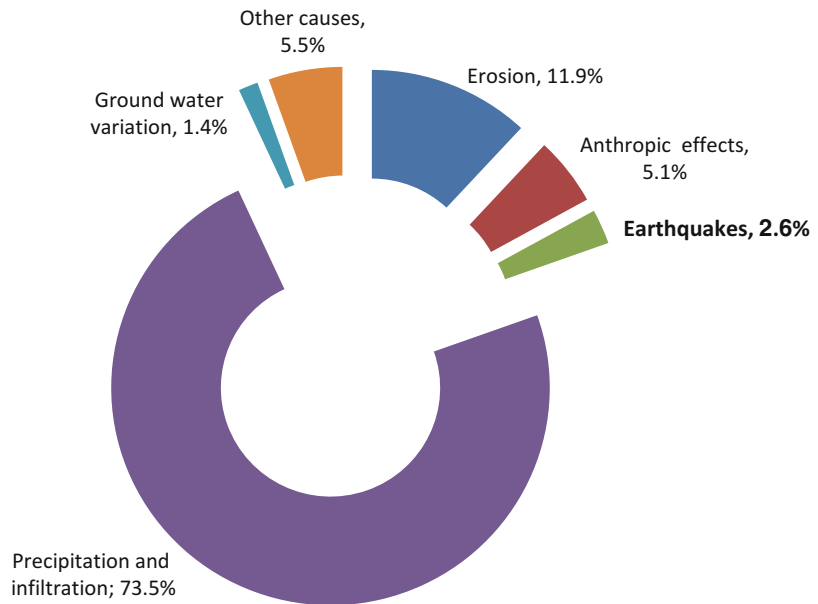
triggered events; and M_c is the completeness magnitude of the data catalog. A space-time generalization of ETAS was later introduced by Ogata (1998). The ETAS model is widely used in different ways, as, for example, to describe short-term seismicity, in the forecasting of aftershock sequences (useful in short-term hazard studies) and also as a declustering method to separate the background seismicity from the triggered one.

From the perspective of considering cascading effects in the seismic risk assessment, the stochastic description of the seismic sequences (in particular by the ETAS model) has important consequences in probabilistic seismic hazard assessments, since it permits to take into account a more realistic behavior of the seismicity to calculate aftershock hazard curves, which is of fundamental importance for short-term hazard assessments.

It is worth noting here that considering the sequences of earthquakes in a chain of triggering effects can be regarded as an interaction at the hazard level (according to the taxonomy provided in this entry). Nevertheless, it may require also taking into account the cumulative effect of the different earthquakes on the exposed elements (e.g., the physical structures), which requires to consider also a possible interaction at the

vulnerability level by using damage-dependent fragilities.

Earthquakes as Triggering Mechanism of Landslides

Strong earthquakes may cause widespread landslides and other ground failure phenomena as, for example, liquefaction. A comprehensive historical review of outstanding past examples of earthquake-triggered landslides can be found in Keefer (2002) and references therein. In many cases the loss of lives and properties caused by the landslides exceed the losses caused by the earthquake itself (see, e.g., Chen et al. 2012). The occurrence of landslides can be associated with different triggering mechanisms as intense precipitation, erosion, ground shacking caused by earthquakes, groundwater variations, nonnatural effects of anthropic activity, etc. For example, let us consider the database from the "National Group for Prevention of Hydrological Hazards" of areas affected by landslides in Italy (CNR-GNDCI, http://avi.gndci.cnr.it); the most frequent kind of events triggering landslides in Italy are (see Fig. 3): the precipitation and infiltration (73.5 %), erosion (11.9 %), anthropic activity (5.1 %), earthquakes (2.6 %), and groundwater variation (1.4 %).

S

Taking as reference the available landslide catalog in Italy, about a 3 % of the total landslide occurrences in this area can be associated with earthquake-triggering mechanisms (Fig. 3). It is worth noting that this example is representative for the Italian area, and therefore the proportion of events for the different triggering phenomena may considerably change from one region to the other in function of different factors as, for example, the amount of precipitation, topography and land use, size and location of earthquakes, etc.

From the cascading effects assessment perspective, the occurrence of earthquakes in landslide-prone areas may increase the likelihood of landslide occurrence due to different physical mechanisms, as, for example, the inertial forces during the ground shaking itself or the rapid increase in pore water pressure induced by the cyclic stresses and subsequent reduction of the soil shear strength. Furthermore, earthquakes may reduce the soil strength for longer periods, causing the affected area to be more susceptible to post-earthquake rainfall-triggered landslides until the soil has regained its original strength.

Several models have been proposed for assessing the stability of slopes during earthquakes; an interesting overview of methods is provided in Jibson (2011). These methods are conveniently grouped into (1) pseudo-static analyses, (2) stress-deformation analyses, and (3) permanent displacement analyses. One of the most popular models in the latter group is that proposed by Newmark (1965), where the landslide is modeled as a rigid block sliding on an inclined surface. Movement of the block occurs when the sum of the static and dynamic driving forces exceed the shear resistance of the block. Permanent deformation occurs when the induced accelerations exceed some critical acceleration. The sliding block model assumes (1) a rigid-plastic behavior of the boundary between the surface and the sliding block, (2) the existence of a well-defined slip surface, (3) negligible loss in shear resistance during shaking, and (4) permanent displacement occurs only if dynamic stresses exceed shear resistance. A displacement can take place then during a cyclic load whenever the shear stress exceeds the base resistance, and

the block displacement ends when the block-base relative velocity equals zero. Displacement steps are summed over the entire duration of the acceleration-time history to obtain the total displacement. The method has undergone several modifications and improvements, and several relations between seismic ground-motion parameters and computed landslide displacements have been proposed. A particularly interesting application for quantitative hazard assessment of earthquake-triggered landslides is presented in Miles and Ho (1999), in which hazard analyses are performed combining Newmark's approach and Boore's stochastic ground-motion simulations. Likewise, stochastic methods for probabilistic earthquake-triggered landslide hazard assessment have been also proposed. For example, Jibson et al. (2000) performed a comparative analysis of the displacements modeled using the Newmark's permanent-deformation analysis and the landslide observations after the 1994 Northridge, California earthquake and used this information to construct a probability curve relating predicted displacement to probability of failure. Likewise, Refice and Capolongo (2002) presented a procedure that takes into account uncertainties and the spatial variability of geological, geotechnical, geomorphological, and seismological parameters in the Newmark slope stability model and applied it for probabilistic earthquake-triggered landslide hazard assessment.

Earthquakes Triggering Volcanic Eruptions

Deformation of the surrounding rocks is required for the magma movement before volcanic eruptions; therefore, small to moderate local earthquakes (usually called *volcano-tectonic* seismicity) normally precede and accompany eruptive activity. However, volcanic unrest episodes and volcanic eruptions have been observed within days after the occurrence of big earthquakes occurring outside the volcanic area. Evidences of tectonic earthquakes preceding volcanic activity can be found in different tectonic setting and on different time and distance scales (see, e.g., Eggert and Walter 2009; Gasparini 2013). Well-known cases of this

possible triggering effect may include, among many others, the 1990 M7.8 Luzon earthquake and the reawakening of Mount Pinatubo in 1991 (e.g., Bautista et al. 1996); the 1960 M9.5 Chile earthquake, which was followed about 38 h later by an eruption in the Puyehue-Cordón Caulle volcano (e.g., Barrientos 1994); Mount Vesuvius having the largest eruptions shortly after earthquakes in the Apennines (Nostro et al. 1998); and eruptions of different volcanoes in Kamchatka (some of them not considered active) following earthquakes in different periods (e.g., Karymsky and Dzenzur volcanoes erupting after an M8.5 earthquake in 1923, the Karpinsky volcano and the Tao-Rusyr Caldera erupting few days after an M9.0 earthquake, and the Karymsky and Akademia Nauk volcanoes erupting few days after a M7.1 earthquake, Walter 2007 and references therein).

Approximately 0.4 % of explosive volcanic eruptions occur within a few days of large, distant earthquakes, which is a value much greater than expected by chance (Manga and Brodsky 2006; Linde and Sacks 1998). All the available examples are suggestive of a triggering connection, but quantifying that connection is not straightforward. Linde and Sacks (1998) analyzed the historical record of large earthquakes and explosive eruptions and found that within a day or two of large earthquakes, there are many more eruptions within a range of about 750 km than would otherwise be expected. Therefore, they concluded that eruptions occur in the vicinity of a large earthquake more often than would be expected by chance. Additionally, it is well known that volcanoes separated by hundreds of kilometers frequently erupt in unison; the characteristics of such eruption pairs are also consistent with the hypothesis that the second eruption is also triggered by the earthquakes associated with the first (Linde and Sacks 1998). There is also strong evidence that large earthquakes can affect the unrest condition of volcanoes producing transient enhancement of microseismic and hydrothermal activities as well as ground deformation (see Gasparini 2013).

As for the earthquake-earthquake interactions, several mechanisms have been proposed to plausibly explain the volcanic eruption triggering by large earthquakes, being the most widely discussed those associated with transient (or dynamic), viscoelastic (or quasi-static), and permanent (or static) stress and strain changes. In fact, earthquakes can stress magmatic systems either through static stresses (e.g., the offset of the fault generates a permanent deformation in the crust), viscoelastic stresses (e.g., a change in the stress state in the elastic crust associated with slow, viscous relaxation of the lower crust and upper mantle beneath the epicenter of a large earthquake), or dynamic stresses from the seismic waves (Hill et al. 2002; Manga and Brodsky 2006).

Dynamic stresses are proportional to the seismic wave amplitude, which is influenced by different factors as directivity, radiation pattern, and crustal structure. However, the significant difference in dependence on distance is a robust feature distinguishing static and dynamic stresses. The difference in decay rate helps to identify which stresses may be important for triggering eruptions. Static stress changes die off rapidly with distance (as $\sim 1/r^3$, with r the distance from the epicenter); therefore, their influence is limited to distances of a few fault lengths (near field). Quasi-static stress changes die off more slowly with distance (as $\sim 1/r$ or $1/r^2$), similarly to the dynamic stress changes that die off as $\sim 1/r$ for surface waves and $\sim 1/r^2$ for body waves (e.g., Hill et al. 2002; Manga and Brodsky 2006).

To provide a comparative example, Table 1 lists the order of magnitude of earthquake stresses

Seismic Risk Assessment, Cascading Effects, Table 1 Static and dynamic stress changes of an M8.0 earthquake at different distances from the source (From Manga and Brodsky 2006)

M8.0 earthquake	Distance from the source		
	Stress (MPa)		
	10^2 km	10^3 km	10^4 km
Static stress changes	1.0×10^{-1}	1.0×10^{-4}	1.0×10^{-7}
Dynamic stress changes	3.0	6.0×10^{-2}	1.0×10^{-3}

**Seismic Risk Assessment, Cascading Effects,
Table 2** Order of magnitude of other sources of forcing
(From Manga and Brodsky 2006)

Phenomena	Stress MPa
Solid earth tides	10^{-3}
Ocean tides	10^{-2}
Hydrological loading	10^{-3} to 10^{-1}
Glacier loading	10^{1} to 10^{2}

at different distances from the source, and Table 2
lists other external stresses that have been also
proposed to influence volcanoes (example from
Manga and Brodsky 2006). In the near field, both
static and dynamic stresses may exhibit signifi-
cant values; conversely, in the far field only the
dynamic stresses are larger than other potential
triggering stresses proposed; for instance, for dis-
tances greater than 10^{3} km, the static stress
changes are at least one order of magnitude
lower than the other forcing sources presented
in Table 2.

Several models have been proposed to explain
the different triggering mechanisms. Here we
summarize the main examples presented in Hill
et al. 2002 and in Manga and Brodsky 2006 (and
references therein) for the static and dynamic
stress triggering. The reader is invited to consult
these references for more detailed information.

Static stress triggering: Static stress changes
can be a factor triggering volcanic eruptions at
intermediate distances (few hundred kilometers).
Explanations for this kind of interaction are based
on pressure changes in a magma body induced by
the isotropic, compressional component of the
stress field in the vicinity of the volcano.
Increased compressional stress in the crust sur-
rounding an already fully charged magma cham-
ber may squeeze magma upward. Likewise, a
decrease in compressional stress can promote
vesiculation, bubble formation, and unclamping
of conduits above the magma body (Hill
et al. 2002).

Dynamic stress interactions: Several models
have been proposed to explain the phenomenon
of dynamic triggering by distant earthquakes but
the physical processes remain a matter of discus-
sion. Considering in particular those mechanisms
in which fluids are active in the triggering

process, examples of the models proposed
include (Hill et al. 2002):

- Excitation of bubbles in a gas-saturated fluid
  by passing seismic waves, which acts to
  increase fluid pressure through adjective over-
  pressure and/or rectified diffusion (see, e.g.,
  Hill et al. 2002; Manga and Brodsky 2006).
- Hydraulic surge, a model in which the
  seismogenic crust is separated from the under-
  lying plastic crust and an embedded magma
  body by an impermeable transition zone; a
  hydraulic surge occurs when the large
  dynamic strains associated with seismic
  waves from a large, distant earthquake disrupt
  the impermeable seal.
- Rupturing blockages in confined aquifers, in
  which passing seismic waves disrupt block-
  ages in fractures formed by precipitates in a
  confined aquifer allowing local adjustments in
  fluid pressure within the aquifer.
- Sinking crystal plume model, in which rela-
  tively dense, loosely held masses of crystals
  can accumulate on the ceiling and walls of a
  crystallizing magma body and might be
  dislodged by passing seismic waves. As crys-
  tal-rich plumes sink, less dense, crystal-poor
  magma will be forced upward initiating a con-
  vective cycle. Volatiles in the ascending
  plume will form bubbles and a pressure
  increase through advective overpressure.
- Relaxing magma body: A magma body may
  reach a partially crystallized state by either
  (1) gradual cooling of a previously molten
  volume or (2) partial melting of a previously
  solid volume. The solid matrix will partially
  support accumulating tectonic stresses in the
  region. Strong seismic waves may disrupt
  the solid matrix thereby redistributing stress
  to the surrounding crust as the magma body
  deforms through viscous relaxation.

Different observations worldwide indicate
then that the triggering effect between earth-
quakes and volcanic eruptions may exist under
certain conditions. Stress changes, in particular,
seem to be a plausible mechanism capable of
triggering volcanic eruptions in a wide range of

distances and times. Given the complexity of the problem and of the interactions, until now the problem is generally treated statistically. New observations in currently well-monitored volcanoes, statistical significance of the correlations, and physics of the triggering mechanisms are important research topics that need to be addressed in order to better constrain this kind of hazard interaction.

Earthquakes Triggering Industrial Accidents

Damages in industrial plants are known to have triggered a number of severe accidents caused mainly by the loss of containment of hazardous substances. Industrial accidents can be triggered by a number of causes, the most common of which are internal system failures and/or human errors (Rasmussen 1995). However, industrial facilities may also be subject to accidents triggered by external factors as natural events. Analyzing different databases of industrial accidents in Europe and the USA, Kraussmann et al. (2011) outlined that between 2 % and 5 % of the recorded accidents were triggered by natural hazards (so-called Natech accidents), being this probably a low bound because of a reporting bias towards accidents with severe consequences.

From the perspective of cascading effects, earthquakes are one of the possible natural hazards triggering Natech accidents. Earthquakes may cause damages in industrial facilities by the direct impact caused by the ground shaking and also by secondary effects as liquefaction. Considering the results presented by Kraussmann et al. (2011), pipework, pipelines, and tanks for the atmospheric storage of chemicals appear to be particularly vulnerable to excitation and damage by the earthquake forces. In the analyzed cases, about 27 % of the accidents presented only structural damages, whereas about 73 % of the cases exhibited damage with release of hazardous materials. These results seem to suggest that chemical equipment is much more likely to suffer damage with loss of containment than not (Kraussmann et al. 2011). Therefore, scenarios of cascading effects with earthquakes as triggering mechanism causing Natech accidents

(because of damages in industrial facilities) can be followed by other triggered hazards as, for example, the release of hazardous materials, explosions, and fires.

A recent outstanding example of cascading effects leading to industrial accidents triggered by earthquakes is the case of the Tohoku earthquake, Japan. On March 11, 2011, the Tohoku-Oki earthquake generated a tsunami 130 km off the coast of Miyagi Prefecture, northeast Japan. The tsunami first reached the Japanese mainland about 20 min after the earthquake and ultimately affected a 2,000 km stretch of Japan's Pacific coast (e.g., Mori et al. 2011). The tsunami caused nuclear accidents, primarily the damages at three reactors in the Fukushima Daiichi Nuclear Power Plant complex. Accounting for all of the factors relevant during the Tohoku disaster, Khazai et al. (2011) analyzed the cascading effects produced by the main earthquake and both the tsunami and the nuclear power plant incidents and produced a causal map depicting the direct and indirect interactions between the various events.

Considering Cascading Effects in the Seismic Risk Assessment

Taking as reference the general concepts described in Eq. 1, in more quantitative terms, risk is often quantified using an expression of the form (e.g., Cornell and Krawinkler 2000; Marzocchi et al. 2012):

$$\lambda(\ell) = \int_D \int_{I_m} \underbrace{F(\ell|D)dG(D|I_m)}_{\text{damgae \& loss}} \underbrace{d\lambda(I_m)}_{\text{hazard}} \quad (4)$$

where $\lambda(l)$ represents a rate of loss, $F(l|D)$ is the probability that a given loss is reached given the damage state D, $G(D|I_m)$ is the vulnerability (or fragility) term, $\lambda(I_m)$ is the hazard, measuring the exceedance rates of given intensity measures (I_m) characterizing the hazard, l is the loss in one specific metric, and D is a given damage state.

Assessing interactions, or cascading effects, in a multi-hazard problem consists on the identification of the possible interactions that are likely to happen and that may result in an amplification of the expected losses in a given area of interest. This process may be performed in the following steps:

1. Identification and structuring of the plausible scenarios of cascading effects
2. Identification of the kind of interactions according to the proposed taxonomy
3. Quantification of the scenarios of interest

Assuming that the value of the exposed elements remains constant in a given time window of interest, the amplification of the losses may be caused by an increase in the occurrence probability of the hazard or by an increase in the probability of damage by a change in the fragility of the exposed elements. In some cases both elements may be considered in a single scenario. In the first case, the analysis of interactions updates the hazard term in Eq. 4; in the second case, the damage term in Eq. 4 is updated.

Identification and Structuring Scenarios of Cascading Effects

An important initial step towards the assessment of cascading effects is the identification and structuring of the possible scenarios. The term "scenario" is used in a wide range of fields and so different interpretations can be found in practical applications. In general, a scenario may be considered as a plausible and consistent representation of an event or series of actions and events. It must be plausible because it must fall within the limits of what might conceivably happen and must be consistent in the sense that the combined logic used to construct a scenario must not have built-in inconsistencies.

Different strategies can be adopted in order to exhaustively identify a complete set of scenarios. In general, they are based on different complementary approaches ranging from event-tree-like to fault-tree-like strategies. In many applications, an adaptive method combining both kinds of approaches is applied in order to ensure an exhaustive exploration of scenarios. A wide description of practical methods for scenario identification and structuring can be found, for example, in Haimes (2009).

Taxonomy of Interactions

Taking as reference the works of Marzocchi et al. (2012), Garcia-Aristizabal et al. (2013), Selva (2013), and Garcia-Aristizabal et al. (forthcoming-2015), it is possible to define a generalized taxonomy of interactions considering two possible kinds of interactions, namely, (1) interactions at the hazard level, in which the occurrence of a given initial *triggering* event entails a modification of the probability of occurrence of a secondary event and (2) interaction at the vulnerability (or damage) level, in which the main interest is to assess the effects that the occurrence of one event (the first one occurring in time) may have on the response of the exposed elements against another event (that may be of the same kind as the former but also a different kind of hazard). Implicitly, a combination of both kinds of interactions is another possible case; therefore, in the discussion of the interactions at the vulnerability level, both dependent and independent hazards can be considered as possible cases.

Triggering Effects: Interactions at the Hazard Level
In this case, the interaction problem is understood as the assessment of possible "chains" of adverse events in which the occurrence of a given initial *triggering* event entails a modification of the occurrence probability of a secondary event. Even if this typology of problems can be assessed in a long-term basis, their utility can be highlighted in short-term assessments, in which specific scenarios conditioned to the occurrence of given specific events can be assessed and compared among them.

Let us first consider two different threatening events, whose occurrences are E_1 and E_2. In general, the probability of E_2 occurring (generically hazard H_2) can be written as (Marzocchi et al. 2012):

$$H_2 = p(E_2)$$
$$= p(E_2|E_1)p(E_1) + p(E_2|\overline{E}_1)p(\overline{E}_1) \quad (5)$$

where, p represents a probability or a distribution of probability and $\overline{E}_1$ means that the event E_1 does not occur. Equation 5 represents the total probability of E_2 occurring considering generically the occurrence or nonoccurrence of E_1. However, the quantification of risk generally requires the hazard assessed in function of a specific intensity measure (IM) characterizing the hazard, and therefore Eq. 5 can be generalized in order to consider the specific contribution of different IM values. Therefore, the occurrence of the triggered event 2 (E_2) given the occurrence of the triggering event 1 (E_1) representing this kind of interaction can be assessed as (Garcia-Aristizabal et al. 2013):

$$p\left(\geq IM_2^i\right) = \sum_j p\left(\geq IM_2^{i,j}\right) \quad (6)$$

that explicitly conditioned to E_1 can be written as (e.g., Garcia-Aristizabal et al. 2013; Selva 2013; Garcia-Aristizabal et al. forthcoming-2015):

$$p\left(\geq IM_2^i\right) = \sum_j p\left(\geq IM_2^i|IM_1^j\right)p\left(IM_1^j\right) \quad (7)$$

For $j = 1, 2, 3, \ldots, n$, for the n (exhaustive and mutually exclusive) classes of *IM* defined for the triggering event. IM_1 is the intensity measure defined for the triggering event, and IM_2 is the intensity measure defined for the triggered event. Note that Eq. 7 is a generalization of Eq. 5. A key element in this formalization is $p(\geq IM_2^{i,j})$, which we call the *interaction term*. When $i > 1$ and $j > 1$, $p(\geq IM_2^{i,j})$ is represented by a $i \times j$ matrix that defines a probability surface, as shown in Fig. 4. When $i = 1$ or $j = 1$, $p(\geq IM_2^{i,j})$ is a row or column vector, and in the hypothetical limit case in which the problem does not consider the intensity measures of both the triggered and triggering events, the problem is simplified as the case described by Eq. 5 (i.e., considering the "occurrence" and "nonoccurrence" of the event 1). Considering, for example, the earthquake-triggered landslide case, the earthquake is represented by E_1 and the triggered landslides are the E_2; IM_1 can represent the ground-motion intensity and IM_2 the volume of the mass sliding. Therefore, $p(\geq IM_2^i|IM_1^i)$ represents the probability of exceeding a given sliding mass volume given a certain ground shaking intensity, and $p(\geq IM_2^i)$ represents the *total* probability of exceeding a given sliding mass volume considering the full range of possible ground shaking intensities taken into account from the triggering earthquake hazard.

Seismic Risk Assessment, Cascading Effects, Fig. 4 Probability surface represented by $p(\geq IM_2^{i,j})$, for $i > 1$ and $j > 1$
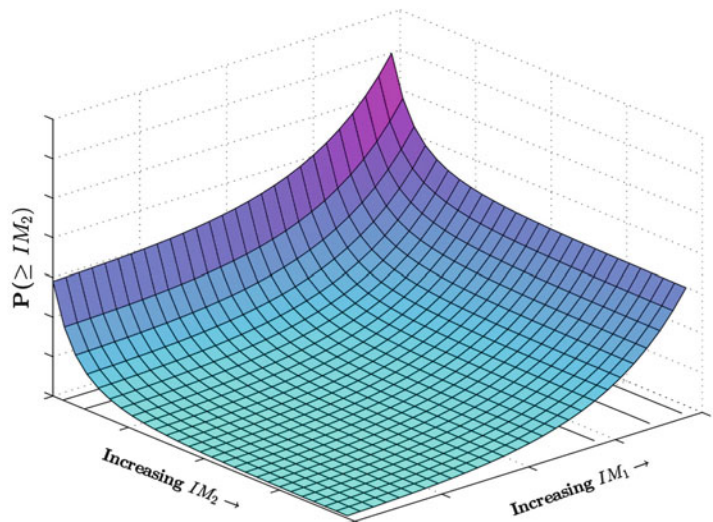
Additive Effects of Different Events: Interactions at the Vulnerability Level

This perspective of the cascading effect problem fundamentally intends to assess the effects that the simultaneous occurrence of two or more events (not necessarily linked among them) may have for the final risk assessment. In this case, the action of different hazards is considered and combined at the vulnerability (or damage) level, and the main interest is to assess the effects that the occurrence of one event (the first one occurring in time) may have on the response of the exposed elements against another event (that may be of the same kind as the former but also a different kind of hazard). Examples of this kind of interactions have been presented in literature, for example, in Lee and Rosowsky (2006), Zuccaro et al. (2008), Marzocchi et al. (2012), Garcia-Aristizabal et al. (2013), and Selva (2013).

In practice, the idea in this case is to quantify how the expected damages in the target area (caused by a given hazard) can be modified if another hazardous event acts on the exposed elements simultaneously or in a short-time window (in general short enough that the system cannot be repaired). In the case of two hazards having additive load effects (i.e., they act simultaneously over the exposed element), the fragility function will depend on the intensities IM_1 and IM_2 of the two hazards, and then it will represent a fragility surface. The probability that a given damage state (D_k) is reached given the occurrence of the i-th value of IM_1 and the j-th value of IM_2 can be represented as (for details see Lee and Rosowsky 2006; Garcia-Aristizabal et al. 2013; Selva 2013):

$$p(D_k) = \sum_i \sum_j \left[P\left(D_k | IM_1^i \cap IM_2^j\right) p\left(IM_1^i \cap IM_2^j\right) \right] \tag{8}$$

The conditional probability $P(D_k | IM_1^i \cap IM_2^j)$, hereinafter referred to as $P(D_k | IM_1^i, IM_2^j)$ for simplicity, is the probability that the damage state D_k is reached at given levels of loads IM_1 and IM_2 (due to hazard events 1 and 2, respectively) acting simultaneously. Therefore, to calculate $p(D_k)$

considering any value of the IM of the events 1 and 2, we need to consider two cases, namely, (1) when events 1 and 2 are independent and (2) when there is dependence between events 1 and 2.

Independent Events In the simpler case, we consider the two hazards as independent events, in the sense that the occurrence of one does not change the probability of the other occurring. In that case, the probability that a given damage state is reached for all the possible values of IM_1 and IM_2 can be defined as (e.g., Garcia-Aristizabal et al. 2013):

$$p(D_k) = \sum_i \sum_j \left[P\left(D_k | IM_1^i, IM_2^j\right) p\left(IM_1^i\right) p\left(IM_2^j\right) \right] \tag{9}$$

which is an expression coherent with the framework defined by Lee and Rosowsky (2006) analyzing the effects of the combined seismic and snow loads on wood structures. In this case, the fragility function $P(D_k | IM_1^i, IM_2^j)$ will represent a probability surface as that shown in Fig. 5. The example presented in Fig. 5 shows the combined effect of seismic (in pga) and volcanic ash (in Kpa) loads acting simultaneously on roof structures in Naples, Italy. The data used to generate this figure was derived from the work presented in Zuccaro et al. (2008).

Dependent Events If the occurrence of one event (E_1) does affect the probability of the other occurring (E_2), then the events are dependent. This is the case described in the interactions at the hazard level. In this case, the term $p(IM_2^j \cap IM_1^i) = p(IM_2^j | IM_1^i) p(IM_1^i)$ and then the probability that a given damage state D_k is reached for the case of dependent events can be defined as:

$$p(D_k) = \sum_i \sum_j \left[P\left(D_k | IM_2^j, IM_1^i\right) p\left(IM_2^j | IM_1^i\right) p\left(IM_1^i\right) \right] \tag{10}$$

In this case, $P(D_k | IM_2^j, IM_1^i)$ is again a fragility surface as that represented in Fig. 5, and the

Seismic Risk Assessment, Cascading Effects, Fig. 5 Fragility surface for the case of seismic and volcanic ash acting simultaneously on the roofs in Naples, Italy (Derived from the data in Zuccaro et al. 2008). Note that when the ash load tends to zero, the fragility function converges towards the pure seismic fragility
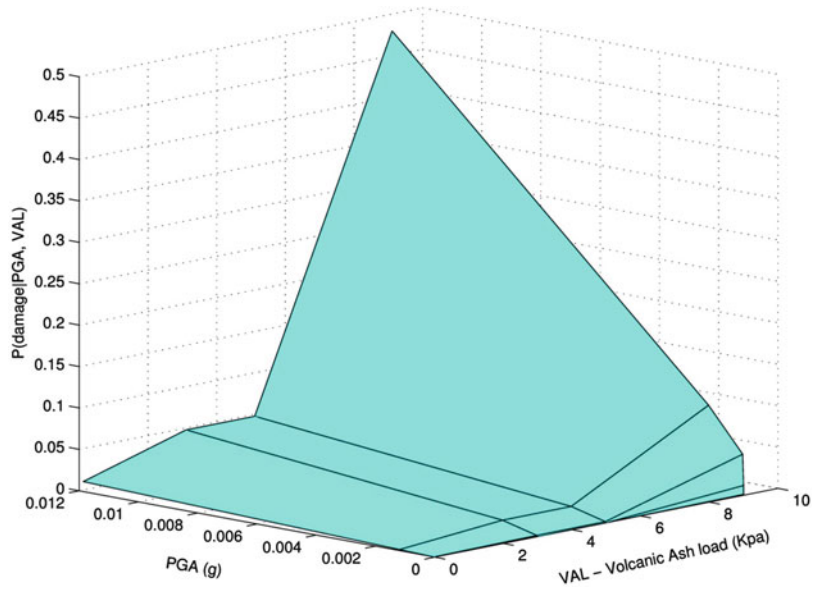
hazard term takes into account the possible dependence between events 1 and 2. For example, considering again the earthquake-triggered landslide case, assessing the "collapse" damage state would require the fragility surface $P(D_k|IM_2^i, IM_1^i)$ for assessing the probability of collapse given the combined effect of the ground shaking intensity (IM_1) and, for example, the depth of the deposit caused by the mass movement (IM_2). On the other hand, the hazard term in this case would be calculated as shown in the example presented for the *interactions at the hazard level*, in which the probability of exceeding a given IM_2 value is assessed considering the full range of likely intensities of the triggering ground motions caused by the earthquakes.

Damage-Dependent Vulnerabilities in Cascading Effects

The damage-dependent vulnerabilities (or fragilities) can be considered as a special case of the interaction at the vulnerability level. In fact, the development of fragility surfaces, as presented in the previous section for that kind of interactions, is intended for general cases in which the additive effect of different loads may act simultaneously over the exposed element, which does not necessarily imply a pre-damage of the element. For example, the action of the first occurring hazard

may just result in a modification of the response of the exposed element (e.g., a structure) to another typology of hazard without causing a permanent deformation or damage that remains in time. Examples of this case can be snow or volcanic ash deposited over a roof; this extra weight in the structure may cause it to behave differently if it is shaken by the ground motion caused by an earthquake (compared with the normal structure's behavior). However, in many cases this interaction may imply that the first acting event produces permanent deformations in the structure (i.e., causing some damages) implying some modification to its response to new solicitations and the need to develop damage-dependent fragility functions.

The most frequent development of damage-dependent fragility functions is found for seismic loads (i.e., to assess the structure's response to new earthquakes after it has been damaged by the past earthquakes). Different works tackling the problem of the response of structures pre-damaged by earthquakes can be found in literature and mainly associated with the assessment of post-mainshock time-dependent risk. Examples can be found in (Bazzurro et al. 2004; Yeo and Cornell 2005; Ryu et al. 2011; Luco et al. 2011) and references therein. In general, the fragility curves used for the post-mainshock

risk assessment accounts for any damage caused by the mainshock. This typically increases the probabilities of collapse for the considered range of potential (future) *IM* values, and the amount of increase depends of course on the extent of the mainshock damage, which is commonly discretized into damage states (DS) as, for example, none, slight, moderate, extensive, and complete (Luco et al. 2011). Following Ryu et al. (2011, and references therein), the fragility for a mainshock-damaged structure can be computed as:

$$
\begin{aligned}
p(D_a > d_a | IM_a = im_a,\ D_m = d_m) \\
= \int \begin{array}{l} P(D_a > d_a | IM_a = im_a,\ EDP_m = edp_m) \times \\ f(EDP_m = edp_m | D_m = d_m) dedp_m \end{array}
\end{aligned}
$$

(11)

where D_a represents the post-aftershock damage state, D_m represents the post-mainshock damage state, EDP_m, the engineering demand parameter (e.g., drift), represents the mainshock building response, and IM_a represents the ground-motion intensity of an aftershock (Ryu et al. 2011). The first term in Eq. 11 can be computed using the fragility for the intact building or that for a mainshock-damaged building whose mainshock response (EDP_m) is edp_m. The second term in Eq. 11 can be computed using the assumed distribution of mainshock response given post-mainshock damage state (for details see Ryu et al. 2011; Luco et al. 2011).

Summary

Different past disasters have highlighted the fact that natural or man-made events can trigger other events, leading to a non-negligible increase of fatalities and damages. As a consequence, there is an increasing interest on implementing multi-type hazard and risk assessments as well as assessing cascading effects, which are one of the fundamental concepts of the multi-risk assessment. As for most of the quantitative risk assessment for natural hazards, the seismic risk assessment is generally carried out separately without considering the effects of possible chains of events triggered by earthquakes. From the risk assessment point of view, the cascades of events occurring after an earthquake may represent a non-negligible source of loss amplifications. For quantitative purposes, we can consider the cascading effects analyzing two major sets of interactions: (1) interactions at the hazard level and (2) interactions at the vulnerability level. In this entry we present and discuss different examples of scenarios of cascading effects triggered by earthquakes.

Cross-References

References

Barrientos SE (1994) Large thrust earthquakes and volcanic eruptions. Pure Appl Geophys 142(1): 225–237

Bautista BC, Bautista MLP, Stein RS, Barcelona ES, Punongbayan RS, Laguerta EP, Rasdas AR, Ambubuyog G, Amin EQ (1996) Relationship of regional and local structures to Mount Pinatubo activity. In: Fire and mud; eruptions and lahars of Mount Pinatubo, Philippines. Philippine Institute of Volcanology and Seismology/University of Washington Press United States, Quezon City/Philippines

Bazzurro P, Cornell CA, Menun C, Motahari M (2004) Guidelines for seismic assessment of damaged buildings. In: Proceedings of the 13th world conference on earthquake engineering, Vancouver

Chen XL, Zhou Q, Ran H, Dong R (2012) Earthquake-triggered landslides in southwest China. Nat Hazards Earth Syst Sci 12:351–363. doi:10.5194/nhess-12-351-2012

Cornell CA, Krawinkler H (2000) Progress and challenges in seismic performance assessment. PEER Center News 3(2), URL http://peer.berkeley.edu/news/2000spring/index.html

Eggert S, Walter TR (2009) Volcanic activity before and after large tectonic earthquakes: observations and statistical significance. Tectonophysics 471:14–26

Garcia-Aristizabal A, Marzocchi W, Di Ruocco A (2013) Probabilistic framework for multi-hazard assessment. Technical report D3.4, MATRIX project (New methodologies for multi-hazard and multi-risk assessment methods for Europe), grant no. 265138. URL http://matrix.gpi.kit.edu/Deliverables.php

Garcia-Aristizabal A, Gasparini P, Uhinga G (forthcoming-2015) Multi-risk assessment as a tool for decision-making. In: Pauleit S, Jorgensen G, Kabisch S, Gasparini P, Fohlmeister S, Simonis I, Yeshitela K, Coly A, Lindley S, Kombe WJ (eds) Future Cities, vol 4 Climate change and urban vulnerability in Africa: a multidisciplinary approach. Springer Netherlands, ISBN 978-3-319- 03984-8

Gasparini P (2013) The physics of volcanoes since the 1960s. Rivista Nuovo Cimento 36:461–500. doi:10.1393/ncr/i2013-10093-0

Haimes Y (2009) Risk modeling, assessment, and management, 3rd edn. Wiley, Hoboken

Harris RA (1998) Introduction to a special section: stress triggers, stress shadows, and implications for seismic hazards. J Geophys Res 103: 24,347–24,358

Hill DP, Pollitz F, Newhall C (2002) Earthquake-volcano interactions. Physics Today 55(11):41–47

Jibson RW (2011) Methods for assessing the stability of slopes during earthquakes – a retrospective. Eng Geol 122:43–50

Jibson RW, Harp EL, Michael JA (2000) A method for producing digital probabilistic seismic landslide hazard maps. Eng Geol 58(3–4):271–289. doi:10.1016/S0013-7952(00)00039-9

Keefer DK (2002) Investigating landslides caused by earthquakes – a historical review. Surv Geophys 23:473–510

Khazai B, Daniell JE, Wenzel F (2011) The March 2011 Japan earthquake: analysis of losses, impacts, and implications for the understanding of risks posed by extreme events. TECHNIKFOLGENABSCHÄTZUNG, Theorie und Praxis. N3, 20 Dec 2011. S. 22–33. http://www.itas.fzk.de/tatup/113/khua11a.htm

Kraussmann E, Renni E, Campedel M, Cozzani V (2011) Industrial accidents triggered by earthquakes, floods and lightning: lessons learned from a database analysis. Nat Hazards. doi:10.1007/s11069-011-9754-3

Lee KH, Rosowsky DV (2006) Fragility analysis of woodframe buildings considering combined snow and earthquake loading. Struct Saf 28(3):289–303. doi:10.1016/j.strusafe.2005.08.002

Linde AT, Sacks IS (1998) Triggering of volcanic eruptions. Nature 395:888–890. doi:10.1038/27650

Luco N, Gerstenberger MC, Uma SR, Ryu H, Liel AB, Raghunandan M (2011) A methodology for post-mainshock probabilistic assessment of building collapse risk. In: Proceedings of the ninth pacific conference on earthquake engineering building an earthquake-resilient society, Auckland

Manga M, Brodsky E (2006) Seismic Triggering of Eruptions in the Far Field: Volcanoes and Geysers. Annual Review of Earth and Planetary Sciences 34:263–291. doi:10.1146/annurev.earth.34.031405.125125

Marsan D, Lengliné O (2008) Extending earthquakes' reach through cascading. Science 319(5866): 1076–1079. doi:10.1126/science.1148783

Marzocchi W, Garcia-Aristizabal A, Gasparini P, Mastellone ML, Di Ruocco A (2012) Basic principles of multi-risk assessment: a case study in Italy. Nat Hazards 62(2):551–573. doi:10.1007/s11069-012-0092-x

Miles SB, Ho CL (1999) Rigorous landslide hazard zonation using Newmark's method and stochastic ground motion simulation. Soil Dyn Earthquake Eng 18:305–323

Mori N, Takahashi T, Yasuda T, Yanagisawa H (2011) Survey of 2011 Tohoku earthquake tsunami inundation and run-up. Geophys Res Lett 38:L00G14. doi:10.1029/2011GL049210

Newmark NM (1965) Effects of earthquakes on dams and embankments. Geotechnique 15(2):139–160

Nostro C, Stein RS, Cocco M, Belardinelli ME, Marzocchi W (1998) Two-way coupling between Vesuvius eruptions and southern Apennine earthquakes, Italy, by elastic stress transfer. J Geophys Res B Solid Earth 103(B10):24487–24504

Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. J Am Stat Assoc 83:9–27

Ogata Y (1998) Space-time point-process models for earthquake occurrences. Ann Inst Stat Math 50:379–402

Plafker G, Ericksen GE, Concha JF (1971) Geological aspects of the 31 May 1970, Peru earthquake. Bull Seis Soc Am 61:543–578

Rasmussen K (1995) Natural events and accidents with hazardous materials. J Hazard Mater 40:43–54

Refice A, Capolongo D (2002) Probabilistic modeling of uncertainties in earthquake-induced landslide hazard assessment. Comput Geosci 28(6):735–749. doi:10.1016/S0098-3004(01)00104-2

Ryu H, Luco N, Uma SR, Liel AB (2011) Developing fragilities for mainshock-damaged structures through incremental dynamic analysis. In: Proceedings of the ninth pacific conference on earthquake engineering building an earthquake-resilient society, Auckland

S

Selva J (2013) Long-term multi-risk assessment: statistical treatment of interaction among risks. Nat Hazards 67(2):701–722

Vere-Jones D, Davies RB (1966) A statistical survey of earthquakes in the main seismic region of New Zealand. N Z J Geol Geophys 9(3):251–284. doi:10.1080/00288306.1966.10422815

Walter TR (2007) How a tectonic earthquake may wake up volcanoes: stress transfer during the 1996 earthquake–eruption sequence at the Karymsky Volcanic Group, Kamchatka. Earth Planet Sci Lett 264(3–4):347–359

Yeo GL, Cornell CA (2005) Stochastic characterization and decision based under time-dependent aftershock risk in performance-based earthquake engineering. Technical report 2005/13. Pacific Earthquake Engineering Research Center (PEER)

Zuccaro G, Cacace F, Spence R, Baxter P (2008) Impact of explosive eruption scenarios at Vesuvius. J Volcanol Geotherm Res 178(3):416–453

Seismic Robustness Analysis of Nuclear Power Plants

Manuel Pellissetti and Ulrich Klapp
AREVA GmbH, Erlangen, Germany

Synonyms

Seismic margin assessment; Seismic PRA; Seismic probabilistic risk analysis; Seismic probabilistic safety analysis; Seismic PSA; SMA

Introduction

Nuclear Safety

Commercial nuclear power plants (NPP) are industrial facilities that represent a significant hazard potential. Consequently, the pertinent safety requirements are particularly stringent. Nuclear safety is assured by deterministic and probabilistic analysis. In deterministic analysis, the safety systems are shown to withstand design basis accidents (DBAs), typically using conservative assumptions. In probabilistic safety analysis (PSA), the frequency of accident progressions leading to core damage (PSA level 1) and large early release (PSA level 2) is quantified and compared to target values that are judged as low enough.

Consideration of Earthquakes in Nuclear Safety

Earthquakes are external hazards on which the duty holder of a nuclear facility, e.g., a nuclear power plant, has no or only little control.

Effects to be expected at a site from an earthquake especially include vibrations induced in structures, systems, or components (SSC) through the civil structures of the plant. These vibrations could affect the plant safety functions directly, e.g., when the induced seismic loads would exceed the capacity of safety relevant equipment items. Indirect failure modes such as mechanical interaction, dropped loads, release of hazardous substances, seismic-induced fire or flooding, impairment of operator access, or unavailability of evacuation and access routes may also be of concern (see IAEA 2003a, §2.11).

Current design practice for a nuclear facility ensures an adequate seismic design from a deterministic perspective.

A safe shutdown earthquake (SSE) is selected based on the results of a seismic hazard analysis of the site. The SSE represents a site-specific extreme seismic event. The attribute "extreme" is justified because the SSE is defined so conservatively that it can be expected with a very high level of confidence that it will not be exceeded during the operational lifetime of the plant. SSCs necessary to transfer the plant into a safe shutdown state are identified and designed to remain functional under SSE loading. Similarly, SSCs whose seismic-induced failure could lead to interaction effects on safety equipment are identified and designed for integrity or stability as required by safety analysis. The demonstration that the design requirements are met consists in mechanical analysis or shaking table testing of the selected systems, structures, and components (SSC), subject to vibrations induced by the SSE.

The described provisions ensure that even in the case of an extreme seismic event, represented by the SSE, the main safety functions (reactivity

control, residual heat removal, confinement of radioactive substances) are met.

Motivation for Evaluating Seismic Robustness

What was described above is – in very simplified terms – the procedure for ensuring a safe initial design of an NPP, as far as seismic loads are concerned. Besides that, nuclear safety is also concerned with ensuring continuously that it is safe to operate the plant, e.g., by means of periodic safety evaluations or if new insights are gained.

Indeed, during the plant's lifetime, new information becomes available which may challenge the initial design basis of the plant. This new information may especially be of seismological nature (e.g., newly discovered seismogenic structures, newly installed seismological networks, new paleoseismological evidence or actual occurrence of damaging earthquakes in the site vicinity). However, there may also be new technical findings such as the indication that certain SSC may be more vulnerable than initially expected (see IAEA 2009, §2.10).

In consequence the nuclear industry and regulatory authorities are frequently faced with the question whether these changes can be accommodated within the seismic capacity of the original design or whether modifications are necessary to maintain an adequate level of safety (see ASCE 2000, p. 41).

When assessing the safety of a nuclear facility, the plant performance in seismic events exceeding the ground motion characteristics of the SSE (typically associated with mean return periods ≤ 10.000 years, see below) is also of interest. See, for example, §2.39 in IAEA (2003a): "*Seismic design should be carried out [...] to provide margins for seismic events that are beyond the design basis and to prevent potential small deviations in plant parameters from giving rise to severely abnormal plant behaviour ('cliff edge' effects). [...]*"

Last but not least, the severe accidents which occurred at Japan's Fukushima-Daiichi plants in response to the Great Tohoku Earthquake of March 2011 have stimulated an increased interest in studies quantifying the actual margin of failure ("Seismic Margin Assessment" – SMA) or providing an insight into the risk associated to earthquakes ("Seismic Probabilistic Safety Analysis" – SPSA, often also referred to as "Seismic Probabilistic Risk Assessment" – SPRA).

Seismic Hazard Assessment

Before introducing the two main approaches for assessing the robustness of NPPs, it is necessary to briefly introduce the notion of seismic hazard.

The assessment of the site-specific seismic hazard – i.e., the characterization of the ground motions to be "expected" at a particular site – is a critical ingredient of both the (initial) design against seismic loads and periodic safety reviews during the plant's lifetime.

Seismic events are unpredictable (stochastic) as far as three of their most basic characteristics are concerned: timing, i.e., when they occur; magnitude; and location of the epicenter. The unpredictability of the latter two properties implies that the ground motion characteristics are also unpredictable. As in many fields of science and engineering, the difficulties associated with the unpredictability can be alleviated – to some extent – by adopting a probabilistic approach: indeed the unpredictability is confined to individual realizations of the stochastic phenomenon, whereas ensemble properties (statistics) are now predictable.

For the purpose of NPP safety, the results of a seismic hazard assessment are essentially pairs of data. Each pair consists of:

- One data item characterizing ground motion (such as the peak ground acceleration, or a response spectrum)
- One data item characterizing how rare it is to observe an event leading to such a ground motion (e.g., return period or annual probability of exceedance)

The latter data item is typically governing the specification of seismic loads. The annual probability of exceedance is prescribed at an

adequately low level, which varies from country to country; typical values of this probability are 10^{-4}/a and 10^{-5}/a. The corresponding ground motion then represents the SSE.

The need to characterize events with extremely low probabilities of exceedance introduces an obvious data problem, as there cannot be – by definition – direct observations within the range of interest. Extensive efforts – which are far beyond the scope of this essay – have been ongoing for decades in order to predict extreme seismic events, mainly by modeling the processes from the source to the origin, i.e., the source mechanisms, the propagation (attenuation), and the site response. For an overview and further reading on seismic hazard refer, for example, to Chen and Scawthorne (2002). The level of uncertainty associated with these models – both due to limited observations for calibrating the model and due to the existence of concurrent models – remains very high.

This large degree of uncertainty associated with seismic hazard estimates is also one of the main reasons for the significant interest in the seismic robustness.

Two Approaches for Quantifying Seismic Robustness

There exist essentially two approaches for quantifying the robustness of NPPs to seismic-induced ground motions, namely, seismic margin assessment (SMA) and seismic probabilistic safety analysis (seismic PSA). It should be noted that in the United States, the term "probabilistic risk analysis" (PRA) is used instead of PSA.

In **SMA** the robustness is determined **explicitly** in terms of the margin between the design acceleration, used for determining the seismic loads on structures and equipment in the design calculations, and the plant capacity, i.e., the limit load that the plant can at most sustain and hence still be brought to a safe state and maintained in that state.

In contrast, in **seismic PSA** the margin is not quantified explicitly. Rather, for a comprehensive range of characteristic ground motions of increasing severity, the corresponding probability of

a core damage is evaluated. Core damage is mainly defined as the failure to meet the core cooling criteria and the associated overheating of the fuel rod cladding.

By taking into account the frequency of occurrence of the characteristic ground motions – expressed by the so-called hazard curve – and the corresponding core damage probabilities, one obtains a concise measure of the seismic risk in the form of the seismic-induced core damage frequency. In this quantity, the robustness is **implicitly** contained. Indeed, for a given hazard curve, a plant with a larger margin will have a smaller core damage frequency than a plant with a smaller margin.

Conversely, the opposite can be said about the methods as far as the notion of risk is concerned: in SMA the risk is contained implicitly, whereas the result of seismic PSA is an explicit expression of risk.

In both approaches – SMA and seismic PSA – the focus is on the **plant-level** robustness, i.e., it is essential that the safeguard systems of the plant be able to fulfill their safety function, even if individual equipment items may have failed. The analysis of the plant-level robustness takes into account the multiple lines of defense and the safety concepts – such as redundancy and diversity – implemented in the engineered safeguards. Ultimately, however, the plant-level robustness can be traced back to the seismic margin of **individual** systems, structures, and components (SSC). Hence, before discussing seismic robustness in terms of plant-level safety functions, the concept of seismic margin is introduced for individual SSC.

Seismic Margin of Systems, Structures, and Components

Load Versus Capacity

Considering the deterministic approach applied for the seismic design of nuclear facilities, a straightforward approach to quantify seismic margins would be to compare (for each important SSC) the load experienced by the SSC in case of

Seismic Robustness Analysis of Nuclear Power Plants, Fig. 1 Margin definition in terms of seismic load
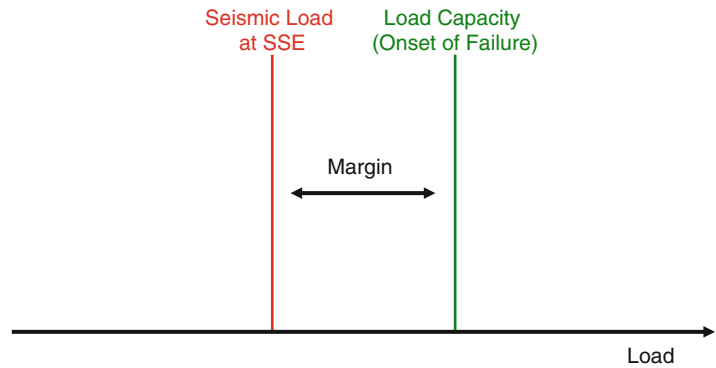
SSE with the load capacity, i.e., with the maximum load the SSC can still handle without failing to meet its requirements (stability, integrity, or operability) (Fig. 1).

However, a statement like "the mechanical analysis of the control rod drive mechanism reveals a seismic margin of x kN" would not be a very comprehensive expression of the existing margin, because it looks only at the quantities in the final step of the analysis chain that needs to be undergone in the seismic design of typical NPP equipment. Indeed, for the standard equipment item, the mechanical analysis is preceded – at least – by the analysis of the response of the floor on which the equipment is mounted.

The above statement is hence an incomplete estimate of the margin, because it neglects additional margins accumulated in the previous links of the analysis chain, such as the dynamic analysis of the civil structures.

Furthermore, it would not be a suitable basis for comparing the margins of different components relative to each other. For these reasons, it has proved useful to define the margin in terms of one characteristic ground motion parameter.

Expressing Margins in Terms of Ground Motion Parameters

Typically an SSE is defined by free-field ground motion response spectra. Seismic margin above the SSE is then most commonly quantified in terms of one parameter representative of the ground response spectra, the peak ground acceleration (PGA) being the most frequently used parameter. Alternative parameters are, e.g., the

peak spectral accelerations or the intensity on the Medvedev–Sponheuer–Karnik (MSK) scale. Frequently, the use of PGA as the representative parameter for expressing seismic margin and seismic fragility is criticized, on the grounds that it is a poor indicator of structural damage. Nevertheless, the PGA has gained the broadest acceptance as the characteristic margin parameter, probably due to its simple physical interpretation that can be grasped also by decision makers not familiar with more sophisticated notions of earthquake engineering or seismology (refer to IAEA 2003b, p. 6).

It should be noted that there are ways for compensating the limitation of PGA as a unique descriptor of a seismic event, either on the hazard definition – relying on notions such as the intensity or cumulative absolute velocity (CAV; refer to EPRI (2006), Campbell and Bozorgnia (2011)) in the hazard curve development – or on the structural response (→spectral shape factor in the fragility definition below).

Variability Due to Randomness and Uncertainty

The load to which an SSC is subjected during an earthquake of a certain PGA is clearly nondeterministic: indeed, there are an infinite number of possible ground motions showing the same PGA, each of which would induce different loads. This inherently aleatory characteristic of seismic ground motions is denoted as randomness.

Moreover, the majority of safety-related SSC is not directly subjected to ground motion but to vibrations induced through the plant's civil

Seismic Robustness Analysis of Nuclear Power Plants, Fig. 2 Nondeterministic character of load demand and capacity
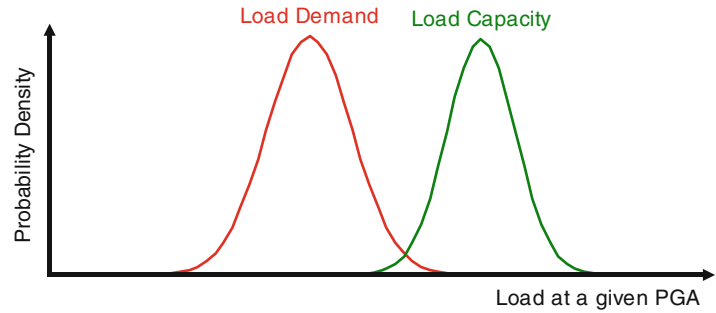
structures. Thus, there is nondeterminism – both inherent (randomness) and epistemic (uncertainty) – associated to effects like soil-structure interaction, damping, load combination, etc. Finally, also the load capacity is nondeterministic, considering, e.g., that material properties cannot be fully controlled.

The nondeterministic character of both load demand and seismic capacity is typically modeled by random variables, as illustrated in Fig. 2.

Seismic failure will occur when the load demand D exceeds the capacity C; the probability of failure P_F is then

$$P_F = P(D > C)$$

In view of the clearly nondeterministic nature of load demand and capacity, a conceptually sound definition of margin must be formulated in a probabilistic framework. A concept that has found widespread acceptance for this purpose is the definition of a HCLPF (pronounced as "hay-clipf") capacity. It is an implicit definition: if a component has a HCLPF capacity of x, then the probability of failing due to a demand of value x is low (5 %), with a high confidence (95 %). For the reasons stated in the previous section, the parameter in terms of which the HCLPF capacity is typically defined is the PGA.

Having captured the nondeterministic nature of the quantities governing failure, by means of the HCLPF capacity, it is now straightforward to quantify margin, e.g., as the distance between the HCLPF capacity and the load demand prescribed in the design of the NPP (i.e., the SSE).

More details on the HCLPF concept, including the approaches for its quantification in practical problems, are presented in section "Sources of Seismic Margin" below.

Sources of Seismic Margin

In principle seismic margin could be introduced on a deterministic basis specifying design basis loads exceeding those determined for the SSE or introducing penalizing stress utilization criteria. However, an overly conservative approach would induce significant extra cost in the entire life cycle of the plant (equipment design and procurement, civil construction, maintenance of heavy equipment items, decommissioning, etc.) and possibly detrimental effects on SSC performance in normal plant operation. In practice there are already various conservatisms in the seismic design of SSC. These conservatisms can be assessed and quantified when performing an SMA or an SPSA to show that adequate margin is available. Seismic margin especially results from the following sources:

– *Definition of the safe shutdown earthquake:*
  SSE ground motion response spectra are enveloping design spectra, i.e., the spectra have usually been subject of smoothing and broadening. In particular this introduces a substantial artificial increase of energy content of the excitation.
– *Conservatisms in the determination of the seismic demand:*
  Earthquakes induce oscillations of a chain of different oscillators at a site: starting from

Fig. 3 Earthquakes and their impact on civil structures and equipment (Translated from Sadegh-Azar and Hartmann 2011)
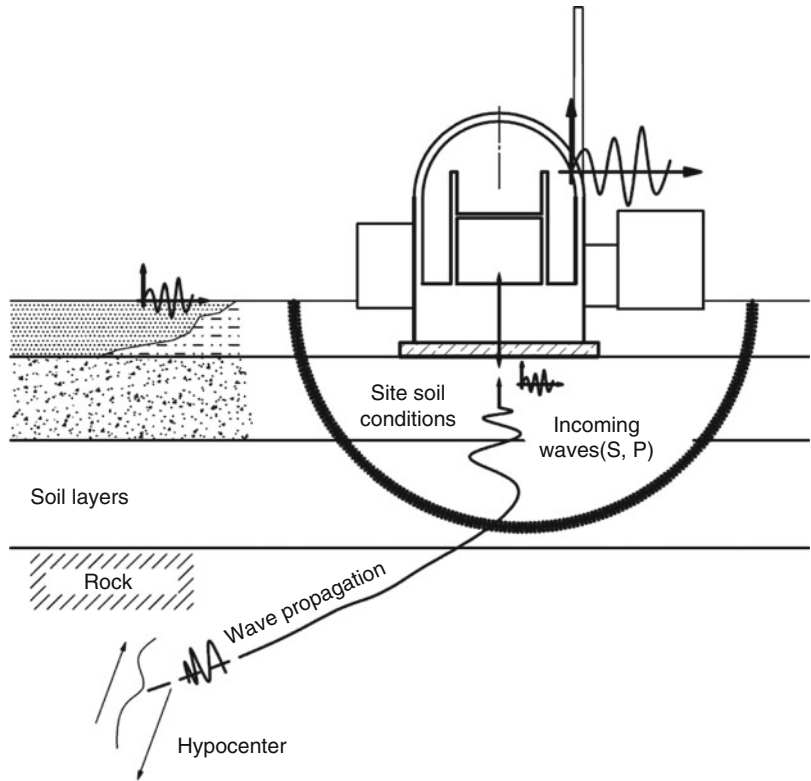
the soil, the base plate, and the upper floors of buildings up to smaller pieces of equipment installed in instrument racks or connected to flexible piping (Fig. 3). These different oscillating subsystems show more or less significant interaction effects; for example, the heavy weight of the reactor building has a damping effect on ground oscillations. Similarly heavy equipment installed inside the containment has a damping effect on the vibrations of the supporting floors. A realistic description of this complex oscillation behavior would therefore require modeling the entire system and considering also nonlinear effects such as plastic deformation.

Since such a global analysis is usually not practicable, the overall system is normally split into several subsystems analyzed separately. The following example illustrates this approach:

Subsystem	Excitation	Response
Ground + buildings	Ground response spectrum and corresponding time histories	Floor response spectra (secondary spectra) and corresponding time histories
Large equipment (e.g., primary circuit), civil substructures	Floor response spectra (secondary spectra) and corresponding time histories	Tertiary spectra and corresponding time histories
Equipment (e.g., pipes)	Tertiary spectra and corresponding time histories	Quaternary spectra and corresponding time histories
Built-in components (e.g., valve)	Quaternary spectra and corresponding time histories	Loads

Such an approach induces several sources of conservatism:

– Excitation characteristics and physical parameters considered for the different models are chosen conservatively to ensure that the conclusions made are valid without high uncertainty. The resulting level of conservatism may be moderate considering one model only. However, combining several models in a calculation chain, these conservatisms are multiplied and result typically in high factors of safety.

– Especially with respect to input characteristics, bounding assumptions are typically made. For example, many relevant codes require smoothing, broadening, or even increasing of input spectra. Similarly, the combination of load cases which can physically not occur simultaneously is often applied for the sake of simplification. Finally, the orthogonal vibration components are typically combined in a conservative manner.

– Interaction effects and in particular damping effects between the subsystems are neglected. The resulting conservatisms are especially important if the substructure is excited with a frequency near its resonance frequency, i.e., when the resulting vibrations are most important.

– Realistic anchorage characteristics allowing small displacements - significantly influencing the vibration behavior of the anchored system (shifting of natural frequencies) and thus avoiding resonant situations - are neglected.

– Other nonlinear effects such as plastic deformation and friction are typically also not modeled, and the corresponding energy dissipation is neglected or considered with simplified approaches only.

– *Conservatisms in the determination of the resistance to seismic loads:*

These conservatisms result especially from:

– Application of safety factors following the relevant codes.

– Use of conservative material properties (typically 95 % fractiles are used). In case of civil structures, the time-dependent concrete strengthening is normally also neglected.

– Neglecting plastic deformation capability.

– Finally, seismic margin results from stress utilization < 1, i.e., if the resulting stresses are smaller than the code allowable stresses. This is particular true when seismic loads are not the governing loads for an SSC.

Quantification of Seismic Margin: Fragility Curves and HCLPF Values

Capacity and Fragility

In order to quantify the abovementioned sources of conservatism and derive an expression of the overall margin, the concept of seismic fragility as introduced by Kennedy and Ravindra (1984) has proven useful.

Seismic fragility is defined as a conditional failure probability, where the condition is represented by a seismic event with a given value of the characteristic ground motion parameter, denoted by a. In the sequel it will be assumed that this ground motion parameter is given by the peak ground acceleration, i.e., the zero period acceleration in the ground response spectrum. In this case the fragility is defined as follows:

$$Fr(a) = P[\text{failure} \,|\, PGA = a]$$

The definition of fragility is flexible in the sense that the meaning of the term *failure* is arbitrary. In the context of seismic PSA, it relates mainly to components of safety-relevant systems or to structures in which such components are installed; the entire set of these components and structures is abbreviated as SSC (systems, structures, and components). Quantitatively, failure

occurs if one of the parameters representing the loads experienced by the analyzed SSC exceeds its limit state (e.g., the tensile stress in a reinforcement bar exceeds the tensile strength of the utilized steel). The definition of the limit states for an SSC requires the identification of the relevant failure modes.

It is useful to express the fragility in terms of a random variable representing the capacity, denoted by A, and defined in terms of the same ground motion parameter used for the fragility definition. Obviously, failure occurs if the capacity is lower than the peak ground acceleration of the assumed seismic event. The definition of the fragility can then be extended as follows:

$$Fr(a) = P[\text{failure} \,|\, PGA = a] = P[A < a]$$
$$= F_A(a)$$

The right portion of the above equation emphasizes that the fragility is equivalent to the cumulative density function of the capacity A. This implies that defining the random capacity A is equivalent to defining the fragility.

Log-Normal Capacity Model

The most widely used model for the seismic capacity is given by

$$A = \breve{A} \, \varepsilon_R \varepsilon_U$$

where $\breve{A}$ is the median of the capacity, while ε_R and ε_U are log-normally distributed with unit median and logarithmic standard deviations of β_R and β_U, respectively. The random variables ε_R and ε_U model the variability due to randomness and due to uncertainty, respectively. In order to preserve the distinction between these two different sources of variability, the variables ε_R and ε_U are introduced in two distinct steps of the analysis. To begin with, it is assumed that only the variability to randomness exists, whereas there is no variability due to uncertainty, i.e., $\varepsilon_U = 1$.

At this stage, the fragility is then (due to the properties of the log-normal distribution of ε_R)

$$Fr(a) = \Phi \left(\frac{\ln \left(\frac{a}{\breve{A}} \right)}{\beta_R} \right)$$

where Φ is the standard normal cumulative distribution function.

In the second step of the fragility analysis, the assumption that $\varepsilon_U = 1$ is relaxed. Instead, ε_U is assumed to be log-normally distributed with logarithmic standard deviation β_U.

Rewriting the capacity as

$$A = \breve{A} \, \varepsilon_R \varepsilon_U = \left(\breve{A} \varepsilon_U \right) \varepsilon_R$$

and replacing $\breve{A}$ with $\breve{A} \varepsilon_U$ in the cumulative probability density leads to

$$Fr(a) = \Phi \left(\frac{\ln \left(\frac{a}{\breve{A} \varepsilon_U} \right)}{\beta_R} \right)$$

Due to the presence of ε_U, the fragility at a given PGA a itself turns also into a random variable. It can be shown that for a given probability level Q, the fractile of this random variable is

$$Fr(a, Q) = \Phi \left(\frac{\ln \left(\frac{a}{\breve{A}} \right) + \beta_U \Phi^{-1}(Q)}{\beta_R} \right)$$

Interpreting Q as a subjective probability, or equivalently as a confidence level, one may use the above equation to obtain the fragility curves associated with specific confidence levels. This is demonstrated in Fig. 4 for the fragility model with the parameters $\breve{A} = 0.5\,\text{g}$, $\beta_R = 0.3$, and $\beta_U = 0.4$.

The median (50 % confidence level) fragility curve is identical to the one obtained by neglecting uncertainty (i.e., $\beta_U = 0$).

Seismic Robustness Analysis of Nuclear Power Plants, Fig. 4 Example of fragility curve in linear scale
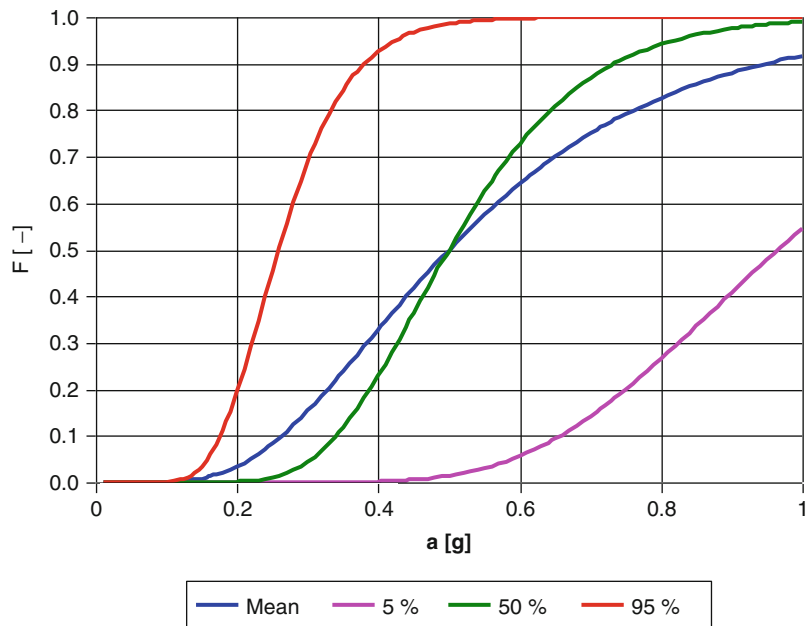
The mean fragility curve is obtained by averaging over the random fragility $F_A(a)$ for each value of the ground motion parameter a, i.e., the averaging is done along the vertical direction.

It can be shown that the mean fragility has the following simple form

$$F_A(a) = \Phi\left(\frac{\ln\left(\dfrac{a}{\tilde{A}}\right)}{\beta_C}\right) \qquad \beta_C = \sqrt{\beta_R{}^2 + \beta_U{}^2}$$

The above model is called the **composite fragility model**, because it considers random variability and variability due to uncertainty in a compact form. It is used for point estimates (best estimate) of the seismic risk.

Methods for Deriving Fragility Curves of SSC

For the quantification of seismic fragility, one can generally distinguish between approaches relying on the assessment of conservatisms in **existing** seismic verification studies and approaches requiring **new** dynamic analyses.

For cost reasons, where possible, preference is given to the former. Within this category, the most widely used approach for deriving fragility curves is based on **separation of variables**, as described in detail in Kennedy and Ravindra (1984) and in EPRI (1994).

The basic concept underlying this approach is the following expression for the capacity A as a product of the (deterministic) value of the PGA, a_{SSE}, adopted in the seismic design of the SSC under consideration, and the scaling factor F:

$$A = a_{SSE} \cdot F$$

The scaling factor F may hence be viewed as the maximum scalar, by which the design ground motion can be multiplied ("scaled") without producing failure. Alternatively, the factor F may however also be viewed as a safety factor, considering that it is the ratio between the actual capacity A and the acceleration used in the design, a_{SSE}:

$$F = \frac{A}{a_{SSE}}$$

It should be noted that the scaling (safety) factor F is a random variable, since the capacity A is also a random variable.

The essence of the separation of variables approach is to break down the scaling (safety) factor F into a product of "partial" safety factors. This factorization process extends over two levels, since some of these partial safety factors are yet again factorized.

For structures, the factor of safety F can be modeled as the product of three random variables:

$$F = F_S F_\mu F_{RS}$$

where F_S is the strength factor, F_μ is the inelastic energy absorption factor, and F_{RS} is the structural response factor.

The strength factor, F_S, represents the ratio of ultimate strength (or strength at loss of function) to the stress calculated for a_{SSE}. In calculating the value of F_S, the non-seismic portion of the total load acting on the structure is subtracted from the strength as follows:

$$F_S = \frac{S - P_N}{P_T - P_N}$$

where S is the strength of the structural element for the specific failure mode, P_N is the normal operating load (i.e., dead load, operating temperature load, etc.), and P_T is the total load on the structure (i.e., sum of the seismic load – due to a_{SSE} – and the normal operating load).

The inelastic energy absorption factor (ductility factor), F_μ, accounts for the fact that an earthquake represents a limited energy source, and many structures or equipment items are capable of absorbing substantial amounts of energy beyond yield without loss of function.

The structural response factor, F_{RS}, is based on the recognition of the fact that in the design analyses, structural response is computed using specific (often conservative) deterministic response parameters for the structure. Since many of these parameters are nondeterministic (often with wide variability), the actual response may differ substantially from the calculated response for a given peak ground acceleration. F_{RS} is modeled as a product of factors influencing the response bias and variability:

$$F_{RS} = F_{SA} F_{GMI} F_\delta F_M F_{MC} F_{EC} F_{SSI}$$

where:

F_{SA} = spectral shape factor representing bias and variability in ground motion and associated ground response spectra

F_{GMI} = ground motion incoherence factor that accounts for the fact that a traveling seismic wave does not excite a large foundation uniformly

F_δ = damping factor representing bias and variability in response due to the difference between actual damping and damping assumed in design

F_M = modeling factor accounting for limitations regarding the resolution and calibration of the model adopted for representing the structural mechanics

F_{MC} = mode combination factor accounting for bias and variability in the estimated response due to the method used in combining dynamic modes of response

F_{EC} = earthquake component combination factor accounting for bias and variability in response due to the method used in combining earthquake components

F_{SSI} = factor that accounts for the effect of soil-structure interaction

Similarly, for equipment and other components, the overall factor of safety is composed of a capacity factor, F_C; a structure response factor, F_{RS}; and an equipment response (relative to the structure) factor, F_{RE}. Thus,

$$F = F_C F_{RE} F_{RS}$$

The individual safety factors F_i are assumed to follow a log-normal distribution. Their characteristic parameters are their median values $\bar{F}_i$ – representing the (often conservative) **bias** – and the logarithmic standard deviations $\beta_{i,R}$ and $\beta_{i,U}$, representing the **variability** due to aleatory and epistemic nondeterminism, respectively. The most widely used method for estimating these characteristic parameters is the approximate second moment procedure.

The overall median safety factor $\breve{F}$ may be found by scaling the ground response spectrum so long, until the limit state corresponding to the analyzed failure mode is reached. All input variables must be set to their median values in this process for the scaling factor to correspond to the median safety factor.

Alternatively, the sources of conservatism of the design calculations – compared to median-centered calculations – may be identified individually for each partial safety factor. For each of the response factors, the response quantity of interest (e.g., stress) is calculated (or estimated) anew, but only with the corresponding input variables set to their median values. The ratio of this newly computed response with the one from the design calculations gives the corresponding median partial safety factor. For the median capacity factor, the margin between the median strength and the stress obtained in the design calculation is quantified. The overall median safety factor is then simply given by the product of the individual safety factors.

In order to estimate the logarithmic standard deviations $\beta_{i,R}$ and $\beta_{i,U}$, it is necessary to perform an additional calculation of the safety factor, in which all model variables are median centered, with the exception of the input variables corresponding to the analyzed partial safety factor. The latter ones are instead perturbed by a multiple κ of their standard deviation. The input variables should be perturbed to the side which leads to a smaller safety factor (since we are more interested in quantifying the effect of variables in the more unfavorable cases). The resulting partial safety factor is denoted here as $F_{\kappa\sigma_i}$.

The variability parameter (either randomness or uncertainty) of the analyzed safety factor is then

$$\beta_i = \frac{1}{|\kappa|}\ln\left|\frac{F_{\kappa\sigma_i}}{\breve{F}}\right|$$

For several safety factors it is common practice to use generic variability parameters, as suggested in EPRI (1994).

Finally, the logarithmic standard deviations of the individual safety factors are combined using square root of sum of squares (SRSS), in order to obtain the corresponding variability parameter of the global safety factor, β_R and β_U, respectively.

Alternative, more refined methods can be used for deriving fragility curves, if necessary. For instance, the simplifying assumptions adopted in the separation of variables method can be relaxed by applying Monte Carlo simulation to the structural dynamic analysis. In this method, all parameters assumed to contribute significant nondeterminism to the overall capacity are sampled simultaneously, and the model used for structural dynamic analysis is perturbed correspondingly. Repeating this exercise for several levels of the PGA leads to a discrete set of points of the fragility curve, which can be interpolated or extrapolated as needed.

The HCLPF Capacity

As already indicated in earlier sections, the HCLPF value or HCLPF capacity is a crucial concept in the context of seismic robustness of NPPs. It is defined – implicitly – as the value of the PGA for which there is a *h*igh *c*onfidence (95 %) that the *p*robability of *f*ailure does not exceed 5 % (i.e., it is *l*ow).

Based on the above definition, the HCLPF capacity can be related to the median capacity and the variability parameters as follows:

$$HCLPF = A_m\exp(-1.65(\beta_R + \beta_U))$$

where -1.65 is the 5 % fractile of the standard normal distribution.

The HCLPF capacity can be approximated by the 1 % fractile of the composite mean fragility curve. This definition of the HCLPF in the composite model is conservative, because it is guaranteed to be equal or smaller than the HCLPFs of all corresponding non-composite fragility models, i.e., those in which β_R and β_U are kept separately (for a specific β_C, there are infinitely many possible combinations of β_R and β_U).

Seismic Robustness Analysis of Nuclear Power Plants, Fig. 5 Example of fragility curve in logarithmic scale

Mean — 5 % — 50 % — 95 %
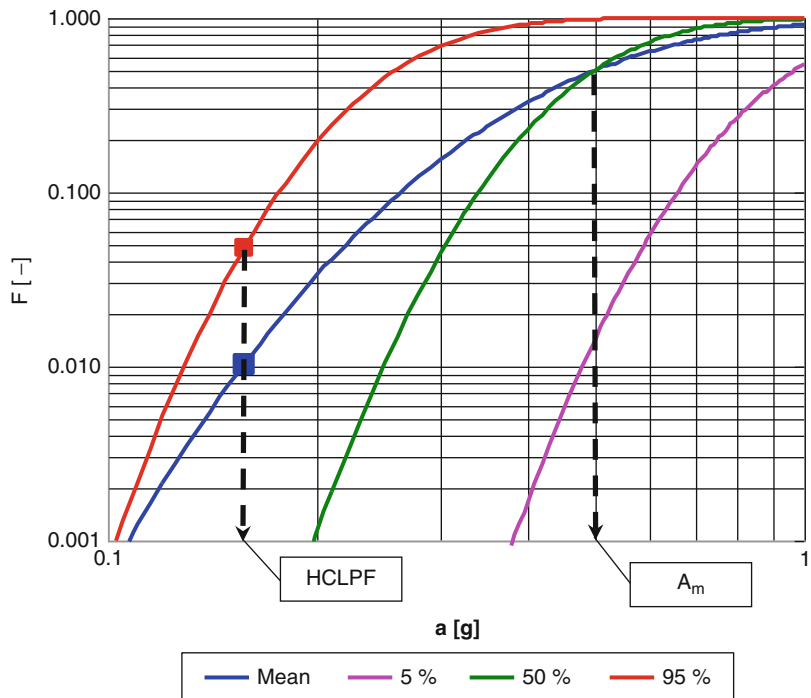
The HCLPF capacity has then the following simple form:

$$HCLPF = A_m \exp(-2.33\beta_C)$$

where -2.33 is the 1 % fractile of the standard normal distribution.

The above-described relationship of the HCLPF value with the fragility curve is exemplified in Fig. 5 for an SSC with a median capacity of 0.5 g and variability parameters β_R and β_U equal to 0.3 and 0.4, respectively.

Simplified HCLPF Estimation Using the CDFM Method

The previous section described the formal definition of the HCLPF capacity based on the corresponding fragility curve. It is clear that once a full-scope fragility analysis has been performed, the HCLPF capacity can also be easily obtained.

Within SMA it is, however, not necessary to have a full description of the fragility curve; the HCLPF capacity is sufficient. As a simplified alternative to estimate the HCLPF value,

a semi-probabilistic approach has been proposed. It is referred to as "conservative deterministic failure margin" (CDFM) method; compare EPRI (1991). The CDFM method can be briefly summarized by the following four steps:

1. For the targeted HCLPF capacity (e.g., the Review Level Earthquake in an SMA), the floor responses (demand) are computed using the 84 % fractile (i.e., an 84 % probability of non-exceedance) of the ground response spectrum. Damping is to be assumed conservatively.
2. The strength parameters are evaluated at the 2 % fractile, i.e., the probability of exceedance of the selected strength is at the 98 % probability level.
3. The inelastic energy absorption factor, F_μ, should be estimated conservatively, i.e., at a 84 % probability of exceedance level.
4. If, based on the above, the following inequality is satisfied, then the targeted HCLPF capacity has been demonstrated:

$$\text{Demand/Capacity} \leq F_\mu$$

Analysis of Seismic Robustness at Plant Level

Seismic Margin Assessment (SMA)

Goals and Approaches

The main goal of an SMA is to identify the seismic capacity of the plant, i.e., the maximum level of seismic ground motion for which the plant can still be safely shut down. In addition, it is the goal of any SMA to identify the weaker components, i.e., the components limiting the seismic capacity of the plant.

There are essentially two approaches that have gained widespread acceptance and that will be described in more detail in the sequel: a success-path-based approach ("EPRI method") and a fault-tree-based approach ("NRC method").

Stated in simple terms, the difference of the approaches is the following:

The success-path approach identifies the components **necessary** to **avoid** a core damage; the plant capacity is that of the **weakest** of these components.

The fault-tree approach identifies the combinations ("minimum cutsets," abbreviated MCS) of component failures **sufficient** to **cause** a core damage. For each of these MCS, the capacity is that of the **strongest** of these components. The plant capacity is then equal to the minimum capacity of any of the MCS capacities.

Having stated the differences between the two approaches, this introductory section deserves to be concluded by noting a fundamental, unifying feature of the two approaches, namely, that they are essentially **deterministic**. This property becomes particularly obvious when comparing the plant-level HCLPF capacity with a **target** value of the PGA. For instance the US Nuclear Regulatory Commission requires 1.67 times the PGA of the design basis earthquake; refer to NUREG (1993). In the European Utility Requirements, a minimum capacity of 1.4 times the PGA of the SSE is recommended; refer to EUR organisation (2001). The result of this comparison is a **binary and hence deterministic expression**, indicating either sufficient margin or lack thereof, but no intermediate answers.

Success-Path-Based Approach

This approach is also referred to as the EPRI (Electric Power Research Institute)-approach, because it is best described in EPRI NP-6041.

The basic concept underlying this approach is that of a success path, indicating the transition from a disturbance to the normal plant operation – caused by a design-exceeding earthquake – to a safe state, in which the fundamental safety functions (reactivity control, residual heat removal, and radioactivity confinement) are durably ensured.

An example of a success path is shown for a pressurized water reactor (PWR) in Fig. 6.

It is important to note that a reasonable success-path definition should take into account the seismic design of the plants' safety systems. The preferred success paths should be those along which the safety functions are fulfilled by seismically designed systems.

The success-path definition should make sure that the most likely scenarios of plant disturbance after a design-exceeding seismic event are covered. At least two success paths – leading to a durable safe state of the plant – need to be defined, and one of the two success paths should include the mitigation of a small break loss of coolant accident (LOCA), unless even a small break LOCA can be excluded, e.g., due to a particularly high quality of the supports of small-bore primary piping such as measurement lines.

The above mentioned Fig. 6 describes the success path at the level of the safety functions. The goal is to ensure that all the safety functions along the success path have a sufficient margin over the SSE. This requirement obviously needs to be cascaded down to the components making up the systems fulfilling the safety functions. For example, in order to ensure steam generator (SG) feed following the earthquake, it is necessary for the emergency feedwater pump to remain operational.

The requirement of sufficient margin does not only apply to the so-called "frontline" systems, i.e., the fluid systems specifically performing the respective safety function, but also to the so-called support systems, such as electrical power supply, heating/ventilation/air-conditioning (HVAC), and instrumentation and control (I&C).
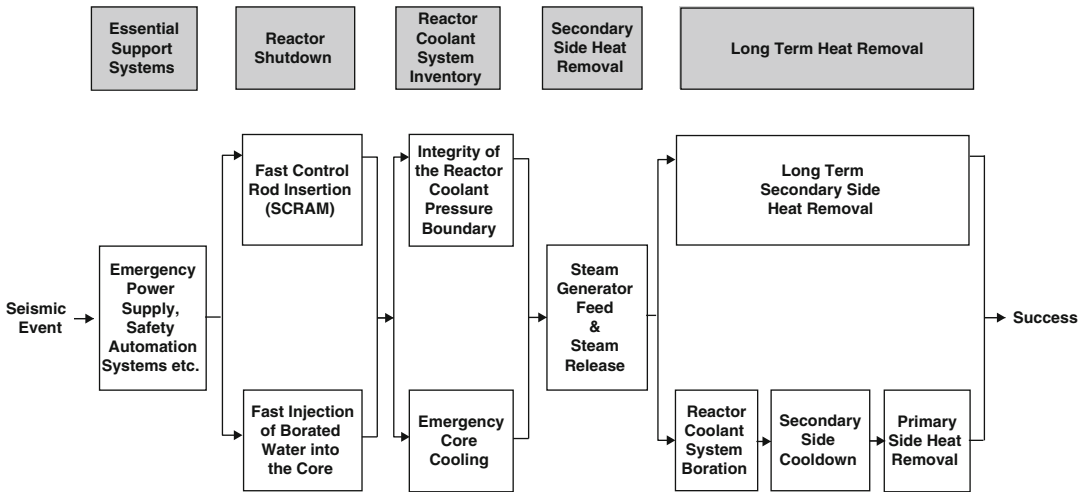
Seismic Robustness Analysis of Nuclear Power Plants, Fig. 6 Example of a success path to a safe shutdown state for a pressurized water reactor

The result of cascading the requirement for sufficient margin to the individual components consists in a list of components which are needed to fulfill the safety functions credited in the success-path definition. This list is referred to as Seismic Equipment List (SEL) or also Safe Shutdown Equipment List (SSEL).

The metric used to measure the margin is the HCLPF value introduced in earlier chapters. Since the success-path-based approach is meant to be a predominantly deterministic approach, the CDFM method is preferred.

Considering that – as a rough estimate – the size of the SEL is in the order of 10^3 components, it should become clear that it is not feasible to perform dedicated HLCLF calculations for all components of the SEL. Instead, a walkdown-based **screening process** is to be implemented. Since this screening process is common to all approaches for seismic robustness and risk assessment, it is discussed in the final chapter of this essay on practical issues.

A critical step preceding the screening process is the definition of the so-called Review Level Earthquake (RLE). It consists in a site-specific characterization of a seismic ground motion (typically a response spectrum) that is more severe than the ground motion considered in the design of the plant (i.e., the SSE).

The result of the screening process is a, typically large, set of components which are judged – based on a structured and standardized walkdown procedure – to have a generic HCLPF capacity **equal or larger** than the RLE. For the remaining components, individual HCLPF values need to be derived.

The fact that the walkdown-based screening provides only a lower bound of the *generic* HCLPF capacity, instead of a sharp value, implies that the RLE should be set high enough, so that some components for which a non-generic HCLPF is evaluated are below the RLE. Only in this way, a sharp plant HCLPF can be determined, whereas otherwise one obtains only a lower bound, because the generic HCLPFs can only be stated as being equal or higher than the RLE.

For those components for which an individual HCLPF is to be evaluated, the responses due to the RLE are needed. These may be either obtained by performing new floor response spectra and equipment response calculations based on the RLE or by scaling existing responses to the ground motion of the SSE.

Once both generic and individual HCLPF capacities have been evaluated, the overall capacity of the plant results as the lowest HCLPF of the considered equipment. This implies that both success paths must be sufficiently robust.

As a concluding remark, the deterministic nature of the success-path-based approach is emphasized, both at the individual component level (CDFM used for HCLPF estimation) and at the system modeling level (no systematic generation of minimum cutsets; see the following section).

Fault-Tree-Based Approach

This approach is also referred to as the NRC approach, because it has been proposed and developed for the NRC in NUREG (1985, 1986).

In the fault-tree-based approach, several aspects are unchanged with respect to the success-path-based approach, such as the walkdown-based screening process. The main difference compared to the success-path-based approach is that a fault-tree-based system analysis is relied upon in order to determine **minimum cutsets**, i.e., combinations of failures leading to core damage. The attribute "minimum" indicates that no individual failure event can be removed from the set without altering its property of leading to core damage. The **minimum** cutsets have the highest risk relevance, because their probability is higher than non-minimum cutsets.

The components screened out on the basis of the walkdown observations are not included in the system modeling, i.e., only components which cannot be generically stated as having a HCLPF capacity equal or higher than the RLE are modeled in the fault trees. Individual HCLPF capacities of these components are to be estimated, either based on a fragility analysis or at least based on the CDFM method.

Performing a cutset analysis for the event trees judged as relevant for the SMA (typically at least a seismic-induced transient, such as loss of off-site power, and a small break LOCA) results in a list of minimum cutsets to be further processed as follows.

For each of the cutsets, a characteristic HCLPF capacity is obtained by taking the **maximum** HCLPF capacities of the component failure forming the cutset. From the resulting list of cutset HCLPFs, the **minimum** is then taken, in order to obtain the fault-tree-based plant HCLPF capacity. Fig. 7 is meant to illustrate the process.

Seismic Robustness Analysis of Nuclear Power Plants, Fig. 7 Min-Max operation in fault-tree-based SMA

Individual seismic-induced failure events are denoted by $SEIS_i$, $SEIS_j$, $SEIS_k$, etc.

If the fault-tree-based plant HCLPF capacity is higher than the RLE, then the plant HCLPF can only be judged as being at least equal to the RLE, since the components screened out on the basis of the walkdown observations were not included in the system modeling.

If a PSA model is available, then the effort for performing the system modeling is greatly reduced, since the seismic-induced failures of the non-screened components merely need to be added to the existing fault trees. A major benefit of reusing an existing PSA model is also the possibility to analyze mixed cutsets, i.e., combinations of seismic-induced and non-seismic-induced failures, such as random component failures or human (operator) errors.

As a final remark it is emphasized that – even though the probabilities of the cutsets are not used in the quantification of the plant HCLPF capacity – the fault-tree-based approach to SMA resembles a **probabilistic** approach, because cutsets are generated, i.e., a systematic exploration of the combinations of individual component failures is performed. Nevertheless, the subsequent steps of the procedure and – in particular – the final result are **deterministic**, since a single number (the HCLPF capacity) is used to characterize the capacity of each individual SSC.

Seismic PSA

Basic Notions

While sharing numerous aspects with SMA, seismic probabilistic safety analysis (PSA) is distinct from its deterministic counterpart (SMA) in several respects. The most prominent difference is – of course – due to the attribute "probabilistic," which implies nondeterminacy for all the possible states of the system (i.e., the model of the NPP). Indeed, for a seismic event of a given PGA, a component may fail or may not fail, and both possibilities are assigned a **probability**. This probability is called fragility and has been introduced at length in section "Quantification of Seismic Margin: Fragility Curves and HCLPF Values." The same property of nondeterminacy applies to the event "seismic-induced core damage," which may or may not occur in case of a seismic event of a given PGA.

The event "seismic-induced core damage" is a central notion of both seismic PSA and SMA and hence a unifying aspect. Clearly, since both methodologies (seismic PSA and SMA) aim at analyzing the plant-level resistance to seismic events, one is – in general – not concerned with individual failures, but only with combinations of failure leading to core damage.

However, there is a significant difference: in the seismic PSA, it is assumed that the event "seismic-induced core damage" can – at least theoretically – occur due to seismic events of **any** intensity (and hence PGA).

This is where the concept of seismic hazard at the NPP site comes into the picture: it is the characterization of the seismic ground motions "to be expected" at the site (see section "Motivation for Evaluating Seismic Robustness"). For the purpose of seismic PSA, the most important characterization is afforded by the so-called hazard curve, in which the probability of observing a seismic event exceeding a PGA equal to a during a predefined time interval (typically one year) is expressed as a function of a. Thus, the hazard curve is a monotonously decreasing function $H(a)$.

In Fig. 8, a set of hazard curves is shown, each curve being associated with a probability level (50 %, 5 %, and 95 %). This characterization of the site hazard in terms of a population of hazard curves, rather than a single hazard curve, reflects the uncertainties in the hazard estimation procedures.

Recall that for each level a of the PGA, core damage may or may not occur and denote the corresponding probability by $F_A^{CDF}(a)$. The probability that a seismic event leading to core damage occurs during the time interval to which the hazard curve refers (typically one year) can be shown to be:

$$CDF = \int_{a=0}^{\infty} F_A^{CDF}(a) \frac{dH}{da} da$$

The acronym CDF denotes core damage frequency, which is – in quantitative terms – almost
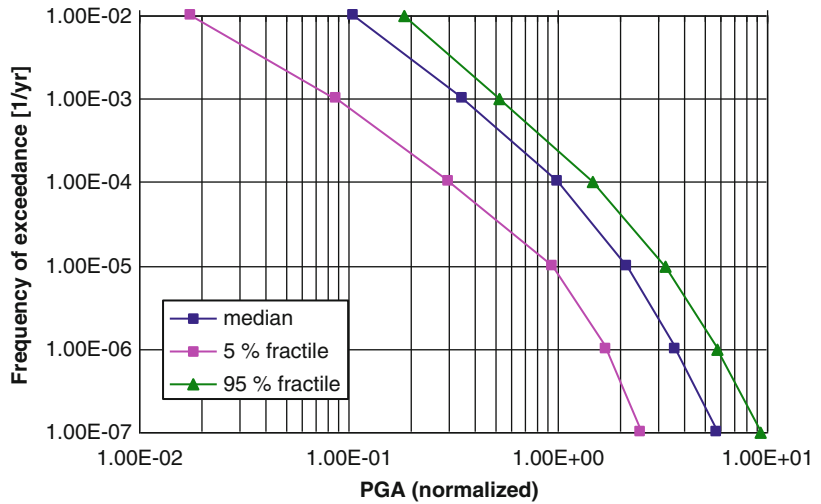
identical to the annual probability of core damage.

In the practical implementation of seismic PSA, the evaluation of the CDF is however more involved. The reason is that the conditional probability of core damage $F_A^{CDF}(a)$ cannot by obtained in a closed form, because of the enormous number of component failure combinations potentially leading to core damage.

This is where the specific application, seismic PSA, becomes reminiscent of its more general "parent," namely, level 1 PSA for internal events. In level 1 PSA risk-relevant accident progression sequences are modeled by event trees, and the potential combinations of component failures leading to unavailability of safety systems called upon during the accident progression sequences are modeled by fault trees.

These event trees and fault trees represent a (very complex) Boolean expression of the event "core damage."

A well-established procedure for evaluating the probability of this event is to express the (complex) Boolean expression as a union of (much simpler) intersections of component failures; these intersections are called "minimum cutsets" (MCS).

For each minimum cutset, MCS_i, the contribution to the core damage frequency is

$$CDF^{(i)} = \int_{a=0}^{\infty} F_A^{(i)}(a)\frac{dH}{da}da$$

where $F_A^{(i)}(a)$ is the probability of the MCS_i, conditional on a given PGA a, and $H(a)$ is the frequency of occurrence of a seismic event with a PGA equal or higher than a.

The probability of the MCS_i is

$$F_A^{(i)}(a) = \prod_{j=1}^{n^{(i)}} f_j^{(i)}(a) \prod_{k=1}^{m^{(i)}} p_k^{(i)}$$

where $f_j^{(i)}(a)$ is the jth fragility appearing in MCS_i and $p_k^{(i)}$ is the failure probability of the kth non-seismic basic event in MCS_i. The integers $n^{(i)}$ and $m^{(i)}$ are the number of fragilities and the number of non-seismic basic events in MCS_i.

In practice the above integral is approximated by

$$CDF^{(i)} \approx \sum_{l=1}^{n} f\left(a_{l-\frac{1}{2}}\right)(-\Delta H_l)$$

where a_0 and a_n are the limits of the risk-relevant PGA range and the remaining a_i's are selected at convenient points, e.g., the evaluations of the inverse hazard function at the values

10^{-1}, 10^{-2}, etc. The corresponding mid-interval fragilities $f\left(a_{l-\frac{1}{2}}\right)$ are logarithmically interpolated,

$$f\left(a_{l-\frac{1}{2}}\right) = 10^{(\log_{10} f(a_{l-1}) + \log_{10} f(a_l))/2}$$

The probability of occurrence associated with each interval is $\Delta H_l = H(a_l) - H(a_{l-1})$. The total CDF is then obtained by evaluating the union of the individual MCS.

Correlation Between Seismic Failures

The issue of mutual dependence (or alternatively, correlation) between seismic-induced failures relates to the question whether the seismic-induced failure of a component, e.g., an emergency diesel generator (EDG), is more probable, under the condition that another component (e.g., the EDG of another safety train) has experienced seismic-induced failure. Several studies indicate that in seismic PSA, the effect of correlations on the seismic-induced core damage frequency (CDF) may be very significant; refer to Pellissetti and Klapp (2011) and to references 1 through 4 therein.

The most straightforward and – in general – most conservative approach consists in assuming full correlation between identical and "seismically similar" components of different, redundant safety trains, i.e., components connected by "AND"-logic. Therefore, with this model the redundancy represented by the different trains of the plant is not credited, as far as seismic-induced failures are concerned. Furthermore, for components of a single train combined under the same "OR" gate in the fault tree, the failures are assumed to be independent.

In this approach failures are thus assumed to be either fully correlated or totally uncorrelated; it is hence a binary model (black/white, zero/one) which cannot accommodate gradually varying levels of correlation. In fact, for redundant components connected by "AND"-logic, this model negates the existence of distinct failure events for the individual redundant trains and will hence also be referred to as "deterministic correlation model." This binary, deterministic model is conservative with respect to the probability of the top event, because correlation is detrimental (i.e., the failure probability increases) for intersection of events and beneficial for unions of events (Fig. 9).

An alternative approach, described in NUREG (1990), consists in modeling the dependencies between seismic failures in terms of the correlation coefficients between the variables characterizing the seismic demand and capacity of the SSCs. The adopted model makes full use of the possible (positive) range of correlation coefficients (between 0 and 1).

The separation of the correlation coefficients of demand and capacity reflects the difference in their respective cause: the correlation of the demand variables is due to "similarity" of the excitation experienced by distinct components (e.g., due to spatial proximity), whereas the correlation of the capacity variables is due to "similarity" in terms of the physical properties of the components (e.g., due to the fact that they have been manufactured by the same supplier). The word "similarity" is placed between quotes because it is to be understood in the statistical sense.

In this approach failure is defined in terms of the following limit-state function:

$$X = \ln\left(\frac{R}{S}\right)$$

where R and S represent the response and the strength for a given failure criterion. Failure takes place if

$$R > S \quad \Leftrightarrow \quad X > 0$$

It is convenient to model both R and S by log-normal random variables with the parameters μ_R, β_R and μ_S, β_S. According to the definition of log-normal variables, the respective companion normal variables are then $C_R = \ln R$ (with mean μ_R and standard deviation β_R) and $C_S = \ln S$ (with mean μ_S and standard deviation β_S). In view of its definition, X is then normally distributed and its mean and standard deviation are
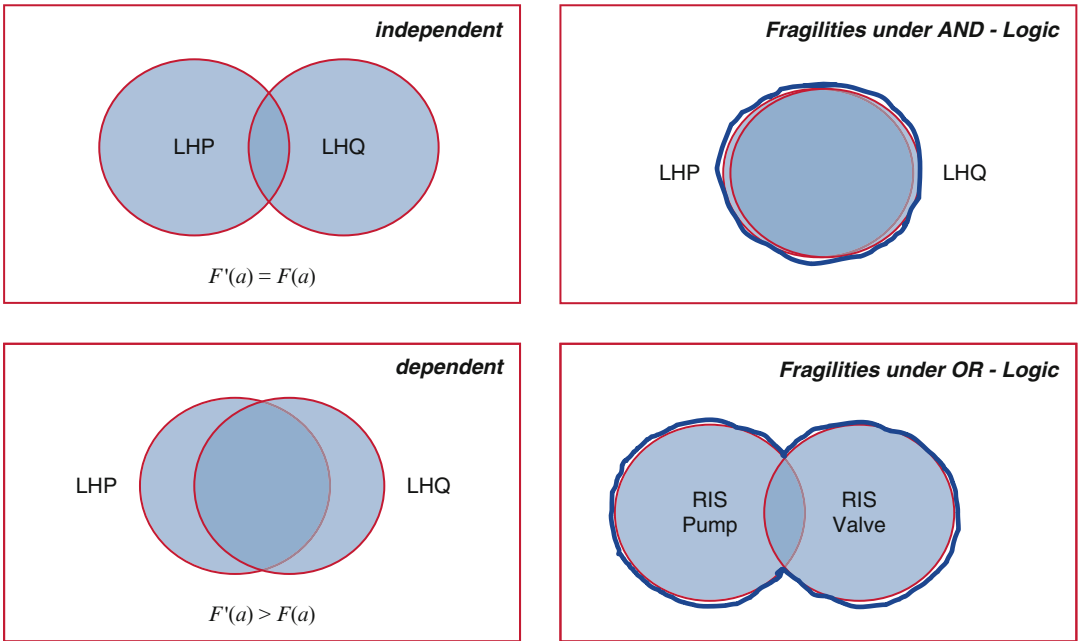
Seismic Robustness Analysis of Nuclear Power Plants, Fig. 9 Schematic representation (Venn diagrams) of dependence between individual seismic-induced failures

$$\mu_X = \mu_R - \mu_S, \quad \beta_X = \sqrt{\beta_R^2 + \beta_S^2}$$

The probability of failure is then given by

$$P[X > 0] = 1 - \Phi\left(\frac{-\mu_X}{\beta_X}\right) = \Phi\left(\frac{\mu_X}{\beta_X}\right)$$

where Φ is the cumulative probability density function of the standard normal distribution. Considering two individual components, one may then define the random variables R_1, S_1 and R_2, S_2, respectively, and X_1, X_2 according to the above definition of X. Assuming no cross-correlation between R_i and S_j, after some arithmetics, the correlation coefficient for X_1, X_2 may be expressed as

$$\rho_{X_1 X_2} = \frac{1}{\beta_{X_1}\beta_{X_2}}\left(\beta_{R_1}\beta_{R_2}\rho_{R_1 R_2} + \beta_{S_1}\beta_{S_2}\rho_{S_1 S_2}\right)$$

where $\beta_{X_i} = \sqrt{\beta_{R_i}^2 + \beta_{S_i}^2}$.

It should be noted that in the above expression, the correlation coefficients $\rho_{R_1 R_2}$ and $\rho_{S_1 S_2}$ relate to

the companion normal variables of R_i and S_j, i.e., ln (R_i) and ln (S_i). The advantage of calibrating the correlation model in terms of the companion normal variables is that the meaningful range of the correlation coefficient is always $[-1, 1]$. This is not the case for log-normal random variables. It should be noted that according to its definition, the correlation coefficient $\rho_{X_1 X_2}$ depends on both the correlation coefficients and the standard deviations; this is necessary to prevent a pair of variables with a high correlation but with very small standard deviation from introducing a high overall correlation. With the above definitions, the joint probability of failure of the two components $Q_{1\cap2}$ is defined in terms of the following two-dimensional integral:

$$Q_{1\cap2} = P[(X_1 > 0) \wedge (X_2 > 0)]$$
$$= \iint\limits_{0 \leq x_1, x_2 \leq \infty} f_{X_1 X_2}(x_1, x_2) dx_1 dx_2$$

In the above equation, $f_{X_1 X_2}$ is the bivariate normal probability density function (PDF):

$$f_{X_1 X_2}(x_1, x_2) = \frac{1}{2\pi \beta_{X_1} \beta_{X_2} \sqrt{1 - \rho_{X_1 X_2}{}^2}}$$

$$\cdot \exp\left(-\frac{1}{2(1 - \rho_{X_1 X_2}{}^2)}\left(\frac{(x_1 - \mu_{X_1})^2}{\beta_{X_1}{}^2} + \frac{(x_2 - \mu_{X_2})^2}{\beta_{X_2}{}^2} - 2\rho_{X_1 X_2}\frac{(x_1 - \mu_{X_1})(x_2 - \mu_{X_2})}{\beta_{X_1}\beta_{X_2}}\right)\right)$$

Accordingly, the probability of simultaneous failure of n components has the form

$$Q_{1 \cap \dots \cap n} = P[(X_1 > 0) \wedge \dots \wedge (X_n > 0)]$$
$$= \int_{0 \le x_1, \dots, x_n \le \infty} \dots \int f_X(\boldsymbol{x}) d\boldsymbol{x}$$

The application of the above correlation model in the context of fault-tree-based seismic PSA – using general purpose PSA software – is exemplified in Pellissetti and Klapp (2011).

Operator Actions

Earthquakes not only may impact the reliability of safety systems but also influence operator performance:

- The workload for the shift team may be higher due to a failure of operational systems.
- Spurious signals, e.g., caused by relay chatter, may make appropriate decision-making more difficult.
- Accessibility of buildings may be degraded (e.g., due to a failure of lighting or automatic locking systems) and may hence lead to smaller grace times for the necessary actions.
- Finally, earthquakes may have a direct impact on human behavior (shocks).

Past seismic PSA experiences have shown that the choice of probabilities for seismic-induced operator actions can have a significant impact on the overall PSA result. Thus, a short overview of seismic human reliability models discussed in the literature is presented in the following.

According to the **shock model**, it is conservatively assumed that all human actions fail if the peak ground acceleration exceeds a certain limit. For instance, Swiss regulations define this limit at two times the PGA of the DBE; refer to Klügel (2007).

In the **time-dependent model** described in NUREG (1987), the seismic-induced human error probability (HEP) is expressed as a function of the time t between the earthquake and the action:

$$HEP(t) = HEP_{t=1\,\min} \cdot t^{-a}$$

In the **ramp model** presented by Yokobayashi et al. (1998), the operator performance is modeled by a linear relationship between the seismic-induced HEP and the PGA until a maximum ground acceleration level L is reached.

Finally, it is worth mentioning a **model based on the CAV** (cumulative absolute velocity) value presented by Klügel (2007). No adverse effect on operator performance is expected if the CAV is below 0.16 gs. In case of stronger earthquakes,

$$HEP = HEP_{Baseline} + HEP_{GF,\,direct} + HEP_{Shock}$$

$HEP_{Baseline}$ is the HEP used in the PSA for internal events. $HEP_{GF,direct}$ describes structural failure of equipment (e.g., instruments, main control room boards) as well as accessibility restrictions leading to a guaranteed failure (GF) of the considered action. Finally, HEP_{Shock} models a failure of all operator actions due to operator shock, depending on the CAV: shock-induced failure is assumed with probability one if the CAV is greater than 0.16 gs and strong-motion duration exceeds a threshold value; otherwise, the probability of shock-induced failure is zero.

Uncertainty Analysis

In the earlier sections on seismic hazard analysis and – in greater detail – on fragility analysis, two sources of nondeterminism have been introduced

for the parameters relevant in a seismic PSA, i.e., the frequency of occurrence of a specific PGA (hazard) and the capacity of a component or structure (again in terms of the PGA).

These two sources of nondeterminism are random variability and (epistemic) uncertainty.

Random variability is the nondeterminism due to the inherently unpredictable nature of earthquakes and, consequently, of the earthquake-induced vibrations of the ground, the NPP civil structures, and the NPP equipment. The degree of scatter induced by random variability is termed as irreducible, since additional collection of data can improve its statistical representation, but not the scatter itself (expressed for instance by the variance or standard deviation).

Uncertainty, on the other hand, denotes the nondeterminism due to limited accuracy of models used to characterize the structural response. This lack of accuracy can be caused by inherent limitations of the model when representing the phenomena it is meant to capture (including issues such as model resolution). Also, it can be caused by limited data availability for model calibration. Another type of nondeterminism described by the term "uncertainty" is the subjective variability associated with the process of selecting a model from a set of concurrent models involving different experts. This source of uncertainty is particularly pronounced in probabilistic seismic hazard analysis.

The uncertainty in the seismic hazard is typically characterized by deriving different hazard curves for different confidence levels. In the following figure this is shown in the upper left subfigure; in this case the confidence levels are expressed by multiples of the standard deviation. For a defined value of the PGA, the uncertainty is schematically highlighted by the blue, double-sided vertical arrow.

As discussed in section "Log-Normal Capacity Model," the uncertainty associated with the seismic capacity of SSC is characterized by the uncertainty variability parameter β_U. Similarly as for the hazard curve, the uncertainty expressed by β_U can be shown graphically in terms of the fragility curves associated with the various confidence levels, as shown in the lower left subfigure below. In this subfigure, too, for a defined value of the PGA the uncertainty in the fragility is schematically highlighted by a blue, double-sided vertical arrow.

Finally, in the right portion of the figure, it is indicated how the uncertainties in the seismic hazard and in the fragility propagate – across the convolution integral introduced previously – to the core damage frequency.

The result of this uncertainty analysis – or uncertainty propagation – is then a distribution function of the core damage frequency, as shown in the lower right subfigure. The uncertainty in the core damage frequency is here emphasized by a blue, double-sided *horizontal* arrow (Fig. 10).

Practical Issues

Seismic Equipment List

The assembly of the so-called Seismic Equipment List (SEL) is a major task in both SMA and seismic PSA. To begin with, this list includes all those systems, structures, and components (SSC) which contribute to bringing the plant to a safe shutdown state after a seismic event (this is why in EPRI (1991), this list is referred to as Safe Shutdown Equipment List (SSEL)).

For success-path-based SMA, EPRI (1991) provides detailed guidance on how to assemble the SEL. For fault-tree-based SMA and seismic PSA, it is a common situation that a PSA model already exists for the level 1 PSA of internal events. The basic events of this level 1 PSA model represent a good starting point for the SEL in this case; however, the list is to be extended by including passive components (e.g., tanks) and structures, which are typically not modeled in level 1 PSA for internal events.

Walkdown-Assisted Use of Generic Data and Screening

The number of components in the initial SEL (in the order of magnitude of 10^3) prevents a consistent degree of detail to be applied in the analysis of all of them. It is hence a common feature of SMA and seismic PSA that the initial SEL has to undergo a more or less extensive
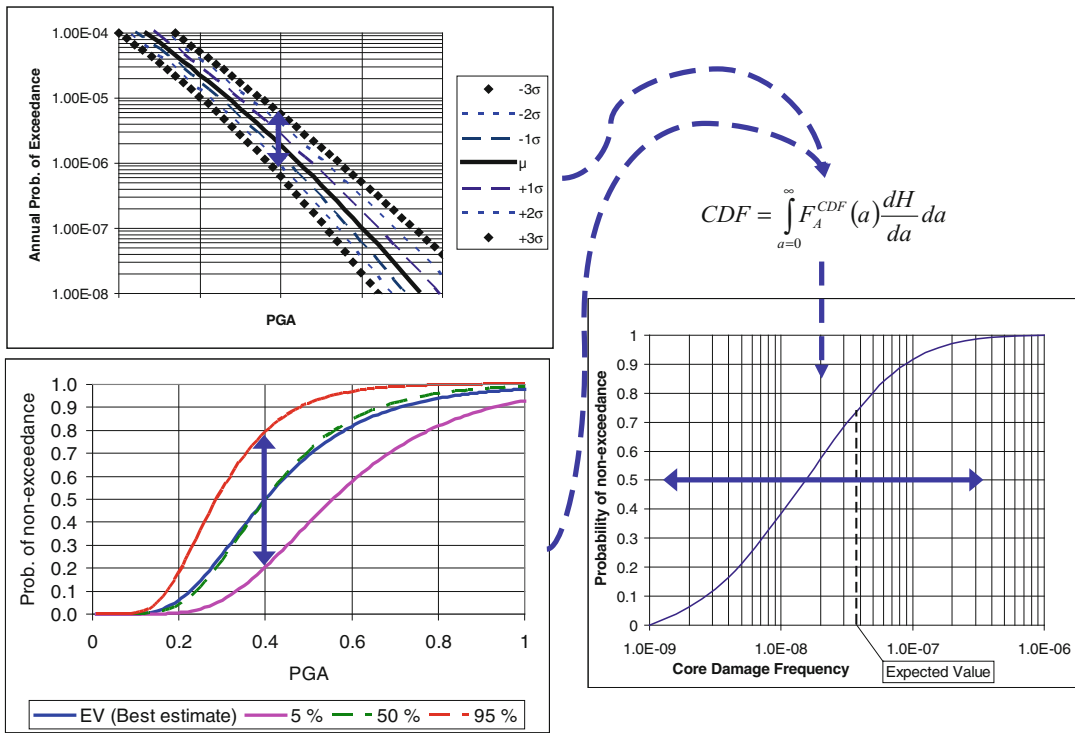
Seismic Robustness Analysis of Nuclear Power Plants, Fig. 10 Schematic illustration of uncertainty quantification in seismic PSA

screening process in order to reduce the number of equipment items on the list. The main purpose of doing so is to enable the analysts to focus the majority of their subsequent efforts on equipment with relatively low capacity and consequently with the highest risk relevance.

The screening process takes its legitimacy from extensive earthquake experience accumulated over the years. A crucial ingredient of the process is the execution of a seismic walkdown during which it is verified, whether the equipment in the experience database is representative of the components encountered in the plant under investigation.

Summary

Seismic robustness of nuclear power plants (NPP) denotes the ability of the plant to withstand severe seismic ground motion – possibly even beyond the seismic loads against which the

plant is designed – without experiencing the degradation of equipment to a point that the plant cannot be brought and maintained in a safe shutdown state.

By necessity, seismic robustness assessment focuses on the systems, structures, and components (SSC) that are essential to the goal of a safe shutdown in the wake of an extreme seismic event. For these SSC the actual capacity – or a lower bound thereof – is evaluated; in order to ensure a common measure of this capacity for all SSC, it is quantified in terms of a characteristic ground motion parameter, typically the peak ground acceleration.

Based on the above, evaluating the seismic robustness requires both system analysis – accounting for the response of a plant to a seismic-induced accident from a process engineering point of view – and structural analysis of the essential SSC.

There are two approaches that have been found to be viable for fulfilling the above

requirements, namely, seismic margin assessment (SMA) and seismic probabilistic safety analysis (seismic PSA; in the United States, the term seismic PRA is preferred, with "risk" replacing "safety").

The main difference between the two methods is that SMA is essentially a deterministic approach, leading to an estimate of the seismic robustness of the plant in terms of a single number, namely, the plant-level capacity.

In contrast, seismic PSA is an explicitly probabilistic approach, analyzing those combinations of component failures that prevent the plant from being transferred to a safe shutdown state, thus resulting in a damage to the reactor core. Compounding the probabilities of these combinations with the frequency of occurrence of seismic events of various intensity results in an estimate of the seismic-induced core damage frequency. A small value of the latter indicates a large seismic robustness of the plant.

Cross-References

▶ Analytic Fragility and Limit States [P(EDP| IM)]: Nonlinear Dynamic Procedures
▶ Code-Based Design: Seismic Isolation of Buildings
▶ Seismic Risk Assessment, Cascading Effects
▶ Site Response for Seismic Hazard Assessment
▶ Structural Reliability Estimation for Seismic Loading

References

American Society of Civil Engineers (ASCE) (2000) Seismic analysis of safety-related nuclear structures and commentary. ASCE 4-98, Reston

Campbell KW, Bozorgnia Y (2011) Evaluation of the use of cumulative absolute velocity (CAV) in the shutdown of nuclear power plants after an earthquake. In: Transactions, SMiRT 21, New Delhi

Chen W-F, Scawthorn C (eds) (2002) Earthquake engineering handbook. CRC Press, Boca Raton

Electric Power Research Institute (EPRI) (1991) A methodology for assessment of nuclear power plant seismic margin (revision 1). Report EPRI NP-6041-SL, Palo Alto

Electric Power Research Institute (EPRI) (1994) Methodology for developing seismic fragilities. Report TR-103959, Palo Alto

Electric Power Research Institute (EPRI), U.S. Department of Energy (2006) Program on technology innovation: use of cumulative absolute velocity (CAV) in determining effects of small magnitude earthquakes on seismic hazard analyses. Report 1014099, Palo Alto/Germantown

EUR Organisation (2001) European utility requirements for LWR nuclear power plants

International Atomic Agency (IAEA) (2003a) Seismic design and qualification for nuclear power plants. Safety guide NS-G-16, Vienna

International Atomic Agency (IAEA) (2003b) Seismic evaluation of existing nuclear power plants. Safety series 28, Vienna

International Atomic Agency (IAEA) (2009) Evaluation of seismic safety for existing nuclear installations. Safety guide no. NS-G-2.13, Vienna

Kennedy RP, Ravindra MK (1984) Seismic fragilities for nuclear power plant risk studies. Nucl Eng Des 79:47–68

Klügel J (2007) An improved methodology for the evaluation of human error probabilities in a seismic PSA. In: Transactions, SMiRT 19, Toronto

Pellissetti MF, Klapp U (2011) Integration of correlation models for seismic failures into fault tree based seismic PSA. In: Transactions, SMiRT 21, New Delhi

Sadegh-Azar H, Hartmann H-G (2011) Grundlagen der seismischen Auslegung von Kernkraftwerken und Einfluss der Boden-Bauwerk Wechselwirkung. Bauingenieur (D-A-CH-Mitteilungsblatt)

U.S. Nuclear Regulatory Commission (NUREG) (1985) An approach to the quantification of seismic margins in nuclear power plants. NUREG/CR-4334, UCID-20444, Washington, DC

U.S. Nuclear Regulatory Commission (NUREG). (1986). *Recommendations to the* Nuclear regulatory commission on trial guidelines for seismic margin reviews of nuclear power Plants. NUREG/CR-4482, UCID-20579, Washington, DC

U.S. Nuclear Regulatory Commission (NUREG) (1987) Relay Chatter and operator response after a large earthquake. NUREG/CR-4910, Washington, DC

U.S. Nuclear Regulatory Commission (NUREG) (1990) Procedures for the external event core damage frequency analyses for NUREG-1150. NUREG/CR-4840, SAND88-3102, Washington, DC

U.S. Nuclear Regulatory Commission (NUREG). (1993). Subject: policy, technical, and licensing issues pertaining to Evolutionary and Advanced Light-Water Reactor (ALWR) designs. SECY 93–087, Washington, DC

Yokobayashi M et al (1998) Consideration of the effect of human error in a seismic PSA. In: Proceedings of the 4th international conference on probabilistic safety assessment and management, Berlin

Seismic Sources from Landslides and Glaciers

Ewald Brückl[1], Daniel Binder[2] and Stefan Mertl[3]
[1]Department of Geodesy and Geoinformation,
TU Wien, Vienna, Austria
[2]Zentralanstalt für Meteorologie und
Geodynamik, Vienna, Austria
[3]Mertl Research GmbH, Vienna, Austria

Synonyms

Earthquakes generated by landslides and glaciers; Seismic activity of landslides and glaciers

Introduction

With few exceptions, the discovery and analysis of global earthquakes with landslides and glaciers as sources did not begin until the twenty-first century. This may come as a surprise considering that the magnitudes of such earthquakes have reached and even exceeded M = 5. In this essay, we describe the specific characteristics of these seismic signals and their source locations. We describe the source mechanisms and discuss the information that can be derived from interpretation of the data. In the case of landslides, the source process represents a severe hazard, and the same may be true for glaciers.

Although seismic signals from landslides and glaciers are observed at intermediate scales, we will jump from magnitudes of M ~ 5 to M < 0. The deployment of seismic stations on the target landslide or glacier is necessary to observe such weak signals. Systematic research on these microearthquakes also did not start until the twenty-first century. We offer an overview on landslide and glacier processes relevant to the generation of microearthquakes. Using selected case studies, we present the diversity of seismic signals generated by landslide and glacier sources and emphasize the close relations to geodetic and hydrometeorological data. We relate the seismic data to landslide and glacier processes as far as possible at the current state of knowledge. The potential usage of this seismic monitoring data for hazard estimation is also addressed.

A New Class of Earthquakes

Seismic Signals from the 1980 Mount St. Helens Eruption and Landslide

Mount St. Helens (46.2°N, 122.2°W) has been the most active volcano in the Cascade Range during the Holocene. Its most impressive active phase in historical times started with a series of small earthquakes in March 1980. During the first half of the following May, seismic activity increased dramatically, and a bulge over 100 m grew outward at the mountain's north flank. The cataclysmic Plinian-style eruption, VEI 5, occurred on May 18, 1980. Within 15–20 s of a magnitude 5.1 earthquake, the volcano's bulge and summit slid away in a huge landslide (Fig. 1a). Volcanic explosions interacted with the landslide process (Fig. 1b). The landslide, comprising a volume of 2.8 km^3, rushed first down the slope to the north and bent later to the west-northwest along the North Fork Toutle River with a speed briefly exceeding 200 km/h. The landslide covered an area of nearly 60 km^2, with the most distal deposits traveling about 25 km (Fig. 1c). The height of Mount St. Helens was reduced from about 2950 to 2550 m (Brantley and Myers 2000).

Seismic signals from the May 18, 1980, eruption and landslide of Mount St. Helens have been recorded by seismic stations at epicentral distances from 23° up to 133° (Fig. 2a). The seismograms show strong surface waves in the period range of 100–260 s corresponding to Ms = 5.9. The spectra fall off very rapidly at periods shorter than 75 s. This makes these recordings significantly different from typical tectonic earthquake spectra (Fig. 2b). A cutoff period of ~5 s would be expected for a tectonic earthquake of similar magnitude. The low-frequency content of the Mount St. Helens recordings suggests a very slow source process. Kanamori and Given (1982) analyzed the
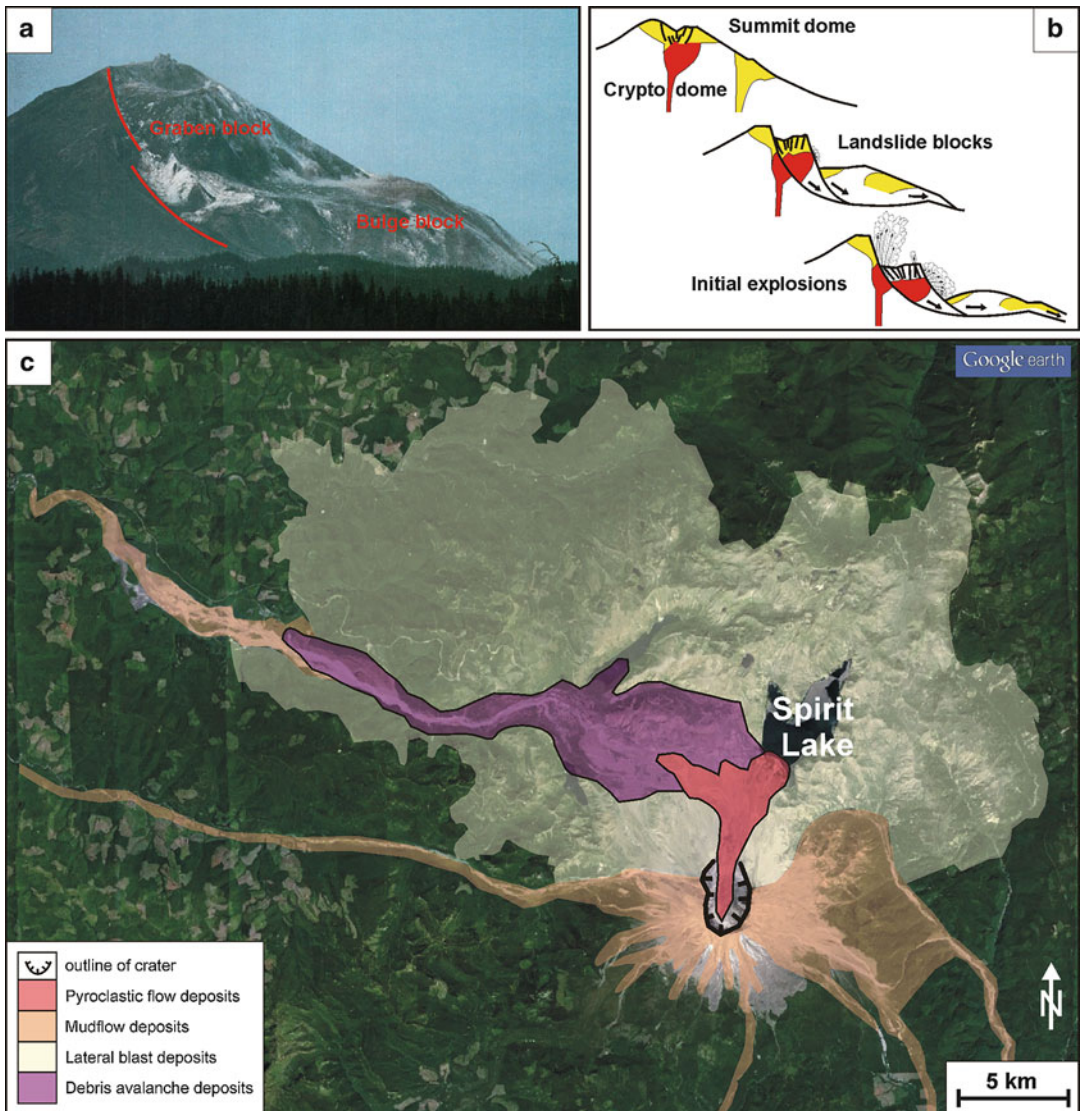
Seismic Sources from Landslides and Glaciers, Fig. 1 The Mount St. Helens eruption and landslide 18 May 1980; (**a**) the bulge at the north flank; (**b**) development of the landslide during eruption; (**c**) landslide deposit area (Modified after Voight 1981; Moore and Albee 1981, http://pubs.usgs.gov/gip/msh/mudflows. html)

radiation pattern of the Love and Rayleigh waves and concluded that a subhorizontal, about southward-directed single force is an adequate source mechanism and describes the radiation pattern much better than a conventional DC. The earthquake attributed to the Mount St. Helens landslide resembled the first example of a new class of global earthquakes.

The Single-Force Seismic Source

Earthquakes of the new class produce long-period surface waves detectable on a global scale with wavelengths at least ten times the extent of the source area. The spatial migration of the source during the event cannot be resolved by long period data, therefore we consider a single force **F(t)** varying in time but fixed in space as

Seismic Sources from Landslides and Glaciers, Fig. 2 Recordings from seismic observatories in Japan (*MAJ*), Norway (*KON*), and Australia (*NWA*), 30 s high-cut filtered of (**a**) Mount. St. Helens eruption and landslide, (**b**) Mammoth Lakes earthquake (Modified after Kanamori and Given 1982); V, R, and T are the vertical, radial and transverse components, R1 and G1 mark the first Rayleigh and Love phases
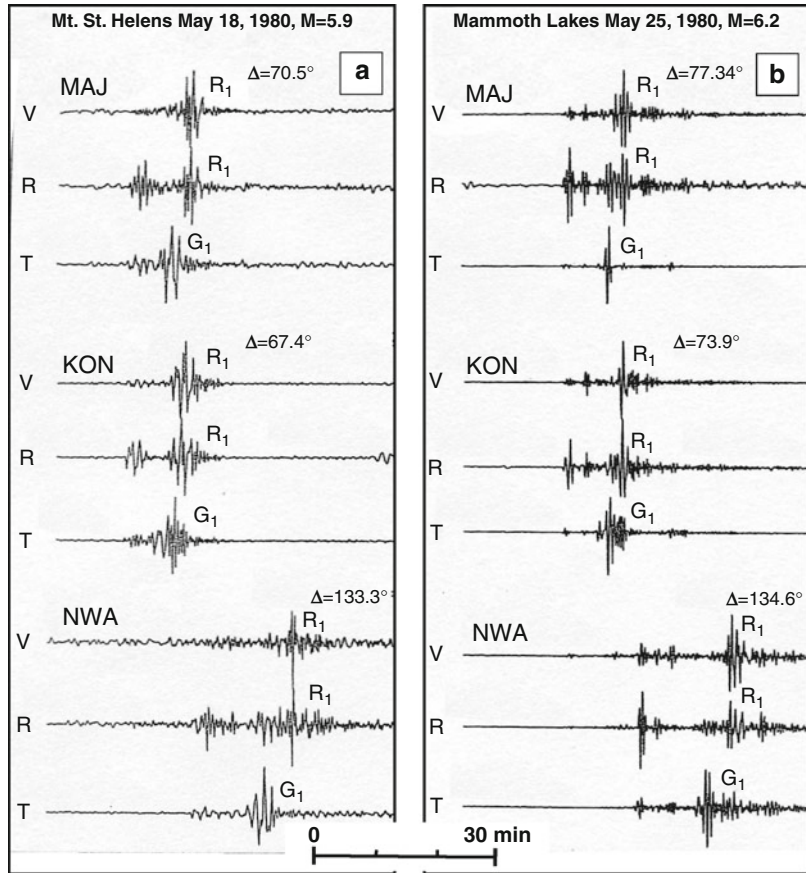
the seismic source. We refer to $\mathbf{F(t)}$ as force history. $\mathbf{F(t)}$ is the reaction force to the acceleration or deceleration $\mathbf{b(t)}$ times a mass $\mathbf{M}$ on the Earth's surface (Eq. 1).

$$F(t) = -b(t)M \quad (1)$$

The acceleration is caused by the drop of the frictional force at a sliding plane from static equilibrium to its value during sliding. The excess gravitational forces result in the acceleration. When the slope of the landslide trajectory reduces, frictional forces exceed the driving gravitational forces, and the landslide decelerates. The momentum $\mathbf{I(t)}$ and velocity of the landslide $\mathbf{v(t)}$ follow from integration of Eq. 1 over time (Eq. 2).

$$I(t) = v(t)M = \int_0^t F(s)\,ds \quad (2)$$

We have $I(T) = v(T) = 0$, where $\mathbf{T}$ is the duration of the landslide. This condition imposes a constraint on the derivation of the force history $\mathbf{F(t)}$. Kanamori and Given (1982) found that the data of the Mount St. Helens eruption can be fitted by a sinusoidal F(t) with a period of 240 s.

CSF: Centroid Single Force

The displacement multiplied by the landslide mass at time t follows from an integration of $\mathbf{I(t)}$ (Eq. 2). The total displacement vector $\mathbf{D}$ times $\mathbf{M}$ is given by Eq. 3.

$$DM = \int_0^\tau I(t)\,dt \quad (3)$$

The product of $\mathbf{D}$ and $\mathbf{M}$ can be derived from seismic data. Satellite images or other topographic information may also be used to break

Seismic Sources from Landslides and Glaciers, Fig. 3 Recordings of the Alaska glacial earthquake and a M 4.2 earthquake (Modified after Ekström et al. 2003)
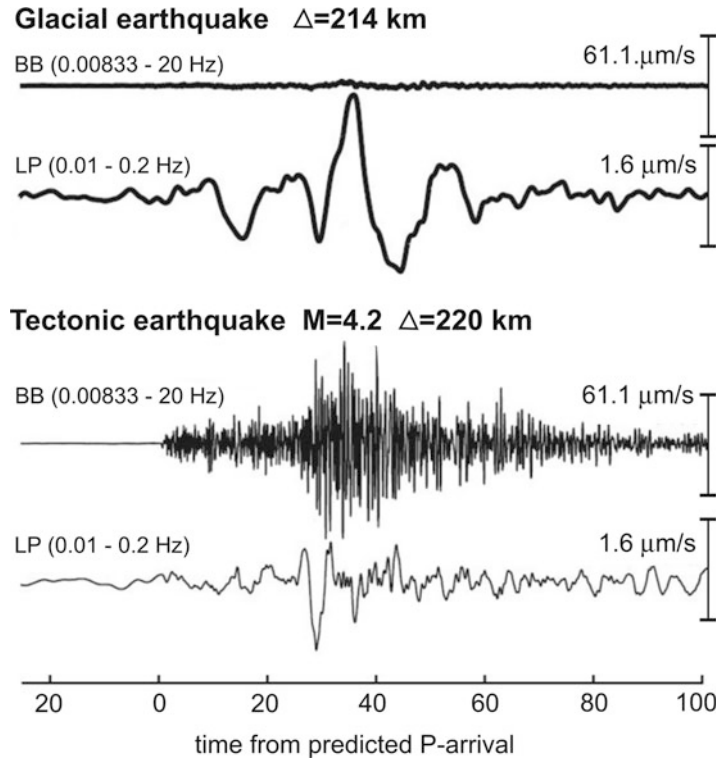
down the product into its factors **D** and **M**. Equations 1, 2, and 3 are valid for any surface mass accelerated or decelerated by either gravitational, inertial, or frictional forces, stress relief, or volcanic processes. Kawakatsu (1989) showed that **D M** represents the spatiotemporal centroid single force vector CSF = −D M, the analogue of the CMT for moment tensor seismic sources. He also provided the formalism for normal-mode inversion.

Global Glacial Earthquakes

A search for earthquakes of the new class in Arctic and Antarctic areas in data provided by global networks was carried out by Ekström et al. (2003) for the period 1999–2001 and extended to the period 1993–2008 by Tsai and Ekström (2007) and Nettles and Ekström (2010). About 243 earthquakes were found in Greenland, 14 in Antarctica, and 1 in Alaska. The one in Alaska (Fig. 3) exemplarily shows the difference

of the new-class earthquake to a nearby tectonic earthquake of similar magnitude and may explain why the new class of earthquakes has so long been undiscovered. CSF inversion was applied to the events in Greenland and Antarctica by Tsai and Ekström (2007) and Nettles and Ekström (2010), resulting in an improved location with an uncertainty of approximately 20 km. These events were clustered at or near eight fast-moving outlet glaciers with speeds >800 m/year (Fig. 4). An antisymmetric boxcar function (25 s of constant acceleration before the centroid focal time followed by 25 s of constant deceleration) was assumed in order to derive reactive forces transmitted to the Earth's crust by the displacement of a hypothetical mass. The magnitudes of centroid single force vectors were in the range $10 \cdot 10^{12} \leq CSF \leq 200 \cdot 10^{12}$ m kg, and the magnitudes were $4.6 \leq Ms \leq 5.1$. Most CSF vectors point in the opposite direction of the general flow of the glaciers.

The frequency of glacial earthquakes in Greenland is subject to seasonal variations and

Seismic Sources from Landslides and Glaciers, Fig. 4 Glacial earthquakes in Greenland (Modified after Tsai and Ekström (2007) and Nettles and Ekström (2010))

synchronous with maximum glacier flow velocities in July–August and a corresponding minimum in January and February (Nettles et al. 2008). A power law does not properly describe the magnitude/CSF – frequency relation. It seems that each outlet glacier has its typical earthquake, which depends on glacier size, hydrological conditions, and the calving rate (Tsai and Ekström 2007).

Chen et al. (2011) detected 13 more new-class glacial earthquakes in Antarctica with epicenters at or near the Ronne Ice Shelf, Ninnis Glacier, and Vanderford Glacier (Fig. 5). At the Ronne Ice Shelf, the direction of the CSF vector was normal to the ice flow direction and parallel to the ice front; force direction was parallel to the rift propagation direction. At Vanderford and Ninnis glaciers, CSF and local ice flow directions coincided with only one exception. The magnitudes of the glacier earthquakes in Antarctica were $4.2 \leq Ms \leq 4.9$.

Global Landslide Earthquakes

Ekström and Stark (2013) applied the detection and location method developed for the glacial earthquakes to known large landslides or areas where a potential for giant landslides can be assumed. The landslide force history **F(t)** was

Seismic Sources from Landslides and Glaciers, Fig. 5 Glacial earthquakes in Antarctica (Modified after Nettles and Ekström (2010) and Chen et al. (2011))
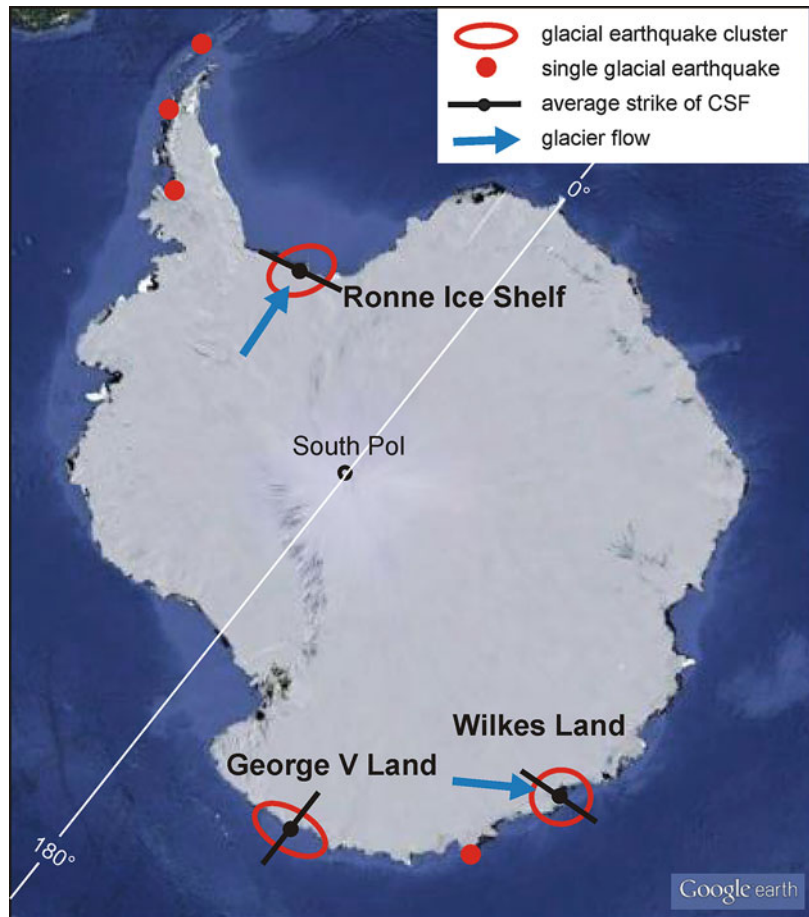
discretized into a sequence of overlapping isosceles triangles with a half-width of 10–15 s. Momentum, displacement, and CSF were derived from **F(t)**. Satellite images and other topographic information were used to separate the factors D and M out of the product CSF. Figure 6 shows, as an example, the Hunza-Attabad landslide in North Pakistan with a total mass of 140 10^9 kg. Ekström and Stark (2013) analyzed seismic recordings of 29 giant landslides in the Aleutian and Coast Ranges, the Rocky Mountains, the Andes, the Alpine – Himalayan orogeny, Taiwan, New Guinea, and Antarctica, with the Mount St. Helens 1980 landslide included. The magnitudes of these events range from Ms = 4.6–5.6. The CSF can be calculated from the data provided by Ekström and Stark (2013). CSF or log10(CSF) significantly

correlates with the landslide mass **M** and the magnitude Ms (Fig. 7).

Mechanisms of Giant-Single Force Sources

Landslide: Rotating Slider Block Model

Planar slider block models have been proven to support understanding of earthquake processes considerably. When considering the application of a planar slider block to earthquake processes, the slider block is loaded by tectonic strain, and after strain release the slider block comes to rest. For landslides, the situation is different; on an inclined planar sliding plane the slider block accelerates constantly and the travel distance is not confined. A rotational slider block overcomes
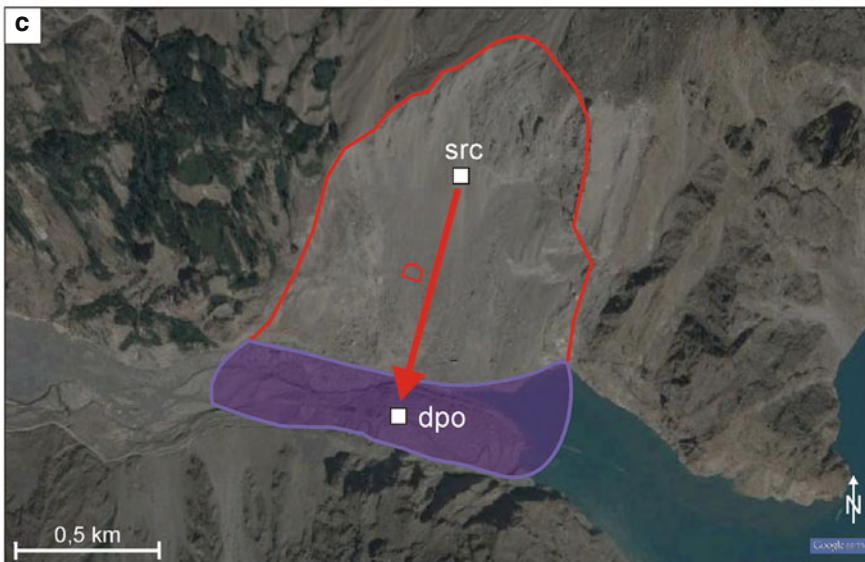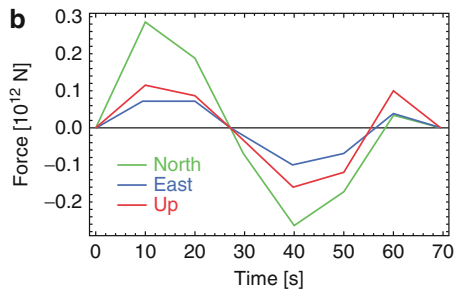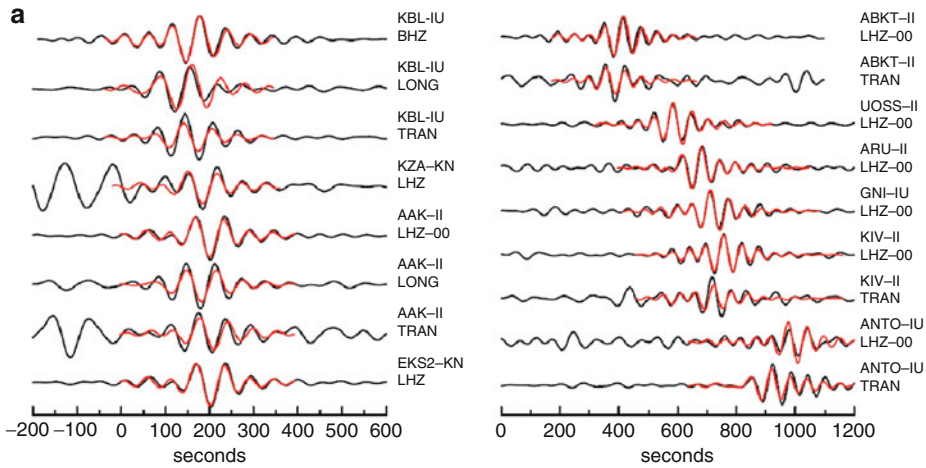
Seismic Sources from Landslides and Glaciers, Fig. 6 Hunza-Attabad landslide 04 January 2010 (Modified after Ekström and Stark 2013); (**a**) observed and synthetic traces; (**b**) landslide force history; (**c**) source (*red outline*) and deposition area (*blue*) of the landslide; src and dpo are centroids of the landslide mass M before and after the slide, D the trajectory derived from satellite image; given CSF = M D = $14*10^{13}$ kg m the landslide mass is M ~ $14*10^{10}$ kg; note the damming of the Hunza river
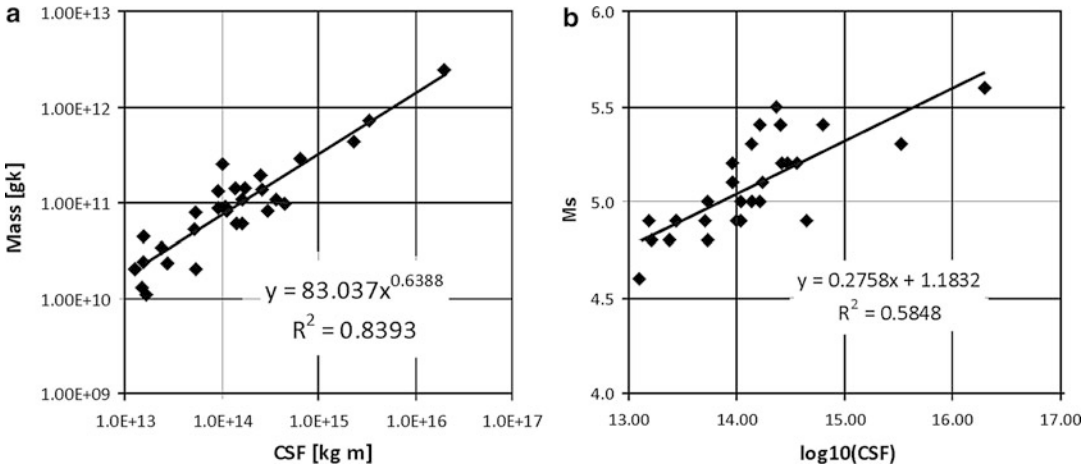
Seismic Sources from Landslides and Glaciers, Fig. 7 Relations between landslide data (Ekström and Stark 2013); (**a**) Mass – CSF, (**b**) Ms – log10(CSF)
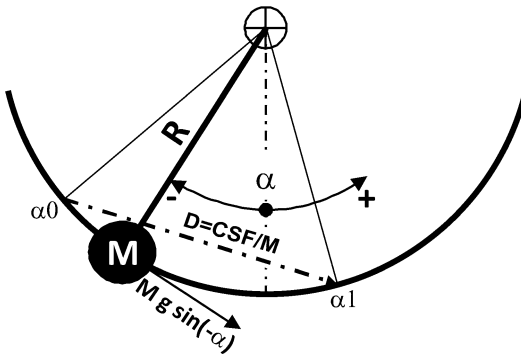
Seismic Sources from Landslides and Glaciers, Fig. 8 Rotational slider block model (for description of symbols see section "Landslide: Rotating Slider Block Model")

this problem and confines the travel distance (e.g., Brückl and Parotidis 2005).

Figure 8 shows the geometry of a rotational slider block. The dynamic system of a landslide is approximated by a mathematical pendulum with the length R and the mass M concentrated to a point. The position of the mass is defined by the angle α times R. Before the landslide happens at $\alpha = \alpha0$, the driving gravitational force M g sin $(-\alpha)$ is in equilibrium with the frictional force $\mu0$ M g cos(α), which always acts in opposition to the direction of movement (Eq. 4). A decrease of the friction coefficient from $\mu0$ to μ initiates the landslide. The excess driving force is

compensated for by the inertial force M $d^2\alpha/dt^2$ R (Eq. 5). We only look at one half oscillation of the slider block.

$$M\,g\,\sin{(-\alpha0)} - M\,g\cos{(\alpha0)}\mu0 = 0 \qquad (4)$$

$$M\,g\,\sin{(-\alpha)} - M\,g(\cos\alpha)\mu = -M\,R\,d^2\alpha/dt^2 \qquad (5)$$

We consider a first-order approximation for small amplitudes (sin(α) ~ α, cos(α) ~ 1) and travel distance D ~ R ($\alpha - \alpha0$). Introduction of these approximations and subtraction of Eq. 4 from Eq. 5 gives

$$(R/g)d^2\alpha/dt^2 + (\alpha - \alpha0) = \mu0 - \mu = \Delta\mu \quad (6)$$

The solution of the inhomogeneous differential equation for the initial conditions $\alpha = \alpha0$ and $d/dt(\alpha) = 0$ is given by Eqs. 7 and 8.

$$\alpha(t) = \alpha0 + \Delta\mu(1 - \cos{(\omega0 t)}) \qquad (7)$$

$$\omega0 = 2\pi/T0 = \text{sqrt}(g/R) \qquad (8)$$

The position angles α of the slider block at start time t0, stop time t1 = T0/2, and the time of maximum velocity tm = T0/4 are $\alpha0$, $\alpha1 = \alpha0 + 2\,\Delta\mu$, and $\alpha m = (\alpha0 + \alpha1)/2 = \mu = -\text{atan}(CSF_z/CSF_h)$. The travel distance D and

maximum magnitude of the landslide force (positive for t = t0, negative for t = t1) are given by Eqs. 9 and 10.

$$D = CSF/M = R(\alpha 1 - \alpha 0) = 2\ R\Delta\mu$$
$$= 2\Delta\mu\ g(T0/2\pi)^2 \tag{9}$$

$$Fmax = M\ g\ \Delta\mu = g\ CSF/(2R)$$
$$= \pi^2 CSF/\left(2(T0/2)^2\right) \tag{10}$$

The half-period T0/2 of the rotational slider block is the duration of the slide or the landslide force history T. We introduce a calibration factor k < 1 (Eq. 11) to account for the observation of the landslide force history F(t) not starting and ending with the extreme value predicted by the slider block model.

$$T0/2 = k\ T \tag{11}$$

A calibration factor of k = 0.68 brings Fmax, calculated in Eq. 10, into a nearly perfect 1:1 relation with the Fmax derived from the inverted force–time function (Ekström and Stark 2013; Fig. 9a). Figure 9b shows the dependence of αm and $\alpha 0$ on CSF. Over the range of CSF (approximately three decades), the average sliding angle reduces from ~25° to ~10°. There is also a systematic increase of ratio ($\alpha 0$–αm) to $\alpha 0$ with CSF from ~28 % to ~55 %.

Calving Glacier: Toppling of an Iceberg

The large outlet glaciers of Antarctica and Greenland lose mass mainly by calving. One of the best-monitored sites where calving takes place is the front of the Jakobshavn Glacier near Ilulissat, West Greenland. On 28 May 2008, a gigantic sequence of calving events was filmed by a team of Chasing Ice (http://earthsky.org/earth/video-largest-glacier-calving-ever-caught-on-film). An area of the size of Manhattan calved successively over a time span of 75 min. One can observe a great number of slabs of about 1000 m in height toppling from unstable to stable equilibrium, thus pushing the iceberg mélange of

previously calved icebergs away from the ice front out to sea. This toppling of large icebergs has been considered a candidate for a seismic source of global scale (Tsai et al. 2008).

Figure 10 illustrates the geometry, kinematics, and acting forces during the toppling of an iceberg, leaving it in unstable equilibrium after calving from the front of an ice stream. The calving front may or may not be grounded. The side lengths of the cross section are H (the height before toppling) and L. The width of the iceberg is W. With H > L, the iceberg is initially in an unstable position. Toppling of the iceberg implies not only a rotation of ~90° but also acceleration off the glacier front. This inertial force causes a reaction force at the contact area of the iceberg with the glacier, which is then transmitted via the glacier ice to the glacier bed and represents the main part of the glacial seismic source. A first-order approximation of CSF:

$$CSF \sim M(H - L)/2,\ \text{with}\ M$$
$$= H\ L\ W\ d_i \tag{12}$$

The density of ice is d_i and that of seawater d_w. The characteristic time scale of toppling can be estimated from the period of an iceberg interpreted as a mathematical pendulum:

$$T \sim 2\pi\ \text{sqrt}(d_i H/((d_w - d_i)g)) \tag{13}$$

Equations 12 and 13 do not take into consideration the remaining kinetic energy the iceberg has after reaching its stable position at the horizontal travel distance (H-L)/2. Furthermore, the friction between the iceberg and the glacier ice front along with the interaction of the iceberg with seawater and the mélange of smaller icebergs and fragments in front of the calving front is not taken into account. However, assuming the very reasonable dimensions H = 800 m, L = 100 m, and W = 400 m, CSF ~ 10^{13} kg m (Eq. 12), T ~ 150 s (Eq. 13), and Ms = 4.8 according to the relation between log10(CSF) and Ms (Fig. 7b).
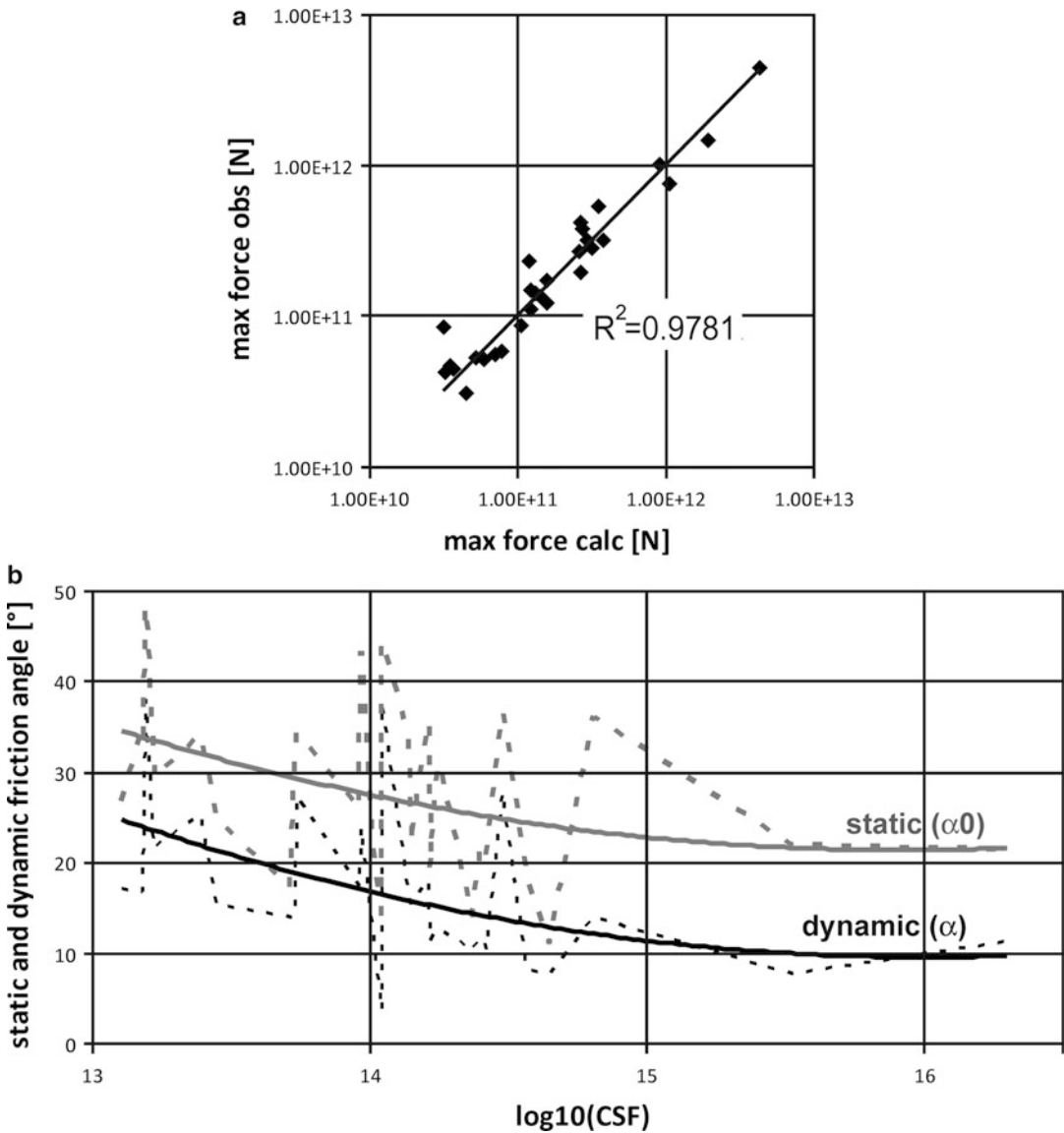
Seismic Sources from Landslides and Glaciers, Fig. 9 Relations derived from the rotational slider block model; (**a**) maximum force observed – maximum force calculated; (**b**) α, $\alpha0$ – log10(CSF)

Microearthquakes Generated by Landslides

The dynamic collapse of a slope within several tens of seconds is only representative of how a landslide can end in the most extreme cases. However, the majority of landslides do not develop into dynamic and catastrophic failure but reach more or less stable final states after phases of varying levels of activity. A deeper understanding of landslide processes and evolution is of fundamental importance for early warning and hazard estimation. Seismic activity connected to these processes can be observed by local seismic networks deployed on the landslide mass. We term microearthquakes generated by landslides LMEs (landslide microearthquakes) in the following.

Seismic Sources from Landslides and Glaciers, Fig. 10 Toppling (capsizing) of an iceberg acting as single force seismic source; (**a**) kinematics and acting force (toppling may also occur by counterclockwise rotation); (**b**) toppling of an iceberg during the 28 May 2008 calving event at Jakobshavn glacier
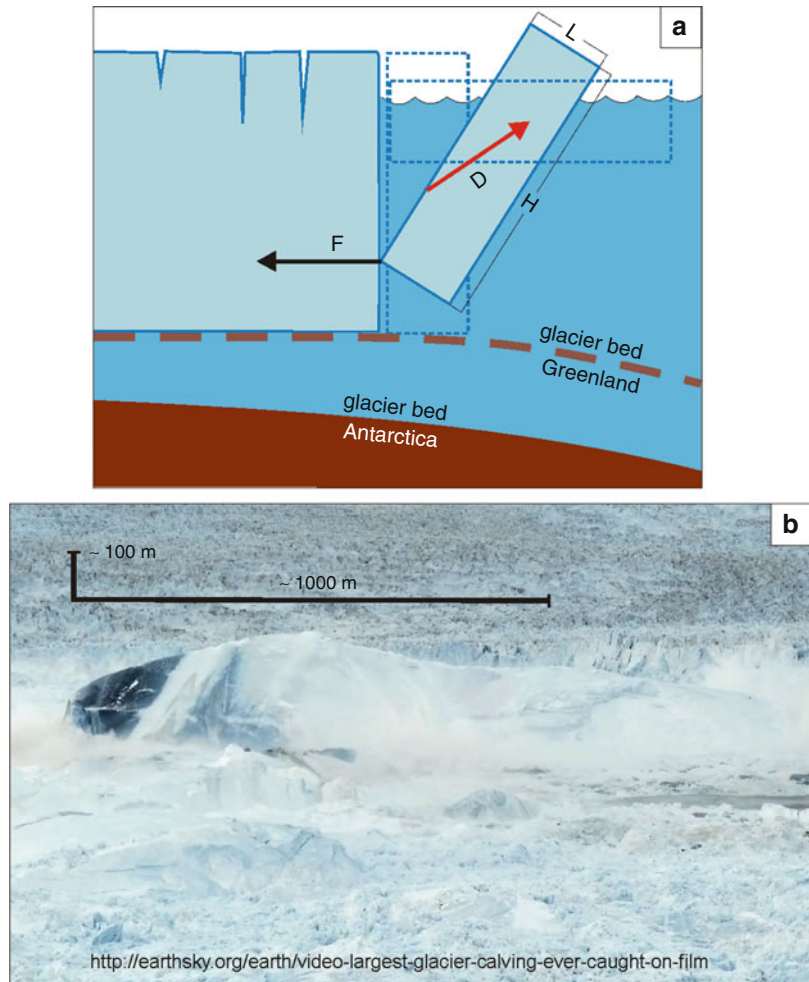
Landslide Processes and Evolution

We assume brittle rheology of the rock material constituting a landslide as a precondition for seismic activity. However, under some circumstances, the landslide mass may behave as ductile, and even a brittle deformation will not necessarily act as a seismic source. In the following, we present a simplified scheme of landslide processes and possible evolutions to dynamic failure or stabilization (Fig. 11). This schema should provide a backbone for the interpretation of the seismic data provided by the case studies presented in sections "Pre- and Postfailure Processes at Steep Slopes" to "Flow".

Slopes of a given inclination and height can only exist if the stability is guaranteed by the initial strength of the rock material. The possible

physical states are depicted in Fig. 11a in a Mohr diagram by the three Mohr's circles (M1, M2, M3) and Mohr-Coulomb failure criteria representing the initial and degraded strengths of the rock mass forming the slope. The Mohr's circles symbolize the states of stress, which are relevant for the stability of the slope. The initial rock strength is given by a Mohr-Coulomb failure criterion for the shear stress, with C0 and ϕ0 the initial cohesion and the initial friction angle. Increasing pore pressure p shifts the Mohr's circles to the left and nearer to the initial rock strength. In an extreme case of high pore pressure, a dynamic rupture could be induced; however, this is not the only situation that reduces the slope stability. Subcritical crack growth (Brückl and Parotidis 2005) supported by deep chemical
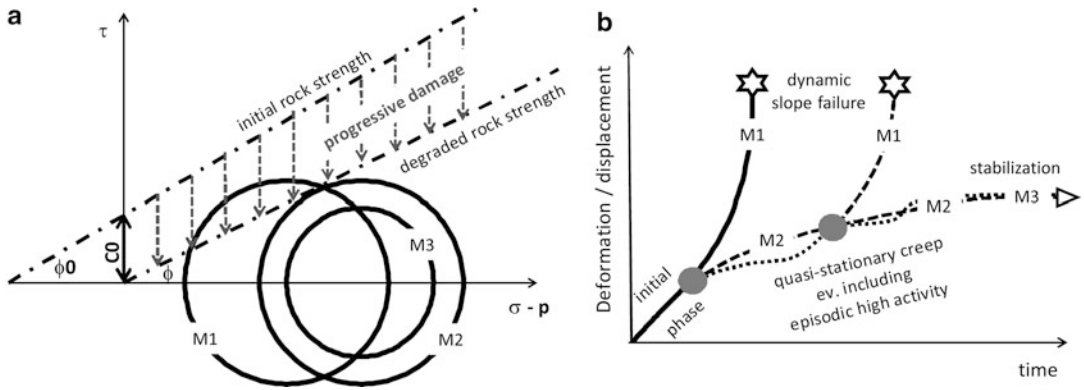
Seismic Sources from Landslides and Glaciers, Fig. 11 Landslide processes and evolution; (**a**) reduction of rock mass strength; (**b**) temporal evolution of displacement

or thermomechanical weathering (Gischig et al. 2011) may progressively damage the rock material and degrade the rock strength. Ultimately, cohesion (C0) is completely lost and the friction angle diminished considerably from φ0 to φ. This failure criterion of the degraded rock strength intersects Mohr's circle M1 and is tangential to M2, which means that in the first case a dynamic failure has already occurred. In the second case, the slope is in an equilibrium state of driving (gravitational) and resistive (frictional) forces. Creep of the whole slope may reduce the inclination and therefore the driving forces, and this situation is symbolized in Mohr's circle M3. It is well below the failure criterion of the degraded rock. The slope comes to an ultimate stabilization by this process.

The temporal evolution of the deformation or displacement of a landslide mass is schematically shown in Fig. 11b. The initial phase corresponds to the progressive damage from initial to degraded rock strength (Fig. 11a), where the whole rock mass of the landslide is degraded and deformed. However, deformation concentrates successively to one basal or several main sliding zones or surfaces. Toward the end of the initial phase, the further development of the slope comes to a bifurcation. In the case where Mohr's circle M1 represents the critical state of stress in the slope, the initial phase leads directly to a dynamic slope failure. In the case of Mohr's circle M2-type critical state of stress, a phase of

quasistationary creep, eventually interrupted by episodic phases of high creep or sliding velocities, follows. However, we cannot exclude that episodic high activity is the precursor for a later bifurcation between quasistationary creep and dynamic failure, as indicated in Fig. 11b by a second grey dot. The longer the quasistationary phase lasts, the more the driving forces and the radius of the Mohr's circle reduce, so that in the final state of stress represented by M3 the slope stabilizes. Decrease of the pore pressure p by improved drainage of the slope due to increased permeability of the rock mass may support stabilization.

Rockfall

Rockfall is a phenomenon occurring on slopes where, at least locally, the inclination equals or exceeds that of talus of the same rock material. It may be the consequence of weathering but also an indication of increased activity in a landslide. The seismic signature of rockfall may be a sequence of many irregular impulsive signals moving down the slope. Figure 13 shows seismic recordings from the tumbling of a medium-sized boulder (~2 m diameter) at the Gradenbach site (Fig. 12), images of the boulder itself and the site of heaviest impact, and the seismic location of that place (Brückl et al. 2013).

As a second example, the seismic signature of a rockfall at Steinlehnen (Fig. 12) is documented in Fig. 14. This rockfall was more of a small rock

Seismic Sources from Landslides and Glaciers, Fig. 12 Monitoring sites in the European Alps

avalanche, which left behind a path resolved by ground-based InSAR monitoring. The whole event lasted around 30 s, and location applied to several time windows allowed for the tracing of the path of this rockfall. There is fair agreement between both tracing methods (Weginger 2012).

Pre- and Postfailure Processes at Steep Slopes

We denote "steep slopes" as slopes that are steeper than the initial angle of friction or the inclination of talus built up by the same rock material. The rock material of these slopes needs cohesion to balance the gravitational forces. In the case of cohesion dropping to zero because of progressive damage (Fig. 11a), the slope will immediately transit from this initial phase to dynamic failure (Fig. 11b). Identification and monitoring of precursors of this kind of dynamic rupture would be of great scientific and practical importance.

The retreat of cliffs along the British and French Channel Coasts by sea erosion–driven rockfalls is an ongoing process. Amitrano et al. (2005) report seismic monitoring of a cliff collapse at Mesnil-Val on the Normandy coast, France. Precursory seismic activity was recorded

about 1 h before the collapse. Frequency of detected seismic signals and the seismic energy obey inverse Omori's laws. The data suggests that a warning based on increased activity could be given ~10 min before the collapse and a reliable prediction of the rupture time about 1 min in advance, provided data acquisition and continuous updating of calculations are done in real time.

The village of Randa, Matter Valley, Swiss Alps (Figs. 12, 15), was threatened by rock- and icefalls several times in its history. We anticipate a look from rock to ice and icefalls because of the similar phenomenology behind these processes in rock and ice. Icefalls were released by the Weisshorn hanging glacier five times in the last 35 years. The break-offs in 1973 and 2005 were monitored by geodetic and seismic methods. The velocity of the glacier front developed according to an inverse Omori law (Faillettaz et al. 2011) toward the instant of rupture. Log-periodic oscillations appeared superimposed on the general trend. Starting 30 days before rupture, three phases characterized by different cumulative size-frequency distribution of the signal energies and interevent waiting-time distribution behavior were identified and a correlation of seismic activity with the log-periodic oscillations established.
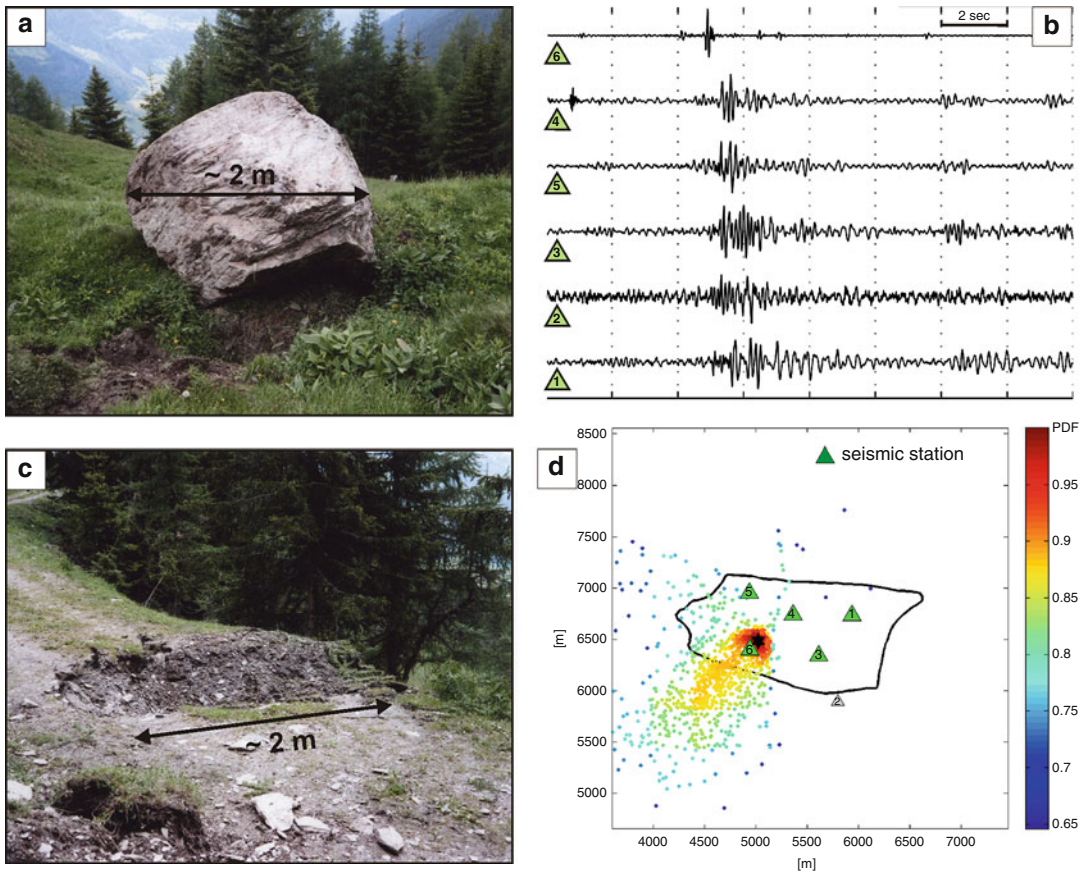
Seismic Sources from Landslides and Glaciers, Fig. 13 Rockfall at the Gradenbach landslide (Brückl et al. 2013); (**a**) main boulder; (**b**) recordings by the local network; (**c**) main impact; (**d**) location by the seismic data (normalized PDF) and actual position of the impact (*star*)

These phases represent the development from random crack generation over crack concentration at a fracture zone to the final rupture. The findings could be the basis of an efficient in-time warning.

From the detour to ice and a hanging glacier, we revert to rock and landslides. On 18 April and 9 May 1991, two successive rockslides took place near Randa (Fig. 15). The failure volumes involved about 20 and 10 10^6 m^3 massive para- and orthogneisses (Eberhardt et al. 2004). The average slopes of the failure planes were roughly 60° and 50° (Fig. 16). The rock masses dropped down 200 and 300 m and produced Ms 4.9 and 4.8 earthquakes detected by Ekström and Stark (2013). Deformation of the slope above the actual scarp continued after the rockfall sequence of the

year 1991 with ~10–20 mm/year (2004–2008). A seismic monitoring campaign was started in January 2002 and continued until July 2004 (Spillmann et al. 2007), where a monitoring network was deployed above the scarp (Fig. 15). Two types of waveforms classified as LMEs were observed: 94 single events comparable to the recordings of small detonations and 129 multiple events with a more complex source time function. Frequency content of the signals reached up to 100 Hz for the surface sensors and up to 200 Hz for the borehole sensors. Most of the seismic activity was concentrated on a volume of low seismic velocities ($<=$ 1500 m/s) near the scarp, but also deeper-reaching zones of increased seismic activity were identified, possibly outlining the scarp of a
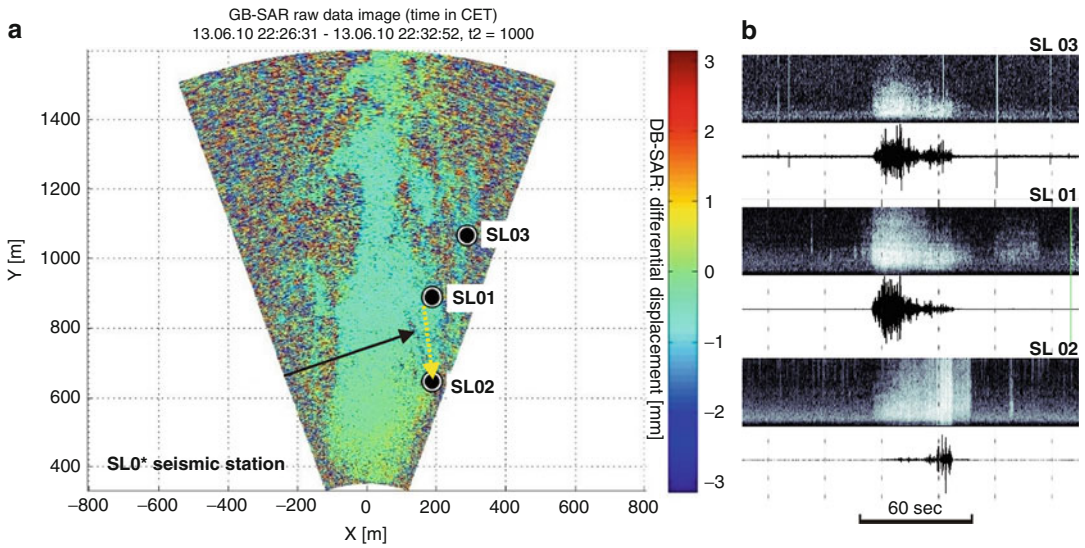
Seismic Sources from Landslides and Glaciers, Fig. 14 Rockfall at the Steinlehnen landslide (Weginger 2012); (**a**) differential ground based InSAR image shows the trace of a rockfall (*black arrow*); *yellow arrow* represents the path of the rockfall from the seismic data shown in (**b**)

Seismic Sources from Landslides and Glaciers, Fig. 15 View from east to the glacier break-off and rockslide sites near Randa; location of the profile in Fig. 16 is shown by a *dashed red line*
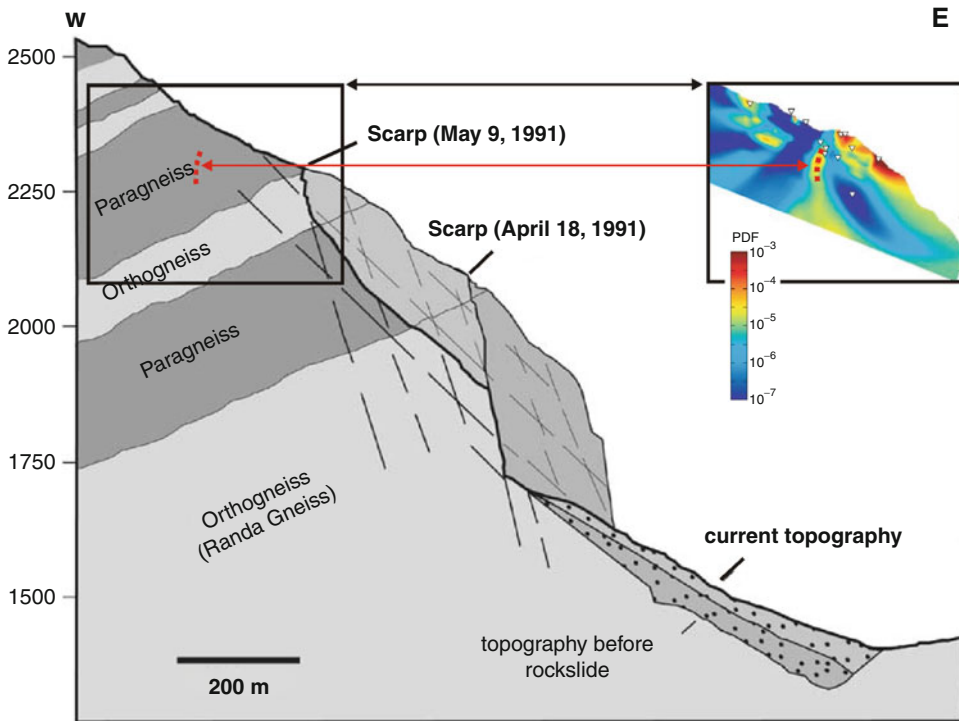
Seismic Sources from Landslides and Glaciers, Fig. 16 Cross section through the Randa rockslide site (Modified after Eberhardt et al. 2004); cumulative PDF of LMEs is shown within the detail marked by *black rectangle* (Modified after Spillmann et al. 2007); the location of a possible next scarp is marked by *red dots*

future rock slide (Fig. 16). Hydrological and meteorological conditions had no impact on the seismic activity and the geodetically observed displacement rates. Most magnitudes were in the range $-2.0 < Mw < -0.5$. It is questionable if the observed seismic activity and geodetically observed displacements are postfailure relaxation processes or if they belong to the initial phase of a third, larger rock slide at Randa.

DSGSD: Deep-Seated Gravitational Slope Deformation

Deep-seated gravitational slope deformation (DSGSD) is a phenomenon frequently observed in the hard rock slopes of alpine regions. Typically, the deformation comprises the whole flank of a valley, and the maximum thickness of the deformed rock mass exceeds 100 m. The mechanisms of DSGSD are manifold, and deformation styles such as "Bergzerreißung," "Talzuschub," sagging or Sackung (Zischinsky 1969), and kink band slumping (Goodman and Kieffer 2000) can be summarized under DSGSD. The initiation of DSGSD at alpine slopes mostly dates back to the retreat of the ice age glaciers, when the lateral support of the valley flank by the ice pressure was lost. This conception implies a long initial phase of mass movement, during which progressive damage degraded the rock strength almost to its residual value (see Fig. 11a; Brückl and Parotidis 2005). DSGSDs may be in a quasistationary creep phase or may exhibit episodic high deformation rates in parts of the slope or its entirety interrupting longer phases of little activity (Fig. 11b). There is currently no general answer as to whether these processes eventually lead to stabilization or a transition to dynamic rupture.

In the following, we consider the seismic signals recorded at three DSGSD examples in crystalline rocks of the European Alps: Séchilienne (Helmstetter and Garambois 2010), Steinlehnen (Weginger 2012), and Gradenbach (Brückl

Seismic Sources from Landslides and Glaciers, Fig. 17 View from south to the Séchilienne DSGSD and the active zone (area colored in *red*); the locations of the 2 seismic micro-arrays RUI and THE are marked by symbols
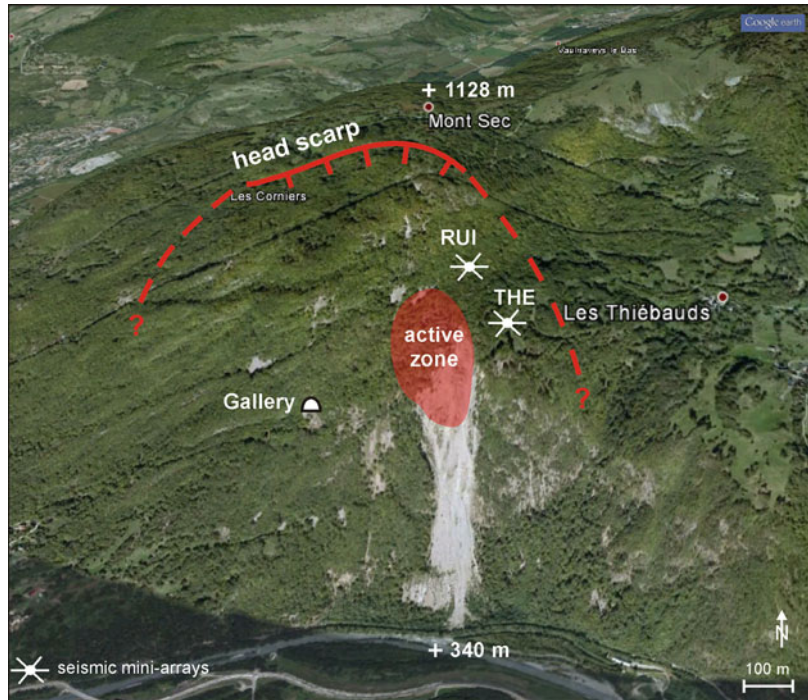
et al. 2013; Mertl 2015). The locations of these landslides are shown in Fig. 12; Figs. 17, 18, and 19 supply information about the morphology, the seismic monitoring network, and partly the geology and structure of these landslides. The average slopes of these DSGSDs vary between 28° (Gradenbach) and 35° (Steinlehnen).

At the Séchilienne DSGSD, a presently very active zone is located on top of the steeper lower part of the slope (marked in Fig. 17). The velocity of this active zone reached 1.4 m/year in 2008. Until April 2009, rockfalls were the most frequent seismic sources related to the dynamics recorded at the Séchillienne DSGSD. The cumulative frequency of >3000 events obeys a power law distribution over 2.5 decades of the amplitude range with an exponent corresponding to a b-value of b = 1.1.

Figure 20 shows typical waveforms of LMEs at the Séchilienne DSGSD. One type corresponds to impulsive sources similar to small explosions, while another event type shows mostly emergent onsets and represents source-time functions lasting ~5 s, either relatively broadband or low frequency. The LMEs cluster either in the

high-velocity area or the head scarp. It is also possible that some hypocenters are located at or near a basal shear zone. Rockfall and LME frequencies precede accelerations of the active zone by 4–5 days. Some but not all rainfall events trigger increased rockfall and seismic activity (Helmstetter and Garambois 2010).

Seismic monitoring campaigns at Steinlehnen and Gradenbach had the opportunity to observe episodic phases of high landslide velocities interrupting longer periods of quasistationary creep. During June 2010, the Steinlehnen DSGSD was monitored by a ground-based InSAR system installed opposite the Steinlehnen DSGSD. An acceleration phase of the active slab was observed with total displacements reaching up to 0.9 m. Up to five seismic monitoring stations were operating at the Steinlehnen slope during this period (Fig. 18a). Besides the rockfall (see Fig. 14), Weginger (2012) identified three classes (n, ln, i in Fig. 21a) of LMEs that had their sources within the area of the accelerated slab. This classification was based on characteristic properties of the seismic signals (e.g., duration,

S

frequency spectral characteristic, amplitude, envelopes, cross correlations, polarization). Figure 21b shows velocity and displacement of the highly active slab for the period of 10–28 June 2010. The high-velocity phase starts on 20 June and lasts until 25 June. The cumulative frequencies for LME types n, ln, and i rise from 13 June onward, and some rockfalls also occur (Fig. 21c). A sudden increase of the cumulative rockfall frequency accompanies the onset of the high-velocity phase between 20 and

21 June, and rockfall ceases completely thereafter. The cumulative frequencies of the events n, ln, i increase more or less continuously with the highest increments during the high-velocity phase. No extraordinary rainfall was reported to precede or accompany the period of high velocities of the active slab, and no external trigger of this phase is known so far. The incipient seismic and rockfall activity on 13 June was most likely the only precursor of the high-velocity phase in the following week.
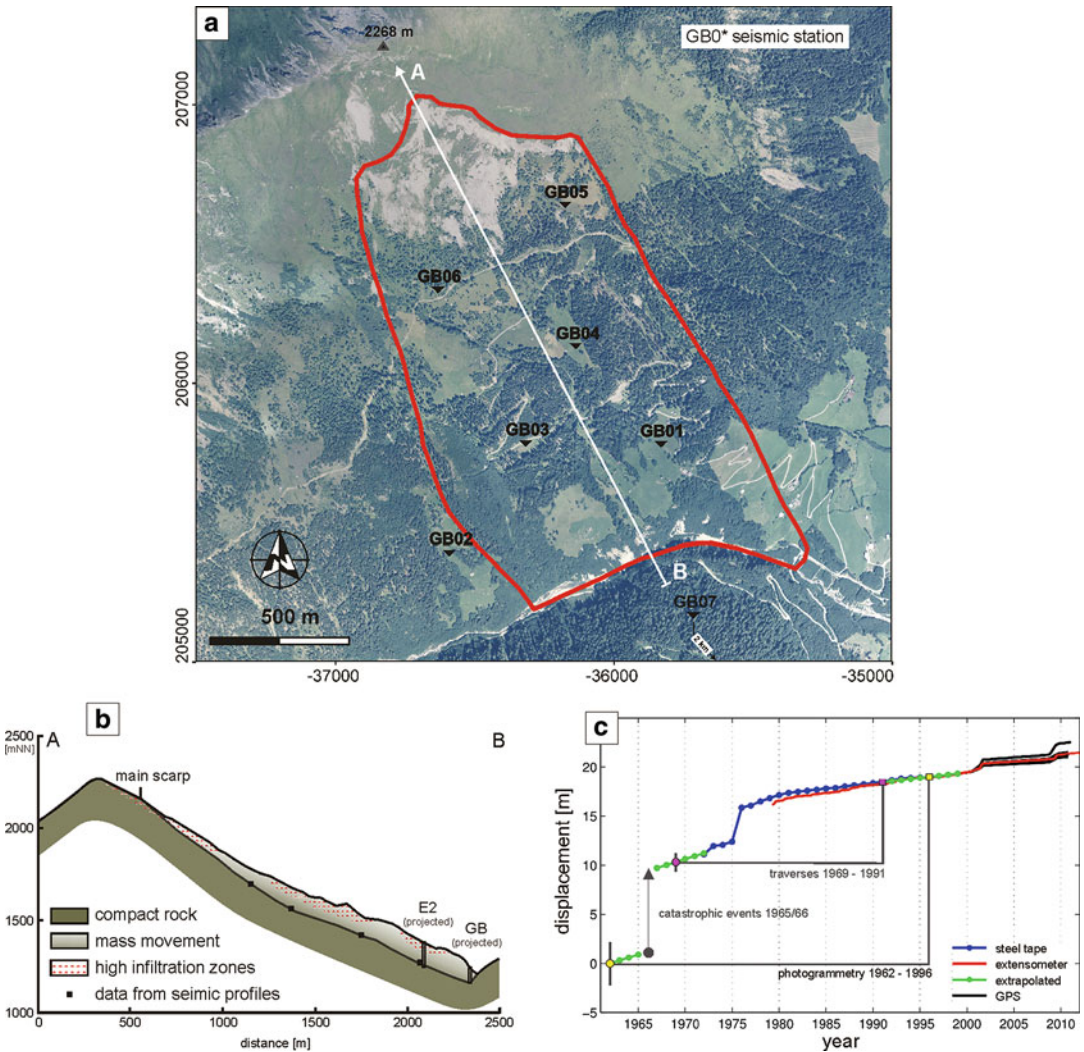
Seismic Sources from Landslides and Glaciers, Fig. 19 Gradenbach DSGSD (Brückl et al. 2013): (**a**) Orthophoto with contour of the active zone and seismic stations GB01–GB07; the main scarp, the highly fractures zone in the uppermost area of the mass movement, and the erosional zone at the toe of the mass movement are well visible in the orthophoto; (**b**) cross section A-B; the depth of the basal sliding surface or zone was determined by drilling (boreholes E2 and GB) and refraction seismic measurements (**c**) displacement history derived from terrestrial geodetic measurements 1969–1991, photogrammetric evaluation of the aerial photos 1962 and 1996, and GPS since 1999

The episodic high–sliding velocity phase at the Gradenbach DSGSD occurred during spring of 2009. In contrast to the Steinlehnen event, the whole landslide became active due to infiltration of melting snow and precipitation (Brückl et al. 2013). Five different LME types (AA, A, B, D, tremor; Fig. 22a) besides rockfall (rf) were identified using criteria comparable to those used for the data from the Steinlehnen site (Mertl 2015). The cumulative frequencies of these events behave distinctly. Event types A and D start with the onset of snow melt (Fig. 22b–d) and high infiltration. The tremor type was only observed at that time. Thereafter, a quiet phase that lasted over 1 month followed. The beginning of the episodic phase of high creep velocity was
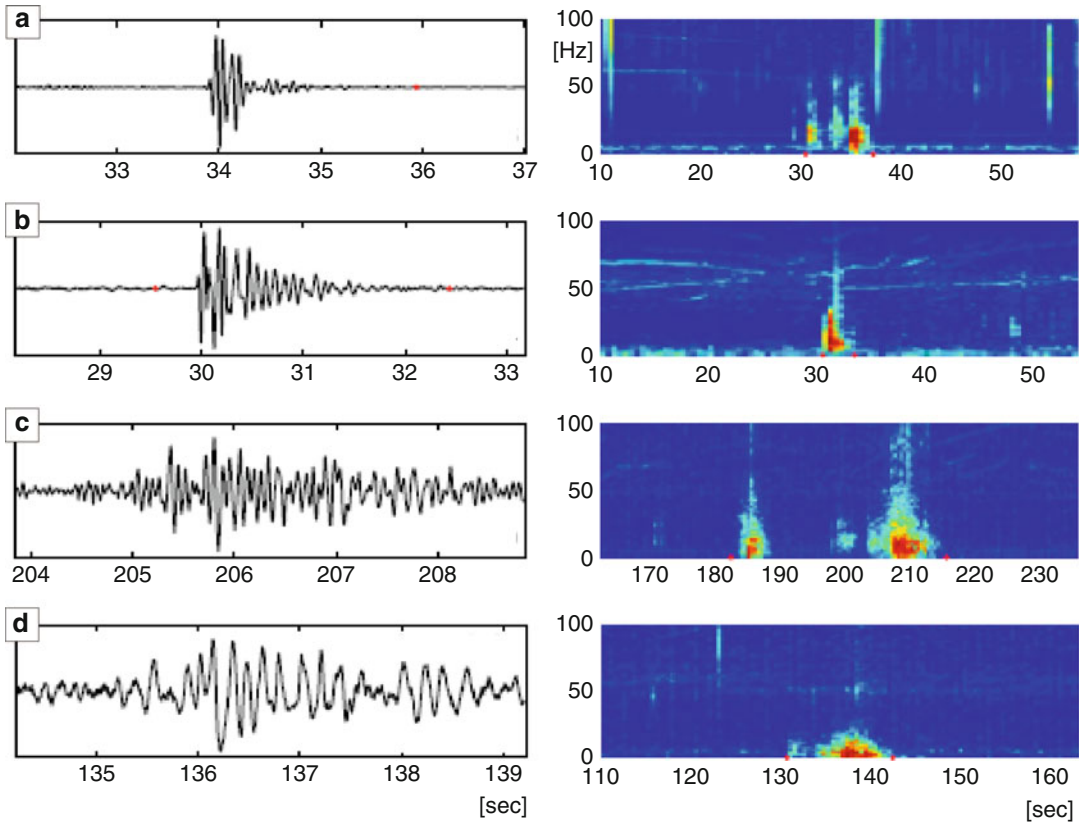
Seismic Sources from Landslides and Glaciers, Fig. 20 Typical waveforms and spectrograms from LMEs recorded at the Séchilienne DSGSD (Modified after Helmstetter and Garambois 2010); (**a**) impulsive LME; (**b**) small explosion; (**c**) emergent and broad band LME; (**d**) emergent and low frequency LME

accompanied by type B microearthquakes and some rockfall. The frequency of type A and D events rose 1–2 weeks later and lasted for the rest of the episodic high-creep phase. Magnitudes of the LMEs range from −2.8 to −0.9 with a b-value of b = 1.7.

Though the geodetically observed movements covered the whole landslide mass, the spatial distribution of the seismic activity during the episodic high-creep phase is concentrated to limited areas (Fig. 23). Event types A, D, and tremor, which coincide with the onset of snow melt and high infiltration, are most likely very shallow events. The waveforms and coincidence in time with the onset of high creep velocities may be indications that the hypocenters of B-type LMEs are at the level of the basal sliding zone. Asperities must have been overcome before the whole

landslide mass could increase in velocity (Mertl 2015).

Flow

A flow is a spatially continuous movement in which surfaces of shear are short lived, closely spaced, and not usually preserved. Whereas in DSGSD the original structure of the rock mass is partly preserved, in a flow it is completely destroyed and mixed up. The ratio of mean thickness to length of a flow is <0.01 and significantly lower than for DSGSDs. When considering the scheme of landslide processes and evolution presented in Fig. 11a, it is evident that the initial rock (or soil) strength has been completely reduced to its residual values. Changing pore pressure may keep the flow either active or temporarily inactive (Mohr's circles M2 and
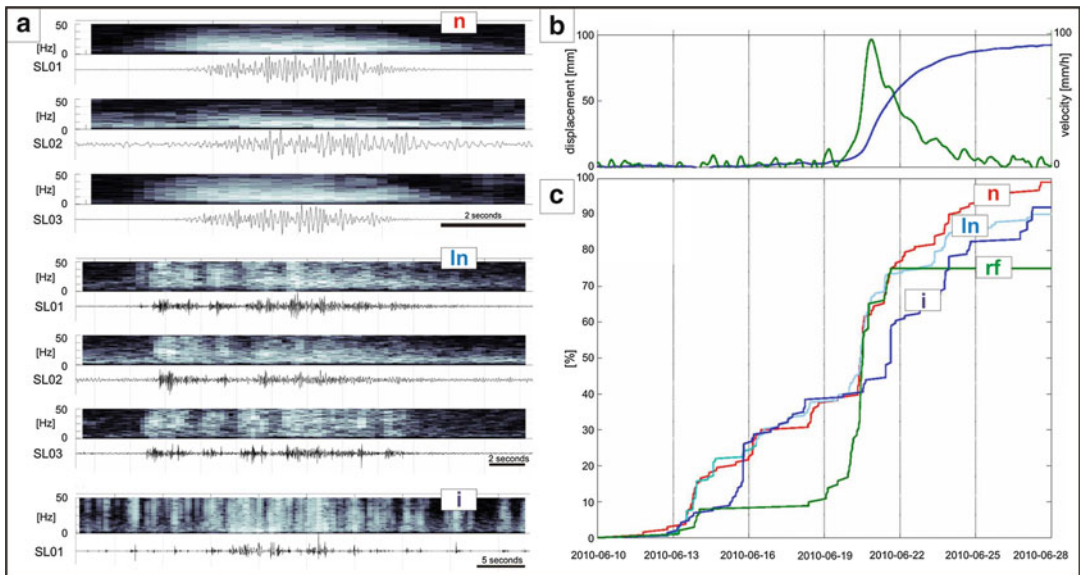
Seismic Sources from Landslides and Glaciers, Fig. 21 LMEs observed at the Steinlehnen DSGSD, 10–28 June 2010 (Weginger 2012); (**a**) typical waveforms and spectrograms of LME types n (emerging signal, few seconds length), ln (irregular sequence of n-type signals), i (irregular sequence of impulses over a period of several seconds); (**b**) displacement and velocity; (**c**) development of cumulative frequency of event types n, ln, i, and rf (rockfall)

M3 in Fig. 11a). Flow here represents quasistationary creep including episodic high activity, depending on infiltration and the pore pressure (Fig. 11b). The distribution of velocities in space and time within the landslide mass justifies its macroscopic description as a ductile medium. However, as the following two examples will prove, brittle deformations or interactions between pore fluid and the rock/soil mass also exist and resemble seismic sources.

The Super-Sauce landslide is a flow-type mass movement situated in the Barcelonnette basin in the southern French Alps (Fig. 12). It developed from heterogeneous soft Jurassic black marks (Fig. 24). The average slope of the flow is ~23°. The landslide triggered around 1960, ruling out a causal connection to the retreat of ice age glaciers. From 1996 to 2007, the velocities covered the range of 2–30 mm/day.

Seismic monitoring at the Super-Sauce landslide was carried out between 14 and 24 July 2008 (Walter et al. 2012). Besides rockfalls, 34 LMEs were located and labeled as "slide quakes." They have an impulsive onset, last 2–5 s, and cover a wide frequency band (10–80 Hz). Later-arriving phases have lower frequencies (Fig. 25a). The epicenters are concentrated around the high-velocity area of the landslide. Magnitudes between $-3.2 \leq M \leq -1.3$ were detected following a power law distribution with a b-value of $b = 0.84$. The hypocenters were probably located within the uppermost few meters and not at the basal sliding zone or surface. The highest magnitudes of slide quakes followed few hours after a strong rainfall event; the temporal distribution was otherwise uniform. Single miniarrays also detected very high-frequency events (5–150 Hz) with durations of 2–20 s (Fig. 25b). These LMEs were concentrated around a stable crest within the high-velocity zone of the flow and related to fissure development in the uppermost layer of the flow. Laboratory tests proved that the flow material can have brittle rheology.

The second example of a flow is the Slumgullion landslide (Fig. 26). The landslide occurs in
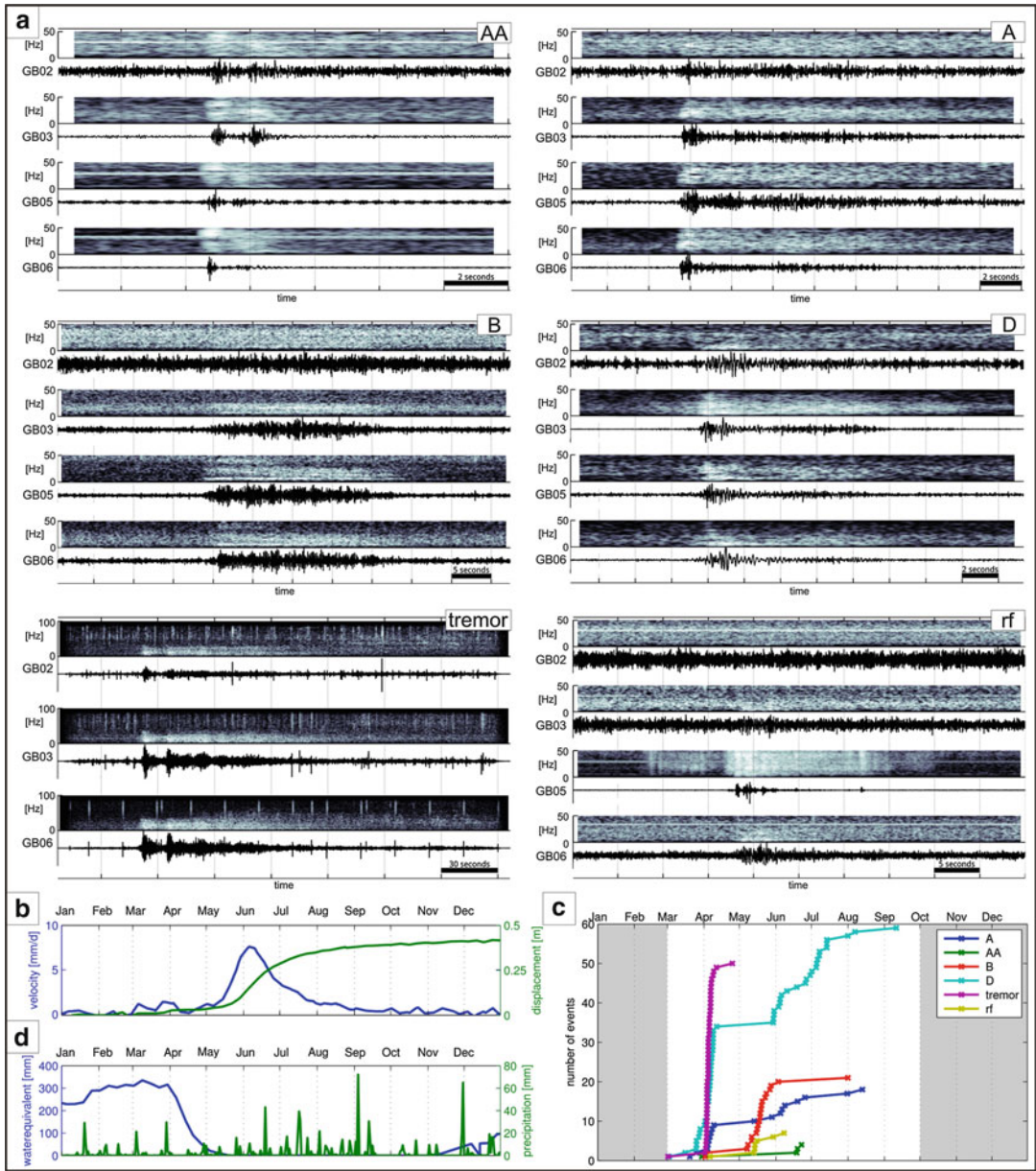
Seismic Sources from Landslides and Glaciers, Fig. 22 LMEs observed at the Gradenbach DSGSD, March–October 2009 (Brückl et al. 2013; Mertl 2015); (**a**) typical waveforms and spectrograms of LME types AA (impulsive onset with short lower frequency coda), A (like AA but with long coda), B (emergent onset, low frequency, typical duration 20–30 s), D (emergent onset but significantly higher frequency content and shorter duration than B), tremor (emergent low frequency signal, duration >1 min, superimposed by irregular spikes), rf (rockfall); (**b**) displacement and velocity; (**c**) snow cover water equivalent and precipitation; (**d**) development of cumulative frequency of event types

Seismic Sources from Landslides and Glaciers, Fig. 23 Spatial distribution of seismic activity March–October 2009 at the Gradenbach DSGSD; colored contour plot visualizes the frequency of possible epicenters with normalized PDF > 0.7, event types A, AA, D; *blue* area marks area of event type B epicenters (Brückl et al. 2013; Mertl 2015)
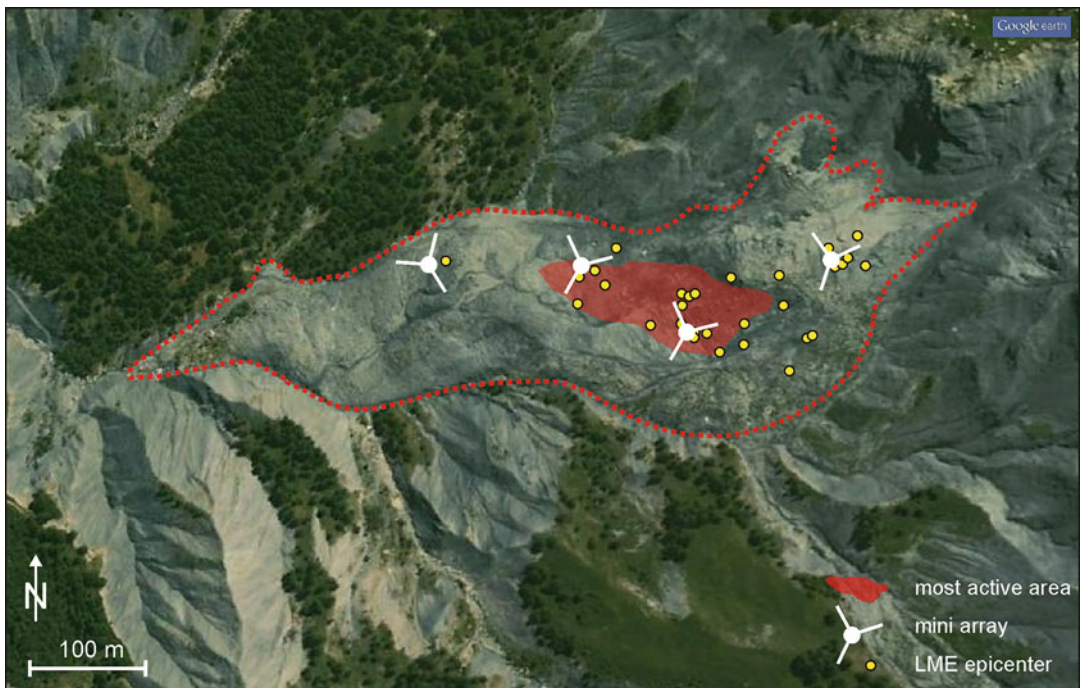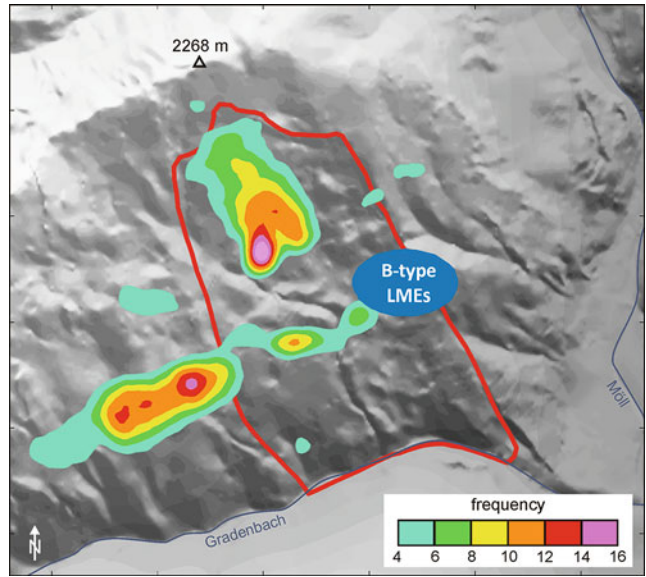
Seismic Sources from Landslides and Glaciers, Fig. 24 Super-Sauce flow, *red dotted line* outlines landslide area (Walter et al. 2012)

Tertiary volcanic rocks and comprises sandy, silty clay with scattered patches of boulder debris, clay, and pond and stream sediments. The landslide consists of an active upper part, which moved over an older, inactive lower part in its path. The average slope of the active part is 13°. The Slumgullion moves primarily by sliding along discrete bounding faults. Additionally,
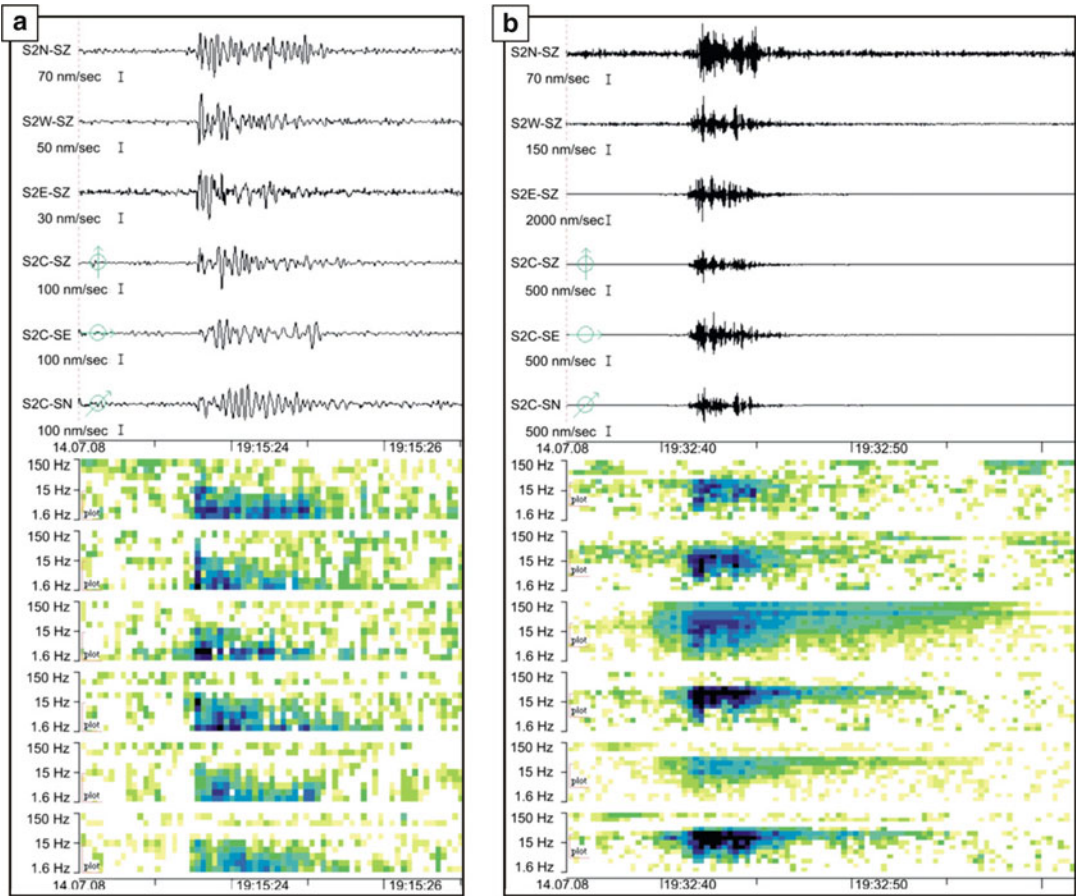
Seismic Sources from Landslides and Glaciers, Fig. 25 Typical waveforms and sonograms of LMEs observed at Super-Sauce debris/mud flow (Modified after Walter et al. 2012); (**a**) "slide quakes"; (**b**) "fissure development"

differential internal processes take place at faults and fractures within the landslide body.

During 18–26 August 2009, a seismic monitoring campaign was carried out at the Slumgullion landslide (Gomberg et al. 2011). Geodetic and hydrometeorological monitoring supplemented the campaign. During the monitoring campaign, the displacement trend was ~10 mm/day. One stronger period of rainfall occurred during the observation period. Several seismic signals were detected and related to LMEs: "tremor" and "harmonic slide quakes" (Fig. 27a, b). The dominant energy of tremor is distributed broadly above 30–50 Hz and

diminishes toward 125 Hz. High amplitudes of tremor envelopes last <10s and follow irregularly in intervals of several tens of seconds. Harmonic slide quakes last ~2 s, their fundamental frequency is ~12 Hz, and no harmonics are visible throughout the whole duration. This LME type occurs more frequently during daytime, and an increase in sliding velocity is also observed during the day. Tremor and harmonic slide quakes most probably originate at the left side-bounding strike-slip faults of the landslide. The stick–slip behavior of the side-bounding strike-slip faults is explained by a transient dilatant strengthening.
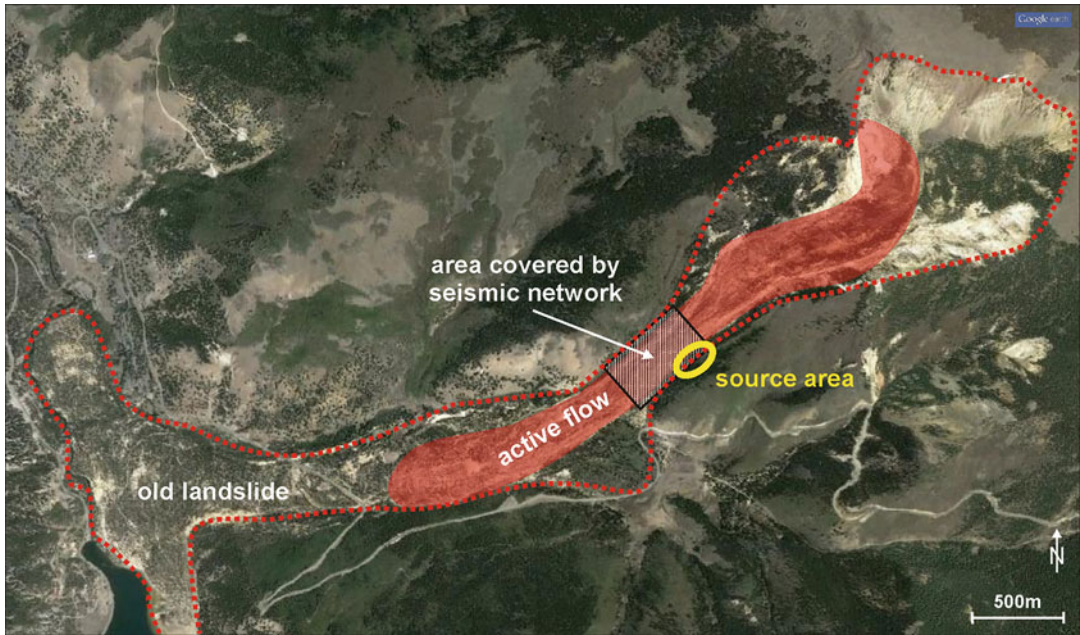
Seismic Sources from Landslides and Glaciers, Fig. 26 Slumgullion landslide (37° 59′ 30″ N, 107° 15′ 25″ W), San Juan Mountains, Colorado, USA (Gomberg et al. 2011)
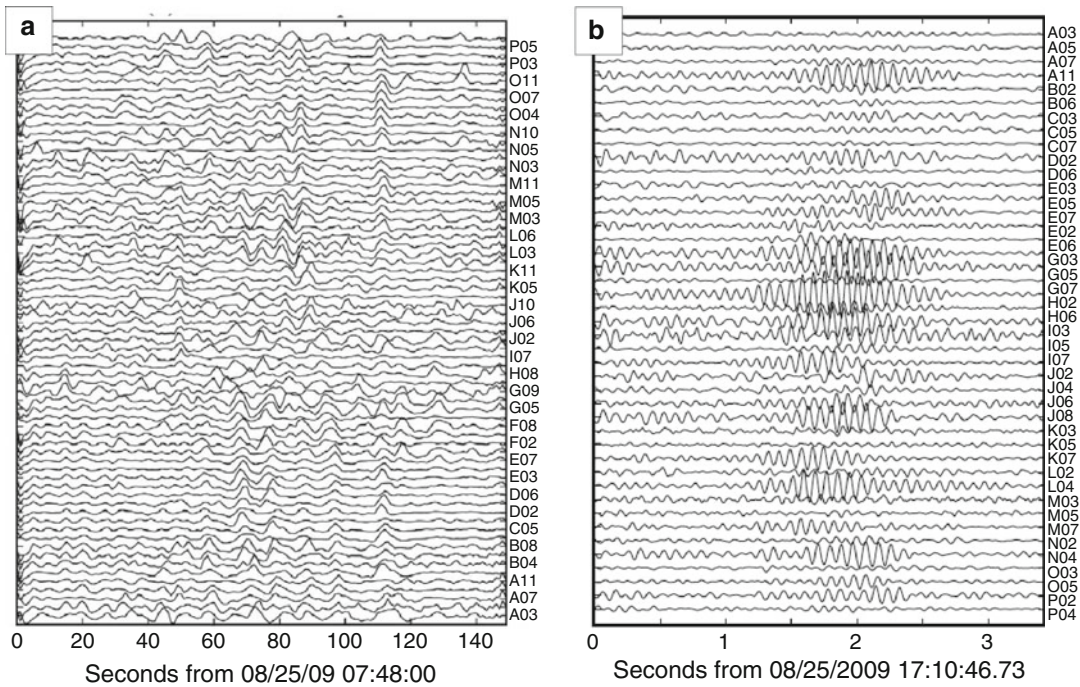
Seismic Sources from Landslides and Glaciers, Fig. 27 Typical LME waveforms observed at Slumgullion flow (Modified after Gomberg et al. 2011); (**a**) envelopes of tremor; (**b**) harmonic slide quake

Microearthquakes Generated by Glaciers

Ductile and Brittle Deformations of Glaciers

In contrast to landslides, which are intrinsic transient phenomena, glaciers or ice sheets can reach a long-lasting stationary state in steady climatic conditions. Snow is transformed to ice and added to the glacier in the accumulation area and removed in the ablation area. Glacier flow transports ice from the accumulation area to the ablation area, thereby keeping the total mass as well as the thickness of the glacier constant under stationary conditions. The glacier flow is driven by gravitational forces, controlled by the rheology of the ice and physical conditions that allow for sliding of the glacier ice over its bed (Cuffey and Paterson 2010).

We consider an infinite slab of ice with constant thickness H resting on a plane glacier bed with the slope α. Shear stress parallel to the glacier bed is proportional to $\sin(\alpha)$ and increases linearly with the overburden (Fig. 28a). According to Glen's law, the corresponding shear strain rate is proportional to the nth power of the shear stress, with n ~ 3.

Integration yields the velocity profile, which is an (n + 1)th-order parabola. For cold glaciers, the ice temperature is below melting point at the base, and the ice may be frozen to the glacier bed (Fig. 28b). For temperate glaciers, a thin coat of water between ice and the bedrock is assumed, and sliding occurs over the thin viscous water layer (Fig. 28c). Sliding at the glacier bed is controlled by obstacles, which are overcome either by enhanced flow due to stress concentration at the obstacles or regelation. All processes related to glacier flow described so far are either ductile (internal deformation of the ice, viscous flow of the water coat at the glacier bed, enhanced flow over obstacles) or correspond to another steady-state process (regelation) and cannot produce earthquakes.

However, the rheology of ice changes significantly in the case of the hydrostatic pressure component p of the state of stress becoming negative. In this case, the ice behaves in a brittle manner and may radiate seismic energy. First, we consider tensile crack nucleation and propagation without the influence of water. Glacier thickness adapts to varying slopes of the glacier bed by keeping the basal shear stress approximately constant. Therefore, the thickness is inversely proportional to $\sin(\alpha)$. Neglecting accumulation or ablation over a relatively short distance of an increasing slope, the velocity must increase to keep the flux of ice constant (Fig. 29). Longitudinal tensile stress must develop to extend the glacier ice moving from the gentle to the steeper slope. Surface crevasses develop

Seismic Sources from Landslides and Glaciers, Fig. 28 Planar slab of ice resting on a sloping plane; (**a**) shear stress parallel to surface; (**b**) velocity profile of a cold glacier – no basal sliding; (**c**) velocity profile of a temperate glacier – basal sliding

Seismic Sources from Landslides and Glaciers, Fig. 29 Crevasse formation due to thickness adaption to increasing slope; *stars* mark possible locations of GMEs

to the depth at which the hydrostatic pressure equals the longitudinal tensile stress component. On alpine glaciers, surface crevasse depths seldom exceed 25–30 m. After reaching the zone of constant steeper slope, the tensile stress component disappears and the crevasses close by ductile deformation.

Water significantly influences the mechanics of landslides and glaciers. The fundamental difference between landslides and glaciers with respect to the effect of water is that buoyancy is lower than hydrostatic pressure in landslides and higher in glaciers due to the comparative densities of water (~1000 kg m^{-3}), rock (~2500 kg m^{-3}), and ice (~900 kg m^{-3}). Therefore, a water column of sufficient height can introduce a tensile stress regime and lead to crack development and propagation at all glacier depths. Possible scenarios are shown in Fig. 30.

Thus far, we have only considered hard-bedded glaciers where the glacier bed consists of hard rock and is not deformed by the glacier flow. However, it is well known that fast-moving outlet glaciers in Alaska, Antarctica, and Greenland rest on a basal layer of soft glacial till, which makes up a considerable part of the basal sliding

(Alley et al. 1987). The mechanical strength of glacial till obeys the Mohr-Coulomb failure criterion and may be comparable to the masses of flow-type landslides treated in section "Flow." Although steady shear deformation may dominate, experience from seismic monitoring on landslides shows that stick–slip deformation may also occur. At the glacier beds of polythermal and temperate glaciers, there may also be asperities or "sticky spots," which induce stick–slip basal sliding under the otherwise nonseismic deformation conditions at the glacier bed.

In the following, we term seismic activity connected to the various processes of glacier flow and observed by local networks GMEs (glacial microearthquakes).

Glacial Lake Outburst Floods at an Alpine Glacier

As an example of seismic activity of an alpine glacier, we review observations made at the polythermal Gornergletscher, Monte Rosa Massif, Switzerland (Fig. 12). At the confluence of Gornergletscher and Grenzgletscher, there is the ice-dammed Gorner lake (Fig. 31). The lake

Seismic Sources from Landslides and Glaciers, Fig. 30 Hydraulically induced fracturing in a glacier (Modified after Walter 2009; Walter et al. 2010); *stars* mark possible locations of GMEs; (**a**) opening of internal and basal crevasses due to high water pressure; meltwater influx from active moulins or basal channels (**b**) closure of internal and basal crevasses due to decreasing water pressure; the collapse of a basal cavity generated during the high water pressure phase acts as seismic source
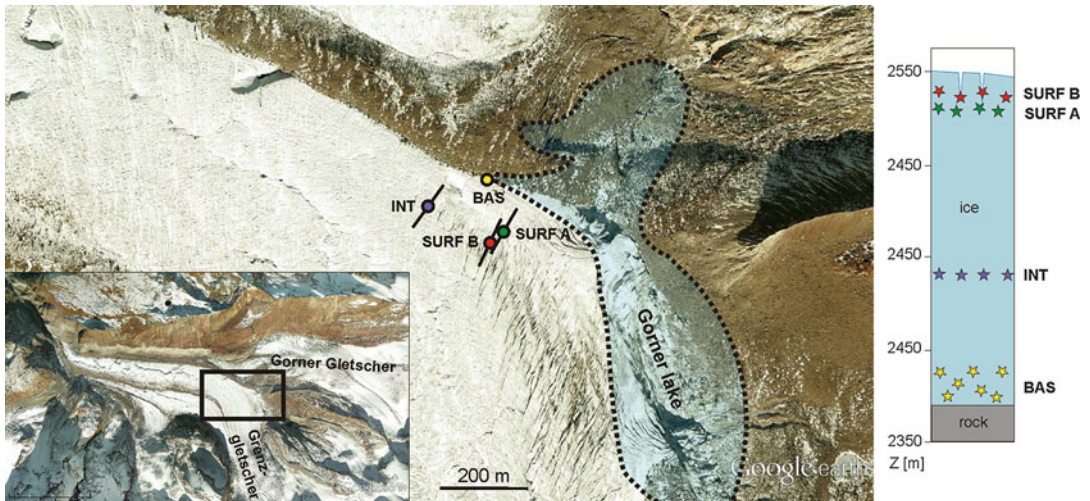
Seismic Sources from Landslides and Glaciers, Fig. 31 GMEs observed at the Gornergletscher before and during the GLOFs 2004 and 2006 (Walter 2009); (**a**) location of the GME clusters; fault plane strike of the SURF A, SURF B, and INT GME clusters agree with the orientation of crevasses; *dotted line* outlines approximately the Gorner lake at the maximum water levels; (**b**) focal depths of the different GME clusters

drains during annual glacial lake outburst floods (GLOFs), which were monitored in the years 2004 and 2006 with seismic and geodetic networks (Walter 2009; Walter et al. 2010). During the 2004 and 2006 field seasons, about 35000 and 50000 events were detected. Generally, diurnal variations in the seismicity were observed, with the peak in the early afternoon and much lower

GME activity during the nighttime and morning hours. GME clusters were recorded at three different depth levels: SURF_A and SURF_B at the depth of surface crevasses, INT at ~100 m depth, and BAS just above the glacier bed (Fig. 31b). The SURF_A, SURF_B, and INT GMEs have magnitudes between $-2.3 \leq Mw \leq -1.6$; the BAS GMEs reach $Mw = -0.7$. With the exception of BAS GMEs, fault planes are steeply dipping, and strike directions of the interpreted fault planes follow the general trend of surface crevasses (Fig. 31a). The radiation pattern of the SURF_A and INT LMEs can best be explained by the generation of tensile cracks, but the SURF_B events more likely result from a DC mechanism. The BAS GMEs occur on subhorizontal fault planes and are well explained by tensile cracks. The source mechanisms of the SURF_A, INT, and BAS cluster correspond to the processes schematically outlined in Figs. 29 and 30. The BAS events occurred mainly during otherwise quiescent night and early morning hours. This observation supports the idea of a collapsing basal cavity due to the reduced meltwater access and basal water pressure (Fig. 30b). The occurrence of DC-source mechanisms with the SURF_B GMEs may be due to the particular dynamic situation near the Gorner lake and the GLOF.

Arctic Glacier Prone to Surges

The temperate Bering Glacier flows from the St. Elias mountain range to its terminus on the south-central coast of Alaska and has a history of dramatic surges. Seismic monitoring was carried out between the equilibrium line altitude and the terminus (Fig. 32a) and covered the early melting period from 20 April until 19 June 2007 (West et al. 2010). Yet another surge started in 2008 (Burgess et al. 2012).

Around 160,000 events were detected during the observation period. Location of the GMEs was not attempted. The waveforms of the GMEs (Fig. 32b) show a bimodal distribution and range from high-frequency signals (20–35 Hz) to low-frequency signals (6–15 Hz). The high-frequency GMEs have impulsive source functions and were interpreted as brittle failures.

In an analogy to waveforms of volcanic earthquakes, the low-frequency tremor GMEs were explained as a fluid-driven crack model with a resonant water-filled cavity as the seismic source. Hybrid GMEs were interpreted as hydrofracturing followed by the rush of water into the new opening.

Relative maxima of the daily rate of low-frequency events precede the corresponding maxima of the high-frequency events by 1–6 days (Fig. 32c). The interpretation of this pattern is that the peaks of the low-frequency events reflect the subglacial rush in of water leading to a decoupling of the glacier from its bed and consequently enhanced flow velocities. The enhanced glacier motion is accompanied by an increase of brittle failures. Velocity data would be essential to further develop this interpretation.

Stick–slip Movement of an Antarctic Ice Stream

Bindschadler et al. (2003) reported on major West Antarctic ice streams discharging in sudden periods of rapid motion, in particular Whillans Ice Stream (WIS) (Fig. 33a), which has an area of 200 km × 100 km with an average ice thickness of ~600 m. During December 2010 and January 2011 and again during December 2011, Winberry et al. (2013) deployed a GPS and seismic network on the WIS in order to monitor the slip nucleation and the subsequent period of rapid motion (Fig. 33b). Figure 33c shows the displacements for a time span of ~1 h, covering the end of a preceding interevent period, the nucleation phase, the rupture, the main-slip phase, and the transition to the following interevent period. Clear evidence for the nucleation phase is shown only at one GPS station (station 1), where the rupture occurs within ~5 min at all stations according to the rupture propagation speed of >500 m/s. The main-slip phase lasts ~30 min and terminates almost simultaneously at all stations.

GMEs recorded at the individual stations correlate well in time and frequency domain with the ice stream velocities (Fig. 34). At station 1, seismicity increases at the beginning of the
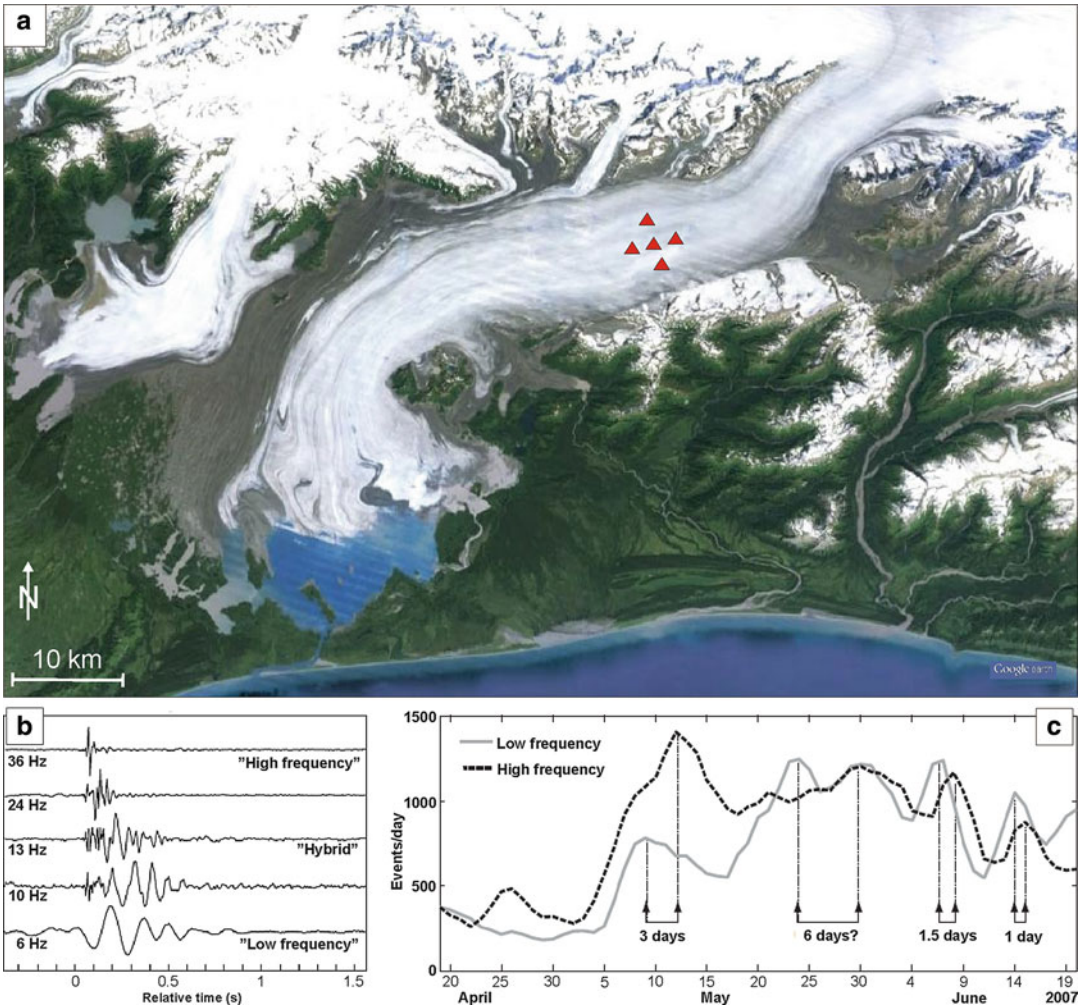
Seismic Sources from Landslides and Glaciers, Fig. 32 Seismic monitoring at the Bering glacier (Modified after West et al. 2010); (**a**) location of seismic stations (*red triangles*); (**b**) waveforms; (**c**) temporal evolution of daily rates of high and low frequency event

nucleation phase but only begins at the other stations around the rupture time. Two different styles of GMEs or microseismicity can be distinguished. One type (Fig. 34a) is built by a sequence of fewer than a hundred to several thousand individual impulsive events. The wide distances between the seismic stations did not allow for a precise location; however, the relative timing of P and S wave arrivals and the lack of crevasses is consistent with source locations at or near the glacier bed. The individual events can be grouped into families with a high degree of similarity between each other. The same event families persist in subsequent high-slip-rate phases. This observation is an indication of the temporal stability of subsurface asperities representing the source regions of the individual events. Other locations show in general a similar behavior but are characterized by lower frequencies, indicating larger source areas. The lack of microseismic activity or GMEs at some locations may be explained by the presence of more continuous fault gouge and higher water pressure at the glacier bed providing basal lubrication.

The second style of microseismicity is the emergent tremor (Fig. 34b). It is continuous and
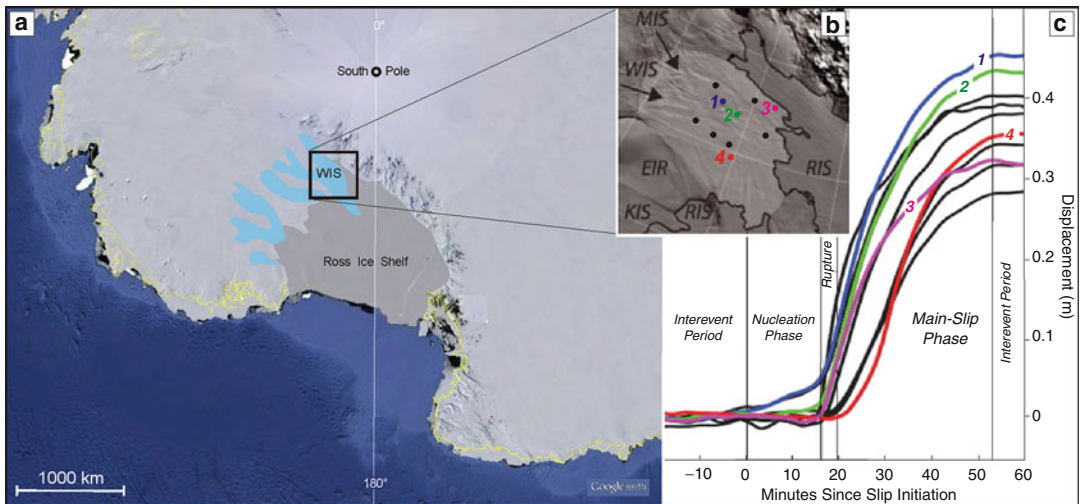
Seismic Sources from Landslides and Glaciers, Fig. 33 Geodetic and seismic monitoring of an episodic high flow velocity phase at the Whillan Ice Stream (WIS); (**a**) location map; (**b**) zoom of the monitoring area with locations of monitoring-stations; (**c**) displacements versus time since slip initiation; numbers and colours assign displacement data to stations in (**b**) (Modified after Winberry et al. 2013)

cannot be resolved into individual events. The overall spectral power of this tremor correlates well with the glacial flow velocities. Seismic energy during the tremor is concentrated in the horizontal components. A source, radiating shear energy from near the glacier bed, is consistent with this observation. The spectra reveal coherent gliding spectral lines, some of which are overtones of fundamental frequencies. Periodic stick–slip failure of small asperities could be the source mechanism of these seismic signals. Winberry et al. (2013) argue that the highly repeatable WIS system represents an excellent natural laboratory which could provide insight into friction and sliding complementary to laboratory or other field studies (e.g., Gomberg et al. 2011).

Summary

A new class of global earthquakes (Ms > 4.5) was discovered and related to giant and catastrophic landslides or phenomena occurring near the front of fast-moving outlet glaciers in Greenland and Antarctica. The unique characteristics of these earthquakes are the depletion of short-period energy, the elongated source time functions, and their radiation patterns, which can be explained by a single surface force better than by seismic moments. Efficient detection and location methods were developed and force histories F(t) along with the centroid single force (CSF) derived. The CSF is the product of the total moving mass M and the displacement D of the centroid. It represents a major constraint on the whole source process.

Originally, the new-class global earthquakes with sources on glaciers were interpreted as stick–slip flow behavior of large portions of outlet glaciers in Greenland and Antarctica. However, the postulated movements of the ice masses could not to date be confirmed geodetically. An alternative source mechanism which generates sufficient CSF and provides the appropriate duration of the source function is the capsizing (or toppling) of large icebergs at the front of the outlet glaciers during major calving events. The basic mechanics of this process are well understood, and correlations with glacier flow, seasonal variations of iceberg production, and glacier retreat support the explanation of the
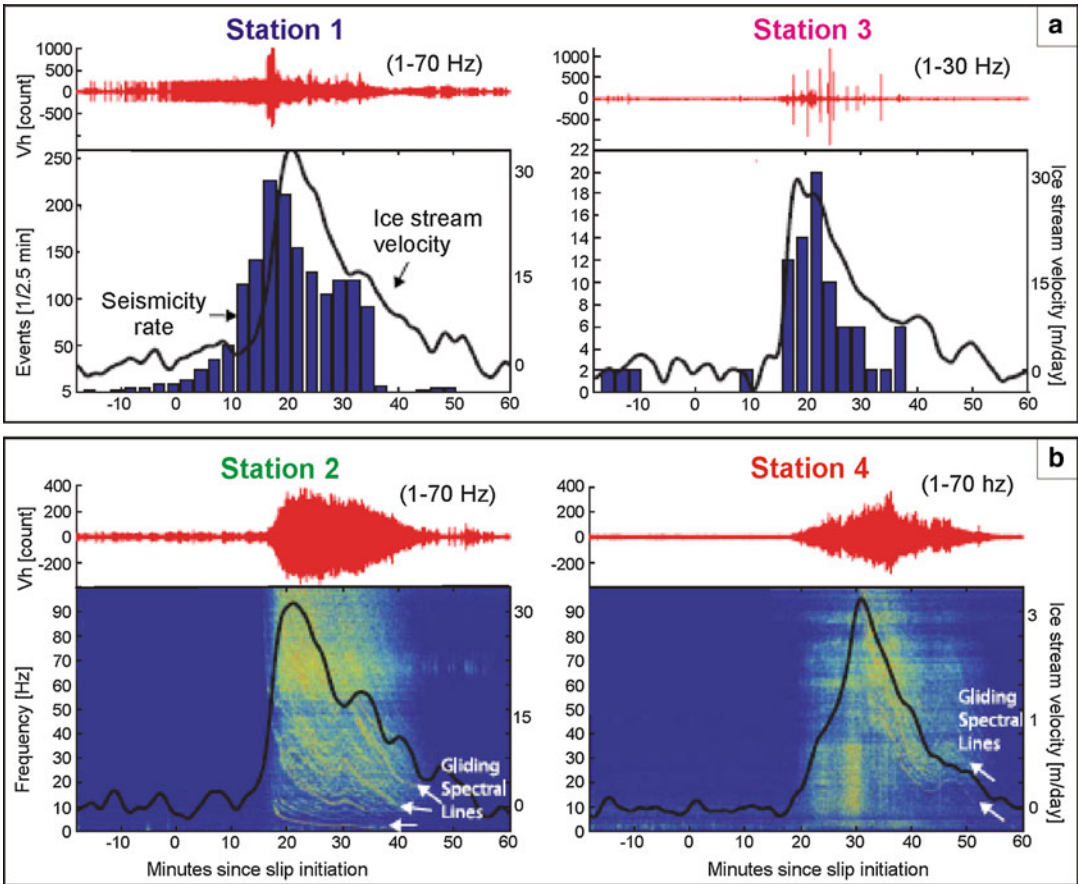
Seismic Sources from Landslides and Glaciers, Fig. 34 Waveforms observed at stations 1–4 at WIS; (**a**) impulsive, high frequency waveforms; (**b**) low frequency (Modified after Winberry et al. 2013)

majority of these earthquakes coming about through this source process.

Related to landslide-generated global earthquakes, additional information from satellite images or local reconnaissance helps to separate CSF into mass and total displacement. The inclination of the CSF vector corresponds to the Fahrböschung; the duration of the landslide force history constrains the sliding velocity. Application of a rotational slider block model allows, e.g., for estimates of the reduction of the friction coefficient from static to dynamic conditions.

Microearthquakes related to the processes in landslide masses (LMEs) that evolve in a more steady form or those in a post- or prefailure state have magnitudes of M < 0. Seismic networks

deployed on or very near to the landslide mass are necessary to capture these signals. A rich variety of waveforms can be observed and classified by attributes derived from the frequency content, the duration, or the shape of the envelope. Currently, no definite assignment of the different waveforms to landslide types and processes can be established. Impulsive events have been observed on slopes with steep or medium inclinations (>20°); tremor or harmonic LMEs dominate the seismic activity of gently sloping flows. Most seismic activity on landslides is rather shallow. However, the location of a few events at the basal sliding zone cannot be excluded. Rockfall and LME frequency correlate frequently with precipitation and precede geodetically observed acceleration. An increase

in seismic activity preceding acceleration without discernible external triggers has also been observed. The general pattern of rockfall and LME activity suggests that the DSGSD- and flow-type landslides are near their yield stress not only at the basal sliding zone but also in their entire moving mass. Further progress in the understanding of LMEs would be supported by denser seismic networks including borehole seismometers as a standard. Long-term seismic monitoring covering seasonal variations and episodic high-velocity phases combined with complementary monitoring (e.g., geodetic, meteorological, hydrological, geomorphologic) are preconditions for testing refined hypotheses about LMEs and their relation to landslide processes and evolution.

Monitoring of glacier flow-related microearthquakes (GMEs) by seismic networks deployed on glaciers and supplemented by geodetic and meteorological devices provides a deeper insight into the stress distribution within the glacier ice, glacier hydraulics, and processes at the glacier bed or within a basal layer of soft sediments. The latter aspect can be exemplarily studied at the Whillans Ice Stream (WIS) in Antarctica. The ice stream exhibits a stick–slip behavior, controlled by tides with a CSF corresponding to $Ms = 5.6$. The duration of this process is above the seismically observable range, but GMEs generated by this process yield unique information about the spatial distribution and mechanical behavior of asperities or "sticky spots" in the basal layer of soft sediments. Research in this field does not only support the understanding of glacier sliding but addresses generally the friction behavior of fault gouges, e.g., the basal sliding zones of landslides or active tectonic faults.

Finally, we touch on how the potential monitoring and analyzing of landslide- and glacier-generated earthquakes contribute to prediction and early warning of catastrophic events. Global landslide or glacial earthquakes always represent the finite state of a development and may message a disaster. Prediction and early warning is not provided by the detection of these earthquakes; however, instantaneous location and quantification by seismic methods may support mitigation measures and warnings against subsequent hazards like damming of the valley (see Fig. 6c) or increased density of icebergs. LMEs have been proven to indicate a dangerous acceleration of a landslide, usually earlier than the geodetically observable displacement. GMEs precede glacier break-off and glacial lake outburst floods (GLOFs). The results from basic research on LMEs and GMEs lead to a recommendation of seismic monitoring as an essential component of an integrated early warning system for landslides and glacial hazards.

Cross-References

▶ Earthquake Location
▶ Earthquake Mechanism Description and Inversion
▶ Mechanisms of Earthquakes in Iceland
▶ Non-Double-Couple Earthquakes
▶ Rockfall Seismicity Accompanying Dome-Building Eruptions
▶ Volcanic Eruptions, Real-Time Forecasting of

References

Alley RB, Blankenship DD, Bentley CR, Rooney ST (1987) Till beneath Ice Stream B. Till deformation: evidence and implications. J Geophys Res 92(B9):8921–8929

Amitrano D, Grasso JR, Senfaute G (2005) Seismic precursory patterns before a cliff collapse and critical point Phenomena. Geophys Res Lett 32, L08314. doi:10.1029/2004GL022270

Bindschadler RA, King MA, Alley EB, Anandakrishnan S, Padman L (2003) Tidally controlled stick–slip discharge of a West Antarctic ice stream. Science 301(5636):1087–1089. doi:10.1126/science.1087231

Brantley SR, Myers B (2000) Mount St. Helens – from the 1980 eruption to 2000. U.S. Geological Survey Fact Sheet, FS-036-00, 2 p. http://pubs.er.usgs.gov/publication/fs03600

Brückl E, Parotidis M (2005) Prediction of slope instabilities due to deep-seated gravitational creep. Nat Hazards Earth Syst Sci 5:155–172. SRef-ID: 1684-9981/nhess/2005-5-155

Brückl E, Brunner FK, Lang E, Mertl S, Müller M, Stary U (2013) The Gradenbach observatory-monitoring deep-seated gravitational slope deformation by geodetic,

S

hydrological, and seismological methods. Landslides. doi:10.1007/s10346-013-0417-1

Burgess EW, Forster RR, Larsen CF, Braun M (2012) Surge dynamics on Bering Glacier, Alaska, in 2008–2011. Cryosphere 6:1251–1262. doi:10.5194/tc-6-1251-2012

Chen X, Shearer PM, Walter F, Fricker HA (2011) Seventeen Antarctic seismic events detected by global surface waves and a possible link to calving events from satellite images. J Geophys Res 116, B06311. doi:10.1029/2011JB008262

Cuffey KM, Paterson WSB (2010) The physics of glaciers, 4th Edition. Academic Press, ISBN 978012333694614, p 704

Eberhardt E, Stead D, Coggan JS (2004) Numerical analysis of initiation and progressive failure in natural rock slopes – the 1991 Randa rockslide. Int J Rock Mech Min Sci 41:69–87

Ekström G, Stark CP (2013) Simple scaling of catastrophic landslide dynamics. Science 339(6126):1416–1419. doi:10.1126/science.1232887

Ekström G, Nettles M, Abers GA (2003) Glacial earthquakes. Science 302:22–624. doi:10.1126/science.1088057

Faillettaz J, Funk M, Sornette D (2011) Icequakes coupled with surface displacements for predicting glacier break-off. J Glaciol 57(203):453–460

Gischig V, Moore JR, Evans KF, Amann F, Loew S (2011) Thermomechanical forcing of deep rock slope deformation: 2. The Randa rock slope instability. J Geophys Res 116:F04011. doi:10.1029/2011JF002007

Gomberg J, Schulz W, Bodin P, Kean J (2011) Seismic and geodetic signatures of fault slip at the Slumgullion Landslide Natural Laboratory. J Geophys Res 116, B09404. doi:10.1029/2011JB008304

Goodman RE, Kieffer DS (2000) Behaviour of rock in slopes. J Geotech Geoenviron 128(8):675–684

Helmstetter A, Garambois S (2010) Seismic monitoring of Séchilienne rockslide (French Alps): analysis of seismic signals and their correlation with rainfalls. J Geophys Res 115, F03016. doi:10.1029/2009JF001532

Kanamori H, Given JW (1982) Analysis of long-period seismic waves excited by the May 18, 1980, eruption of Mount St. Helens – a terrestrial monopole? J Geophys Res 87:5422–5432. doi:10.1029/JB087iB07p05422. ISSN: 0148–0227

Kawakatsu H (1989) Centroid single force inversion of seismic waves generated by landslides. J Geophys Res 94:12,363–12,374. doi:10.1029/89JB01118. ISSN: 0148–0227

Mertl S (2015) Characterization of local seismic events on the deep-seated events gravitational slobe deformation Gradenbach. PhD thesis, Vienna University of Technology, p 176

Moore J, Albee W (1981) Topographical and structural changes, March-July 1980 – photogrammetric data. In: Lipman P, Mullineaux D (eds) The 1980 eruptions of Mount St. Helens Washington. Geological Survey Professional Paper 1250, pp 123–134

Nettles M, Ekström G (2010) Glacial earthquakes in Greenland and Antarctica. Annu Rev Earth Planet Sci 38:467–491. doi:10.1146/annurev-earth-040809-152414

Nettles M, Larsen TB, Elósegui P, Hamilton GS, Stearns LA, Ahlstrøm AP, Davis JL, Andersen ML, de Juan J, Khan SA, Stenseng L, Ekström G, Forsberg R (2008) Step-wise changes in glacier flow speed coincide with calving and glacial earthquakes at Helheim Glacier, Greenland. Geophys Res Lett 35, L24503. doi:10.1029/2008GL036127

Spillmann T, Maurer H, Green AG, Heincke B, Willenberg H, Husen S (2007) Microseismic investigation of an unstable mountain slope in the Swiss Alps. J Geophys Res 112, B07301. doi:10.1029/2006JB004723

Tsai VC, Ekström G (2007) Analysis of glacial earthquakes. J Geophys Res 113:F03S22. doi:10.1029/2006JF000596

Tsai VC, Rice JR, Fahnestock M (2008) Possible mechanisms for glacial earthquakes. J Geophys Res 113, F03014. doi:10.1029/2007JF000944

Voight B (1981) Time scale for the first moments of the May 18 eruption. In: Lipman P, Mullineaux D (eds) The 1980 eruptions of Mount St. Helens Washington. Geological Survey Professional Paper 1250, pp 68–86

Walter F (2009) Seismic activity on Gornergletscher during Gornersee outburst floods. PhD thesis, DISS. ETHNo 18184, ETH Zürich, p 135

Walter F, Dreger DS, Clinton JF, Deichmann N, Funk M (2010) Evidence for near-horizontal tensile faulting at the base of Gornergletscher, a Swiss Alpine Glacier. Bull Seismol Soc Am 100(2):458–472. doi:10.1785/0120090083

Walter M, Arnhardt C, Joswig M (2012) Seismic monitoring of rockfalls, slide quakes, and fissure development at the Super-Sauze mudslide, French Alps. Eng Geol 128:12–22

Weginger S (2012) Detektion und Lokalisierung seismischer Signale zur Überwachung der Massenbewegung Steinlehnen. Master thesis, Vienna University of Technology, p 96

West ME, Larsen C, Truffer M, O'Neel S, LeBlanc L (2010) Glacier microseismicity. Geology 38(4):319–322. doi:10.1130/G30606.1

Winberry JP, Anandakrishnan S, Wiens DA, Alley RB (2013) Nucleation and seismic Tremor associated with the glacial earthquakes of Whillans Ice Stream, Antartica. Geophys Res Lett 40:312–315. doi:10.1002/grl.50130

Zangerl C, Eberhardt E, Schönlaub H, Anegg J (2007) Deformation behavior of deep-seated rockslides in crystalline rock. In: Rock mechanics: proceedings of the 1st Canada-US rock mechanics symposium, Vancouver, 27–31 May 2007. doi:10.1201/NOE0415444019-c112

Zischinsky U (1969) Über Sackungen. Rock Mech 1:30–52

Seismic Strengthening Strategies for Existing (Code-Deficient) Ordinary Structures

Kosmas-Athanasios Stylianidis
Civil Engineering Department, Aristotle
University of Thessaloniki, Thessaloniki, Greece

Synonyms

Intervention procedures to upgrade existing ordinary structures which are designed and executed according to older codes

Introduction

Scope of the Entry
This entry presents basic principles, criteria, and procedures in regard to the conceptual design of seismic strengthening of existing ordinary structures and its treatment in current codes. In this context, certain code provisions constitute a significant part of the present entry.

Although it can be read by nonexperts, this entry is mainly addressed to structural engineers involved in the design of structures and interventions in buildings.

Notes on Terminology Used
The terms used in the entry are standard and familiar to engineers; nevertheless some clarifications are in order:

Seismic: The entry is restricted to seismic upgrading only. Certain aspects of the methodology may also apply to upgrading against other actions. However, it is pointed out that the main characteristic of a seismic strengthening strategy is the fact that an earthquake is an accidental load. In this sense, it is possible to define target performance levels, dependent on the level of the design earthquake, which in general can be related to the accepted risk. This probability leads to a compromise between "desirable" and "feasible," i.e., instead of a practically unfeasible (due to cost or disturbance of use) high target, the adoption of a lower, yet feasible, target is more pragmatic. This option is typically not available in the case of upgrading against permanent loads.

Intervention: The term structural intervention implies any operation that results in the desired modification of existing mechanical characteristics of a member or a structure and has as a consequence the modification of its response.

Repair: The term implies the process of intervention to a structure damaged by any cause that reinstates the mechanical characteristics of its structural members to their pre-damage level and restores its original structural capacity.

Strengthening: The term implies the process of intervention to a structure with or without damage, which increases the capacity and/or ductility of a member or the entire structure to a level higher than that prescribed in the original design.

Strategy: It is a high-level plan to achieve certain objectives, usually under uncertain conditions.

Technique: It is a procedure to be implemented to complete a task (see also Wikipedia 2014).

Code Deficient: The term refers to buildings designed according to older seismic codes or without any code. It is typically the case that contemporary codes introduce significantly higher demands than older codes and this is the main reason that upgrading is needed.

Ordinary Structures: This term includes mostly above surface R/C building structures. Underground, special structures (bridges, chimneys, etc.) and historical buildings fall beyond the scope of this entry. The same principles can be applied to structures made of different structural materials, such as steel, masonry, and wood, after suitable adaptation.

Sources
The entry is primarily based on two sources, the European Standard EN 1998–3 (2005) and the Greek Code of Structural Interventions (2012), which standardize the existing knowledge on upgrading of ordinary structures.

S

Assessment

General

The assessment of an existing structure is not strictly mandatory, but it is valuable, since it reveals the weaknesses of the structural system and indicates the strengthening measures to be taken. In this sense, the assessment is integrated in the design strategy and is strongly recommended.

The assessment of existing structures follows the steps below:

- Collection of data (investigation of structural history)
- Analysis (see Section Methods of Analysis for Assessment and Redesign of the present entry)
- Verification against limit states (see Section Compliance Criteria for Assessment and Redesign of the present entry)

Scope

(a) The purpose of the assessment of an existing structure is the evaluation of its available bearing capacity and the verification that the minimum mandatory requirements imposed by the existing codes are met.

(b) To estimate the available bearing capacity of the structure, the data from the structural history survey should be taken into account (see Greek Code of Structural Interventions (2012), Chap. 3, and European Standard EN 1998–3 (2005), Chap. 3).

(c) The designer should prescribe and supervise a series of investigating works in order to document and justify the assumptions on which the assessment will be based.

(d) The process of assessment differs depending on the existence or otherwise of damage in the building assessed.

(e) In case of no damage, the result of the assessment, depending on the selected redesign objective, will dictate the decision for potential retrofit.

(f) In the presence of damage, due to any past actions (prescribed or not by the standards applicable up to then), the assessment process includes two phases:

    (i) First, the structure is assessed as it is, taking damage into account. Depending on the selected redesign objective, the result of the assessment might lead to a decision for intervention (repair and/or retrofit).

    (ii) In case that intervention is required, the structure is assessed at its pristine condition, i.e., assuming that damage will be repaired. Depending on the selected redesign objective, the result of this assessment will lead to the decision for repair only or for repair and retrofit.

Collection of Data

The collection of the data required for the assessment should be governed by the following principles:

(a) The data required for the assessment of the bearing capacity of existing structures should be cross-verified and properly calibrated, wherever possible.

(b) It is recommended that the program of field and laboratory investigations be designed, executed, and supervised by the designer of the assessment, according to the specific design requirements.

(c) The reliability of the data collected should be properly taken into account in the assessment of the existing structure and the development of the intervention strategy.

Data Reliability Level (DRL)

General

(a) The reliability level of data (DRL) related to actions or resistances indicates the adequacy of the information regarding the existing building and is taken into account in the assessment and redesign. In existing structures, the numerical values of the data required for the assessment and redesign are usually characterized by a higher uncertainty than in the case of new structures.

(b) DRL is not necessarily the same for the entire building. Individual DRLs for the various subcategories of information can be determined. For the selection of methods of analysis, the most unfavorable among the individual DRLs are used.

(c) The concept of DRL is also applied in assessing the completeness of the survey of the structure and infill walls, especially in case of hidden elements. The effects of uncertainties can be taken into account in actions or resistances depending on the case (e.g., uncertainty in the thickness of the flooring of the slab will be taken into account in actions; uncertainty in the thickness of the slab itself will be considered mainly in the resistances).

Description of DRLs

(a) According to the European Standard EN 1998–3 (2005), three knowledge levels of data are adopted: (KL) 1–3 (limited, normal, full). According to the Greek Code of Structural Interventions (2012), Chap. 3, these levels correspond to high, satisfactory, and tolerable data reliability level.
(b) Regarding the self-weight, the characteristic value considered must be the most unfavorable value that is consistent with the geometry of the structure and/or applies for such structures.
(c) Regarding the resistances, their values can be determined from the dimensions, reinforcement, and material characteristics that lead to the justification of prior behavior of the structure. So, for example, a strength value that corresponds to the ultimate resistance of a cross section under the existing acting loads can be used. Similarly, dimensions of inaccessible foundations can be estimated with the assumption that they correspond to the ultimate soil bearing capacity under acting loads, etc.

Impact of DRL on the Assessment and Redesign

Depending on the reliability of the data:

(a) The appropriate safety factors γ_f for certain actions with uncertain values are selected (see Greek Code of Structural Interventions 2012, Chap. 4). Such may be the case for the representative values of some indirect actions (e.g., soil pressure) and the weight of inaccessible infill walls or coating/plastering.

In certain cases of high uncertainty (and if it is considered that the influence of the magnitude of the corresponding action is significant), consideration of two "reasonably extreme" representative values ($S_{k, \, min}$ and $S_{k, \, max}$) is recommended.

(b) The appropriate safety factors γ_m are selected according to the data for existing materials (see also Greek Code of Structural Interventions 2012, Chap. 4). TMaterial data are the dimensions of members and strengths of concrete and reinforcing steel, as well as the actual reinforcement detailing, anchoring, starter bars etc. that determine the resistances.

Assessment Principles

General

Assessment of existing structures follows the principles listed below:

(a) When the existing structural system is expected to contribute to the redesigned structural system solely by resisting vertical loads, its assessment may be performed based on simple, yet conservative, methods.

In this case, the accuracy of the assessment method used should be adjusted to the desired goal. For instance, an approximate, yet conservative, assessment method is sufficient to demonstrate the adequacy of the existing structural system against vertical loads. When the existing structural system is apparently inadequate and is expected to be fully demolished, its assessment is not necessary.

(b) However, when the existing structural system is expected to contribute to the redesigned structural system by resisting both vertical and seismic loads, it should be assessed based on the principles below. For the assessment of the structure against vertical loads, it is possible to use the methods prescribed by Eurocode 2 (EN 1992-1-1 2004), appropriately adapted:
   (i) The assessment is made using analytical methods (see Section Methods of Analysis for Assessment and Redesign

S

of the present entry). Especially in the case of structures without damage, for which the approved design study (which has been applied) is available, the assessment could be based on it.

(ii) The numerical models to be used for the assessment may represent the entire structure or individual members. Different numerical models may be used, depending on the type of the imposed actions. In general, the types of numerical models should be selected consistently with the calculation methods to be applied. Advanced methods of analysis should be combined with detailed numerical models.

(iii) It is recommended that the accuracy of the methods used be compatible with the accuracy of the data. High-quality numerical analysis is reliable only if it is based on equally reliable data.

(iv) The use of empirical–analytical or purely empirical methods is allowed only in cases covered by relevant special provisions issued by the competent public authority. Such specific provisions can be issued, provided that they refer to a building stock with common, known, features and that they are always preceded by a proper investigation.

(v) In cases of structures that have already suffered damage or deterioration, the applied assessment method must be able to interpret, as a rough approximation, both the type and the location of this significant damage. In structures of high importance, where damage has been identified, parametric analyses may be required to achieve the interpretation of damage based on their type and location.

The possible interpretation of damage, in terms of type and location, constitutes an acceptance criterion for the analytical methods used. Possible parameters may include non-visible geometrical data, mechanical characteristics that have not been investigated, random

combinations of actions assumed to have been applied in the past, etc.

(vi) For analysis, limit state control, verification of the adopted behavior factor, control of the imposed displacements, and local ductility indices, Sections Methods of Analysis for Assessment and Redesign and Compliance Criteria for Assessment and Redesign of the present entry are applied. In the case of masonry walls, paragraph Consideration of Masonry Infill Walls is applied.

(vii) In many cases, a quick assessment of the loss of bearing capacity of a damaged or deteriorated structure may be useful and/or necessary. This estimate can be made based on the intensity and extent of damage, as derived according to well-established (refined or approximate) methods.

Consideration of Masonry Infill Walls

(a) It is not permitted to consider masonry infill walls as part of the system that carries non-seismic actions (such as gravity loads). To calculate the internal forces of the structure due to non-seismic actions, numerical models are used that do not include masonry infill walls or do not impose stresses (induced by vertical loads) to them.

(b) It is recommended to consider masonry infill walls as part of the seismic action resisting system.

(c) It is mandatory to consider masonry infill walls as part of the system resisting seismic actions, when this decision has an adverse effect on the results obtained for the structural system at a global or local level.

The assessment of whether the influence of infill walls is favorable or unfavorable has to be made by the designer; however, the difficulty of the assessment has to be noted, particularly in the case that analysis data and calculations are not available. As a result, the above assessment will be on the safety side, if the masonry infills are introduced in advance to the numerical analysis models.

Redesign

General

The redesign of existing structures follows these steps:

- Conceptual and preliminary design
- Analysis (see Section Methods of Analysis for Assessment and Redesign of the present entry)
- Verification against limit states (see Section Compliance Criteria for Assessment and Redesign of the present entry)

Conceptual and Preliminary Design

(a) According to the criteria and the types of structural interventions, to be presented in Section Selection Criteria and Types of Structural Interventions, an intervention strategy is devised and the type and extent of interventions is decided.

Decisions on the appropriate strategy and the subsequent type of interventions for each case should be made on the basis of all information obtained during the assessment stage of the existing structure. The perception of the overall behavior of the building and the identification of its weaknesses, such as inadequate strength or stiffness or ductility, the unfavorable structural system, individual weaknesses, etc., should be dominant in the decision-making process.

Regardless of the analysis method that will be eventually adopted for the redesigned structure, inelastic static analysis may provide substantial assistance in identifying these weaknesses. Furthermore, with the aid of the above method, it is feasible to preliminarily select the characteristics of the types of intervention that will be prioritized.

(b) In all cases this selection should be justified (compared with other possible options), while the anticipated post-intervention behavior of the building should also be described qualitatively.

(c) A preliminary estimate of the dimensions and strength of the materials used and the modified stiffness of the structural elements undergoing intervention should be made.

(d) Preliminary estimate is made of the ductility class that the structure will fall into after the intervention, or (in case of application of inelastic static analysis) preliminary estimate is made of either the amplitude of the target displacement or the tolerable rotation angle of all structural members after intervention.

Assessment and Redesign Objectives

General

(a) In order to satisfy broader socioeconomic needs, various "performance levels" (target behaviors) are stipulated under relevant prescribed design earthquakes.

(b) The objectives of the assessment or redesign (Table 1; see Greek Code of Structural Interventions 2012, Chap. 2.2) are combinations of both a performance level and a seismic action, given an "acceptable probability of exceedance within the technical life cycle of the building" (design earthquake). The European Standard EN 1998–3 (2005) specifies that these levels should be defined in the National Annex (for each country).

(c) In the present entry, objectives that refer solely to the structural system are prescribed. In contrast, no objectives are set for the nonstructural system.

The relevant fundamental requirements, set by Chap. 2.1 of European Standard EN 1998–3 (2005), are satisfied through Table 1. In the case of two reassessment objectives, the possible pairs are B1 and A2 or C1 and B2.

The term "structural system" is used here in the classical sense and corresponds to the system bearing vertical loads. Accordingly, the term "nonstructural system" corresponds to the system that does not participate in bearing vertical loads. It is noted that the above conditions are not associated with the terms "primary" and "secondary" structural elements that are used in subsequent paragraphs.

(d) The objectives of the assessment or redesign are not necessarily identical. The objectives of redesign may be higher than those of the assessment.

S

Seismic Strengthening Strategies for Existing (Code-Deficient) Ordinary Structures, Table 1 Assessment or redesign objectives of the structure

Probability of exceedance of seismic action within a conventional life cycle of 50 years (%)	Performance level		
	Damage limitation (DL)	Significant damage (SD)	Near collapse (NC)
10	A1	1	C1
50	A2	2	C2

(e) The minimum acceptable assessment or redesign objectives for the structural system of existing buildings are defined ad hoc by the public authority. In special cases, the public authority may designate additional objectives for assessment or redesign of the nonbearing system as well. In this case, the same authority also defines the criteria for meeting the respective objectives.

(f) In any case, the reassessment objective (assessment or redesign) is chosen by the project owner, provided that it is equal to or higher than the above minimum acceptable objectives. In defining these objectives, the following criteria (among others) should be taken into account:

- Importance of the building (e.g., temporary structure, ordinary residential house, area of public gathering, area of crisis management, high-risk facility)
- Available financial resources during the given period

(g) The owner of the project or the public authority has to define the time frame within which the relevant interventions will be conducted, wherever required.

(h) A nominal life cycle equal to the conventional lifetime of 50 years is generally accepted, regardless of the estimated "actual" remaining life of the building. An exception to this rule is permitted only under special circumstances where the remaining lifetime is fully guaranteed, based on the judgment and approval of the public authority; in such a case, the seismic actions are modified accordingly.

It is noted that according to Table 1, design objective B1 is the objective normally set for new structures.

The adoption of an assessment or redesign objective associated with a probability of exceedance of the seismic action of 50 % will generally lead to more frequent, more extensive, and more severe damage compared to a corresponding objective associated with a probability of exceedance of seismic action equal to 10 %.

The probability of exceedance of 50 % (maximum tolerable) in 50 years corresponds to an average return period of about 70 years, while a probability of exceedance of 10 % in 50 years corresponds to an average return period of approximately 475 years.

Structural Performance Levels

The performance levels of the structure are defined as follows [see also Chap. 2.1 of European Standard EN 1998–3 (2005)]:

(a) "Damage Limitation" (A) is a condition wherein the structure is only lightly damaged, with reinforced concrete elements prevented from significant yielding and retaining their strength and stiffness properties. Nonstructural components, such as partitions and infills, may show distributed cracking, but the damage could be economically repaired. Permanent drifts are negligible. The structure does not need any repair measures. No building operation is interrupted during and after the design earthquake, with the possible exception of minor importance functions. A few hairline cracks may occur in the structure.

(b) "Significant Damage" (B) is a condition wherein the structure is significantly damaged, with some residual lateral strength and stiffness, and vertical elements are capable of sustaining vertical loads. Nonstructural components are damaged, although partitions and infills have not failed out of plane. Moderate permanent drifts are present. The structure can sustain aftershocks of moderate intensity. The damage to the structure, expected to occur during the design earthquake, is

repairable but possibly uneconomic, without causing loss or serious injury of people.

(c) "Near Collapse" (C) is a condition wherein the structure is heavily damaged, with low residual lateral strength and stiffness, although vertical elements are still capable of sustaining vertical loads. Most nonstructural components have collapsed. Large permanent drifts are present. The structure is near collapse and would probably not survive another earthquake, even of moderate intensity. The damage to the structure is in general unrepairable. Injuries of certain individuals due to structural damage or falling elements of the nonbearing structure or other objects are not excluded.

The term "unrepairable damage" refers to serious or severe damage, for which strengthening (and not just repair) or replacement or substitution of the component or the entire structure is required.

General Principles for Pre- and Post-earthquake Intervention Decisions

Selection Criteria and Types of Structural Interventions

(a) Based on the conclusions drawn during the assessment of the structure, and the nature, extent, and intensity of the damage or deterioration (if any), intervention-related decisions are made, aiming to (a) meet the basic requirements of the seismic code, (b) minimize the cost, and (c) serve the needs of the society.

(b) The selection of the type of structural intervention should be made primarily on the basis of general cost- and time-related criteria, the availability of the resources required, architectural or other needs, etc. In this selection, the financial (or other) value of the structure should also be taken into consideration, both prior and subsequent to the intervention.

Such general criteria include the following:

- The cost, both initial and long term (i.e., the cost of maintenance and possible future damage or deterioration), compared

to the importance and age of the building under consideration.

- The available quality of the work (it is very important that intervention measures be compatible with available resources and available quality of work).
- The availability of an adequate quality control system.
- The use of the building (possible consequences of the intervention works to the use of the building).
- The design, from an aesthetic point of view (the intervention scheme may vary between a fully invisible intervention and a deliberately distinctive set of new or added members).
- The conservation of the architectural identity and integrity of historical buildings and the consideration of the degree of reversibility of the interventions.
- The duration of works.

(c) The selection of the type, technique, scale, and timing of the intervention should be based on technical criteria related to the observed current state of the building, as well as to a provision to maximize the ability of the structure to absorb seismic energy (ductility) after the intervention.

Such technical criteria include the following:

- All identified serious deficiencies must be remedied.
- All identified serious damage (and deterioration) in primary structural members must be repaired properly.
- In the case of highly irregular buildings (mainly in terms of distribution of their overstrength, both in plan and in height of the building), structural regularity should be improved to the maximum possible extent.
- All resistance requirements in critical regions of primary structural members (i.e., the required resistance and plastic deformation capacity) must be satisfied after the intervention (distinguishing between primary and secondary members).
- Where possible, the increase of local ductility in critical regions should be pursued.

S

Particular provision should be taken, to the largest extent possible, so that the local repair and strengthening do not diversely affect the available ductility within the critical region.

- In special cases, the durability of both new and original structural members and the potential acceleration of the deterioration should be taken into consideration.

Types of Intervention and Their Consequences

(a) Based on the foregoing criteria and the results of the assessment of the structure, appropriate types of intervention should be ad hoc selected for individual structural members or the entire building and the nonstructural system (if required), always taking into account the side effect of the interventions on the foundations. This selection is part of an intervention strategy, which aims to improve the seismic behavior of the building by modifying or controlling the basic parameters that affect its seismic behavior. In order to achieve a reduction of seismic risk, strategies of technical or managerial nature, or combination of the two, can be adopted.

A number of technical and managerial strategies are indicatively given herein:

Technical strategies:
- Enhancement of the building strength
- Enhancement of the building stiffness
- Enhancement of the deformation capacity of the structural members
- Reduction of seismic demand

Managerial strategies:
- Limitation or change of use of the building.
- Partial or global demolition (i.e., of a number of storeys).
- "Rigid body" transfer of the entire structure to another location.
- Decision for "no intervention". In such a case, a reduction of the nominal life cycle of the structure can be accepted, under the condition that upon expiry of this period, demolition of the structure is guaranteed.

Some types of interventions in structural elements associated with specific strengthening strategies of technical nature are listed below:

- The enhancement of strength and stiffness is alternatively achieved by selective or large-scale strengthening of structural members or by the addition of new elements that can resist either partially or fully the seismic actions (e.g., reinforced concrete shear walls, steel trusses, infill walls, etc.). In this case, particular attention should be given to the design of the foundation due to the increase of both the structural mass and the seismic loads.

- The enhancement of post-elastic deformation capacity (ductility) is achieved by improving the confinement of existing members, e.g., with external connectors, strips of steel or fiber-reinforced polymers, etc.

- The remedy of critical deficiencies refers to alleviating those features that lead to unfavorable seismic behavior. Indicatively:
  - Modification of the structural system (abolition of certain expansion joints, replacement of sensitive members, alteration actions aiming at a more regular and ductile system)
  - Addition of special links to connect the brittle masonry and surrounding member, whenever this is permitted by the strength of masonry
  - Local or global modification of members with or without damage
  - Full replacement of inadequate members or members that have suffered extensive damage
  - Redistribution of demand (e.g., through external prestressing)

- The reduction of seismic demand is achieved by reducing the mass of the structure and modifying the structural system toward a favorable shift of the fundamental period of the structure [e.g., through seismic isolation systems or absorption of seismic energy; see

Chap. 10 of European Standard EN 1998–1 (2004)].

(b) In cases where, for the redesign objective set, the seismic behavior of nonbearing structural members might endanger the lives of the occupants (or third persons) or might have consequences to stored goods, measures should be taken to repair or strengthen the particular members.

In such cases, local or global collapse should be prevented by:

- Appropriate links to the structural members or by taking supportive measures to prevent possible fallout of parts of those members
- The improvement of the mechanical characteristics of nonstructural members

(c) The potential impact of repairs and strengthening of nonstructural members should be taken into account.

(d) The side effects of all structural interventions on the local and global capacity of the building to absorb seismic energy should be taken into account.

The enhancement of strength usually leads to a reduction in ductility, unless special measures are taken (e.g., in reinforced concrete elements, the increase of the tensile reinforcement should be in principle accompanied by a sufficient increase of the compression reinforcement and of confinement).

Methods of Analysis for Assessment and Redesign

General

(a) The action effects and/or the required plastic rotations of all structural members of the building under the design earthquake and other combinations of actions are derived using appropriate analysis methods.

To determine the internal forces and displacements, it is permissible to ignore proximity to other buildings.

(b) The selection of the appropriate method of analysis should be based on the importance of the building and its potential damage or deterioration, as well as on the available data regarding the cross sections and strength of its structural members.

Whenever possible, it is recommended to calibrate such methods through comparison with the behavior of buildings that have already been studied with the particular methods.

(c) Where appropriate, additional partial factors γ_{Sd} will be applied to account for the uncertainties related to the numerical analysis models.

Consideration of Masonry Infill Walls

Consideration of the masonry infill walls in the redesigned structure can be carried out, subject to the conditions of paragraph Consideration of Masonry Infill Walls

As part of the redesign process, it is desirable to make every effort to mitigate the potential deficiencies imposed by the masonry infills. Addition or upgrading of masonry infills can be used for the improvement and strengthening of existing buildings, subject to conditions of paragraph Consideration of Masonry Infill Walls

Methods of Analysis

For the assessment and redesign of a building, one of the following analysis methods may be used. The field of application of each analysis method depends on satisfying a number of conditions, primarily related to regularity (see Greek Code of Structural Interventions 2012, Chap. 5):

(a) Elastic (equivalent) static analysis with global (q) or local (m) behavior (or ductility) factors, regardless of the data reliability level.

(b) Elastic dynamic analysis with global (q) or local (m) behavior (or ductility) factors, regardless of the data reliability level.

(c) Inelastic static analysis. In this case, it is recommended to ensure, as a minimum, a "satisfactory" data reliability level.

(d) Inelastic dynamic (response history) analysis. In this case, it is again recommended to ensure, as a minimum, a "satisfactory" data reliability level.

S

(e) In special cases, solely for the assessment of existing buildings, it is permitted to estimate the demand approximately, without detailed analysis involving a finite element model of the entire building.

(f) Apart from the above analytical methods, solely for the assessment of existing buildings, in special cases and for specific objectives, it is possible to use empirical methods.

(g) It is permitted to apply the elastic methods provided that all the following conditions apply:

  (i) The failure index (λ) of each primary member is in general lower than 2.5.

    The adopted threshold value of the failure index (λ) generally implies that the available strength of each primary structural member is at least 40 % of the demand resulting from an elastic seismic analysis without reducing the seismic action, i.e., for $q = 1$.

  (ii) The average failure index ($\overline{\lambda}_k$) in each storey does not exceed 1.50 times the average failure index of the storey above or below.

    It is deemed that the average failure index ($\overline{\lambda}_k$) identifies the regularity in the resistance (strength) along the building height, whereas its adopted threshold value ensures that no weak, in flexure and/or shear, intermediate storey exists.

  (iii) The failure index (λ) of each primary structural member that is located on one side of the building, for a given direction of seismic action, does not exceed 1.50 times the average failure index (λ) of a primary member that is located on any other side of the same storey.

    It is deemed that with this provision, issues of torsionally sensitive storeys are addressed.

Principal (or Primary) and Secondary Structural Members

The individual parts of the structure of a building and the individual structural elements (members) affecting the stiffness and demand distribution within the building, or the members that are loaded due to lateral displacements of the building, can be distinguished during assessment or redesign into "principal" (or "primary") and "secondary."

In general, the structural members or substructures that contribute to the strength and stability of a building under seismic loading will be characterized as principal. The remaining structural elements or substructures will be characterized as secondary. See also related Chap. 4.2.2 of European Standard EN 1998–1 (2004).

The main consequence of classifying a structural member (or substructure) as a secondary is that for these members, different performance criteria apply, i.e., they are permitted to undergo larger displacements and suffer higher damage compared to the primary elements.

In cases where damage limitation after the earthquake has been set as the assessment or redesign objective, the above distinction between primary and secondary data is not permitted.

For the masonry infill walls, which do not carry vertical loads (see paragraph Consideration of Masonry Infill Walls), the distinction between primary and secondary members does not apply. Where those members are considered as part of the seismic action resisting system, they are addressed and verified separately.

Compliance Criteria for Assessment and Redesign

General

(a) Compliance with the selected performance level for assessment and redesign is achieved by adoption of the seismic action, method of analysis, verification, and detailing procedures, appropriate for the different structural materials (concrete, steel, masonry) within the scope of the code used.

(b) Except when using the q-factor approach, compliance is checked by making use of the seismic action as defined in Table 1.

(c) For the verification of structural elements, a distinction is made between "ductile" and "brittle" ones. Except when using the q-factor approach, the former is verified by checking

that demands do not exceed the corresponding capacities in terms of deformations. The latter is verified by checking that demands do not exceed the corresponding capacities in terms of strengths (see Greek Code of Structural Interventions (2012), Chap. 9).

(d) Alternatively, a *q*-factor approach may be used, where use is made of a seismic action (response spectrum) reduced by a *q*-factor. In safety verifications all structural elements are verified by checking that demands due to the reduced seismic action do not exceed the corresponding capacities in terms of strength.

(e) For the estimation of the capacities of ductile or brittle elements, where these will be compared with demands for safety verifications, mean value properties of the existing materials are used as directly obtained from in situ tests and from additional sources of information. Nominal properties are used for new or added materials.

(f) Some of the existing structural elements can be designated as "secondary seismic" (see Chapter Principal (or Primary) and Secondary Structural Members), in accordance with the definitions in European Standard EN 1998–1 (2004), Chap. 4.2.2 (1)P, (2) and (3). "Secondary seismic" elements have to be verified according to the same compliance criteria as primary seismic ones, but using less conservative estimates of their capacity than for elements considered as "primary seismic."

Safety Verifications

(a) The available resistance in the critical regions of all structural members (i.e., the resistance quantities and/or the tolerable plastic rotations) should be calculated on the basis of rational numerical models, which are widely accepted by the international scientific community, especially in terms of force transfer between existing and added materials or members (see Greek Code of Structural Interventions (2012), Chap. 6 for the numerical models, Chap. 7 for the determination of the behavior of structural members, and Chap. 8 for the design of the interventions).

(b) The partial safety factors of the existing and added materials should take into account the geometrical uncertainties, the dispersion of material properties, the relevant information available on site, as well as any uncertainties due to the nature of works and the difficulties of effective quality control (see Greek Code of Structural Interventions 2012, Chaps. 4, 7–9).

(c) Where appropriate, additional factors γ_{Rd} are applied to account for the uncertainties arising from the numerical modeling of the resistance in critical (or noncritical) regions.

(d) In cases of structural interventions against seismic actions, the damage limitation verification is made in accordance with appropriate provisions (see Greek Code of Structural Interventions 2012, Chap. 9). The damage limitation verification generally includes the primary and secondary structural members, infills, and appendages.

Verification of the Adopted Behavior Factor

After the verifications above, it is required to approximately reevaluate the predefined behavior factor for the repaired–strengthened building, taking into account all the criteria favoring energy absorption, such as:

(a) The sequence of failure of horizontal and vertical structural members

(b) The type of failure in critical regions of each structural member (i.e., the ratio of the ultimate shear force to the effective shear at the time of flexural failure, as imposed by capacity design)

(c) The local available ductility in critical regions

(d) The available secondary resistance mechanisms at large relative displacements

(e) The potential consequences of the brittleness of a limited number of structural members on the ductility of the entire structure

It is pointed out that in existing structures the requirements of capacity design, limitation of the axial force, local confinement, etc. have not been in general met. The implication of this is the difficulty in assessing a global behavior factor.

Summary

This entry deals with the seismic strengthening strategies for existing (code-deficient) ordinary structures. Reference to two regulatory texts that are compatible with each other, namely, the European Standard EN 1998–3 (2005) and the Greek Code of Structural Interventions (2012), has been selected as a basis for the presentation of the issue. These texts condense existing knowledge concerning the upgrading of ordinary structures and at the same time determine the procedure of the normative treatment of the issue.

First, clarifications on basic terms are given. Then, the general principles that rule the two aspects on which the issue of upgrading is based, i.e., assessment and redesign, are analyzed. Subsequently, the assessment and redesign objectives, which in general differ from those of the design of new structures, are presented.

The possibility of a variety of redesign objectives is introduced. High targets are proposed for a small, yet critical in terms of safety, group of buildings. On the contrary, there is a possibility, under certain conditions, to adopt lower targets for common existing buildings, in comparison to the targets for new buildings, with a view to a realistic treatment of the need for upgrading.

Furthermore, the general principles for pre- and post-earthquake intervention decisions are set, the methods of analysis and the conditions of their application are briefly presented, and, finally, the compliance criteria for assessment and redesign are described.

Cross-References

References

Greek Earthquake Planning and Protection Organisation – EPPO (2012) Greek code of structural interventions – harmonized text http://ecpfe.oasp.gr/sites/default/files/files/%CE%9A%CE%91%CE%9D%CE%95%CE%A0%CE%95_EN2013_FINAL.pdf

European Standard EN 1992-1-1 (2004) Design of concrete structures – general rules and rules for buildings

European Standard EN 1998–1 (2004) Design of structures for earthquake resistance – general rules, seismic actions and rules for buildings

European Standard EN 1998–3 (2005) Design of structures for earthquake resistance – assessment and retrofitting of buildings

Wikipedia, the free Encyclopedia (2014) Terms "strategy" and "technique"

Seismic Strengthening Strategies for Heritage Structures

Dina D'Ayala[1] and Sara Paganoni[2]
[1]Department of Civil, Environmental and Geomatic Engineering, University College London, London, UK
[2]Ziegert|Roswag|Seiler Architekten Ingenieure, Berlin, Germany

Introduction

The global seismic behavior of historic masonry buildings is highly influenced by the integrity of the connections among vertical and horizontal structural elements, to ensure the so-called box behavior. This, providing the transfer of inertial and dynamic actions from elements working in flexure out-of-plane to elements working in in-plane shear, leads to a global response best suited to the strength capacity of the constitutive

LEVEL OF KNOWLEDGE

SEISMIC INTENSITY MEASURE
Peak Ground Acceleration
Linear spectra
Nonlinear spectra
Accelerogrammes

STRUCTURAL ANALYSIS
Simplified models/limit analysis
Elastic FE. models /static analysis
Non linear FE. models /pushover
Non linear FE. models/ time history

RISK
Ratio of demand/capacity
Performance based criteria, drift
Life cycle analysis

INTERVENTIONS
Strength
Ductility
Energy dissipation
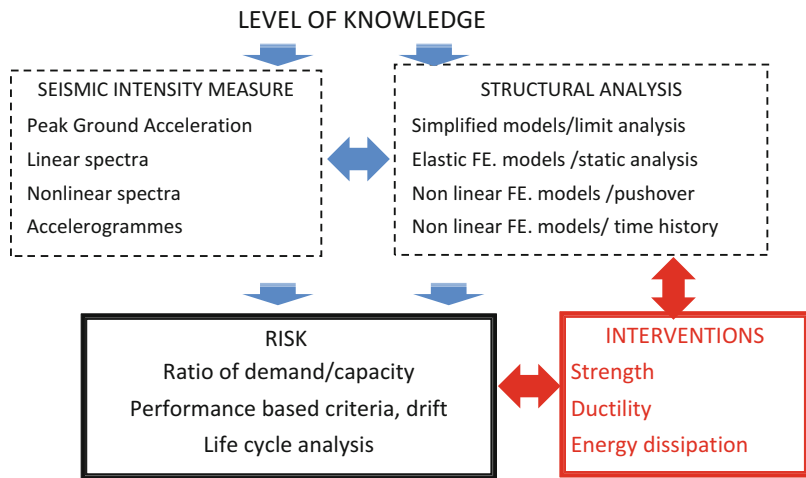
materials and hence enhanced performance and lower damage level. Notwithstanding the importance of connections' integrity, analytical checks of existing connecting elements, or design of new elements to strengthen existing connections, are generally based on qualitative rules or simplified overall checks, rather than rigorous analytical approach. The "Guidelines for Earthquake Resistant Non-Engineered Construction" were first published by the International Association for Earthquake Engineering (IAEE) in 1986 specifically with the objective of improving the seismic safety of non-engineered housing constructions.

A wide range of techniques and products for the seismic strengthening of heritage buildings are reported in the scientific literature and used in current practice to ensure the enhancement of existing connections. Heritage buildings require far more attention, especially when dealing with issues such as the compatibility between the chemical and mechanical properties of the strengthening system and the parent material. Many strengthening techniques, after an initial success and a strong commercial promotion, underperformed and showed unexpected drawbacks when put to the test of real seismic loading outside the controlled conditions of the laboratory environment. On the other hand, some existing strengthening systems can provide highly flexible applications and meet the expected requirements in terms of performance; some of these systems in fact draw on traditional reinforcement techniques, with the addition of innovative materials and a deeper insight in the laws governing the dynamics of structures.

The choice of the most suitable strengthening system for a given building, however, is determined not only by the set of constraints of the specific project but, very importantly, by the framework used for its assessment. The EC8 part 3 or ASCE41-06, dealing with assessment, repair, and strengthening, introduce the concept of knowledge level as the determining factor for the choice of alternative assessment procedures involving diverse levels of resources. The diagram in Fig. 1 shows that there is a strict correlation not just between the representation of the seismic input and the type of analysis that can be carried out with it but also and most importantly between the input and the way risk is quantified and measured and between such measures and the principle on which possible strengthening interventions operate. Hence the level of knowledge and availability of data to carry out the assessment becomes critical in determining and fully designing the most appropriate strengthening solution.

Moreover, although European and national codes (e.g., Italian Ministry of Cultural Heritage and Activities 2006) suggest the use of various systems for the strengthening of connections, for example, ring beams, no detailed reference is made to specific procedures for the dimensioning and checks of such elements. The only indication

in this sense can be found in Section 6.1 Retrofit Design Procedure for existing building of Eurocode 8 (EN 1998-3:2005), which states that the design process should cover:

1. Selection of techniques and/or materials, as well as of type and layout of intervention.
2. Preliminary sizing of additional structural parts.
3. Preliminary calculation of stiffness of strengthened elements.
4. Analysis of strengthened structure by linear or nonlinear analysis. The typology of analysis is chosen depending on the level of knowledge regarding the geometry detailing and materials of the structure.
5. Safety verifications for existing, modified, and new structural elements carried out by checking that the demand at three different limit states – damage limitation, significant damage, and near collapse – is lower than the structural capacity.

In the safety verifications, mean values of mechanical properties of existing materials derived from in situ tests and other sources of information, like available documentation or relevant sources, shall be used, taking into account the confidence factors (CFs) specified in 3.5 of Eurocode 8 (EN 1998-3:2005). Conversely, for new materials, nominal properties shall be used without modification by confidence factor. The code also states that in case the structural system, comprising both existing and new structural elements, can be made to fulfill the requirements of EN1998-1-2004, the verifications may be carried out in accordance with the provisions therein.

This last sentence indicates that for systems such as reinforced concrete (RC) ring beams or corner confinement, reference can be made to the specifications for RC members in the relevant sections of EC8 and other Eurocodes. However, this leaves open the problem of quantifying the interaction between original and new structural elements, and the assessment of the global seismic performance of the strengthened structure will still be affected by a large number of uncertainties.

The last 15 years have seen a steep increase in the use of new technologies for strengthening heritage buildings. Such fast development, as well as the high level of expertise and financial resources required for their application, often results in lack of standardization. Innovative technologies have not been extensively applied and validated in real-life situations yet, and the retrofit of a complex, precious building by means of unconventional systems is a difficult task that goes beyond the standard conservation practice. In fact, looking at the current scientific literature, it is clear that many projects of restoration and upgrade of monumental buildings are carried out by research organizations within the framework of specific projects or by large enterprises that specialize in the production and design of strengthening devices.

On the other hand, the lack of appropriate standards and procedures can be blamed for the incorrect application of some innovative strengthening techniques, notwithstanding recognized effectiveness and other benefits. One exception to this situation is the case of fiber-reinforced polymers for which guidelines and standards have been produced for selected countries (see, for instance, CNR-DT 200/04, CNR-DT 200R/13).

In the following an overview of methods to structurally strengthen the masonry elements of historic buildings is given, pointing out the advantages and pitfalls of each system and the fundamental physical parameters and design process required. Available design procedures are reviewed and referenced, when available, to the writers' knowledge. The systems reviewed are applicable mainly to stone or brick masonry structures with timber floors and roof or with vaulted structures. The concepts underlying these methods are also usually applied to the strengthening of earth structures, although some of the details of the specific applications might differ due to the specific characteristics of the parent material.

The effective strengthening of a heritage building also relies on a correct redistribution of the inertia forces among vertical elements. For this to occur, the floor structures need to be acting

as a whole and ensure diaphragm action. While there is a wealth of research and methods to assess and strengthen historic and traditional floor structures to ensure they develop such action, these are not currently included in this chapter.

Base isolation is also increasingly becoming a proposed option for the retrofitting of historic buildings; however, actual implementations are still very rare and very expensive. This technique is also not treated in this current edition.

In this current edition, only methods suggested in the Eurocodes or methods related to recent technological developments are included. Many traditional "vernacular" methods of strengthening historic building exist in earthquake-prone countries and some have been studied in detail. These are not addressed here either.

Strengthening systems are reviewed depending on the type of enhancement achieved, i.e., whether there is an increase in strength, a control of deformation and displacement, or a dissipation of energy.

Connections Between Vertical and Horizontal Macroelements

The Eurocode 8 states that to improve connection between intersecting walls, use should be made of cross-bonded bricks or stones. The connection can be made more effective in different ways (EN 1998-3:2005):

 I. Through construction of a reinforced concrete belt
 II. By addition of steel plates or meshes in the bed joints
III. Through insertion of inclined steel bars in holes drilled in the masonry and grouting thereafter
IV. Through post-tensioning

The addition of steel ties, along or transversely to the walls, external or within holes drilled in the walls, is an efficient means of connecting walls and improving the overall behavior of masonry buildings (EN 1998-3:2005).

In the following we review:

• Stiffening of the wall system by corner and T-junctions confinement of wall panels achieved by inserting vertical concrete columns or steel meshes, fiber-reinforced horizontal strips, and polypropylene meshes
• Stiffening of the wall and floor system and connection to the horizontal structures by ring beams
• Connection of the walls and floor system by anchorage systems
• Improvement of the performance of the wall system by including energy absorbers and dissipating devices to connect the walls

The design of effective connections between various structural elements of a masonry structures is a critical step for the achievement of a good structural response in case of seismic events. Current codes provide for carefully designed and detailed connections; for instance, the Eurocode states that: "Floor systems and the roof should be provided with in-plane stiffness and resistance and with effective connection to the vertical structural systems (EN 1998-1:2004)."

This rule is also applicable to interventions on existing structures as the Eurocode indeed states that: "The connection between the floors and walls shall be provided by steel ties or reinforced concrete ring beams (EN 1998-1:2004)" and "Specifically for masonry structures: non-ductile lintels should be replaced, inadequate connections between floor and walls should be improved, out-of-plane horizontal thrusts against walls should be eliminated (EN 1998-3:2005)."

Moreover, in relation to the connections between walls and floors, EN 1998-3:2005 states that: if existing tie-beams … are damaged, they should be repaired or rebuilt. If there are no tie-beams in the original building structure, such beams should be added.

Improvement of connections between vertical and horizontal structures can be achieved through a variety of methods, which are described in the following.
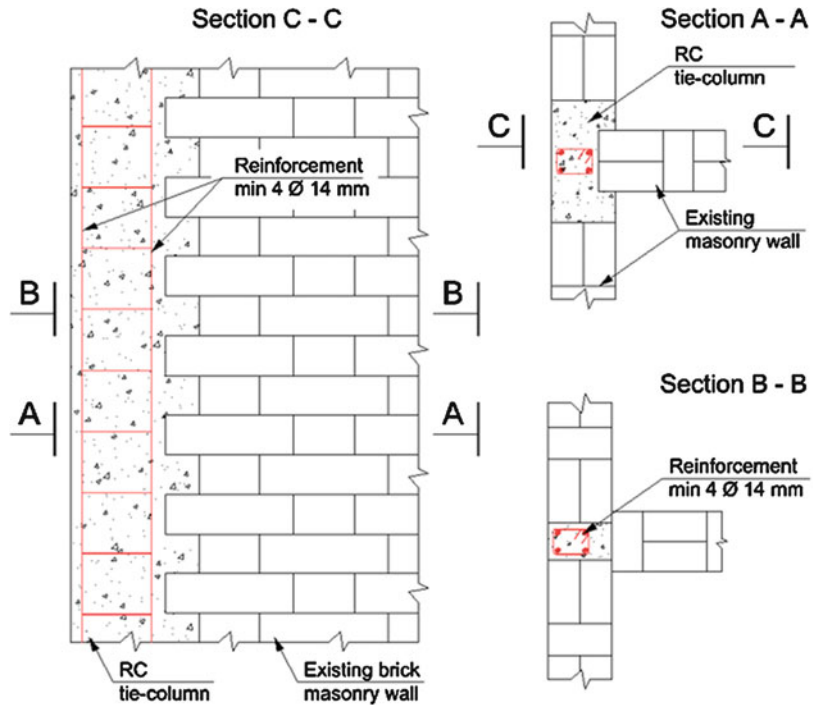
Confinement at Walls Junctions: Strength Enhancement

Column Ties

By confining plain masonry walls with vertical elements placed at all corners and wall intersections, as well as along the vertical frame of large openings, the seismic performance of a masonry building is improved as a result of the enhanced integrity of the structural system (Tomaževič, 1999). This effectively changes an unreinforced masonry structure in a confined masonry structure. The disruption is significant and can affect very substantially the heritage value of the building. Moreover, such a technique is suitable only when horizontal RC tie beams in the bearing walls at floor level (see this IS section "Confinement at Wall to Floor Junctions: Ring Beams") and stiff, monolithic floor diaphragms are in place; otherwise, the effect of the vertical confinement is scarce, or null, as shown by analysis conducted by Karantoni and Fardis (1992). Creation of corner confinement by column ties is not recommended for stonework masonry (Tomaževič 1999).

For the construction of the confinement elements, all the bricks in the intervention area are removed; the concrete of the existing tie beams is also removed to allow for the connection between existing and new reinforcement. Vertical rebars and stirrups are placed and concrete cast (Fig. 2).

Tie columns should be as thick as the wall where they are located, although in many cases of repair of existing buildings, they are of reduced dimensions due to on-site constraints. Furthermore, they may sometimes be replaced by sets of reinforcing bars placed in holes drilled in the masonry and connected to the surrounding substratum by stirrups.

The confinement prevents disintegration and improves ductility and energy dissipation of URM buildings but has limited effect on the ultimate load resistance (Chuxian et al. 1997). However, the real confinement effect mainly depends on the relative stiffness between the masonry wall and the surrounding resulting concrete frame. Before cracking, the confinement effect can in general be neglected (Chuxian et al. 1997; Karantoni and Fardis 1992). For very squat URM walls (geometrical aspect ratio

of 0.33 and double fixed boundary conditions), the confinement increased the cracking load by a factor of 1.27 and the ultimate lateral capacity by a factor of 1.2 (Chuxian et al. 1997). For walls with higher aspect ratio, the confinement increased the ultimate capacity by a factor of 1.5. In addition, the confinement improved the lateral deformations and energy dissipation by more than 50 %.

The dimensioning procedure for column ties can be derived from the prescriptions specific to confined masonry structures. The process can be summarized as follows:

- Calculation of the appropriate combination of static and seismic loads acting on the structural elements (EN 1991-1-1:2002 and EN 1998-1:2004)
- Dimensioning and verification of the column ties for both static and seismic loading following the prescriptions provided for confined masonry structures (EN 1996-1-1:2005 and EN 1998-1:2004)

Special care should be taken in realizing the connection between the column ties and the horizontal structures in order to ensure a monolithic behavior of the resulting RC frame.

A much cheaper solution to confining masonry walls is achieved with Welded Wire Mesh (WWM). This consists in deploying a steel mesh on the two sides of a wall and around the corners, connecting it with bolts and then concrete over, by either shotcreting or other forms of plastering. The anchorage of the mesh to the foundation and horizontal structures is fundamental to achieve proper action transfer during shaking. Although less destructive than the construction of column ties, the system seriously affects the breathability of the walls and may cause deterioration due to moisture entrapment. The thickness of this reinforced plaster may vary between 25 and 100 mm depending on the regularity of the walls underneath. It cannot be applied in the presence of historic plasters or other valuable finishes. Its effectiveness highly depends on the relative stiffness between the original wall and the added layers.

Confinement by Fiber-Reinforced Plastic Strips

The application of fiber-reinforced polymer (FRP) material to the strengthening of masonry structures has to an extent mimicked application to RC structures. As a result, there are numerous examples in the literature relating to the confinement of masonry columns (e.g., Di Ludovico et al. 2010; Corradi et al. 2007). In the case of junctions between walls, horizontal strips of FRPs bonded at various levels along a masonry panel and anchored to the side walls can be used to restore corner connections, thus preventing overturning of façade walls. Optimal application is achieved when the whole perimeter of the structure is confined. Application of such techniques entails the removal of existing plaster and other superficial finish as a strong bond of the FRP strips to the masonry is necessary to ensure effectiveness of the system. The strips also need to be protected from UV radiations to prevent material deterioration. Their implementation might cause substantial loss of heritage value, and it is not recommended when valuable finishes are in place.

For the confinement to be effective, strips should be laid both horizontally and vertically. Tests by Hamoush et al. (2001) on walls strengthened with vertical and horizontal strips showed the effectiveness of the FRP systems as out-of-plane flexural strengthening elements, while Tumialan et al. (2001) have investigated the potential of near surface-mounted GFRP rods embedded into epoxy-based paste in the bed joints as shear enhancement. Cyclic testing of walls strengthened with vertical and horizontal GFRP or CFRP strips was performed by Marcari et al. (2003) and by Krevaikas and Triantafillou (2006). Those tests have highlighted a general decrease in strength and ductility of the confined members with the increase of the aspect ratio of the confined cross section. This is due to a reduction in ultimate strain at failure when increasing the aspect ratio. Glass fibers are more effective than carbon fibers, owing to lower stiffness. However, its effectiveness in connections between orthogonal walls has not been tested to the authors' knowledge. Issues of debonding should be thoroughly investigated on a case by

Seismic Strengthening Strategies for Heritage Structures, Fig. 3 Static scheme for the calculation of the tensile capacity of FRP strengthening to prevent overturning of front wall (CNR-DT 200/04)
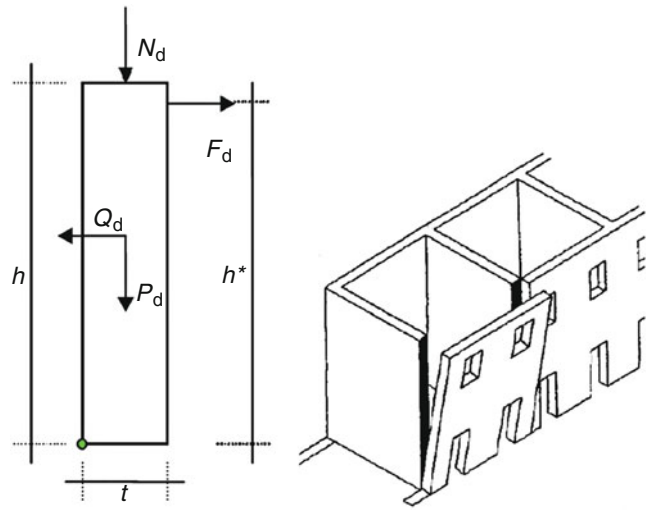
case basis as failure mode and extent will be highly affected by the integrity and mechanical characteristics of the parent material external strata in relation to the composite characteristics.

FRP confinement is dimensioned by verifying:

1. The tensile capacity of the composite material
2. The capacity of the anchorage area

In detail this is done by (CNR-DT 200/04):

1. Checking that:

$$F_d \leq 2F_{Rd} = 2\left(A_f \cdot f_{fd}\right) \quad (1)$$

where

F_d is the design action applied to the FRP stripes by the front wall undergoing overturning caused by seismic equivalent action Q_d; this is calculated by:

$$F_d = \frac{1}{2h^*}\left(Q_d h - N_d t - P_d t\right) \quad (2)$$

Dimensions and weight involved in the overturning mechanism are shown in Fig. 3:

F_{Rd}: design tensile capacity of the FRP strengthening

f_{fd}: design ultimate strength of the FRP strengthening

h^*: distance between level where FRPs are bonded and bottom hinge

A_f: FRP area

2. Cheking the rip-off of FRP stripes from side walls:

$$F_d < 2F_{pd} = 2\left(A_f \cdot f_{pd}\right) \quad (3)$$

where

F_d is calculated as above.

F_{pd} is maximum anchorage capacity of FRP to one of the two side walls.

f_{pd} is design debonding strength of FRPs.

A_f is FRP area.

The second check is generally more demanding than the first, but it is not needed when full confinement of the structure is achieved by wrapping FRP stripes around the whole perimeter of the building and a suitable anchorage length is ensured (CNR-DT 200/04). Anchorage length should be at least 300 mm; otherwise, mechanical fixings can be used; however, the guideline does not specify suitable types and the procedure for sizing and positioning the fixings.

Additional calculations of the strengthened structure are necessary to ensure that the strengthened structure will be able to withstand other actions, such as the increased in-plane shear loading, now transferred from the walls

Seismic Strengthening Strategies for Heritage Structures, Fig. 4 Corner confinement of masonry walls by use of SRP strips anchored with SRP connectors (Photo Paolo Casadei 2008, product FIDSTEEL)
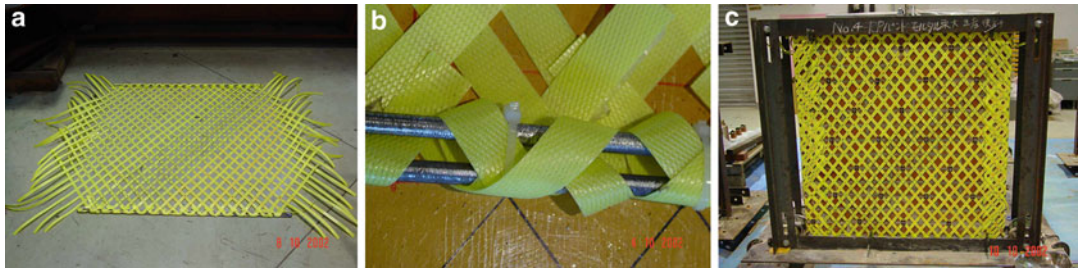
Seismic Strengthening Strategies for Heritage Structures, Fig. 5 Procedure for installation of PP meshing: (**a**) PP-band mesh, (**b**) detail of top/bottom connection, and (**c**) retrofitted wall before application of final layer of mortar (Mayorca and Meguro 2004)

orthogonal to the walls parallel to the action. This should be done according to the specific indications of codes dealing with masonry structures in general.

Of course in order to prevent damage in the historic building, the capacity of the strengthening system needs to be less than the shearing capacity of any masonry course in the façade wall. Given the absence of ductility in the fiber composite, this system also presents the fundamental drawback of brittle failure, modest ductility, and relying on the damaging process of the masonry for energy dissipation. An improved version in this respect is provided by SRP strips, strips reinforced with steel, which show a greater ductility (see Fig. 4).

Confinement by Polypropylene (PP) Mesh

PP meshing uses common packaging straps (PP-bands) to form a mesh, which is then used to encase masonry walls, preventing both collapse and the fall of debris during earthquakes. PP-bands are commonly used for packaging and are therefore cheap and readily available while the retrofitting technique is simple and suitable for local builders. PP meshing was developed by Meguro Lab, Tokyo University, and has had application in Kashmir and Nepal.

The retrofitting installation procedure is as follows (Mayorca and Meguro 2004):

1. The PP-bands are arranged in a mesh and connected at their crossing points (Fig. 5a).
2. Two steel rods are placed at the edges of the mesh; these bars are used to anchor the mesh at the foundation and at the top edge of the wall (Fig. 5b).
3. The walls are cleaned and, if possible, the paint is removed. Any loose brick is removed and replaced.
4. Six millimeter diameter holes are drilled through the wall at approximately 250–300 mm distance. The holes are cleaned with water spray or air.
5. The meshes are installed on both sides of the wall and wrapped around the corners and wall edges. An overlapping length of approximately 300 mm is needed.

6. Wire is passed through the holes and used to connect the meshes on both wall sides. In order to prevent the wires from cutting the PP-band mesh, a plastic element is placed between the band and the wire. Connectors are placed in proximity of the wall intersections and of the wall edges.
7. The top and bottom edges of the mesh are glued to the foundation and top of the wall by epoxy resin. The epoxy is used to connect the bars and the wall and it is not directly applied to the mesh. The bands, which are rolled around the bars, transfer their load through friction.
8. The overlapping parts of mesh are glued together so as to ensure the continuity of the strengthening (Fig. 5c).
9. A layer of mortar is laid on top of the mesh to protect it from UV radiation and rain and also to provide further bond.

Shear tests on walls strengthened by PP mesh (Mayorca and Meguro 2004) showed that although the strengthening does not increase the peak strength, it contributes to improving the structural performance after crack occurrence. The strengthened walls exhibit larger post-peak strength, while the mesh helps to spread the diagonal cracks over a wide region. The presence of connectors and the mortar layer, ensuring bond to the masonry, are critical to the performance of the wall.

Confinement at Wall-to-Floor Junctions: Ring Beams

A ring beam is a structural element built on top of the masonry structure to the purpose of:

- Creating a continuous connection between the roof structure and walls to better distribute the vertical loads
- Improving the connection between orthogonal walls and the three-dimensional behavior of structure

Reinforced concrete (RC) ring beams have been widely used for the retrofit of masonry structures from the aftermath of the Friuli, Italy 1976 earthquake onwards. The construction of RC ring beams involves:

- Consolidating the top part of the existing masonry by injections
- Drilling vertical holes and inserting steel rods for the vertical connection between the masonry and new element
- Building the formwork and placing the reinforcement of the beam
- Casting the concrete

Such system improves the seismic capacity of structures by:

- Distributing the vertical loads
- Transferring the horizontal loads from the floors to the bearing walls
- Connecting the bearing walls so as to create a boxlike behavior and prevent out-of-plane failure

However, during a seismic event, the high stiffness of RC beams may induce out-of-plane bending of the portions of wall between the restrained floors (Borri et al. 2009) with consequent bulging of walls and activation of out-of-plane damage mechanisms. Furthermore, such technique is often time-consuming, not cost-effective, and adds mass to the structure, which increases the earthquake-induced inertia forces and consequently requires strengthening at the basis of the walls.

All these collateral effects represent major drawbacks for heritage buildings due to the large deformability and scarce cohesion of historic masonry as opposed to the high stiffness and weight of reinforced concrete. Moreover, the construction of a ring beam has a high impact from the aesthetic point of view and may require the removal of large portions of the original material of the walls and of part of the structure of the roof.

Indeed, several surveys carried out in European historic centers after earthquakes (e.g., Spence and D'Ayala 1999; D'Ayala and Paganoni 2011) highlighted how RC beams with poor connections to the underlying material

negatively affected the buildings rather than improving its structural response. Damage generally consisted of a shear failure at the interface between roof ring beam and wall, with more or less extensive damage to the masonry and eventual collapse of walls.

Nevertheless, the use of RC ring beam is allowed for the connection between walls and roof structure (Italian Ministry of Cultural Heritage and Activities 2006), providing that the dimensions of the ring beam itself do not cause an excessive increase of mass. However, conversely to steel ring beam, the use of RC beams for the strengthening of the connection between walls and intermediate floor structures is advised against (Ministero deri Beni Architettonici e Culturali, Italia, 2006), as it is far too invasive and deeply affects the performance of the structural system.

The 2007 edition of the California Historical Building Code (2007), instead, generally recommends the use of tie beams but does not specify the use of reinforced concrete.

In the scientific literature, references to procedures for the sizing of reinforced concrete ring beams can be hardly found, and it is highly likely that in common practice a large majority of additional concrete elements are simply dimensioned by the rule of thumb, using standard geometry and amount of reinforcement of ring beams in RC frames.

An improvement on RC ring beams are reinforced masonry ring beams, combining good quality brickwork (generally made of solid bricks) to steel or composite reinforcement bonded together by a binder-like grout (Ministero deri Beni Architettonici e Culturali, Italia, 2007).

In the case of steel reinforcement, the brickwork of the ring beam is laid leaving an internal cavity where the reinforcement bars and stirrups are positioned before casting the mortar-crete.

The creation of a reinforced masonry ring beam aims to:

- Provide a continuous connection between the roof structure and walls to better distribute the vertical loads.

- Improve the connection between orthogonal walls so as to enhance the three-dimensional behavior of structure.

The advantage of the technique in respect to RC ring beams is that stiffness and mass are closer to the original masonry. A RM ring beam has also good vertical deformability, and therefore it spreads the vertical load to the masonry beneath. Furthermore, it has a lower aesthetic impact, as the additional element can match the appearance of the original masonry.

RM beams are realized by:

- Building two wallets of brickwork or stonework on top of the existing wall, leaving a cavity between the two
- Placing reinforcement bars within the cavity
- Casting the concrete within the cavity, as to create a beam

Steel plate and transversal bricks connect the outer wallets.

In some cases, consolidation of the masonry under the ring beam by injections is needed to ensure a sufficient shear capacity and similar stiffness.

Like in the case of RC ring beams, examples of the sizing of RM ring beams can be hardly found in the scientific literature.

A systematic dimensioning procedure for RM ring beams could be derived from the prescriptions specific to RM structures. The process can be detailed as follows:

- Calculation of the appropriate combination of static and seismic loads acting on the structural elements (EN 1991-1-1:2002 and EN 1998-1:2004).
- Dimensioning and verification of the beam for both static and seismic loading following the prescriptions provided for reinforced masonry structures (EN 1996-1-1:2005 and EN 1998-1:2004). This is done because a RM ring beam is in fact a bearing element and should therefore be calculated for vertical loading as beams are. Furthermore, the RM ring beams should be dimensioned for tensile load as the confining
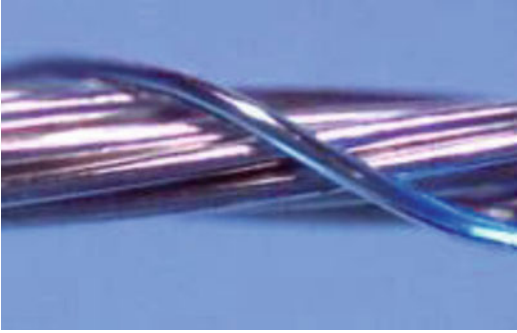
Seismic Strengthening Strategies for Heritage Structures, Fig. 6 Example of steel cord used in SRG (Borri et al. 2007)

Seismic Strengthening Strategies for Heritage Structures, Fig. 7 Assembly of LATLAM ring beam (Borri et al. 2007)

action of the beam consists indeed in resisting the thrust generated by walls that separate under horizontal actions; this is done by ignoring the cross-sectional area of concrete and masonry and sizing the steel reinforcement so that it can bear the whole in-plane action.

- Dimensioning and checks of metallic connectors for both static and seismic loading (EN 1993-1:2005).
- Dimensioning and checks of connection of the metallic fasteners to the concrete element of the beam for both static and seismic loading (EN 1992-1-1:2004 – Section 8 Detailing of Reinforcement; DD CEN/TS 1992-4-1:2009).
- Dimensioning and checks of connection of the metallic fasteners to the masonry of walls for both static and seismic loading.
- Dimensioning and checks of connection of the metallic fasteners fixed to the timber elements of the roof for both static and seismic loading (EN 1995-1-1:2004 – Section 8 Connections with Metal Fasteners).

Borri et al. (2009) propose a system combining FRPs and steel-reinforced grout (SRG), the latter made of high strength steel wires forming cords (Fig. 6) embedded in a cementitious grout. The ring-beam system, called LATLAM, is obtained by overlapping several layers of bricks and laminates embedded within a polymeric matrix or a cementitious grout (Fig. 7). By superposition of blocks and composite sheets, it is possible to achieve a structural element of any size and length,

similarly to glue-laminated timber beams. The assembly acts as a single structural element and combines the compressive strength of masonry units with the good tensile properties of composite materials, avoiding the problem of the increase of mass typical of RC ring beams. A simplified model for the evaluation of the ultimate strength of a LATLAM section under bending stresses is available in Borri et al. (2007). The collapse is assumed to occur by crushing of masonry.

Strengthening by Way of Ties: Displacement Control

The insertion of ties to connect walls to walls, walls to floor, or walls to vaults is a traditional system, still in large use nowadays.

Post earthquake observation, experimental evidence and computational results, all show that crossties installed at the intersection of perpendicular sets of walls are able to prevent the overturning of whole façades without interfering with the original structural layout (Tomaževič 1999) (D'Ayala and Yeomans 2004; D'Ayala and Paganoni 2011).

For the ties to be effective, proper anchorage within or against the masonry is necessary. This might be achieved, either by using end plates or by using grouted ends. The tie is usually a passive element which becomes active when cracks open between orthogonal walls or timber beams tend to slide off their seat.
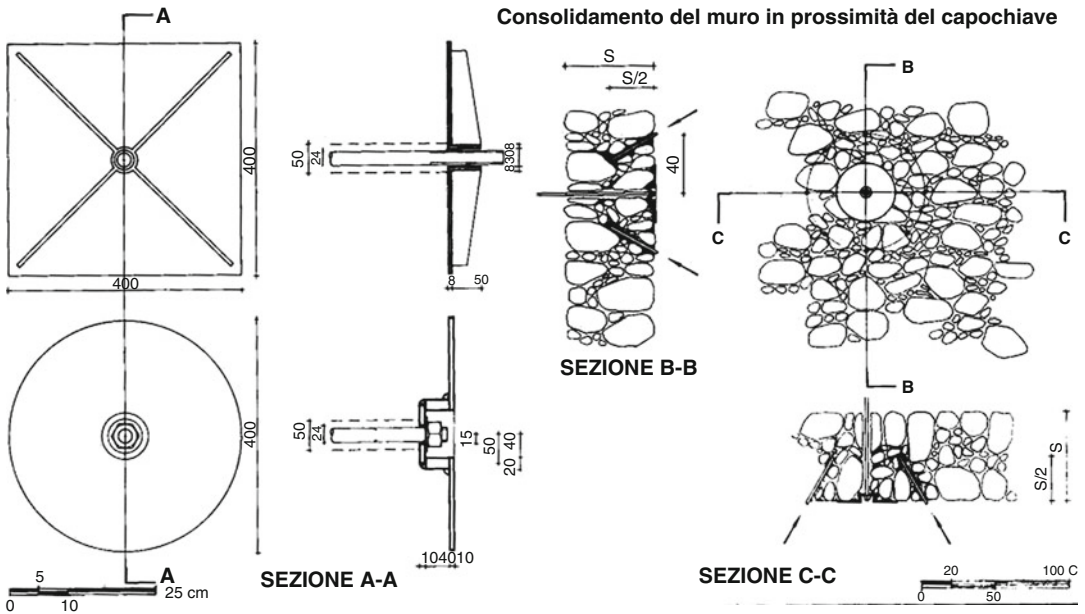
Seismic Strengthening Strategies for Heritage Structures, Fig. 8 Examples of metallic ties with end plates (Giuffrè 1993)

The dimensioning procedure of metallic ties connecting timber joist to walls can be derived from the prescriptions specific to steel structures. The process can be detailed as follows:

(a) Calculation of the appropriate combination of static and seismic loads acting on the structural elements (EN 1991-1-1:2002 and EN 1998-1:2004).
(b) Dimensioning and checks of the ties following the prescriptions provided for steel structures (EN 1993-1-1:2005).
(c) Dimensioning and checks of connection between metallic elements and timber elements of the horizontal structure (EN 1995-1:2004).
(d) Dimensioning and checks of connection of the metallic elements to masonry. Depending on the type of connection – by binder or by mechanical locking – checks should be performed in a different way as discussed in the next two subsections.

In the case of ties connecting masonry to masonry point c is not relevant. However, it might be important to proceed to a local consolidation of the masonry by use of mortar grout before implementing the tie. The major pitfall of crossties is the possibility of pullout damage at the head of the anchorage due to the different deformabilities of anchor and masonry. The use of high-ductility systems recommended by codes would overcome this issue, thus achieving the objective of protecting culturally valuable finishes and preserving life and safety.

Ties with End Plates

The dimensioning process of metallic ties with end plates (Fig. 8) consists in:

- Sizing the cross section of the metallic rod to resist the axial load deriving from the pulling action of the portion of wall constrained by the tie in case of out-of-plane action. The minimum diameter of the rod is calculated by (Tomaževič 1999):

$$D_{\min} = \sqrt{\frac{H_{u,\,\text{seg}}}{n}\frac{4}{\pi}\frac{1}{f_y}} \qquad (4)$$

where

$H_{u,\mathrm{seg}}$: ultimate seismic resistance of a critical part of a building,

n: number steel ties,

f_y: yield stress of steel.

The horizontal load that activates the overturning mechanism in a portion of wall, $H_{u,\mathrm{seg}}$, needs to be equilibrated by the action of crossties and can be calculated by limit analysis. The position of hinges and the collapse load factor, for the portion of wall involved in the out-of-plane mechanism, are calculated so as to satisfy rotational equilibrium and the distribution of stress assumed in the masonry section at collapse as shown in Fig. 9. In deciding the static scheme for the calculation of the mechanism, the type of connections should be considered as they influence the constraints of the ideal beam that represents the wall: for instance, a wall with no positive connections to the floor structures can be modeled as a cantilever, while the positive effect of well-connected horizontal structures should be accounted for by using a simply supported beam scheme.

- Sizing the end plate to prevent punching failure. This is done by verifying that the tensile strength of the parent material is sufficient to withstand the 45° component of the axial load acting on the tie. The rupture surface to consider for the calculation of the acting stress is a conical/pyramid surface with sides inclined at 45° in respect to the plain of the wall (Fig. 10).

Ties Without End Plates

Anchors without end plates (Fig. 11) consist of a metallic profile that is embedded in the masonry and can transmit loads to the substratum by mechanical locking, friction, bond, or a combination of these three (Eligehausen et al. 2006), rather than use of an element such as a plate, a key, or a peg. Mechanical locking can be obtained, for instance, by undercut, namely, shaping the end part of the hole and introducing an anchor with an end shape so as to result larger than the rest of the shaft. Friction systems, instead, consist of torsion or displacement-
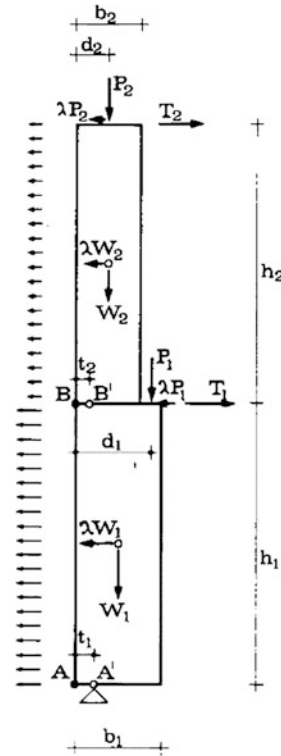
Seismic Strengthening Strategies for Heritage Structures, Fig. 9 Static scheme of multistory wall panel for determination of hinge position and collapse load multiplier (D'Ayala, Speranza, 2003)

expansion anchors, which are designed to introduce a radial pressure between the anchor and the hole. Finally bond systems rely on the use of a binding agent, which is either injected or released through a capsule system (Fig. 12).

A fairly common commercial product consists of metallic profiles shaped as a coil; such profiles can be dry screwed in the masonry, thus mainly relying on a mechanical and frictional mechanism, or injected with resins/grout, in which case they work through the bond established between masonry and binder. Another popular system, increasingly used in conservation, is formed by steel sections provided with a fabric sleeve; profiles are installed in holes and then grout/resin is injected in the sleeve, so that the sleeve molds itself to the spaces and voids in the wall and creates a system combining mechanical locking and bond. A further advantage of this latter system is the
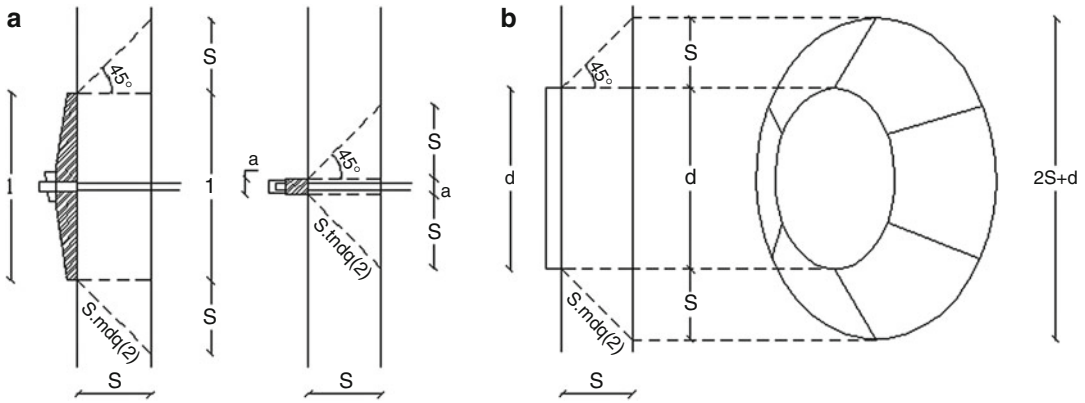
Seismic Strengthening Strategies for Heritage Structures, Fig. 10 Rupture surface for the calculation of tensile strength acting on parent material and sizing of the end plate in case of: (**a**) rectangular key or (**b**) circular end plate (http://postterremoto.altervista.org/Dati/i/i.html, last accessed 7th March 2011)
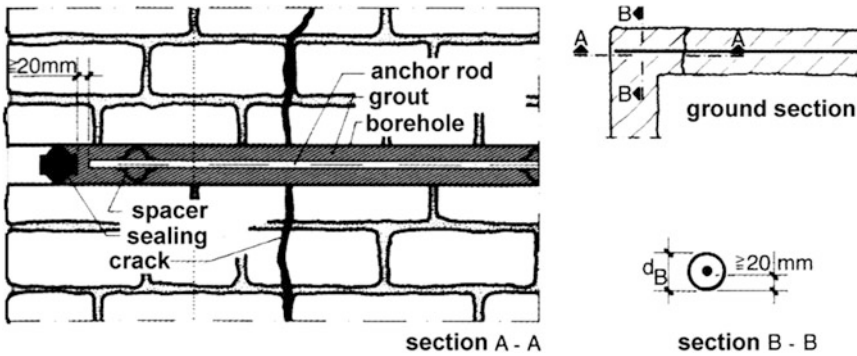
Seismic Strengthening Strategies for Heritage Structures, Fig. 11 Grouted anchor injection scheme (Gigla, Wenzel 2000)
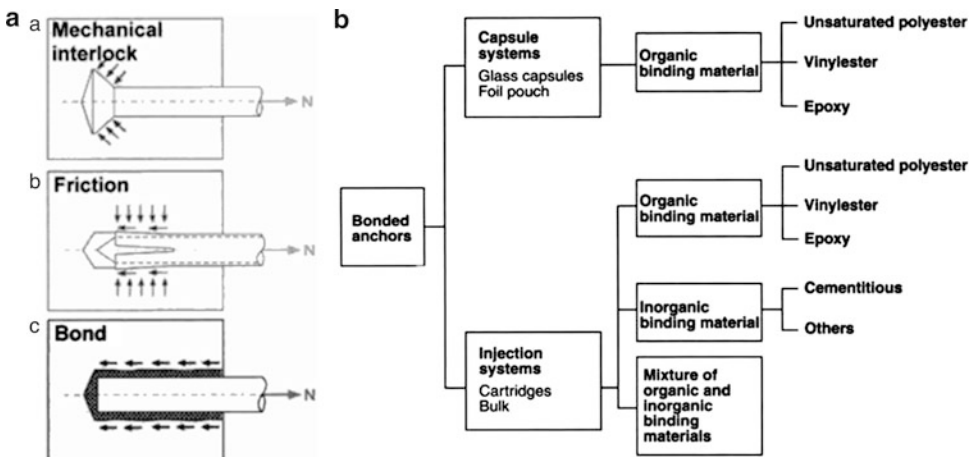
Seismic Strengthening Strategies for Heritage Structures, Fig. 12 Post-installed anchors: (**a**) systems for the transmission of load to substratum. (**b**) Typologies of binding materials for bond anchors (Eligehausen et al. 2006)

control over the diffusion of the binder within the parent material, while the grout can expand to the cavity around the ties, particularly useful for rubble double leaf masonry.

A variety of materials, ranging from mineral binders to polymers, can be used as binder. However, materials based on polymers, such as epoxy or polyester resins, which are often used for concrete structures, should undergo careful assessment before use in historical masonry, due to possible issues regarding the mechanical and physical compatibility. Much more common for the repair of historical buildings are mineral binder systems based on cement or hydraulic lime with the addition of admixtures and fillers or aggregate. To inject bore holes, usually pure water/binder systems are used with typical w/b values of 0.8–1.0. However, the w/b ratio has to be adjusted according to the volume to be injected and to the moisture content of the substrate.

Post-installed anchors are a very common method for the strengthening of historic masonry since they allow for a wide range of applications, from restoring the through-thickness cohesion of multilayered masonry to the increase of tensile and shear capacity of panels; furthermore, they can be used at the joints between structural elements either in the form of short connectors or as longitudinal elements. These latter have the same function as metallic crossties, but, rather than relying on an end plate for the connection to the wall, pullout loads are transmitted by a shear mechanism – for friction and bond anchors – or through mechanical locking within the masonry itself.

The diameter of the anchors is chosen as a function of the expected pullout force applied, the tensile strength of the masonry, and the bond strength of the grout to masonry interface. Bonded anchors can be passive or active, i.e., prestressed or posttensioned. Passive anchors are used for up to 4 m in length; prestressed anchors can be used up to 35 m in length (Gigla and Wenzel 2000).

The installation of anchors is carried out by:

- Drilling holes in the parent material; due to the weakness and preciousness of the parent

material, dry/wet diamond rotary drilling rather than percussive drilling is recommended.
- Removing all cores from the bore hole and checking the depth. Removing dust and debris.
- Placing anchors and injecting the binder.

The anchors utilized are usually stainless steel, threaded rods, or special prestressing bars. The diameter of the hole drilled to accommodate the anchor is a function of the strength differential between parent material and grout and this and the steel itself. The bond of the anchor inside the masonry unit depends on the type and properties of the grouting material and on the type of masonry block units.

The failure mechanisms of injection anchors in masonry are discussed in Gigla (2004, 2010), based on an extensive test program of over 500 pullout tests. The studies consider five different failure modes including failure mechanisms related to the (a) poor strength of the injected grout, (b) exceeded tensile strength of surrounding parent material, (c) bond failure between the outer surface of the grout and the parent material, (d) a combination of b and c, and (e) failure of the steel or grout in tension (under-designed).

Recommendations for the design of injection anchors are given in Gigla (2004). The design value for the bond strength is computed as follows:

$$X_{A,d} = \frac{\Phi_J}{\gamma_m}\left(\frac{f_{G,c}^2}{500} + X_{B,W}\right) \qquad (5)$$

$X_{A,d}$: Design value of bond strength, as a function of the compressive strength of grout and independent of bond length.

Required minimum bond length: $L_b = 150$ mm inside monolithic stone, $L_b = 190$ mm in bed or head joints of brick, and $L_b = 430$ mm in bed and head joints of blockwork.

$f_{G,c}$: Compressive strength of grout. Minimum value: $f_{G,c} = 16.6$ N/mm^2. Maximum value covered: $f_{G,c} = 38.7$ N/mm^2. A minimum

bending tensile strength of $f_{G,ct} = f_{G,c}/8 = 2.0$ N/mm^2 is required.

$f_{G,ct}$: Flexural strength of grout obtained from standard.

Φ_J: Reduction factor for bond in bed or head joints, $\Phi_J = 0.5$.

$X_{B,W}$: Term indicating the increase of bond strength inside water-absorptive stone material. $X_{B,W} = 0.15$ N/mm^2.

γ_M: Partial factor for property, recommended: $\gamma_M = 1.35$

The term $X_{B,W}$ considers the water absorption of the masonry substrate. Pullout studies revealed that materials with higher water absorption capacity yielded higher pullout strength than materials with low or no water absorption capacity.

Furthermore, the anchor capacity can be designed by first computing $R_{a,d}$ which refers to the bond capacity between grout and rod. The equation considers the ratio of bed and head joints of the borehole surface across bond length and limits $R_{A,d}$ to bond inside full stone sections ($A_B/A_{G,d} = 1$ completely in stone; $A_B/A_{G,d} < 1$ in joint):

$$R_{A,d} = X_{A,d} \cdot \frac{A_B}{A_{G,d}} \cdot A_{A,d} \qquad (6)$$

with $R_{A,d}$, design capacity of the injected anchor; A_B, portion of the cylindrical surface of injected mortar grout made of stones or units; $A_{G,d}$, total cylindrical surface of injected mortar grout; and $A_{A,d}$, interface of tensile element and injected mortar plug (cylindrical surface of steel bar), calculated with nominal bar diameter and bond length. The design strength of the anchor shall be smaller than the tensile capacity of the surrounding masonry so as to cause failure by yielding of the metallic element and prevent brittle failure of the masonry. This threshold can be computed as

$$F \leq \frac{1,9 \cdot f_{B,t} \cdot L_b \cdot \pi \cdot d_B \cdot (h_S{}^2 - d_B{}^2)}{\gamma_M \cdot \tan(\varphi) \cdot (d_B{}^2 + h_S{}^2)} \qquad (7)$$

where

F: anchor tensile force
$f_{B,t}$: tensile strength of surrounding stone
d_B: borehole diameter

h_S: minimal distance to edge of surrounding masonry

L_b: bond length

$\tan(\varphi)$: tangent of angle of friction for force transmission between anchor's grout and borehole, about 50° in water-absorptive material and about 60° in non-water-absorptive material

γ_M: partial factor of safety for stone tensile strength; recommendation, 1,5

While these values are appropriate for an axial pullout test, they do not apply to failures of anchors loaded in shear or bending as a result of relative movements between the structural elements that they reconnect. Furthermore, as the family of dowel anchors also includes short fixings such as pins and nail-like fixings, failures connected to loads transmitted from other elements and in proximity to edges or openings should also be analyzed, these being highly relevant to various typologies of strengthening, such as confinement or end plates, where the use of fixings is required but rarely regulated.

Whereas the dimensioning procedure of bonded anchors in concrete has been extensively studied and commented (Eligehausen et al. 2006) and European guidelines do exist (EOTA TR029 2010; EOTA TR045 2013; DD CEN/TS 1992-4-1:2009), anchors in masonry lack specifically dedicated codes or recommendations. Modes of failure and hence the procedure for dimensioning could be partly derived from DD CEN/TS 1992-4-1:2009. Possible modes of failure due to axial load may be bond between metal and binder or between binder and parent material; cone pullout of parent material, due either to the failure of bond in the mortar joints or to tensile failure of masonry units; and failure of steel, although this is highly unlikely unless the metallic section is severely under-designed. Possible modes of failure due to loading transversal to the anchor axis may be crushing failure of either the binder or the parent material, edge failure of masonry when the load is applied in proximity of openings and corners, pryout failure of parent material, or failure of steel for combined bending and shear action. Each of these failures is considered by the guidelines and the elements of the anchorage assembly dimensioned accordingly.

Other factors considered in the guidelines are the performance of anchors when undergoing fatigue loading and the dimensioning criteria for seismic loading. For anchors embedded in concrete, ductile failure is achieved by controlling the dimensioning of steel so as to avoid the fragile failure of the substratum, and this concept should be extended to masonry structures. However, prescriptions in EOTA TR029, EOTA TR045, and DD CEN/TS 1992-4-1:2009 rely on a series of parameters, such as bond strength or minimum distance to edge, that should be either provided by the producer of the anchors or derived experimentally. In case of anchors for masonry, due to the lack of standardization as well as the variability of parent materials and masonry fabrics, not every producer is able to provide all the required information and extensive tests covering the full range of parameters needed are missing from the scientific literature.

Energy Dissipation Systems

Systems like crossties have been and are still commonly applied in rehabilitation practice throughout Europe for providing connection at the joints of perpendicular sets of masonry walls where out-of-plane damage is most likely to occur. Nonetheless pullout damage at the head of the anchorage and increased in-plane diagonal cracking may affect valuable finishes and precious frescoes.

The current codes encourage the use of these conventional stiffness-based systems (EN 1998 Eurocode 8; Italian Ministry of Cultural Heritage and Activities 2006) because innovative techniques drawing on performance-based principles (Priestley 2000), despite their effectiveness in new structures, rarely meet the requirements of reversibility and low impact required for historic structures.

As widely experienced in many applications carried out in recent years, an effective alternative to conventional strengthening techniques is the use of passive energy control and dissipation techniques.

Energy dissipation systems indeed comply with the concept expressed in EC 8 (EN 1998-3:2005) that:

The selection of the type, technique, extent and urgency of the intervention shall be based on the structural information collected during the assessment of the building. The following aspects should be taken into account:

[..]

d) Increase in the local ductility supply should be pursued where required;

e) The increase in strength after the intervention should not reduce the available global ductility.

In the field of cultural heritage, such techniques must aim to enhance the seismic performance of structures as well as limit the aesthetic impact on the building and protect elements that, despite not being relevant to the structural behavior, have cultural and historic value.

Energy control systems can either provide the structure with additional dissipation capacity and/or reduce the amount of ground input energy transferred to the structure. However, base isolation systems require heavy interventions on the bearing structure (base cut, new foundation structure, etc.) and are therefore rarely suitable for the retrofit of heritage buildings.

Energy dissipation devices can instead be placed in a number of key locations within the structure where relative displacements between members are expected to occur; thus, from the point of view of intrusiveness, they do not differ from many other standard techniques. A suitable position for the installation of dissipative devices is at the wall-to-floor interface of masonry buildings, where relative motion can occur without impairing the global structural integrity.

Energy Absorbers

Benedetti (2004) developed a series of energy-absorbing devices for existing masonry buildings drawing on experimental results that had showed that the more energy is absorbed through damage by the noncritical elements of the structure, the later global failure occurs. The main requirements set for the devices were:

- Activation of the device for very low level of damage and cracking
- Sensitivity to both in-plane and out-of-plane movements
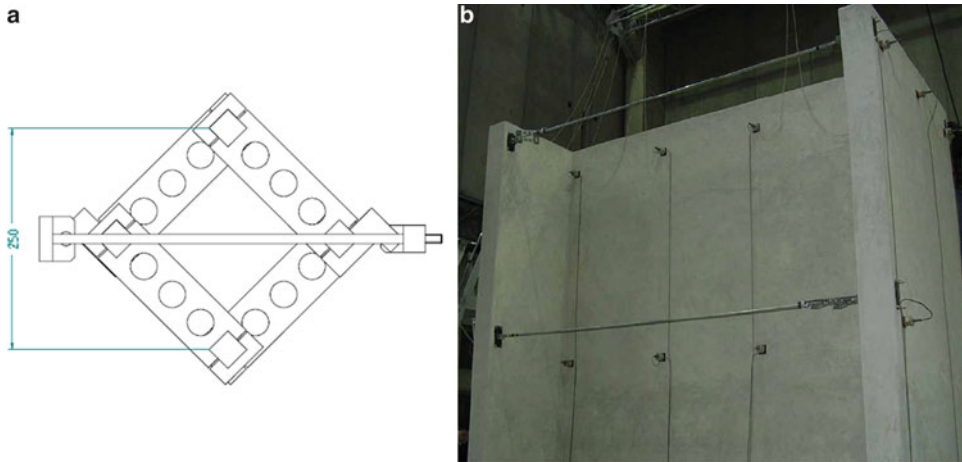
Seismic Strengthening Strategies for Heritage Structures, Fig. 13 RAG energy absorber: (**a**) prototype (Benedetti 2004); (**b**) setup in series with metallic crosstie installed on a masonry specimen (Benedetti 2004)

- Limitation of forces transmitted to the parent material at the fixings of the dissipative devices, so as to avoid localized damage and detachment of the devices

The RAG energy absorber (Fig. 13a) consists of four arms hinged at their end so as to form a square element. Hinges are made of lead cylinders press-fitted into the arms. When two (or four) of the corners of the device, which are connected to the masonry wall, displace relatively to each other as a consequence of damage in the parent material, the hinges deform plastically in torsion (in-plane action of the device) or bending (out-of-plane action of the device), thanks to the low yield strength of lead. Conversely, the arms are made of a stiffer material so as to remain in the elastic range. Besides a full characterization of the device, RAG energy absorbers were also tested inserted in full-scale masonry specimens on shaking table in an out-of-plane configuration (Fig. 13b), where the devices were used in series with traditional metallic crossties to connect two parallel walls. The holes that can be observed in the arms of the RAG device aim at reducing the weight of the element and avoiding a delay in the activation of the yielding mechanism. A pretensioning element was positioned in series with the anchor to further ensure the lack of any initial deformation in the crosstie.

Shock Transmission Units

STUs, or viscous dampers, consist of a hollow cylinder filled with fluid, this typically being silicone based. As the damper piston rod and piston head receive an impact, the fluid is forced to flow through orifices either around or through the piston head. The resulting differential in pressure across the piston head can produce very large forces that resist the relative motion of the damper, while the input energy is dissipated in form of heat due to friction between the piston head and fluid particles flowing at high velocities. Conversely, since this type of devices is velocity dependent, slow movements such as thermal expansions are allowed.

The viscous devices are characterized by a nominal strength: for high load rates and for input forces below their capacity, they show high stiffness; above it they feature a perfect plastic behavior, with dissipation of hysteretic energy in case of cyclic loads. The main drawback of this type of devices is the risk of leakage of the fluid, which involves the need for regular inspections and further costs involved with the maintenance of the dissipative system. Mandara and Mazzolani (2001) investigated two different methods for the dimensioning of viscous devices. In the Plastic Threshold Approach (PTA), devices are conceived and sized to limit the magnitude of force transmitted across connected

S

members to a maximum value, determined according to the design resistance of structural elements involved. Beyond this threshold, the hysteretic energy dissipation takes place, while below the threshold the behavior of structural members is virtually rigid, which ensures the maximum degree of redundancy to the structure under serviceability load conditions. According to the Optimal Viscous Approach (OVA), the interaction between connected members is controlled by the viscous properties of the devices, which are dimensioned so as to minimize the magnitude of the force acting on them independently from its value. Contrary to PTA, the connection between elements is never fully rigid, so that energy can be dissipated under moderate intensity earthquakes too.

Shape Memory Alloys

Shape Memory Alloys (SMAs) are metallic materials that show special thermomechanical properties due to a reversible transformation between two crystalline configurations, austenite and martensite, without degradation of the crystal structure.

For some SMAs, the phase transformation is temperature-correlated, while for others, such as nitinol (NiTi SMA), the phase change can be stress induced at room temperature.

The stress–strain curve, measured during a monotonic tension test on a NiTi SMA wire, shows:

(a) An almost linear portion corresponding to the elastic deformation of the material in its austenitic phase.
(b) A loading plateau due to the transformation in martensite.
(c) A new elastic phase that initiates after the complete transition to the martensitic phase (whose upper boundary is called maximum superelastic strain).
(d) A final plateau, related to the true yielding of the alloy, which ends with failure.
(e) Stress removal causes a reverse phase transformation, where the material goes back to the austenitic phase, stable for lower load's values.
(f) Strains are almost completely recovered. This property is known as superelasticity.

These characteristics make SMAs particularly suitable for use in seismic dampers, especially considering that some alloys, such as nitinol, have very good corrosion-resistance characteristics.

Shape Memory Alloy Devices (SMADs), which consist of metallic ties and groups of wires of SMAs, were developed within the framework of the ISTECH project (Indirli et al. 2001) and patented by Fip Industriale, Padova, Italy. Their application has featured in high-profile projects such as the repair and strengthening of the Upper Basilica of St. Francis in Assisi, San Feliciano cathedral in Foligno, and San Serafino Church in Montegranaro, Italy. In the Upper Basilica of St. Francis, 47 SMADs were used to connect the roof to the tympanum of the transept; the SMADs were connected on one end to a threaded bar attached to an anchorage plate in the façade and on the other end to a plate bolted to a counter plate embedded in the RC rib that was cast to stiff the existing concrete roof (dating back to the 1950s). SMADs were three-plateau self-balanced devices of three different sizes, with design capacity ranging from 17 to 52 kN and maximum allowable displacement between ±8 and ±25 mm.

Dissipative Anchor Devices

Given the difficulty of ensuring ductile failures of anchors embedded in masonry, D'Ayala and Paganoni (2014) in collaboration with CINTEC International developed two prototypes of dissipative anchor devices to address the problem of out-of-plane mechanisms.

The devices are conceived to be inserted at the joint between perpendicular walls, as part of longitudinal steel anchors grouted within the thickness of the walls. Devices are designed to work in tension, similarly to traditional crossties, which experience axial loading when one wall panel tilts outwards as a consequence of ground acceleration. This type of installation ensures a low impact on the aesthetics of the building as it does not affect the finishes.

One device relies on the plastic behavior of steel and consists of a metallic element designed to achieve lower capacity than the standard tie to
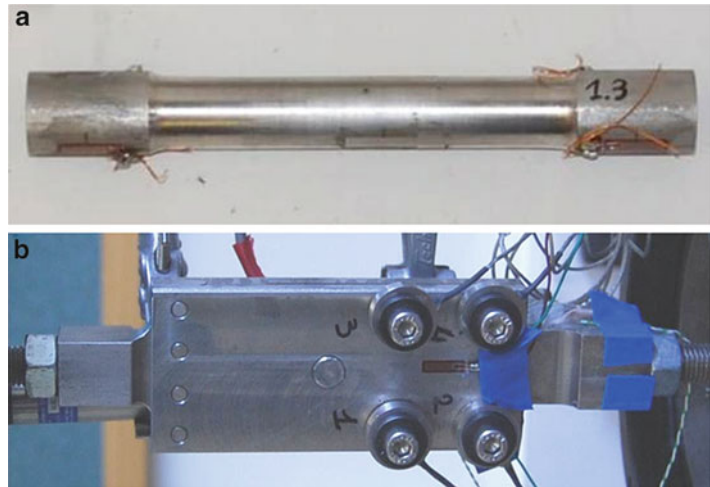
which it is connected (Fig. 14a). The second device (Fig. 14b) is made of a set of small plates; bolts are used to apply a perpendicular pressure creating friction and adjusting the level of slip load beyond which relative sliding occurs. Specially designed stops ensure that the sliding motion remains in the desired range.

While the metallic profiles improve the box-like behavior of the building, contributing to an increase in stiffness that improves the structural response to low excitations, the devices allow small relative displacements between orthogonal sets of walls; additionally, for higher horizontal loads, they dissipate part of the energy input into the structure so that problems of localized damage can be avoided. Therefore, the design focuses on control of displacements and reduction of accelerations and stress concentration.

The two dissipative devices have been validated by cyclic pseudo-static and dynamic tests on the isolated devices and by pullout tests on the devices anchored in series to a conventional post-installed anchor embedded in low-shear capacity masonry (D'Ayala and Paganoni 2014).

Experimental results confirmed that both the yielding and frictional elements are able to provide the anchorage with ductility and prevent damage to occur in either the grouted sleeve or the masonry, while traditional anchors, which were tested for comparison purposes, display a high stiffness and fail at the interface between grout and parent material or in the masonry units. In the case of friction devices, it was observed that, due to tolerances of the pieces composing the assembly, a locking phenomenon occurred for the higher levels of perpendicular pressures; consequently, load-deflection curves feature additional stiffness in respect to the flat branch typical of a friction mechanism. Such behavior will need to be corrected so that the device can provide homogeneous performance over the required range of perpendicular forces.

The dimensioning procedure is based on the idea that a strengthening system can be divided into subcomponents to which one type of failure controlled by a single parameter can be associated. These parameters can be used in the formulae prescribed by design codes to calculate the capacities of components and hence determine the hierarchy of failure. The parameter values can be obtained via a number of sources: producers' specifications, recommended or limit values provided by codes, and, if any, laboratory and/or on-site tests.

Table 1 exemplifies the procedure for the calculation of the tensile capacity of the dissipative anchoring devices in series with a grouted metallic anchor. Maximum capacity is reached if one of the components fails or when the dissipative device is activated. Further checks can also be

S

Seismic Strengthening Strategies for Heritage Structures, Table 1 Design of dissipative anchors: parameters that identify the tensile capacity, value range achieved during experimental validation, and range prescribed/recommended by design codes (From D'Ayala and Paganoni 2014)

Performance parameters	Achievable range			Expected range		
1a) F_{yield}: yielding capacity of hysteretic dissipative device [kN]	F_{yield}=33 kN (for hysteretic device of size suitable to coupling with M16 threaded bar)			F_{yield}=27.8 kN; calculated as: $F_{yield}=f_{y,yield}A_{yield}$ with $f_{y,yielding}$ yielding strength of steel of hysteretic element and A_{yield} net cross sectional area of hysteretic element (EN 1993-1-1:2005)		
1b) $F_{//}$: slip-load of frictional dissipative device [kN]	Considering: ✓ $F_\perp$: initial value imposed on devices. Variations recorded during tests are not considered; ✓ Slip load is given as range of values between maximum and minimum recorded values at constant level of $F_\perp$.			Calculated as: $F_{//}=\Phi n F_\perp$ with Φ coefficient expressing the ratio between $F_\perp$ and $F_{//}$, n=2 number of frictional surfaces and $F_\perp$ applied perpendicular pressure.		
	$F_\perp$ [kN]	$F_{//}$ min [kN]	$F_{//}$ Max [kN]	$F_\perp$ [kN]	$F_{//}$ (Φ=0.15)	$F_{//}$ (Φ=0.55)
	12.5	3.25	14.5	12.5	3.75	13.75
	15	5.7	18.3	15	4.5	16.5
	17.5	6.65	22.4	17.5	5.25	19.25
2) F_{steel}: tensile capacity of metallic bar at yielding [kN]	F_{steel}=71 kN (for M16 threaded bar - values stated by producer)			F_{steel}=71 kN; calculated as: $F_{steel}=f_y A$ with f_y yielding strength of steel and A net cross sectional area of metallic profile (EN 1993-1-1:2005)		
3) $f_{b\,a/b}$: bond strength anchor/binder [MPa] calculated on cylindrical surface of embedded bar	Calculated as: $f_{b\,a/b}=F_{s/b\,bond}/A_{steel}$ with $F_{s/b\,bond}$ recorded load at failure and A_{steel} cylindrical lateral surface calculated as: $A_{steel}=\pi l d_{pitch}$ with l embedment length and d_{pitch} pitch diameter of steel bar. For pull-out tests of M16 threaded bars from 550 mm long grouted socks: $f_{b\,a/b}$=2.07 MPa (CoV 4%)			$f_{b\,a/b}=$ 3.4 MPa – design value suggested in BS 5268-2 for tested binder, bar diameter and type of bar 2 MPa – design value suggested in EN 1996-1-1:2005 for tested binder and type of application		
4) $f_{b\,b/p}$: bond strength binder/parent material [MPa] calculated on cylindrical surface of grouted socket	Calculated as: $f_{b\,b/p}=F_{b/p\,bond}/A_{hole}$ with $F_{b/p\,bond}$ recorded load at failure and A_{hole} inner cylindrical surface of drilled hole of length l. For pull-out tests with vertical load on masonry specimens σ_d:			Calculated as: $f_{b\,b/p}=f_{vk}=f_{vk,0}+0.4\sigma_d$ with $f_{vk,0}$ initial shear strength (calculated through experimental results) and σ_d vertical load (EN 1996-1-1:2005).		
	l [mm]	σ_d [MPa]	$f_{b\,b/p}$ [MPa]	σ_d [MPa]	$f_{b\,b/p}$ [MPa]	
Brick masonry, f_c=6.7 MPa, f_w=0.7 MPa	350	0.70	0.67 (CoV 8%)	0.7	0.52	
		0.07	0.57 (CoV 18%)	0.07	0.27	
Brick masonry f_c=3.1 MPa, f_w=0.33 MPa	220	0.10	0.26 (CoV 34%)	0.10	0.08	
		0.05	0.4	0.05	0.06	
5) $f_{masonry}$: Shear strength of parent material [N/mm²]	This type failure, although expected, did not occur during experimental campaigns			Calculated as: $f_{masonry}=f_{vk}=f_{vk,0}+0.4\sigma_d$ (EN 1996-1-1:2005). In the tested case it would be expected: 0.52 MPa 0.27 MPa The failure surface, A_f, is a truncated cone with smallest base corresponding to the drilled hole, apothem inclined at 45° and height equal to the wall thickness		
6) $f_{masonry}$: Tensile strength of parent material [N/mm²]	A "wrench" failure occurs instead of the expected "cone pull-out" failure. Failure surface, A_f, develops along vertical joints. $f_{masonry}=f_w$=0.67 MPa (from wrench test)			No mention about this type of failure has been found in the technical literature or design codes.		

performed in terms of displacement and ductility, as discussed in the following.

The demand in terms of tensile capacity of metallic anchors is calculated as:

$$F_D = M \cdot a_i \qquad (8)$$

where

M: mass of structure that bears on the ith anchor of the strengthening system. M depends on the geometry and construction arrangement of the building, including horizontal structures, and on the layout of the set of anchors to be designed.

a_i: horizontal acceleration experienced by the mass M. An estimate of the natural period of the system can be used to determine the correct design spectral ordinate and the distribution of amplification over the height of the structure. An accurate estimate of the acceleration at the height of the anchor can be obtained either by using the first natural modal shape or the procedure proposed by Miranda and Taghavi (2005).

The reference acceleration is calculated as function of the three limit states defined in EN 1998-3:2004, so that the design demand is:

- F_{DNC}: near collapse, calculated for a seismic action with a probability of exceedance of 2 % in 50 years
- F_{DSD}: significant damage, calculated for a seismic action with probability of exceedance of 10 % in 50 years
- F_{DDL}: damage limitation, calculated for a seismic action with probability of exceedance of 20 % in 50 years

All the subcomponents of the strength-only portion of the anchor assembly are brittle or, in the case of the grouted steel elements, are not supposed to experience large deformations; therefore, they are dimensioned in terms of strength for near collapse limit state, according to EN 1998-3:2004. The minimum capacity in the assembly must be:

$$\text{Min} \left(F_{steel}, \ F_{a/b \ bond}, \ F_{b/p \ bond}, \ F_{masonry} \right) > F_{DU} \qquad (9)$$

Capacities are calculated using information in the relevant row in Table 1. It is important to notice that while certain values in Table 1 are generally applicable, others are not: the standardization of steel production ensures highly repeatable performance, whereas bond strength, for instance, largely varies depending on the substratum. Therefore, it is advisable that tests be performed each time; if this cannot be done, lower limit values provided by the code and reported in Table 1 can be used.

Table 1 does not include the connections' anchor/dissipative devices and the stops of the friction device. However, these components should also be dimensioned to resist F_{DNC} as also required by EN 15129:2009.

The dissipative elements of the devices, either hysteretic or frictional, are designed to be activated at the threshold of damage limitation, when cracks start opening and allow for the dissipative elements to become active. When the hysteretic devices enter the plastic field, or the friction devices start sliding, the connection between wall panels is still maintained, but the pullout of the head of the anchorage is prevented and drift controlled.

Hence, from the point of view of force design, depending on which dissipative device is used, the following requirements apply:

$$\begin{aligned} F_{yield} &= F_{DD} \text{ or} \\ F_{//} &= F_{DD} \end{aligned} \qquad (10)$$

where F_{yield} is the yielding capacity of the hysteretic device and $F_{//}$ is the slip load of the frictional device.

The dissipative devices should also comply with requirements for interstorey drift of buildings undergoing seismic action. The chosen value of maximum allowable drift, $d_r = 0.003$, for damage limitation is taken from OPCM (2005) which provides drift limits for damage limitation in walls of masonry buildings. This limit is also in line with the expected drift stated in FEMA 356 (BSSC 2000) for unreinforced masonry buildings at the limit state of immediate occupancy.

S

For historic buildings, the allowable drift is increased by using the reduction factor $v = 0.4$, which is taken from EN 1998:2004 and accounts for the fact that devices are designed to be used in heritage structures, which fall in the importance category III of Eurocode 8 (EN 1998:2004).

It is therefore

$$d_r = \Delta_e < 0.003 \cdot h/v \qquad (11)$$

where

Δ_e: the elongation of the device before yielding, in case of the hysteretic device, or before activation of the friction mechanism, in case of the frictional device,

h: interstorey height, or vertical distance of installation of anchors. A standard distance of 3 m has been assumed in the calculations, but anchors might need to be spaced more closely along the height of the wall to prevent substratum failures.

The first threshold of the hysteretic device, identified at 0.5 % elongation of the dissipative element, coincides with 46 % of the hysteretic device load capacity. The second threshold, at 5 % elongation and 72 % maximum load, should instead be used to verify the capacity of the dissipative device for the limit state of significant damage, so that dissipation of energy is ensured during low-to-medium seismic excitations. Beyond this limit, the device has a further margin of safety given by buckling, meaning that the device can reach the limit of near collapse with damage, but not complete failure, and it could still be substituted, as it has been proved by experimental testing and in response to the requirements of EN 1998:2004.

In the case of a frictional device, the drift limit is ensured by default because: before activation of the friction mechanism, deformations are negligible and, beyond activation of sliding, the device displacement is limited by the assembly stops. The device can therefore perform for all limit states, as long as the connections and stops in the assembly are designed to resist up to the state of near collapse.

Summary

In the last decades, a number of technical solutions for the improvement of structural connections have been developed in response to the increasing demand for strengthening systems specifically designed for the protection of heritage assets from earthquake-induced damage.

Notwithstanding recommendations from seismic standards and conservation codes of practice, and the availability of alternative technical solutions on the market, much emphasis when choosing strengthening solutions of connections is still placed on their capacity to enhance the overall strength of the system, rather than improve its ductility or ability to dissipate energy. When connections between vertical elements, and between vertical and horizontal elements, as well as vaults and domes have sufficient capacity, vertical and horizontal loads are better distributed, out-of-plane mechanisms of damage are prevented, and, hence, failure of floors and roofs can be avoided, substantially reducing the probability of collapse and fatalities in an earthquake.

Historically, builders mastered the art of careful detailing of joints from a process of trial and error; today's engineers possess the necessary insight into physical and mechanical laws governing the dynamics of structures to control the process of structural upgrading at the design level. Thus, when connections, in existing buildings, lack adequate capacity, further elements can be added to the original structure to achieve the desired performance.

The review of modern techniques to enhance connections presented in the previous section highlights aspects specific to each applications as well as advantages and pitfalls typical of each system. A large majority of the discussed methods are applicable both at the local and global scale for repair and upgrade of single structural elements or of whole structures.

In respect to the past, strengthening systems for structural connections rely on a more accurate design, modern structural systems, and innovative and more durable materials, for example, stainless steel or titanium, which substituted iron in crossties, with considerable advantages

Seismic Strengthening Strategies for Heritage Structures, Table 2 Performance parameters typical of each strengthening technique

Typology of strengthening		Performance parameters
Ring beams		Flexural, tensile, and shear capacity of beam (kNm/kN)
		Capacity of fixings between ring beam and parent material (crushing or pryout of parent material, shearing off of fixings, shearing of parent material, etc.) (kN)
		Horizontal sliding shear capacity of the masonry underneath the level of the fixings of the ring beam into the parent material (−)
Wall plates		Flexural, tensile, and shear capacity of wall plates (kNm/kN)
		Capacity of fixings between wall plates and parent material (crushing or pryout of parent material, shearing off of fixings, shearing of parent material, etc.) (kN)
		Horizontal sliding shear capacity of the masonry underneath the level of the fixings of the ring beam into the parent material (−)
Steel connectors for floor/roof structures		Tensile capacity (kN) as minimum depending on
Anchoring with end plates		Combined tensile and bending strength of connector/tie (N/mm^2)
Anchoring without end plates and with grouting		Tensile capacity of fixings in timber elements (kN)
Nailing		Bond strength (for bonded systems) connector/binder (N/mm^2)
		Bond strength (for bonded systems) binder/parent material (N/mm^2)
		Tensile capacity of fixings (frictional or mechanical fixings) in parent material (kN)
		Tensile strength of parent material for cone pullout or punching shear (N/mm^2)
		Reduction of tensile capacity due to distance to edge (%)
		Reduction of tensile capacity due to distance between anchors, nails, and connectors (%)
		Compressive strength of parent material under action of end plate (N/mm^2)
		Flexural capacity of end plate (kN)
		Shear capacity (kN) as minimum depending on
		Shear strength of anchor (N/mm^2)
		Tensile strength of anchor for combined axial and bending load (N/mm^2)
		Bearing capacity of parent material (N/mm^2)
		Bearing capacity (for bonded system) of binder (N/mm^2)
		Shear capacity of fixings in timber elements (kN)
		Reduction of shear capacity due to distance to edge (%)
		Reduction of shear capacity due to distance between anchors (%)
Corners confinement	Column ties	Flexural, tensile, and shear capacity of columns (kNm/kN)
		Capacity of fixings between ring beam and parent material (crushing or pryout of parent material, shearing off of fixings, shearing of parent material, etc.) (kN)
	FRPs	Tensile capacity of FRPs (kN)
		Bond strength between FRP/parent material (N/mm^2) and the tensile and compressive capacity of the parent material (N/mm^2)
	Polymer mesh, PP mesh, steel mesh	Tensile capacity (kN) as minimum depending on
		Tensile strength of mesh (N/mm^2)
		Tensile strength of fixings (N/mm^2)
		Bond strength fixing/binder (N/mm^2)
		Bond strength binder/parent material (N/mm^2)
		Shear strength of parent material for cone pullout (N/mm^2)

(*continued*)

S

Seismic Strengthening Strategies for Heritage Structures, Table 2 (continued)

Typology of strengthening		Performance parameters
		Shear capacity (kN) as minimum depending on
		Shear strength of parent material at the interface strengthened/unstrengthened material
		Shear strength of fixings (N/mm^2)
		Bearing capacity of binder (N/mm^2)
		Bearing capacity of parent material (N/mm^2)
Energy dissipation systems	Yielding anchor devices	Yielding load of device (kN)
		Allowable displacement after yielding (mm)
		Dissipated energy (kJ)
		Change in material properties during cyclic loading (e.g., hardening of metal, progressive crushing of parent material)
		Buckling of device
		Ultimate tensile capacity (kN) as minimum depending on
		Tensile strength of device (N/mm^2)
		Other checks same as traditional anchors
		Shear capacity (kN) as minimum depending on
		Shear strength of device (N/mm^2)
		Other checks same as traditional anchors
	Frictional anchor devices	Load for activation of frictional mechanism (kN)
		Ultimate tensile load of device (kN)
		Ultimate allowable displacement for sliding (mm)
		Dissipated energy (kJ)
		Change in material properties during cyclic loading (e.g., hardening of metal, deterioration of frictional surfaces, effect of thermal expansion, presence of debris on frictional surfaces, progressive crushing of parent material)
		Ultimate tensile capacity (kN) as minimum depending on
		Tensile capacity of device (kN)
		Other checks as traditional anchors
		Shear capacity (kN) as minimum depending on
		Checks as per traditional anchors
Energy dissipation system	RAG energy absorbers	Axial yielding load that leads to yielding of lead pins (kN)
		Allowable displacement after yielding (mm)
		Dissipated energy (kJ)
		Change in material properties during cyclic loading (e.g., hardening of metal, progressive crushing of parent material)
		Ultimate tensile capacity (kN) as minimum depending on
		Tensile strength of anchor (N/mm^2)
		Tensile strength of RAG device
		Tensile strength of parent material to punching failure (N/mm^2)
		Compressive strength of parent material under action of end plate (N/mm^2)
		Flexural capacity of end plate (kN)
	RETE energy absorbers	Yielding load of the lead wire (kN)
		Ultimate tensile load of device (kN)
		Allowable displacement after yielding (mm)
		Dissipated energy (kJ)
		Change in material properties during cyclic loading (e.g., hardening of metal, progressive crushing of parent material)

Seismic Strengthening Strategies for Heritage Structures, Table 2 (continued)

Typology of strengthening		Performance parameters
	Shock Transmission Units (STUs) and viscous dampers	1st threshold – thermal deformations
		Allowable displacement (mm)
		Transmitted load for low-rate loading (kN)
		2nd threshold – low seismic excitation, connections expected to perform as pinned connections
		Allowable displacements (mm)
		Capacity (kN)
		Capacity of connectors between STUs and parent material at 1st threshold (kN)
		3rd threshold – higher seismic excitation, connections expected to have hysteretic response
		Allowable displacement (mm)
		Transmitted load (kN)
		Dissipated energy for hysteretic response (kJ)
	Shape Memory Alloy Devices (SMADs)	1st threshold – thermal deformations
		Allowable displacement (mm)
		Transmitted load for low-rate loading (kN)
		2nd threshold – low seismic excitation, connections expected to perform as pinned connections
		Allowable displacements (mm)
		Capacity (kN)
		Capacity of connectors between STUs and parent material (e.g., pullout or tensile failure of fixings) (kN)
		3rd threshold – higher seismic excitation, connections expected to have hysteretic response
		Allowable displacement (mm)
		Transmitted load (kN)
		Dissipated energy for hysteretic response (kJ)
		ith threshold: same parameters as per 2nd and 3rd threshold repeated for the higher plateaus of Shape Memory Alloys

in terms of issues related to material deterioration (e.g., expansion due to corrosion, failures due to reduced resisting section).

Some of the techniques described have undergone revision after observation of their performance in the aftermath of major earthquakes proved that they are unsuited to certain applications. For instance, RC ring beams need to be especially well connected to bearing walls; otherwise, they tend to create shear failure in walls and collapse; furthermore, it has been shown that they are not a viable solution in case of multileaf masonry not well bonded.

Recent strengthening solutions are evolving to include issues of environmental compatibility while remaining economically feasible, especially for vulnerable communities in developing countries.

The most important shift for future implementation is the increased awareness and technological research for economic solutions for ductility and energy dissipation-based devices, which can allow the improvement of performance of entire building stock in historic centers of earthquake-prone countries.

The performance of each system can be verified by determining the parameters and carrying out the checks summarized in Table 2. The parameters listed refer, when available, to dimensioning procedures prescribed by codes and

guidelines and otherwise from the available technical literature included in this review. Some of the systems in the table are not codified; therefore, the dimensioning procedure is not standardized or sometimes not reported for intellectual property reasons; in these cases parameters are suggested by the authors on the basis of their expertise and of the available information on the behavior of the various devices.

The major advantage of Table 2 is that parameters have been homogenized, using a common nomenclature, to facilitate the comparison of various techniques.

Cross-References

▶ Passive Control Techniques for Retrofitting of Existing Structures
▶ Retrofitting and Strengthening of Structures: Basic Principles of Structural Interventions
▶ Strengthening Techniques: Masonry and Heritage Structures

References

Benedetti D (2004) Increasing available ductility in masonry buildings via energy absorbers. Shaking table tests. Eur Earthq Eng 18(3):3–16

Borri A, Castori G, Grazini A, Giannantoni A (2007) Performance of masonry elements strengthened with steel reinforced grout. In: Fiber-reinforced polymer reinforcement for concrete structures – FRPRCS-8, Patras, 16–18 July

Borri A, Castori G, Grazini A (2009) Retrofitting of masonry building with reinforced masonry ring-beam. Constr Build Mater 23(5):1892–1901

Building Seismic Safety Council (BSSC) (2000) Prestandard and commentary for the seismic rehabilitation of buildings, FEMA, 356. Federal Emergency Management Agency, Washington, DC

California Historical Building Code (2007) California code of regulations. Title 24, part 8. California Building Standards Commission, International Code Council, Washington, DC

Casadei P, Agneloni E (2008) Advance composites applications on historical structures in Italy: Case studies and future developments 6th International Conference on Structural Analysis of Historical Construction

Chuxian S, Guiqiu L, Wenchao W (1997) The design of brick masonry structure with concrete column. In: XI brick and block masonry conference, Shanghai

CNR-DT 200 R1/2013. Guide for the Design and Construction of Externally Bonded FRP Systems for Strengthening Existing Structures, ROMA – CNR October 10th 2013 – release of May 15th 2014

Corradi M, Grazini A, Borri A (2007) Confinement of brick masonry columns with CFRP materials. Compos Sci Technol 67:1772–1783

D'Ayala D, Paganoni S (2011) Assessment and analysis of damage in L'Aquila historic city centre after 6th April 2009. Bull Earthq Eng 9(1):81–104

D'Ayala D, Paganoni S (2014) Protocol for testing and design of dissipative devices for out-of-plane damage Struct Build 167(1):26–40

D'Ayala D, Speranza E (2003) Definition of Collapse Mechanisms and Seismic Vulnerability of Historic Masonry Buildings. Earthquake Spectra 19 (3):479–509. doi:10.1193/1.1599896

D'Ayala D, Yeomans D (2004) Assessing the seismic vulnerability of late Ottoman buildings in Istanbul. In: IV international seminar on structural analysis of historical constructions – SAHC04, Padua

DD CEN/TS 1992-4-1:2009. Design of fastenings for use in concrete

Di Ludovico M, D'Ambra C, Prota A, Manfredi G (2010) FRP confinement of tuff and clay brick columns, experimental study and assessment of analytical models. J Compos Constr 14(5):583–596

Eligehausen R, Mall R, Silva JF (2006) Anchorage in concrete construction. Ernst Sohn, Berlin

EN 1991-1:2002. Eurocode 1- Actions on structures

EN 1992-1:2004. Eurocode 2- Design of concrete structures

EN 1993-1:2005. Eurocode 3- Design of steel structures

EN 1995-1:2004 + A1:2008. Eurocode 5- Design of timber structures

EN 1996-12005. Eurocode 6- Design of masonry structures

EN 1998-1:2004. Eurocode 8- Design of structure for earthquake resistance. Part 1: General rules, seismic actions and rules for buildings

EN 1998-3:2005. Eurocode 8- Design of structure for earthquake resistance. Part 3: Assessment and retrofitting of buildings

EN 15129:2009: Anti-seismic devices

EOTA, European Organisation for Technical Assessments (2010) TR 029, Design of bonded anchors, Sept 2010. Downloaded at www.eota.eu. Accessed 8 Jan 2014

EOTA, European Organisation for Technical Assessments (2013) TR 045, Design of metal anchors for use in concrete under seismic actions. Downloaded at www.eota.eu. Accessed 8 Jan 2014

Gigla B (2004) Bond strength of injection anchors as supplementary reinforcement inside historic masonry. In: XIII international brick and block masonry conference, Amsterdam, pp 1–10

Gigla B (2010). Comparison of failure of injection anchors as supplementary reinforcement inside masonry and concrete. In: VIII international masonry conference, Dresden, pp 2019–2028

Gigla B, Wenzel F (2000) Design recommendations for injection anchors as supplementary reinforcement of historic masonry. In: XII international brick and block masonry conference, Madrid, pp 691–706

Giuffrè A (1993) Sicurezza e conservazione dei centri storici: il caso di Ortigia. Laterza, Roma-Bari

Hamoush SA, McGinley MW, Mlakar P, Scott D, Murray K (2001) Out-of-plane strengthening of masonry walls with reinforced composites. J Compos Constr 5(3):139–145

Indirli M, Castellano MG, Clemente P, Martelli A (2001) Demo-application of shape memory alloy devices, the rehabilitation of S. Giorgio Church Bell-Tower. In: VI international symposium on smart structures and materials – SPIE 2001, Newport Beach

Ministero deri Beni Architettonici e Culturali, Italia, Guidelines for evaluation and mitigation of seismic risk to cultural heritage : (2007) Cangemi Editore S.p.A.

Karantoni F, Fardis M (1992) Effectiveness of seismic strengthening techniques for masonry buildings. ASCE 118(7):1884–1902

Krevaikas TD, Triantafillou T (2006) Masonry confinement with fiber-reinforced polymers. ASCE J Compos Constr 9(2):128–135

Mandara A, Mazzolani FM (2001) Energy dissipation devices in seismic up-grading of monumental buildings. In: III international seminar on structural analysis of historical constructions – SAHC01, Guimarães

Marcari G, Manfredi G, Pecce M (2003) Experimental behaviour of masonry panels strengthened with FRP sheets. In: Tan KH (ed) 6th international conference on fibre-reinforced plastics for reinforced concrete structures, World Scientific, Singapore, pp 1209–1218

Mayorca P, Meguro K (2004) Proposal of an efficient technique for retrofitting unreinforced masonry dwellings. In: XIII world conference on earthquake engineering, Vancouver

Miranda E, Taghavi S (2005) Approximate floor acceleration demands in multistory buildings. I: formulation. J Struct Eng 131:203–211

OPCM (2005) No. 3431, May 3, 2005. Official Bulletin no. 107, May 10, 2005 (in Italian). Gazzetta Ufficiale – Serie Generale n. 107

Priestley MJN (2000) Performance based seismic design. In: XII world conference of earthquake engineering, Auckland

Spence R, D'Ayala D (1999) Damage assessment and analysis of the 1997 Umbria-Marche earthquakes. Struct Eng Int 9(3):229–233

Tomazevic M (1999) Earthquake-resistant design of masonry buildings. Imperial College Press, London

Tumialan G, Huang P-C, Nanni A, Silva P (2001) Strengthening of masonry walls by FRP structural repointing. In: Burgoyne CJ (ed) 5th international conference on fibre-reinforced plastics for reinforced concrete structures, Thomas Telford, Cambridge, pp 1033–1042

Seismic Tomography of Volcanoes

Ivan Koulakov[1,2] and Nikolay Shapiro[3]
[1]Trofimuk Institute of Petroleum Geology and Geophysics, SB RAS, Novosibirsk, Russia
[2]Novosibirsk State University, Novosibirsk, Russia
[3]Institut de Physique du Globe de Paris Laboratoire de Sismologie, Paris, France

Synonyms

Ambient noise tomography; Body-wave tomography; Magma sources; Seismic observations on volcanoes; Seismic properties of crust; Seismic tomography

Introduction

Volcano tomography is a branch of geophysics oriented to studying the deep structures beneath volcanoes by means of seismic tomography. **Seismic tomography** is a method for reconstruction of continuous distribution of seismic parameters in 1D, 2D, 3D, or 4D (space and time) using the characteristics of seismic waves traveling between sources and receivers. Seismic parameters to be found in tomographic inversion are in most cases velocities of P and S seismic waves (P and S velocities). For volcanoes, one of the key parameters appears to be the Vp/Vs ratio which can be used to evaluate the content of fluids and melts. Besides the velocity distributions, seismic tomography may provide the information on the anisotropy of seismic parameters which helps studying regional stresses and space-oriented geological structures. In some tomography studies, the target parameter might be the attenuation of P or S waves which may also give important information on magma sources beneath volcanoes.

At all stages of the human history, understanding the mechanisms causing volcano eruptions was one of the most intriguing problems.

S

Until relatively recent time, scientists had only indirect ways to explore processes beneath volcanoes, mainly by making analogies with geological signatures of ancient magmatic systems. During the last decades, the advancement of geophysical methods made possible direct observations of deep structures beneath presently active volcanoes which give precious information about general mechanism of working the volcanic systems. Among all geophysical methods, seismic tomography is one of the most powerful and effective approaches to look at great depths below the volcanoes. Using powerful artificial and natural sources of seismic signal makes it possible getting the information from depths where the main magma sources are located.

Observation Schemes

The configurations and characteristics of observation systems in seismology depend on the properties and sizes of the target objects, as well as on the type of seismic sources. In case of volcano seismology, the parameters of the target volume depend on the type of volcanoes. For example, the subduction-related volcanoes usually have complex multilevel structures of feeding system which may cover the depths down to 100–200 km. In the case of the hot-spot volcanoes, the initial source of magmatism is related to mantle plumes which may propagate throughout the mantle; however, the magma reservoirs and the related seismicity are fairly shallow in this case (only the first kilometers of depth). When designing an experiment for tomography studies of volcanoes, one should clearly figure out the target properties and define the configuration of the network according to this. In reality, there are many limitations, such as insufficient amount of instruments and difficulties in access to some areas around a volcano that make impossible deploying an ideal network. In the following section, the major components of the observation systems are described, and some recommendations to optimize the designing of the seismic networks are given.

Seismic Sources

Sources used for seismic imaging of volcanoes can be divided in three categories: human-made sources, earthquakes, and ambient seismic noise. Fully controlled artificial sources are expected to provide the highest accuracy of tomographic images. At the same time, their implementation on volcanoes is very sophisticated and expensive. Another shortcoming of artificial sources is that they cannot be produced at depth. In regions with well-distributed volcanic or local tectonic seismicity, earthquakes may give a distribution of sources plausible for high-resolution seismic tomography. However, many volcanoes do not generate sustained seismicity. In this case, the recently developed methods based on correlations of ambient seismic noise become suitable for imaging volcanic edifices.

Human-Made Seismic Sources

Human-made seismic sources include explosions, vibrators, air-gun shots, and weight (or mass) drops and are usually called **active sources**. There are some onshore experiments which use chemical explosions of sufficiently large magnitude to generate seismic signal propagating to the distances of dozens kilometers and depths of a few kilometers. The problem of such explosions is that they are fairly expensive and thus their number is strongly limited. Furthermore, chemical explosions are not ecologically friendly and they are prohibited in most volcanic areas. In most cases, the surveys based on such sources are used for 2D profiling (Fig. 1c).

An alternative method of active-source onshore generation might be using vibrators which are more ecologically safe compared to explosions. However, for the purposes of volcano tomography, the vibrators should be sufficiently powerful and heavy to enable the required propagation of seismic rays. In reality, the transportation of powerful vibrators appears to be not possible in most volcanic areas; therefore, they are not widely used in practice.

In case if a volcano is located close to the seashore, it is possible to use underwater air guns as sources of seismic signal (Fig. 2a). The advantage is that air-gun shots are rather cheap
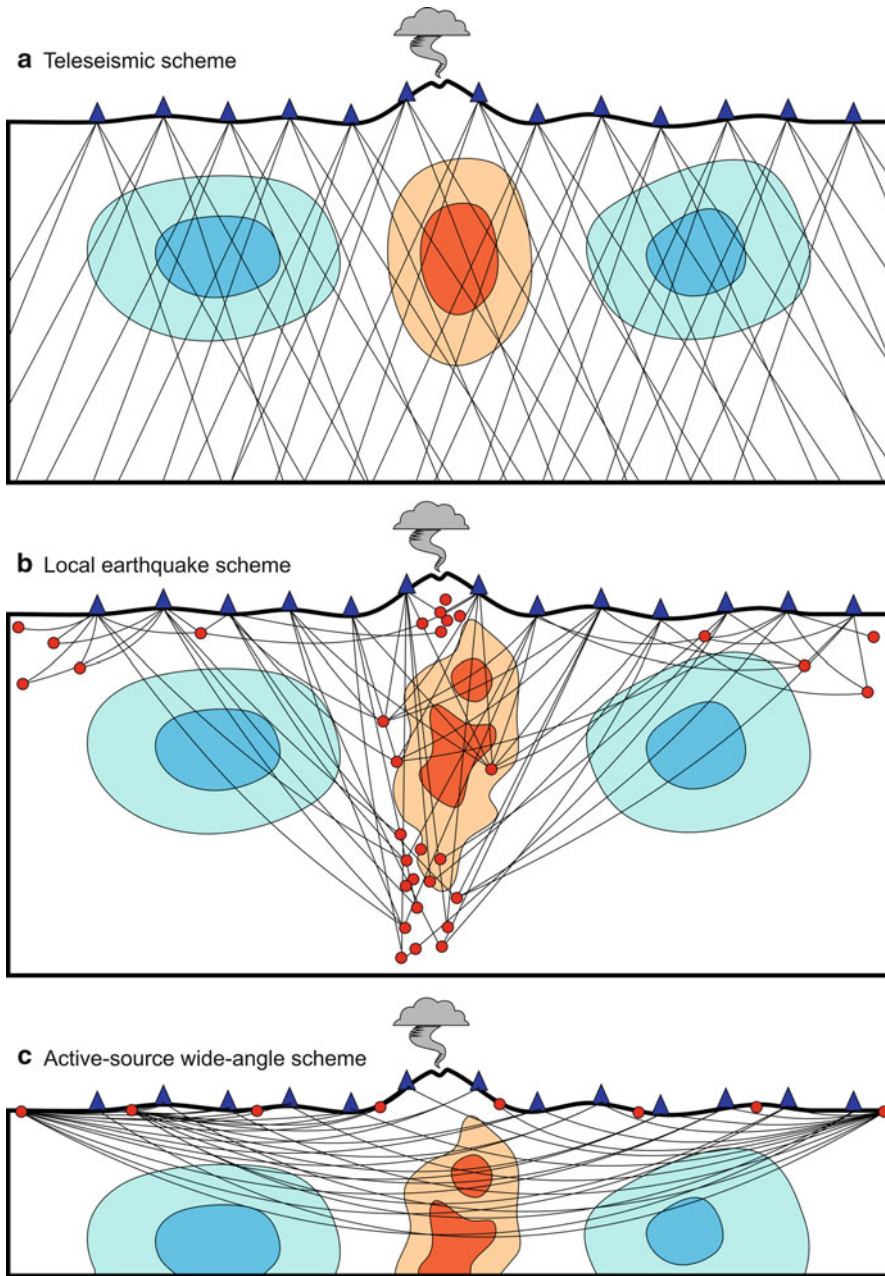
Seismic Tomography of Volcanoes, Fig. 1 Observation schemes used in volcano tomography. *Blue triangles* depict seismic stations; *red dots* are sources; *black lines* are seismic rays

and sufficiently powerful. They can be frequently generated along the ship cruise that gives a huge data base with a clear technology of processing. The problem is that these sources are limited by the shore line and not suitable for most volcanoes.

The major advantage of the active sources is that their parameters are known. The problem is that they are located close to the surface, and seismic signal propagate along the nearly horizontal ray paths. Another problem, which is
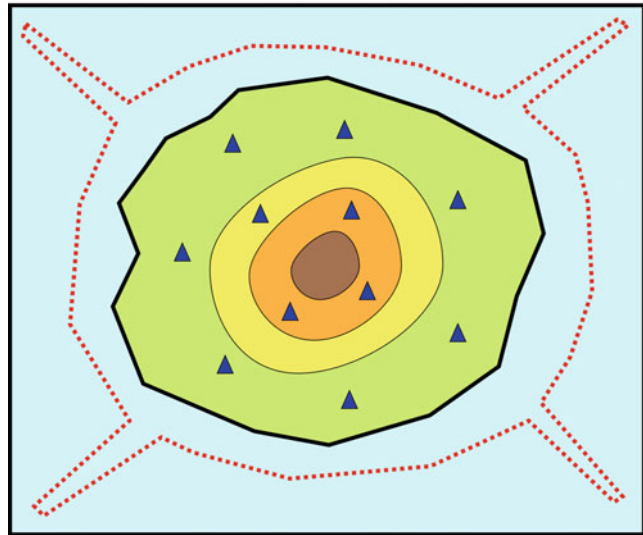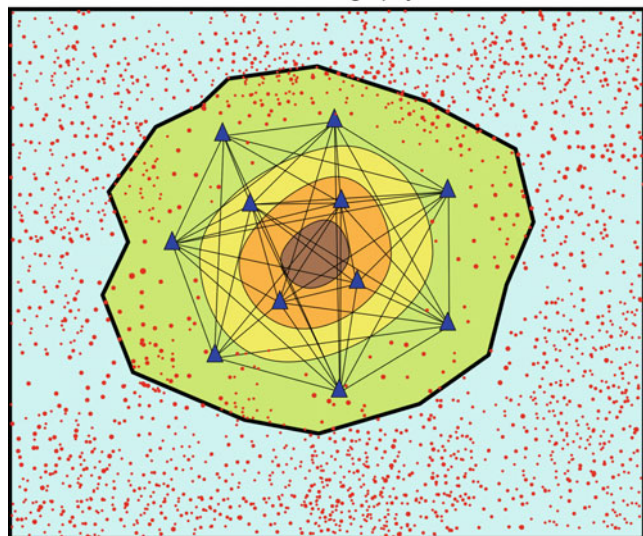
mostly actual for the air-gun shots, is that they generate very weak S waves which cannot be picked and used for tomography studies.

Earthquakes

Volcanic and tectonic earthquakes located in the volcanic area (including some events located outside the perimeter of the network) are regularly used as sources for seismic tomography of volcanoes (Fig. 1b). The major problem with earthquakes is that their location parameters (coordinates and origin times) are not exactly known and must be determined simultaneously with velocity models. This leads to a coupled problem with significant trade-off between velocity and source parameters. However, an important advantage of natural sources compared to the artificial sources is that they are located at some depths which allow illuminating the target object from below at different directions. For many cases, the amount of earthquakes is much larger than that of the artificial explosions. In addition, earthquakes in most cases produce clear P and S waves which can be used for exploring P and

S velocities. In turn, both P and S velocities give much more constraints on the petrophysical state of rocks below volcanoes than just a single P model.

Ambient Seismic Noise

Ambient seismic noise is the permanent motion of the Earth surface that is not related to earthquakes or specific controlled sources (see also "▶ Seismic Noise"). When averaged over long time series, this wavefield can be approximately considered as produced by randomly and homogeneously distributed sources (Fig. 2b). Cross-correlation between records of such random wavefields at two stations yields the impulsive response of the media (Green's function) between these two points (e.g., Gouédard et al. 2008). In the present form, every recorder can be converted into a "virtual" source recorded at all other recorders resulting in **N(N-1)/2** source-receiver pairs for a network of **N** stations, as shown in Fig. 2b. Shortcomings of noise-based approach are (1) dependence on noise sources distribution that in many cases is not perfectly homogeneous possibly preventing successful reconstruction of impulsive responses for a part of station pairs; (2) need of long continuous records (months or years); and (3) the fact that reconstructed virtual sources are located at the surface and are dominated by surface waves.

Receivers

Networks for active and passive surveys on volcanoes use generally the same instruments as for similar seismological studies in other regions. Tomography inversion needs numerous stations in a network (more than ten) which are usually deployed temporarily. Such networks consist of portable seismic stations which are composed of three major components: sensor, recorder, and power supply. In some cases, there is also a device for telemetric transition of signal.

Seismic sensor is a kind of "microphone" which converts the ground movements into electric signal. Modern sensors which are used for volcano surveys are three-component broadband instruments which allow recording oscillations of more than 30–40 s period (see also "▶ Principles of Broadband Seismometry"). It should be noted that for the body-wave tomography based on local seismicity such broad frequency capacity is not really required: the signal from local earthquakes is clearly seen at frequencies of more than 1Hz.

Seismic recorder is a tool which digitizes the analogue electrical signal coming from the sensor and records it onto a flash or another kind of memory. In each recorder, there should be also a GPS receiver which synchronizes the absolute time. In some cases, sensors and recorders are combined in one unit.

All recorders and most broadband sensors consume the electric power, and an important part of the seismic station is **power supply**. It can include chemical rechargeable or single-used batteries, solar panels, or wind generators. Using each type of power source is determined by the conditions of real experiments. For example, in Kamchatkan volcanoes, solar panels cannot be used due to large amount of snow during the most part of the year. Furthermore, there is a risk of destruction of any part located outside the ground by wild animals or vandals. In this case, an only possibility to conserve the station is to place the station deep into the ground together with batteries.

Some Hints on the Network Design

Design of seismological networks for performing tomography studies has some particular features. Here some hints which should be taken into account while deploying seismic networks are given.

- In case of using refracted rays from active (artificial) sources, when both sources and receivers are located near the surface, a typical depth of seismic ray propagation is 1/10–1/15 of the source-receiver distance (Fig. 1c). Thus, to study the depths of ∼10 km beneath the volcano, the sources and receivers should be located in the opposite sizes of the volcano at distances of 50–70 km.
- The spacing between stations normally has the same order as the minimum size of resolved anomalies in the tomographic inversion.

Further decrease or increase of the resolution depends on the distribution of events. If there are a very few events, densification of the network would hardly bring to the resolution increase. In contrary, in case of a large amount of sources evenly distributed throughout the study area, the resolved patterns can be smaller than the station spacing.

- For passive sources (local earthquakes), the depth of seismic model resolved from tomographic inversion cannot be larger than the maximum depths of sources.
- For uncontrolled sources (earthquakes), a critical parameter for tomography is a number of picks per events. If this number is equal to 4, one can theoretically locate the sources but cannot get any information about the velocity model. In practice, the data are noisy, and to get a stable location of source, more than four picks are required. The information on the velocity structure can be extracted only if the number of picks is more than 10. The higher this value is, the more velocity parameters can be retrieved and the higher resolution model will be obtained.
- Using out-of-network events (within reasonable limits) can appear to be useful for tomography as was shown by Koulakov (2009b). Although the location accuracy for such events is fairly low, they improve the ray coverage and help in increasing the spatial resolution beneath the network area.
- In all cases, prior the deployment of the network, it is useful to simulate a series of synthetic models with possible configurations of network, sources, and target objects that are expected in the real case. This can help to optimize the distribution of seismic stations in the real experiment.

Seismic Tomography Methods Based on Body Waves

Basic Principle of Body-Wave Seismic Tomography

In the ray-based approach which is used in most body-wave tomography algorithms, seismic waves are produced by point sources. Along this way, seismic rays accumulate the information on the inner structure of the Earth. The purpose of seismic tomography is to decipher this information and to reconstruct the structure inside the Earth. Most simple and most popular way for many practical applications, including volcano tomography, is the kinematic approach based on using travel times of seismic rays. Travel times of seismic rays can be represented as an integral along the ray paths

$$T = \int_{\gamma} \frac{dS}{V(x, y, z)} \qquad (1)$$

where γ is the ray path and V is seismic velocity. Here, an infinite frequency of the wave is presumed which allows presenting the ray as a curved line. Actually, this is often not true in seismology where the wavelengths are in some cases compatible with the size of the target objects. For these cases, there are methods of finite-frequency tomography which are based on "fat" rays representing a banana-shaped area having a rather complex sensitivity patterns (e.g., Husen and Kissling 2001). Although there are some successful applications of this method in global or in teleseismic tomography (Montelli et al. 2004), it was not yet widely implemented for volcano tomography. Up to today, most of the practical applications for studying the volcanoes are based on the ray-based kinematic approach.

In general, the velocity field in Eq. 1 is a function which is represented by infinite number of mathematical points. In practice, to solve the inverse problem, i.e., to derive velocity distribution based on given travel times, the velocity field should be parameterized with a finite number of parameters. There are different methods of parameterization, such as using regular or irregular cells, nodes with linear interpolation, harmonic functions with unknown coefficients, etc.

Another problem of solving the inverse problem based on the representation in Eq. 1 is that this problem is nonlinear. It means that one should find the velocity distribution based on the ray paths which are strongly affected by the

unknown velocity distribution. The direct solution of this nonlinear problem is almost never used for practical applications. Actually, in most cases, this problem is reduced to the linear representation. It is assumed that the unknown velocity distribution is close to the given reference model, and in this case, the problem in Eq. 1 is reduced to the linear integral

$$\Delta t = -\int_{\gamma_0} \frac{\Delta v\, dS}{V_0^2(x, y, z)} \qquad (2)$$

where Δv is unknown velocity anomalies, Δt is time residuals, V_0^2 is velocity in the reference model, and γ_0 is the ray path in the given reference model. Then, the continuous representation in Eq. 2 can be reduced into the discrete system of linear equations:

$$\Delta t_i = \sum_j A_{ij}\Delta v_j \qquad (3)$$

where Δv is the unknown value of jth velocity parameter and A_{ij} is the first-derivative matrix which represents an effect of a unit velocity change of jth parameter on travel time ith ray. In the case of cell parameterization, each element of this matrix is just the length of jth ray in ith cell. In the case of uncontrolled sources, this matrix also includes the elements for relocation of sources (four parameters, dx, dy, dz, and dt, for each source). In practical tomography studies, where this matrix may appear to be very large, the system in Eq. 3 is solved using the iterative LSQR algorithm (Paige and Saunders 1982) which proceeds only nonzero elements in the matrix and performs the calculations line by line without keeping the entire matrix in the memory. The direct solving of the system as it stands in Eq. 3 normally fails because of strong instability of the solution due to noise in the data. To avoid this problem, the inversion should be regularized using some additional terms. The most popular way of regularization is the amplitude damping which is realized by adding a diagonal matrix block with zero data vector. Another way is damping a gradient between all pairs of neighboring parameters (nodes or cells) that allows controlling the smoothness of the solution.

The quality of the solution and optimal values of the inversion parameters are estimated based on synthetic modeling which should simulate the conditions of a real experiment as close as possible. The synthetic travel times are computed in an artificial model and perturbed by a realistic noise. Then, the model parameters are "forgotten," and the reconstruction procedure is performed identically as in the case of observed data processing. This step is especially important for the volcano tomography where the distributions of sources and events are often far from optimal.

There are different seismic tomography schemes which are used for different purposes in geosciences. Many of them are used for studying the volcano feeding systems on scales from the entire mantle (thousands kilometers) to the uppermost layers (hundreds meters) (e.g., Koulakov 2013). Here, only some examples of tomography studies on scales from first to dozens kilometers are given. They provide the information about magma sources in the crust and uppermost mantle that are directly responsible for eruptions of active volcanoes.

Teleseismic Tomography

Seismic tomography is a relatively young method which started actively developed from the late seventies after the pioneer work by Aki et al. (1977). The method proposed in this study is called **teleseismic tomography**, and it uses the relative time delays from distant (teleseismic) earthquakes recorded by seismic stations located in the study area (Fig. 1a). There were several attempts of using this method for studying the volcanoes. For example, Ellsworth and Koyanagi (1977) have implemented the teleseismic scheme for studying the crustal structure beneath Kilauea volcano in Hawaii. Three years later, Sharp et al. (1980) have published the tomographic model for Mt. Etna. Among later studies, one can mention the work by Stauber et al. (1988) who presented tomographic models of Newberry Volcano, Oregon, based on teleseismic delays.

It should be noted that during the last decades, the teleseismic tomography is almost not used

S

anymore for studies of crustal structures beneath volcanoes because of several reasons. The first problem is that nearly vertical seismic rays in the study area cause poor vertical resolution which leads to strong vertical smearing of the retrieved anomalies. Another problem, which seems to be more serious, is related to the fact that the teleseismic waves are usually of low frequency (around 1 Hz and lower). Thus, the large wavelengths strongly limit the resolution capacity of this method. Theoretically, it cannot provide robust images of anomalies smaller than 8–15 km. Magma reservoirs are usually expected to be smaller in size, and thus, they are hardly resolvable by teleseismic tomography.

Local Earthquake Tomography

Another method, which is called **local earthquake tomography** (**LET**), uses the data of arrival times of P and S waves from earthquakes located within the study area or slightly outside (Fig. 1b). Because of using the uncontrolled sources, the solution in this case is reduced to the coupled problem of simultaneous determination of source parameters and velocity models. Active volcanoes usually produce a lot of seismicity related to processes in magmatic reservoirs and tectonic displacements in the crust. Therefore, deploying stations for the LET experiment is much cheaper than organizing active-source works with the use of explosions or air-gun shots. Furthermore, the earthquakes in passive source schemes are distributed at some depths that enable much more favorable observation system than in the case of active-source studies. In addition, seismic records from earthquakes usually provide the data on P and S waves that allows getting more information on petrophysical state of rocks in volcanoes than just a single P-velocity model which is usually resulted from the active-source studies. The problem of LET studies is an uncontrolled distribution of sources which is not always optimal to get optimal ray coverage in the target areas. An important difficulty of LET is the uncertainty of the source locations and the trade-off between source and velocity parameters. Another technical problem of LET is the necessity of fairly long

deployment of a large number of seismic stations (for more than several months) which is required to accumulate sufficient information on the seismicity in the volcano area. This needs solving some logistical issues related to providing power supply and hiding the equipment.

The first LET algorithm was developed and implemented to real seismological data by Thurber (1983). One year later, the same algorithm was used for studying the crustal structure beneath Kilauea volcano (Thurber 1984). Among the later studies based on this approach, one can single out the studies of Hengill-Grensdalur volcanic complex in Iceland (Foulger and Toomey 1989) and of Mount St. Helens (Lees 1992). There are more than ten different LET studies of Mt. Etna which appears to be one of the best studied volcanoes in the world (e.g., Villasegnor et al. 1998; Chiarabba et al. 2000; Patane et al. 2006). Note that the work by Patane et al. (2006) was a pioneer study presenting the 4D tomography model which revealed temporal variations of seismic structure beneath the volcano having possible relations to the eruption activity (see also "▶ Seismic Monitoring of Volcanoes"). During the last decades, the LET methods have become the most popular for studying crustal structure beneath volcanoes. Nowadays, there are dozens of successful studies of volcanoes in different parts of the world. In the final part of this paper, two examples of recent studies of Mt. Spurr in Alaska (Koulakov et al. 2013a) and of Klyuchevskoy volcano group in Kamchatka (Koulakov et al. 2011, 2013b) will be presented. Both of these studies were performed using the LET code LOTOS (Koulakov 2009a).

Studies with Human-Made Sources

Human-made sources or active sources are used in many seismic studies of volcanoes. The first experiments of three-dimensional active-source investigations of volcanoes were made in two relict calderas located in the Cascade Range, western USA. One of them is the potentially active Newberry shield volcano with the caldera of ∼8 km diameter. In the active seismic experiment, Achaer et al. (1988) used three chemical explosions and a dense receiver network covering the caldera. They obtained a rather clear

low-velocity anomaly at about 3 km depth which was interpreted as a magma chamber. Another study was made in the same year in the Medicine Lake Caldera by Evans and Zucca (1988). They deployed 120 seismometers and made eight chemical explosions which were recorded by all stations. The tomography inversion of travel time data revealed a low-velocity pattern beneath the caldera which was also interpreted as a magma source.

Performing 3D experiments using onshore explosions appeared too expensive and not providing sufficient spatial resolution with the available number of explosions. Therefore, many onshore experiments were mostly organized along one or several profiles (Fig. 1c). For example, Zollo et al. (1996) used four chemical explosions on a profile passing through the summit of the Vesuvius volcano (Italy). Their P-velocity model reveals a rather clear high-velocity anomaly beneath the volcano. A similar experiment, but with five explosions on two crossing profiles, was performed by Aoki et al. (2009) for the Asama volcano in Japan. Similarly, as in the previous case, they found a clear high-velocity body beneath the volcano.

It should be noted that logistically the experiments based on chemical explosions are very difficult. Because of high expenses, the number of shots is strictly limited that makes it problematic to illuminate the target objects from different directions. Furthermore, the 2D approximations are often too rough for the volcanic areas where the expected structures are strongly 3D. In addition, the explosion type sources are not ecologically safe, and in most volcanic areas, they are prohibited now.

Good solutions for areas which are close to the seashore are the underwater air-gun shots which can be used to generate high-quality and sufficiently powerful seismic signal and can surround the target area (Fig. 2a). Nowadays, this scheme is widely used in experiments in many volcanic areas. Among many different studies, three examples are selected here:

- A very interesting site for seismic investigations is the volcanic Deception Island in Antarctic which was studied by Zandomeneghi et al. (2009). This island has a circular shape and is suitable for performing air-gun shots inside and outside the circle. Using onshore and offshore seismometers made this observation system suitable for high-resolution tomography studies. The inversion of travel time data revealed the seismic structures down to ∼2 km which appeared to be linked with the main geological structures. The circular rim of the caldera is associated with high-velocity lineaments. Inside the caldera, a low-velocity anomaly is observed.

- A network with onshore and offshore stations and 4,414 air-gun shots were used by Paulato et al. (2010) to build a tomographic model beneath Montserrat Island in Lesser Antilles. They found that the location of the Island is associated with high P velocity. Especially sharp positive anomaly is located beneath the Soufriere volcano.

- The high-resolution 3D structure beneath the Tenerife Island was studied by Garcia-Yeguas et al. (2012) based on a large amount of air-gun shots and fairly dense onshore seismic network consisting of 140 stations. The main volcanic structure of the Las Cañadas-Teide-Pico Viejo Complex (CTPVC) is characterized by a high P-wave velocity body, similar to many other stratovolcanoes. Furthermore, reduced P-wave velocities are found in a small confined region in CTPVC and are more likely related to hydrothermal alteration, as indicated by the existence of fumaroles.

Imaging and Monitoring of Volcanoes Based on Ambient Seismic Noise

General Principle, Data Processing, and Inversion Approaches

New methods for probing the Earth's interior using noise records only emerged during the last decade (e.g., Campillo et al. 2011). Origin of seismic noise strongly depends on the considered spectral range (see also "▶ Seismic Noise"). At high frequencies (>1Hz), the noise is strongly dominated by local sources that are often related

to strong winds or the anthropogenic activity. Seismic noise at intermediate periods (between 1 and 20 s) that is most often used for studies of volcanoes is dominated by natural sources. In particular, the two main picks in the seismic noise spectra in this so-called microseismic band (1–20 s) are related to forcing from oceanic gravity waves. The interaction between these oceanic waves and the solid Earth is governed by a complex nonlinear mechanism (Longuet-Higgins 1950), and the noise excitation depends on many factors such as the intensity of the oceanic waves but also the intensity of their interferences as well as the seafloor topography. Overall, the generation of seismic noise is expected to be strongly modulated by strong oceanic storms and, therefore, to have a clear seasonal and nonrandom pattern. However, when averaged over long time series, distribution of sources of microseismic noise can be approximated as random and homogeneous. In this case, cross-correlation of noise records between two stations yields the impulsive response (Green's function) between these two points (e.g., Gouédard et al. 2008). In the case of a uniform spatial distribution of noise sources, the cross-correlation of noise records converges to the complete Green's function of the medium, including all reflection, scattering, and propagation modes. However, in the case of the Earth, most of ambient seismic noise is generated by atmospheric and oceanic forcing at the surface. Therefore, the surface-wave part of the Green's function is most easily extracted from the noise cross-correlations.

Seismic surface waves, in contrast with body waves, are waves that propagate in a waveguide near the Earth's surface. There are two types of **surface waves**: **Rayleigh** (vertically polarized waves) and **Love** (horizontally polarized waves transverse to the direction of motion). Their depth sensitivity depends on frequency, with a fair approximation being about a third of a wavelength. As a result, the surface waves are strongly dispersive with longer periods (wavelengths) propagating faster than shorter periods because of the general increase of seismic speed with depth in the Earth.

Using surface waves extracted from correlations of seismic noise for imaging of the Earth's interior is called **ambient noise surface-wave tomography**. This approach generally consists of three steps (e.g., Ritzwoller et al. 2011). First, noise cross-correlations between every pair of stations in the array are computed, and interstation phase or group travel times as a function of period are measured from the resulting waveforms. Second, two-dimensional maps of the speed of Rayleigh or Love waves as a function of frequency are produced. At the last step, these maps are inverted for a 3D distribution of shear speeds.

Applications of Ambient Noise Seismic Tomography (ANST) to Studies of Volcanoes

After recent emergence of the ambient noise surface-wave tomography (e.g., Shapiro et al. 2005), it became clear that this method might be very useful for studying structures within volcanoes because of its relatively low coast (especially comparing with active-source studies) and because of good coverage of the shallow part of edifices compared with earthquake-based body-wave imaging. First "volcanic" application of ANST was performed at Piton de la Fournaise located at La Réunion Island (Brenguier et al. 2007). This hot-spot volcano is very active with on average more than one eruption per year during the past two centuries. The noise-based tomography was based on correlations of 18 months of records by 21 short-period vertical-component stations of the permanent volcano monitoring network. Rayleigh wave group velocities could be measured between 2 and 5 s of period. Inversion of these dispersion curves resulted in a shear velocity model extending between the surface and the depth of 1 km below sea level (while average elevation of the volcano is ∼2 km). The main result of this study was imaging of an intrusive high-velocity body below the main crater moving westward from the surface to 1 km below sea level. This body is located above the shallow magma reservoir and coincides with the region through which the magma ascends during the summit eruptions.

Following the pioneering study of Brenguier et al. (2007), the ANST was applied to image

subduction zone volcanoes: Okmok in Aleutian Islands (Masterlark et al. 2010), Lake Toba in Sumatra (Stankiewicz et al. 2010), and Asama in Japan (Nagaoka et al. 2012). The first study used 12 stations from a permanent monitoring network, and the second study used 40 stations from a temporary deployment, while the third study used 81 sensors combined from permanent and temporary networks. In all these studies, the authors reported very similar results: imaging of low-velocity anomalies corresponding to magma reservoirs at depths below 4 km.

Noise-Based Volcano Monitoring

One of the advantages of using continuous seismic noise records to characterize the earth materials is that a measurement can easily be repeated (see also "▶ Seismic Monitoring of Volcanoes"). First application of this idea to volcano monitoring was done by Sens-Schönfelder and Wegler (2006) who used the repetitive waveforms of seismic noise cross-correlations to track changes in the subsurface material properties on mount Merapi in Indonesia during 2 years. This study was based on high-frequency (>1 Hz) waves and mainly revealed seasonal variations caused by changes of hydrological conditions in a very shallow layer. Brenguier et al. (2008) used noise cross-correlations at intermediate periods (between 1 and 10 s) to detect decreases in seismic velocity a few weeks before eruptions that suggests preeruptive inflation of the volcanic edifice of the Piton de la Fournaise. A comprehensive review of recent advancements in the noise-based volcano monitoring is provided by Brenguier et al. (2011) (see also "▶ Tracking Changes in Volcanic Systems with Seismic Interferometry").

Current Limitations and Prospective

Application of seismic noise cross-correlation methods for studies of volcanoes is a rapidly growing field. However, these methods have some important limitations that are mainly related to the properties of the seismic noise recorded at the Earth surface. So far, the basic theoretical foundation of the cross-correlation method assumes a random and homogeneous distribution of the wavefield sources, while properties of a real seismic noise may deviate from this simplified model. This implies that applications of the noise-based imaging or monitoring should be accompanied by studies of properties of recorded seismic noise to characterize the degree of its spatial and temporal inhomogeneity and to determine the minimal time required for good convergence of noise cross-correlations. Development of advanced methods for the preprocessing of the noise time series in order to improve the quality of the extracted signals and to decrease the convergence time is a very active research area.

Up to date, the noise-based seismic tomography was based on surface waves that are most easily extracted from noise cross-correlations but have a limited resolution in deep layers. Most recent results demonstrate that, when working with large and dense arrays of seismometers, it is possible to indentify body waves in correlation waveforms and to measure their travel times (e.g., Boué et al. 2013; Lin et al. 2013). This gives a hope that noise-based body-wave tomography of volcanoes could be developed in the future.

Examples of Tomography Studies of Volcanoes

Tomography of Mt. Spurr

Mount Spurr is an active predominantly andesitic volcano in Alaska located at the northeastern end of the Aleutian arc. Before a single eruption in 1953, it was considered as a dormant volcano. In 1992, a series of three explosive eruptive events occurred on June 27, August 18, and September 16–17 in a small composite cone forming the Crater Peak on the south flank of Mount Spurr. The structure beneath Mt. Spurr in a period of time nearly corresponding to this eruption was studied by Power et al. (1998) using the data collected from 1991 to 1993 by 7–11 seismic stations of the Alaska Volcano Observatory (AVO).

The most recent episode of Mt. Spurr unrest occurred in 2004–2006 which accompanied with

strong seismicity and fumarolic activity. However, this episode did not result at any magma eruption, and thus, it was called as "lost eruption." For 1 month in June 2005 corresponding to the middle stage of this unrest, the AVO deployed a fairly dense seismic network consisting of 26 broadband three-component stations which operated in addition to 11 short-period stations of the permanent network. During this period, 512 crustal earthquakes were located in the area of Mt. Spurr using P and S picks recorded by portable and permanent stations. The final catalog contained 5,960 P- and 4,973 S-wave arrival times with an average of 20 picks per event. These data were used to perform the tomographic inversion presented in Koulakov et al. (2013a).

The result of tomographic inversion including the 3D distributions of P and S velocities, Vp/Vs ratio, and accurate source locations is presented in Fig. 3. The horizontal section at 5 km depth demonstrates a prominent anomaly of high Vp/Vs ratio which reaches the value of 1.9 located beneath the northern rim of the volcano. This pattern coincides with the distribution of shallow seismicity. In the deeper section at 22 km depth, this anomaly is located at the same place, but it appears to be less intensive than another zone of high Vp/Vs ratio which is collocated with a cluster of deep crustal earthquakes. In vertical section, these two anomalies look like two fingers of high Vp/Vs ratio separated by 8 km of lateral distance. One of the two anomalies nearly reaches sea level beneath the northern border of the caldera, but is separated from the surface with a zone of low Vp/Vs ratio. The top of the second anomaly appears to be much deeper and may reach a depth of 20 km. The earthquakes that occurred during the 3-month span show distinct clusters near the tops of each of these fingers which support the active behavior of these structures. It is important that for both these finger-shaped anomalies, the high Vp/Vs correspond to higher P and lower S velocities.

The observed finger-like features are suggestive of pathways through the crust and possibly correspond to complexes of conduits that transport liquid volatiles and/or melts from the mantle.

The combination of high Vp and low Vs suggests the presence of liquid fluids and/or melts and the rock composition corresponding to the low crustal or even mantle levels. The existence of very low values of Vp/Vs above the NW anomaly might be an indicative for aggregate transition occurred between the top of the plume and the surface. It is important that the NW finger location generally coincides with the presence of fumaroles and active volatile flux during the most recent eruption in 2004–2005. It can be proposed that this liquid fluids brought by the NW conduit are transformed to gases at shallow depths which can escape to the surface without producing any stresses. This explains lack of any explosive activity in this location during the 2004–2005 unrest of Mt. Spurr (Fig. 3).

Another conduit located below the Crater Peak seems to be blocked at ∼20 km depth. This fact is supported by the occurrence of deep seismicity just above this anomaly. In this blocked channel, the accumulation of high pressure is more probable which makes it more dangerous for a potential explosive eruption. Note that the explosive eruption of Mt. Spurr in the SE rim in 1992 occurred just above this anomaly.

Tomography of the Klyuchevskoy Volcano Group

The Klyuchevskoy volcano group is located in Kamchatka, in the far east of Russia. It includes 13 active and dormant volcanoes of various types and compositions located in a relatively small area of approximately 100×60 km. The main volcano of the group, Klyuchevskoy Sopka, has the elevation of ∼4,800 m, and it is the highest volcano in continental Eurasia. Another volcano of the group, Bezymianny, located at approximately 10 km distance from the Klyuchevskoy volcano is an explosive-type andesitic volcano. One of the world's largest explosive eruptions in the twentieth century occurred here in 1957. The Tolbachik complex is situated in the southwestern part of the Klyuchevskoy group, and it consists of several shield and stratovolcanoes of predominantly basaltic compositions. Plosky Tolbachik, which is currently active, is a typical Hawaiian-type volcano with fissure eruptions and
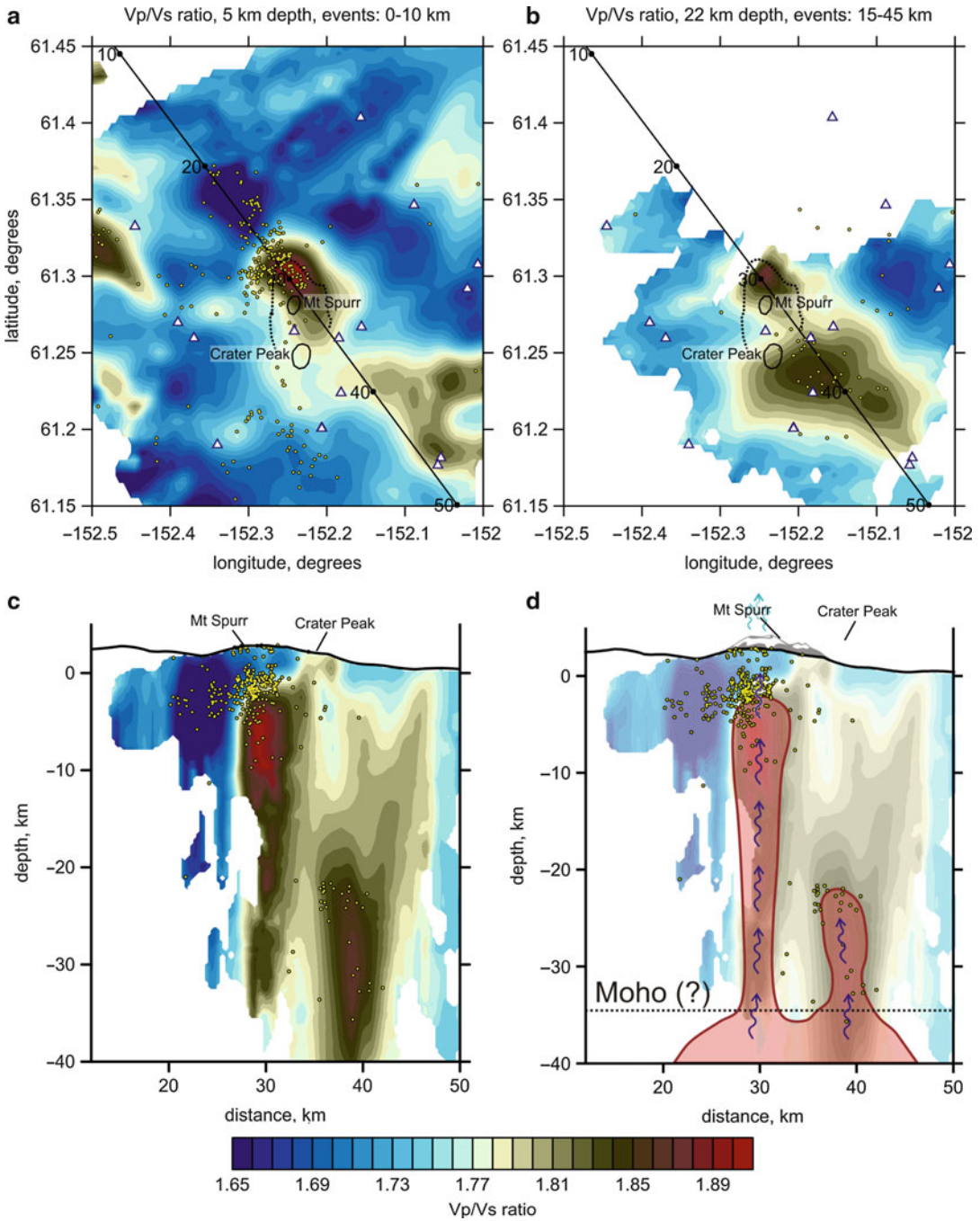
Seismic Tomography of Volcanoes, Fig. 3 Distribution of the Vp/Vs ratio in two horizontal and one vertical section. *Dotted line* depicts the limits of the caldera rim. *Triangles* indicate the locations of stations. *Yellow dots* mark the source locations. *Bottom-right plot* represents the interpretation of the results. *Blue waved arrows* indicate hypothetical migration of fluids and their escape to the atmosphere in 2004–2006 unrest (Reproduced from Koulakov et al. 2013a, JRL)

calderas. The existence of so many different volcanoes located close to each other can most likely be explained by complex processes of fractioning, mixing, and melting in magmatic reservoirs located under the volcanic group at different depth levels.

During the last several decades, the volcano-related seismicity in the Klyuchevskoy group has been monitored by the Kamchatkan Branch of Geophysical Survey. After 1999, a network consisting of 17 digital telemetric stations was installed in the region of the Klyuchevskoy group. From 1999 to 2009, it provided the information on more than 80 thousand events and the corresponding half million of P and S arrival times. The tomographic inversions were performed based on selected yearly subsets. The temporal variations of seismic structure are especially interesting in the context of the volcanic activity during the considered time. In particular, in 2005, there were simultaneous eruptions of the Klyuchevskoy and Bezymianny volcanoes that give a possibility to explore the changes of the crust related to the volcanic activity at stages of preparation, activation, and relaxation (see also "▶ Seismic Monitoring of Volcanoes").

The results of tomographic inversion corresponding to the year 2004 just before the activation phase were presented in a separate study by Koulakov et al. (2011). Similar structures are observed in the precedent 2003 year (Koulakov 2013). The key feature of these models is a strong anomaly of a high Vp/Vs ratio, reaching values of 2.2, located below 25 km depth. Within this anomaly, the P velocity appears to be higher than average, whereas the S velocity is strongly low. This combination of seismic anomalies is a possible indicator for a mantle conduit beneath the Klyuchevskoy group with rocks brought from deep levels and high content of fluids and melts. It was proposed that this conduit is the original magma source that feeds all of the volcanoes in the Klyuchevskoy group.

The tomography result corresponding to 2004 reveals complex contrasted structures in the crust which can be associated with magma sources. The distributions of high values of Vp/Vs ratio

reveal at least two levels of magma sources in the crust: one is located at a depth of 10–12 km and another is just below the Klyuchevskoy volcano close to the surface. This may explain the diversity of the volcanic styles in the group. The distribution of earthquakes appears to link different levels of magma storage that may trace the paths of magma migration.

Time variations of seismic structure beneath the Klyuchevskoy group in a time period from 1999 to 2009 were studied by Koulakov et al. (2013b) and presented in Fig. 4. A common feature which is seen in all yearly windows is the anomaly with an extremely high Vp/Vs ratio located below 25 km depth, which was previously mentioned for the year 2004 and interpreted as a mantle conduit feeding all volcanoes of the group. For the crust, this study demonstrates considerable temporal changes of seismic structures. In particular, prior the 2005 eruptions of the Klyuchevskoy and Bezymianny volcanoes, the seismic structures in the crust look generally similar and contain two anomalies of high Vp/Vs ratio beneath the Klyuchevskoy volcano at a depth of 10–12 km and close to the surface. As mentioned above, these anomalies are interpreted as two levels of magma sources. In 2005, the average value of the Vp/Vs ratio increases, and it may be explained by a considerable increase of the fluid and melt content in the crust. In this year, a strong anomaly of high Vp/Vs ratio is observed beneath the Bezymianny volcano, and it appears to link the volcano with mantle sources. In other years, this anomaly was not observed. The shallow anomaly of high Vp/Vs anomaly, which was previously observed beneath the Klyuchevskoy volcano, disappears in 2005. The deeper anomaly at 10–12 km depth is still presented. In the years after the eruption, 2006 and 2007, the average Vp/Vs ratio lowers, and all the patterns at 10–12 km depth that were interpreted as intermediate magma sources in the crust are not observed anymore. Three years later, in 2008, a similar anomaly at the same location starts reappearing again.

The observed variations of the seismic properties are probably related to changes in the stress and deformation fields, migration of fluids, and
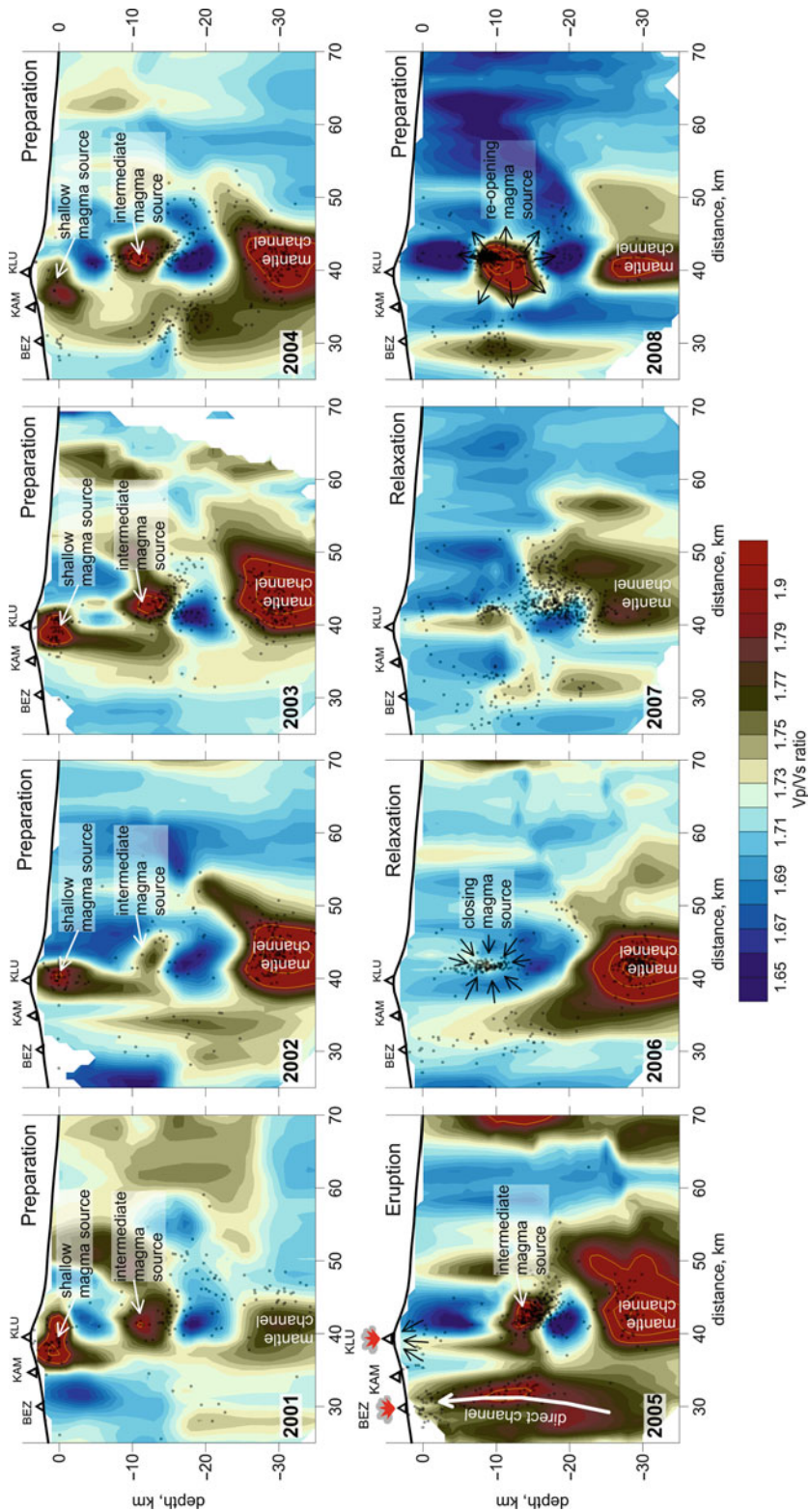
Seismic Tomography of Volcanoes, Fig. 4 Cross sections of Vp/Vs from tomographic inversions for each year from 2001 to 2008 for a profile passing through Klyuchevskoy (KLU) and Bezymianny (BEZ) volcanoes. In areas of high Vp/Vs ratios, yellow lines show contours of Vp/Vs = 2.0, 2.1, and 2.2. *Dots* show earthquakes less than 0.5 km from the profile, and triangles mark the locations of the volcanoes near the profile. The eruptions of Klyuchevskoy and Bezymianny volcanoes in 2005 are highlighted. The main phases and structural elements, which are discussed in the text, are indicated (Reproduced from Koulakov 2013, JVGR)

aggregate transformations in the crust. The mantle conduit, which is visible as an anomaly of high V_p/V_s ratio below 25 km depth at all time periods, should have strong mechanical, thermal, and/or chemical influence upon the bottom of the crust and cause strong seismicity. These stresses lead to the origin of fractures and upward penetration of the fluids to the crust. Reaching intermediate levels, these fluids lead to melting of the overheated rocks. Due to decompression, some of the fluids are transformed to gases that increase the pressure and cause additional fractures. These avalanche-type processes finally lead to the eruptions. During the eruption, most of the fluids escape to the atmosphere, and it causes a decrease in the melt content in the intermediate magma sources in the post-eruption period. Several years later, a new portion of fluids comes from the mantle conduit, and melting of rocks in intermediate reservoirs starts again. An important conclusion that can be drawn from the 4D tomography study is that the magmatic sources appear to be dynamic systems with rapidly variable properties. According to this concept, the magma sources are composed of sponge-type materials with the capacity to change the proportion of the molten partition depending on the presence of fluids. Propagation of cracks and fluids may occur rather rapidly, and this may abruptly change the properties of the "sponge" material over a period of years or even months or weeks.

Summary

This paper presents an overview with the description of several seismic tomography schemes which are used for studying volcanoes. It describes several examples of practical works including two featured case studies of Mt. Spurr and Klyuchevskoy volcano group. It was demonstrated that seismic tomography gives rich material which allows better understanding of the behavior of magmatic systems. A variety of the tomographic solutions show that there is no uniform structure: each volcano has some particular features and appears to be unique.

During the last years, there appear more and more dense networks on different volcanoes of the world designed for passive and active experiments. There are also new algorithms actively developed during the last years, such as scattering tomography or full waveform inversion, which might appear very effective for studying volcanoes. This gives a hope that increasing amount of data and implementing new algorithms will lead in the nearest time to a real breakthrough in understanding the causes of volcanic activity.

References

Achauer U, Evans JR, Stauber DA (1988) High-resolution seismic tomography of compressional wave velocity structure at Newberry Volcano, Oregon Cascade Range. J Geophys Res Solid Earth (1978–2012) 93(B9):10135–10147

Aki K, Christoffersson A, Husebye ES (1977) Determination of the three-dimensional seismic structure of the lithosphere. J Geophys Res 82(2):277–296

Aoki Y, Takeo M, Aoyama H, Fujimatsu J, Matsumoto S, Miyamachi H, Yamawaki T (2009) P-wave velocity structure beneath Asama Volcano, Japan, inferred from active source seismic experiment. J Volcanol Geotherm Res 187(3):272–277

Boué P, Poli P, Campillo M, Pedersen H, Briand X, Roux P (2013) Teleseismic correlations of ambient seismic noise for deep global imaging of the. Earth Geophys J Int 194(2):844–848. doi:10.1093/gji/ggt160

Brenguier F, Shapiro NM, Campillo M, Nercessian A, Ferrazzini V (2007) 3-D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlations. Geophys Res Lett 34:L02305. doi:10.1029/2006GL028586

Brenguier F, Shapiro N, Campillo M, Ferrazzini V, Duputel Z, Coutant O, Nercessian A (2008) Toward forecasting volcanic eruptions using seismic noise. Nat Geosci 1(2):126–130

Brenguier F, Clarke D, Aoki Y, Shapiro NM, Campillo M, Ferrazzini V (2011) Monitoring volcanoes using seismic noise correlations. C R Geosci 343:633–638. doi:10.1016/j.crte.2010.12.010

Campillo M, Roux P, Shapiro NM (2011) Correlations of seismic ambient noise to image and to monitor the Solid Earth. In: Gupta HK (ed) Encyclopedia of solid earth geophysics. Springer, Dordrecht, pp 1230–1235

Chiarabba C, Amato A, Boschi E, Barberi F (2000) Recent seismicity and tomographic modeling of the Mount Etna plumbing system. J Geophys Res Solid Earth (1978–2012) 105(B5):10923–10938

Ellsworth WL, Koyanagi RY (1977) Three-dimensional crust and mantle structure of Kilauea Volcano, Hawaii. J Geophys Res 82(33):5379–5394

Evans JR, Zucca JJ (1988) Active high−resolution seismic tomography of compressional wave velocity and attenuation structure at Medicine Lake Volcano, Northern California Cascade Range. J Geophys Res Solid Earth (1978–2012) 93(B12):15016–15036

Foulger GR, Toomey DR (1989) Structure and evolution of the Hengill-Grensdalur Volcanic Complex, Iceland: geology, geophysics, and seismic tomography. J Geophys Res Solid Earth (1978–2012) 94(B12): 17511–17522

García-Yeguas A, Koulakov I, Ibáñez JM, Rietbrock A (2012) High resolution 3D P wave velocity structure beneath Tenerife Island (Canary Islands, Spain) based on tomographic inversion of active-source data. J Geophys Res 117, B09309, doi:10.1029/2011 JB008970

Gouédard P, Stehly L, Brenguier F, Campillo M, Colin de Verdière Y, Larose E, Margerin L, Roux P, Sanchez-Sesma FJ, Shapiro NM, Weaver RL (2008) Cross-correlation of random fields: mathematical approach and applications. Geophys Prospect 56:375–393

Husen S, Kissling E (2001) Local earthquake tomography between rays and waves: fat ray tomography. Phys Earth Planet In 123(2):127–147

Koulakov I (2009a) LOTOS code for local earthquake tomographic inversion. Benchmarks for testing tomographic algorithms. Bull Seismol Soc Am 99(1):194–214. doi:10.1785/0120080013

Koulakov I (2009b) Out-of-network events can be of great importance for improving results of local earthquake tomography. Bull Seismol Soc Am 99(4):2556–2563. doi:10.1785/0120080365

Koulakov I (2013) Studying deep sources of volcanism using multiscale seismic tomography. J Volcanol Geotherm Res 257:205–226. doi:10.1016/j. jvolgeores.2013.03.012

Koulakov I, Gordeev EI, Dobretsov NL, Vernikovsky VA, Senyukov S, Jakovlev A (2011) Feeding volcanoes of the Klyuchevskoy group from the results of local earthquake tomography. Geophys Res Lett 38: L09305. doi:10.1029/2011GL046957

Koulakov I, West M, Izbekov P (2013a) Fluid ascent during the 2004–2005 unrest at Mt. Spurr inferred from seismic tomography. Geophys Res Lett. doi:10.1002/grl.50674.

Koulakov I, Gordeev EI, Dobretsov NL, Vernikovsky VA, Senyukov S, Jakovlev A, Jaxybulatov K (2013b) Rapid changes in magma storage beneath the Klyuchevskoy group of volcanoes inferred from time-dependent seismic tomography. J Volcanol Geothermal Res (available online). doi:10.1016/j. jvolgeores.2012.10.014

Lees JM (1992) The magma system of Mount St. Helens: non-linear high-resolution P-wave tomography. J Volcanol Geotherm Res 53(1):103–116

Lin FC, Li D, Clayton RW, Hollis D (2013) High-resolution 3D shallow crustal structure in Long Beach, California: application of ambient noise tomography on a dense seismic array. Geophysics 78(4):Q45–Q56. doi:10.1190/geo2012-0453.1

Longuet-Higgins MS (1950) A theory of the origin of microseisms. R Soc Lond Philos Trans Ser A 243:1–35

Masterlark T, Haney M, Dickinson H, Fournier T, Searcy C (2010) Rheologic and structural controls on the deformation of Okmok volcano, Alaska: FEMs, InSAR, and ambient noise tomography. J Geophys Res 115:B02409. doi:10.1029/2009JB006324

Montelli R, Nolet G, Dahlen FA, Masters G, Engdahl ER, Hung SH (2004) Finite-frequency tomography reveals a variety of plumes in the mantle. Science 303(5656): 338–343

Nagaoka Y, Nishida K, Aoki Y, Takeo M, Ohminato T (2012) Seismic imaging of magma chamber beneath an active volcano. Earth Planet Sci Lett 333–334:1–8. doi:10.1016/j.epsl.2012.03.034

Paige CC, Saunders MA (1982) LSQR: an algorithm for sparse linear equations and sparse least squares. ACM Trans Math Soft 8:43–71

Patanè D, Barberi G, Cocina O, De Gori P, Chiarabba C (2006) Time-resolved seismic tomography detects magma intrusions at Mount Etna. Science 313(5788): 821–823

Paulatto M, Minshull TA, Baptie B, Dean S, Hammond JOS, Henstock T, Voight B (2010) Upper crustal structure of an active volcano from refraction/reflection tomography, Montserrat, Lesser Antilles. Geophys J Int 180(2):685–696

Power JA, Villasenor A, Benz HM (1998) Seismic image of the Mount Spurr magmatic system. Bull Volcanol 60(1):27–37

Ritzwoller MH, Lin FC, Shen W (2011) Ambient noise tomography with a large seismic array. Compte Rendus Geosci. doi:10.1016/j.crte.2011.03.007

Sens-Schönfelder C, Wegler U (2006) Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia. Geophys Res Lett 33:L21302

Shapiro NM, Campillo M, Stehly L, Ritzwoller M (2005) High resolution surface wave tomography from ambient seismic noise. Science 307:1615–1618

Sharp ADL, Davis PM, Gray F (1980) A low velocity zone beneath Mount Etna and magma storage. Nature 287:587–591

Stankiewicz J, Ryberg T, Haberland C, Fauzi, Natawidjaja D (2010) Lake Toba volcano magma chamber imaged by ambient seismic noise tomography. Geophys Res Lett 37:L17306. doi:10.1029/2010GL044211

Stauber DA, Green SM, Iyer HM (1988) Three-dimensional P velocity structure of the crust below Newberry Volcano, Oregon. J Geophy Res Solid Earth (1978–2012) 93(B9):10095–10107

S

Thurber CH (1983) Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California. J Geophys Res Solid Earth (1978–2012) 88(B10):8226–8236

Thurber CH (1984) Seismic detection of the summit magma complex of Kilauea volcano, Hawaii. Science 223(4632):165–167

Villasenor A, Benz HM, Filippi L, De Luca G, Scarpa R, Patanè G, Vinciguerra S (1998) Three-dimensional P-wave velocity structure of Mt. Etna, Italy. Geophys Res Lett 25(11):1975–1978

Zandomeneghi D, Barclay AH, Almendros J, Ibáñez JM, Wilcock WSD, Ben-Zvi T (2009) Crustal structure of deception island volcano from p-wave seismic tomography: tectonic and volcanic implications. J Geophys Res 114:B06310

Zollo AEA, Gasparini P, Virieux J, Le Meur H, De Natale G, Biella G, Vilardo G (1996) Seismic evidence for a low-velocity zone in the upper crust beneath Mount Vesuvius. Science 274(5287):592–594

Seismic Vulnerability Assessment: Lifelines

Kyriazis Pitilakis and Sotiris Argyroudis
Department of Civil Engineering, Aristotle University, Thessaloniki, Greece

Synonyms

Fragility curves; Infrastructures; Seismic risk assessment; Transportation networks (railway, roadway, harbor, airports); Utilities (gas, oil, water, wastewater, energy supply)

Introduction

Main Features of Lifelines

Lifelines refer to the complex system and network assets of connected components, usually interacting with other components and systems, which are performing vital functions that are essential to sustain the life and the growth of a community, such as producing, transporting, and distributing goods or services. Their global value for the society and economy is permanently increased in our modern, technologically advanced, highly demanding, and fragile world.

They constitute by themselves a set of critical facilities for the proper and safe functioning of the societies. In case of a strong earthquake motion, their physical damages and the consequent disruption of the services they provide may contribute seriously to the global economic loss. At the same time, their reparation cost may be very high, reaching in several cases 10 % or 15 % of the initial construction cost of the whole system to which they belong.

The term *lifeline* is ambiguous. Lifelines may be distinguished in two major categories: (i) Utility systems including potable water, natural gas, oil, electric power distribution, wastewater, or communication systems and (ii) transportation systems comprising roadways, railways, airport, and port facilities. Sometimes the terms infrastructures or critical facilities are used instead, at least for some of them.

Compared to buildings, lifeline systems present three distinctive features: (a) spatial variability and topology and exposure to different geological and geotechnical hazards, (b) wide variety of component typologies and material used, and (c) specific functionality requirements, which make them highly hierarchical networks. For practical reasons they are generally categorized in links (e.g., pipelines, roads) and nodes (e.g., tanks, power substations). Moreover, for their vulnerability assessment, an important parameter is the presence of synergies between components within the same system (intra-dependencies) or between different systems (interdependencies). Due to these synergies, the physical damage of a component of a system interacting with another one may affect seriously the second system's performance and functionality. Therefore, it is essential to define the taxonomy of each interacting lifeline system that describes the individual components of each system and their role in the network as well as the way other systems affect its performance.

Due to their spatial extent, lifeline systems are exposed to variable seismic ground motions, often presenting important incoherency (shaking effects), and to geotechnical hazards expressed in terms of permanent ground deformations, resulted from fault crossing, landslides, and

liquefaction (e.g., lateral spreading, settlements, buoyancy effects). Various parameters may be used to describe and characterize the strong ground motion and the induced phenomena. The selection of appropriate parameter (i.e., intensity measure), which is efficiently correlated with the response and the expected damages of each exposed elements, depends on the technical and response characteristics of each element.

As an example, a motorway network presents the following features: It includes multiple elements at risk like bridges, overpasses, interchanges, tunnels, culverts, embankments, cuts, slopes, retaining walls, signage and markings, electrical systems, rest areas, and buildings (e.g., tolls, maintenance, traffic management). These structures are usually built with different design methods, construction techniques, and materials. The seismic design requirements, and therefore their vulnerability, depend on the type and the location of the structure, the seismicity, and the tectonic features of the area. In particular, roadways are crossing different geological and topographical conditions such as valleys, rivers, canals, gulfs, or mountainous regions. Obviously, the network components can be exposed to various seismic ground motions in terms of amplitude, frequency, or duration, as well as to different geotechnical hazards (subsidence, liquefaction, landslides, faults, etc.). In this context, the design of a long bridge crossing an active seismic fault is much more demanding than a single span bridge in the same area. On the other hand, nowadays, the design level (and subsequently the resistance to earthquakes) is higher than of 30 years ago, making old structures more vulnerable, introducing a new factor, the aging or time effects. Furthermore, the hierarchy (importance) of roads depends on their traffic capacity, i.e., highways and major arterials and secondary and local urban roads. The importance of the road segments, serving different part of the urban fabric, depends also on the land use and the density of the population. The functional role of roads is related to the specific served areas such as industrial or residential areas, hospital facilities, and transport hubs (e.g., airports and major harbors). Therefore, the importance of each

individual infrastructure and part of the network is variable and at the same time an essential factor in the seismic risk management. Finally, the synergies that are presented in a road network and can affect its overall functionality in case of earthquake include the supply of electric power for the signaling and lighting and the proximity to buildings and utility networks (e.g., water or gas pipes) that may induce the interruption of road traffic.

Experience from Past Earthquakes

Several devastating earthquakes occurred during the last century, causing extensive losses in terms of fatalities, injuries, damages, and disruptions. Knowledge, experience, and lessons learned from past earthquakes have significantly contributed to disaster risk reduction efforts across the globe.

The 1971 San Fernando earthquake in USA ($M_L 6.6$), an event that caused catastrophic damage to almost all types of lifelines, motivated the evolution of the lifeline earthquake engineering as well as many new efforts to better understand the causes of these failures and identify ways to mitigate future earthquake damage and disruption (e.g., the establishment of the Technical Council of Lifeline Earthquake Engineering/ TCLEE, in ASCE). Other strong and damaging earthquakes that contributed to the progress of lifeline earthquake engineering are the 1978 Miyagiken-oki (Japan, $M_s 7.7$), the 1985 Michoacan (Mexico, $M_w 8.3$), the 1989 Loma Prieta (USA, $M_w 6.9$), the 1990 Luzon (the Philippines, $M_s 7.8$), the 1993 Kushiro-Oki (Japan, $M_w 7.6$), the 1993 Hokkaido–Nansei-Oki (Japan, $M_w 7.7$), the 1994 Northridge (USA, $M_w 6.7$), the 1995 Hyogoken–Nanbu (Kobe) (Japan, $M_w 6.8$), the 1999 Chi-Chi (Taiwan, $M_w 7.6$), the 1999 Kocaeli (Turkey, $M_w 7.4$), and the 2001 El Salvador ($M_w 7.6$) events. More recent earthquakes which caused serious damages and losses to utility systems and transportation infrastructures, even in developed countries, include the 2004 Niigata ken Chuetsu ($M_w 6.8$) and 2007 Niigata-Chuetsu Oki ($M_w 6.6$) earthquakes in Japan, the 2008 Wenchuan (China, $M_w 7.9$), the 2009 L'Aquila (Italy, $M_w 6.3$), the 2010 Maule

S

(Chile, M_w8.8), and recently in 2011 the Christchurch (New Zealand, M_w6.3) and the mega 2011 Tohoku (Japan, M_w9.0) earthquake and the tsunami associated with this earthquake.

The experience from major earthquakes proved that most lifeline elements are vulnerable, and the mid- and long-term effects, mainly in terms of financial losses, may be very important. A typical example of long-term effects is the port of Kobe in Japan, one of the largest container cargo ports in the world, which suffered major damages, mostly due to liquefaction-induced phenomena, as a result of the 1995 M_w6.8 Hyogoken–Nanbu earthquake. A year after almost all port infrastructures have been reconstructed, but it is remarkable that in 1998, 3 years after the disaster, cargo traffic remained at roughly half of pre-disaster levels. For the same earthquake, the direct repair cost for the Hanshin Expressway was $4.6 billion (NCEER 1995), while the total cost was actually much higher if the $3.4 million of daily income from tolls and the losses from business interruption and traffic delays was taken into account. In the 1994 M_w6.7 Northridge earthquake, the repair cost for damaged bridges was evaluated at $190 million, almost the 10 % of the total cost of the whole transportation system ($1.8 billion, Basöz and Kiremidjian 1998).

Interactions between systems may have a significant impact exceeding direct consequences. Loss of power or communication can severely impact the emergency response, especially when the disruptions concern critical facilities, such as hospitals, command centers, etc. Loss of power may have severe indirect effects due to the synergies between the lifelines and the dependence of all networks on the power supply system. In 1989 M_w6.9 Loma Prieta earthquake, the damages that occurred in the water network greatly affected the fire-fighting capacity in the area of San Francisco, Oakland, and Berkeley. Representative cases of interactions have been also reported after the 1994 M_w6.7 Northridge and 1995 M_w6.8 Hyogoken–Nanbu earthquakes, when the reparation of the water and gas networks have been delayed due to traffic congestion, street blocking, damaged buildings, and

water that flowed into gas pipelines. A large area was also burned down in Kobe due to the disruption and the inefficiency of the fire-fighting network. Considering that modern societies and built environment are relying on the good performance of their interconnected infrastructures, it is clear that an efficient rescue policy and an optimum recovery strategy after a strong earthquake require the evaluation of the interactions between different lifeline systems.

In addition to ground shaking, ground failures (i.e., landslides, liquefaction, later spreading, fault ruptures), if they occur, may produce even more severe damages and losses to lifelines and infrastructures. As an example, the 2008 M_w7.9 Wenchuan earthquake in China triggered more than 15,000 landslides of various types within an area of 50,000 km^2 and causing more than 20,000 fatalities representing one-quarter of the total fatalities. The landslides produced extensive damage to housing settlements, irrigation channels, highways, bridges, and other infrastructures. The city of Wenchuan and many other towns were isolated from the rest of the country, and the rescue and relief efforts were greatly affected due to blocking and damages in road network (Tang et al. 2011). During the 2010 M_w7.1 followed by the 2011 M_w6.3 Christchurch earthquake sequences in New Zealand, the ground and slope failures, mainly due to liquefaction, produced extensive damages to the water, electricity, and road networks rendering many of them inoperable (O'Rourke et al. 2012). On the contrary the gas system performed rather well. A large number of slope failures, landslides, and rock falls occurred during the 2004 M_w6.8 Niigata ken Chuetsu earthquake on steep natural slopes and artificial cuts in the area, resulting in major damage to road and railway infrastructures (cuts, embankments, pavements, tunnels, abutments, retaining walls), as well as to other lifeline components (water pipes, electric power poles). Similar damages and disruptions, related to geotechnical failures, were recorded after the subsequent 2007 M_w6.6 Niigata-Chuetsu Oki earthquake in Japan and in several recent events in Europe (e.g., 2003 M_w6.4 Lefkas earthquake in Greece, the 2009

Seismic Vulnerability Assessment: Lifelines, Fig. 1 Damages to road, telecommunication, and gas system in 1994 Northridge (M_w6.7) earthquake (Photo Credit: US Geological Survey)

Seismic Vulnerability Assessment: Lifelines, Fig. 2 Damaged equipment in Adapazari substation in 1999 Kocaeli (M_w7.4) earthquake (Photo Credit: Kandilli Observatory and Earthquake Research Institute/KOERI)

M_w6.3 L'Aquila earthquake in Italy, the 1999 M_w7.4 Kocaeli earthquake in Turkey) and worldwide. In Figs. 1, 2, 3, 4, 5, 6, 7, and 8, some representative damages for utility and transportation networks are shown.

Several utility systems and networks are now outdated and lack proper seismic design and protection. For example, there are parts of the water and wastewater systems in several cities built more than a century ago. Hence, their material properties and strength are seriously affected from aging effects and are consequently more vulnerable and exposed to earthquake risk. Currently, most of urban infrastructures are not designed to sustain severe seismic hazard and complex systemic threats, and it is difficult to anticipate what may be the complex socioeconomic losses due to potential strong earthquakes.

The poor performance of elements at risk in a network is likely to make the whole system much more vulnerable and susceptible to strong earthquakes.

Finally, the short- and long-term economic impact due to earthquake damages of lifelines can be significant. In a recent study by CATDAT (Daniell and Vervaeck 2013), the disaggregation of direct economic losses of 61 selected major earthquakes from 1900 to 2012 showed that around 30 % of the direct losses came from infrastructure losses in transport, communications, pipelines, energy supply systems, etc.

Lifeline Earthquake Engineering Research

In the last two decades, several important projects and research efforts related to the vulnerability and risk assessment of different lifelines at

S

Seismic Vulnerability Assessment: Lifelines, Fig. 3 Lateral deformation of train tracks at the Arahama station due to surface waves in 2007 Niigata-Chuetsu Oki (M_w6.6) earthquake (Photo Credit: US Geological Survey)

Seismic Vulnerability Assessment: Lifelines, Fig. 4 Damages to road embankments in 2007 Niigata-Chuetsu Oki (M_w6.6) earthquake (Photo Credit: US Geological Survey)

regional or urban scale have been undertaken. They resulted in the development of relevant methods and guidelines. A short reference to the most important contributions is following.

United States

In 1992, the Multidisciplinary Center for Earthquake Engineering Research (MCEER) initiated the Highway Project, funded by the Federal Highway Administration (FHWA). The project uniquely examined the impact of earthquakes on the highway system as an integrated network, rather than a collection of individual road segments, bridges, embankments, tunnels, etc. The purpose of the project was to ensure the usability of highways following earthquakes, by improving performance of all interconnected components. Overall goals were to deepen the understanding of the seismic hazard impact to highways and to improve and develop analysis

methods, screening procedures and additional tools, retrofit technologies, design criteria, and other approaches to reduce seismic vulnerability of existing and future highway infrastructure. REDARS (Risks due to Earthquake DAmage to Roadway Systems) is a public-domain software package developed under this MCEER project that accounts for how earthquake damage affects post-event traffic flows and travel times and estimates losses from these travel-time and traffic-flow impacts (http://mceer.buffalo.edu/research/redars/).

In 1996, the PEER Lifelines Program (http://peer.berkeley.edu/lifelines) was initiated, under the coordination of the Pacific Earthquake Engineering Research Center, funded by the California Department of Transportation and the Pacific Gas and Electric Company. The aim was to improve the seismic safety and reliability of lifeline systems with particular focus on the characterization of ground motions, local soil and site response, and the performance of bridge structures and electric substation equipment.

In 1997, the Federal Emergency Management Agency (FEMA) released the first edition of HAZUS methodology and software that

Seismic Vulnerability Assessment: Lifelines, Fig. 5 Collapse of a section of the Nishinomiya in 1995 ($M_w6.8$) Hyogoken–Nanbu earthquake (Photo by K. Pitilakis)

Seismic Vulnerability Assessment: Lifelines, Fig. 6 Damage to Trans-European Motorway due to fault rupture in 1999 Kocaeli ($M_w7.4$) earthquake (Photo Credit: Kandilli Observatory and Earthquake Research Institute/Photo by D. Kalafat)

Seismic Vulnerability Assessment: Lifelines, Fig. 7 Blocking of Pan-American Highway due to landslide in 2001 El Salvador ($M_w7.6$) earthquake (Photo Credit: US Geological Survey/Photo by E.L. Harp)

estimates potential building and infrastructure losses from earthquakes, built using GIS technology. The current version is HAZUS-MH V2.0, including also losses from floods and hurricanes. It was released in 2004 (www.fema.gov/hazus). HAZUS-MH loss estimates reflect the state-of-the-art scientific and engineering knowledge and can be used to inform decision makers at all

Seismic Vulnerability Assessment: Lifelines, Fig. 8 Failure of port infrastructures in 1995 (M_w6.8) Hyogoken–Nanbu earthquake (Photo by K. Pitilakis)

levels of government by providing a reasonable basis for developing mitigation, emergency preparedness, response, and recovery plans and policies. The main merit of the HAZUS platform is that it provided for the first time an unparalleled set of fragility models for all the main components and systems of the built environment. However, it must be recognized that many of these models have been derived solely on expert judgment and the level of knowledge at that time. One effect of the sheer size of the HAZUS framework and its set of tools is that it established itself very soon as the reference for all studies in the sector worldwide.

In 1998, FEMA and the American Society of Civil Engineers entered into a cooperative agreement to establish the American Lifelines Alliance (ALA). ALA's objectives were to facilitate the creation, adoption, and implementation of design and retrofit guidelines and other national consensus documents to improve the performance of utility and transportation systems (electric power, telecommunication, water, wastewater, oil, natural gas, rail, and shipping ports) in natural hazard events, including earthquakes. ALA has not attempted to focus efforts on all lifeline systems and components but has instead chosen to place priority on specific topics relevant to selected lifeline systems where a need for improved hazard mitigation practices is identified. The primary function was to bridge the gap between hazard mitigation practices for buildings and lifeline systems; therefore a close collaboration with lifeline

stakeholders, owners/operators, and recognizant regulatory bodies was established to facilitate use of developed guidelines (http://www.americanlifelinesalliance.com/).

ERGO (the new name for the Multi-Hazard Assessment, Response, and Planning software previously known as mHARP and MAEviz) started somewhat later than HAZUS and was the product of the research efforts carried out at the Mid-America Earthquake Center in collaboration with the National Center for Supercomputing Applications (NCSA). It is an open-source software and incorporates many of the design concepts and capabilities motivated by NCSA efforts to develop "cyberenvironments" that span scientific disciplines and that can rapidly evolve to incorporate new research results (Elnashai et al. 2008). An important aspect of ERGO is its extensibility, both in terms of analysis/features modules and of visualization/representation (GIS) modules. The framework has been designed to implement the consequence-based risk management (CRM) paradigm supported by the NCSA and MAE center.

Europe

Until now, no coordinated action for the development of a comprehensive methodology and software tool at European level is made. However, important efforts have been undertaken through research projects mainly funded by European Union. Brief description of the most important is following.

RISK-UE (2001–2004) project involved the assessment of earthquake scenarios based on the analysis of the global impact of one or more plausible earthquakes at city scale, within a European context. The primary aim of these scenarios was to increase awareness within the decision-making centers of a city. The developed methodology for creating earthquake scenarios focused on the distinctive features of European cities with regard to current and historical buildings and lifelines, as well as on their functional and social organization, in order to identify weak points within the urban system. The approach was applied to 7 European cities: Barcelona, Bitola, Bucharest, Catania, Nice, Sofia, and Thessaloniki (Mouroux and Le Brun 2006).

LESS-LOSS (2003–2007) was a European Integrated Project focusing on risk mitigation for earthquakes and landslides that relied on the active participation of 46 partners from both academia and industry. The project addressed research issues on seismic engineering, earthquake risk and impact assessment, landslides monitoring, mapping and management strategies, improved disaster preparedness and mitigation of geotechnical hazards, development of advanced methods for risk assessment, methods of appraising environmental quality, and relevant pre-normative research. Separate sections were devoted on earthquake disaster scenario predictions and loss modeling for urban areas and infrastructures with emphasis on water and gas systems (Faccioli 2007).

SYNER-G (2009–2013, www.syner-g.eu) is the most recent European funded research project that developed an integrated, general methodology and a comprehensive framework for the systemic vulnerability and risk analysis of buildings, lifelines, and infrastructures. The proposed methodology and tools encompass in an integrated fashion all aspects in the chain, from hazard to the vulnerability assessment of components and systems and to the socioeconomic impacts of an earthquake, accounting for most relevant uncertainties within an efficient quantitative simulation scheme and modeling interactions between the multiple component systems. It systematically integrates the most advanced fragility or vulnerability functions to assess the vulnerability of physical assets. The increasing impact due to interdependencies and intra-dependencies between different components and among different interacting systems is treated in a comprehensive way, providing specifications for each network and infrastructure. The proposed socioeconomic model integrates social vulnerability into the physical systems modeling approaches providing to decision makers with a dynamic platform to capture post disaster emergency issues like shelter demand and health impact decisions (Pitilakis et al. 2014a, b).

Other Efforts Worldwide

The Central American Probabilistic Risk Assessment (CAPRA) platform was developed in partnership with Central American governments, the support of the Central American Coordination Centre for Disaster Prevention (CEPREDENAC), the Inter-American Development Bank (IDB), the International Strategy of United Nations for Disaster Reduction (UN-ISDR), and the World Bank. CAPRA is an information platform to enhance decision-making in risk management across various sectors, such as emergency management, territorial planning, public investment, and the financial sectors. Through the application of probabilistic risk assessment principles to threats like hurricanes, earthquakes, volcanic activity, floods, tsunamis, and landslides, CAPRA allows to measure and compare different types of risks and to develop sector-specific applications for risk management (Cardona et al. 2012; http://www.ecapra.org/). It is not directly designed for lifeline vulnerability and risk assessment, but it can be used for this purpose as well.

Key Elements of Seismic Risk Assessment

Risk assessment is the process used to determine risk management priorities by evaluating and comparing the level of risk to specific standards, either defined by a code or a set of target risk levels and criteria. The estimation of losses, both material and immaterial, is the ultimate target of the risk assessment. Then, the risk management procedure should define pre-seismic (retrofitting

and strengthening actions), coseismic, and post-seismic strategies and policies. The risk assessment of lifelines and infrastructures follows the general scheme of seismic risk assessment:

$$[\text{seismic risk}] = [\text{seismic hazard}] \times [\text{vulnerability}] \\ \times [\text{exposure} - \text{elements at risk}]$$

The complexity of elements at risk, their variability from one place and one country to another, and, till recently, the lack of well-validated damage and loss data from strong earthquakes make the vulnerability assessment of each particular component and of the network as a whole a challenging task. Adding to that the spatial extent of lifelines, the synergies between different systems, and the several sources of uncertainties that are inherent in the various aspects (e.g., typology, damage states), models (e.g., seismic hazard, spatial correlation, fragility curves), and finally tools and methods for estimation of losses, it is obvious that the risk assessment of lifelines is indeed a very complex and challenging issue. Therefore, multidisciplinary task and combined efforts are really needed to reach reliable and comprehensive estimates.

This difficulty is amplified due to lack of well-documented data and the partial or poor knowledge of the geometric and other features of the systems; in some cases it is also due to the confidentiality of data (i.e., communication or oil networks). The situation however has been improved the last decade. Important strong earthquakes provided valuable good quality data, while the public awareness and the reported huge direct and indirect losses associated to lifeline damages drew the attention of the scientific community, the governmental authorities, and the insurance sector. Moreover, the development of geographical information systems (GIS) as well as remote sensing technologies offered an excellent platform for the implementation of efficient and innovative techniques to screen, capture, store, manipulate, analyze, manage, and present geographical data.

Few terms employed in this chapter are described in the following. Inventory and taxonomy/typology definitions of all elements at risk and systems are the necessary initial steps to describe the elements and networks exposed to seismic risk; at the same time, the development of detailed inventories is the most expensive task in terms of financial cost and time. The estimation of seismic hazard and site-specific seismic ground motion characteristics provides appropriate seismic intensity measures (IM), which will be used later in the vulnerability and loss assessment modeling. Vulnerability is the propensity of damage of an element (building, bridge, pipeline, roadway, oil tank, etc.) or a network of specific elements (e.g., water or gas system) to a given seismic intensity defined in the seismic hazard analysis. It is commonly assessed based on fragility models, which estimate the probability of damage for a given seismic intensity. Fragility functions are constructed with respect to the taxonomy and typological characteristics of each element at risk. The expected damages and losses of the entire network are estimated through a systemic analysis, which takes into account the physical damages of the various parts and segments of the system, and in some cases, the intra-dependencies between different components within the system and the interdependencies between systems, in order to estimate the serviceability or flow capacity of the damaged network. The results are commonly provided in terms of performance indicators (PI), which are describing the network functional or the expected economic or socioeconomic losses. Efficient risk management is based on the evaluation of the resulting performance indicator values.

Seismic Hazard for Distributed Networks

Lifelines are in most cases spatially distributed systems. Therefore, seismic hazard should meet the specific features of each system considering also its spatial variability. The vulnerability analysis and risk assessment must be evaluated according to the precise typological characteristics of the components and networks, taking also into account the models used to describe vulnerability, usually in terms of fragility curves and relationships. Moreover, due to the spatial extent of utility and transportation networks, the seismic

hazard should describe the spatial variability of ground motion considering local soil conditions, topographic effects, and geotechnical hazards; site-specific seismic hazard analyses are always necessary. Traditional seismic hazard analysis, while effective in translating the hazard into a probabilistic formulation, is limited in the extent to which it can incorporate spatial coherency of the form needed for estimation of loss to spatially distributed portfolios. The extension of seismic risk analysis to multiple systems of spatially distributed infrastructures presents new challenges in the characterization of the seismic hazard input, particularly with respect to the spatial correlation structure of the ground motion residuals (i.e., residuals from empirical intensity models) that form the basis for the systemic risk analysis.

A general procedure entitled "Shakefield" has been recently established in the frame of SYNER-G project (Pitilakis et al. 2014b), which allows for the generation of samples of ground motion fields both for single-scenario-type events and for stochastically generated sets of events needed for probabilistic seismic risk analysis (Weatherill et al. 2014). For a spatially distributed infrastructure of vulnerable elements, the spatial correlation of the ground motion fields for different measures of the ground motion intensity is incorporated into the simulation procedure. This is extended further to consider spatial cross-correlation between different measures of ground motion intensity.

The consideration of hazard from permanent ground deformation (PGD) is essential in modeling the seismic risk to lifeline systems. For pipelines and similar systems with linear elements, fragility models are generally given in terms of PGD, as they are mostly vulnerable to permanent displacement of the ground rather than transient shaking. Four primary causes of permanent ground displacements are commonly considered: liquefaction-induced lateral spread, liquefaction-induced settlement, slope displacement, and coseismic fault rupture. In addition to strong shaking, another transient effect that poses a potential risk to lifeline systems is the transient ground strain (PGS). Several models are available for the estimation of PGD; some of them are

intended to relate the degree of deformation and the probability of the geotechnical hazard occurring to the intensity of the ground motion (Weatherill et al. 2014). However, it should be noted that most theoretical and empirical models relating PGD to strong shaking require a level of geotechnical detail that may be impractical to obtain for a spatially distributed set of sites. HAZUS methodology (FEMA 2003) provides a rather simple "baseline" model that can be implemented in the widest variety of applications.

When the scenario type or probabilistic site-specific seismic hazard is available, the damage assessment of a lifeline system is evaluated using appropriate fragility or vulnerability functions relating the probability of damages to seismic intensity measures. The data and parameters needed to perform this analysis are shortly described in the following.

Taxonomy, Inventory, and Typology

The key assumption of the vulnerability assessment of lifelines is that the structures and components of the systems, having similar structural characteristics (e.g., a bridge of a given typology), being in similar geotechnical conditions (e.g., a segment of a pipeline crossing the same soil conditions and exposed to the same geotechnical hazards), are expected to perform in the same way for a given seismic excitation. Within this context, expected damages are directly related to the structural properties of the elements at risk. *Taxonomy* and *typology* are thus fundamental descriptors of a system that are derived from the inventory of each element and system. Geometry, material properties, morphological features, age, seismic design level, anchorage of the equipment, soil conditions, and foundation details are among usual typology descriptors/ parameters. Buildings, bridges, pipelines (gas, fuel, water, wastewater), tunnels, road embankments, power substations, harbor facilities, and road and railway networks have their own specific set of typologies and different taxonomy.

The taxonomy of any lifeline network is thus an essential step for identifying, characterizing, and classifying all types of lifeline elements according to their specific typology and their

distinctive geometric, structural, and functional features. It allows the classification of the network elements in an ordered classification system adequate for the seismic risk assessment of each component and the network (or system) as a whole. The taxonomy is specific for each lifeline system and it needs a detailed typology definition of all elements at risk. For that, an adequate inventory process is essential. However, several difficulties arise in the collection and archiving of the data related to the absence of well-organized archives in public and often in private organizations managing the systems, the oldness of the networks (e.g., water network), and the high cost to draw up complete inventories. To this respect remote sensing techniques and GIS offer a useful and indispensable instrument and platform to enhance with relatively low cost the collection of the data and implement any inventory inquires. Different types of satellite data can be used to extract information and parameters needed to compile or update inventories for seismic risk assessment. Examples include the width of roads, the geometry or material of buildings, building aggregates, infrastructures like bridges and channels, land use characteristics, and others.

The inventory of a specific structure in a region and the capability to create classes of structural types (e.g., with respect to material, geometry, and design code level) are among the main challenges when carrying out a general seismic risk assessment, for example, at a city or region scale, where it is practically impossible to perform this assessment separately for each single structure. It is necessary to classify all elements at risk, in "as much as possible" homogeneous classes and subclasses presenting more-or-less similar response characteristics to ground shaking. Thus, the derivation of appropriate fragility curves for any type of structure depends entirely on the creation of a reasonable taxonomy that is able to classify the different kinds of structures and infrastructures in any system exposed to seismic hazard. Several uncertainties are inherent to each taxonomy system, which are inevitably propagated through the generic fragility (vulnerability) curves to the final loss estimation. The most coherent and comprehensive taxonomy presently available in Europe is produced in SYNER-G project (Table 1, Pitilakis et al. 2014a). Detailed classifications are also provided in HAZUS (FEMA 2003) and ALA (2001).

Each component is further classified according to its specific features regarding its seismic behavior. For instance, the main typological features of gas pipelines include the material type, material strength, diameter, wall thickness and smoothness of coating, type of connection and joints, as well as the nominal design and actual flow. Among them, material and the connection types (joints) together with the soil type are the most critical parameters for their seismic response. Table 2 outlines common types of pipelines based on the above classification criteria.

Another example is storage tanks, which are usually categorized according to their material type (steel or reinforced concrete), construction type (at grade or elevated), anchorage (anchored or unanchored), roof type and capacity, shape factor (height-on-diameter ratio), and amount of content in the tank (full, half full, empty). Among them, material type, construction type, and anchorage are considered as the most important. Table 3 presents the classification of tanks in USA according to ALA (2001).

For a systematic and detailed description of all classes and subclasses of all elements at risk in most lifeline systems and networks in the European context, the reader should refer to Pitilakis et al. (2014a).

Damage Assessment

The damage patterns and mechanisms are different for each lifeline system and element (component), and they are strongly depending on the typology of each structure (i.e., materials, geometry, structural types). For example, the damage mechanisms of pipelines are generally classified in the form of breaks or leaks; material type (i.e., brittle or ductile) and joints (i.e., flexible or rigid) are also factors affecting the seismic response and damages of pipelines. In case of tunnels, damage patterns include lining cracks (longitudinal or transverse) and spalling, wall

Seismic Vulnerability Assessment: Lifelines, Table 1 SYNER-G infrastructure taxonomy (Pitilakis et al. 2014a)

System	Component (and subcomponents)
BDG: buildings	Force-resisting mechanism (FRM1): moment-resisting frame structural wall, flat slab, bearing walls, precast, confined masonry
	FRM material (FRMM1): concrete, masonry
	Plan (P): regular, irregular
	Elevation (E): regular/irregular geometry
	Cladding (C): regular infill vertically, irregular infill vertically, bare
	Detailing (D): ductile, non-ductile, with tie rods/beams, without tie rods/beams
	Floor system (FS): rigid, flexible
	Roof system (RS): peaked, flat, gable end walls
	Height level (HL): Low rise, mid-rise, high rise, tall
	Code level (CL): none, low, moderate, high
EPN: electric power network	EPN01: electric power grid
	EPN02: generation plant
	EPN03: substation
	EPN04: distribution circuits
	EPN05–09: substation macro-components (autotransformer line; line without transformer; bar-connecting line; bars; cluster)
	EPN10–23: substation micro-components (circuit breaker; lightning arrester or discharger; horizontal disconnect switch or horizontal sectionalizing switch; vertical disconnect switch or vertical sectionalizing switch; transformer or autotransformer; current transformer; voltage transformer; box or control house; power supply to protection system; coil support; bar support or pothead; regulator; bus; capacitor tank)
	EPN24: transmission or distribution line
GAS: natural gas system	GAS01: production and gathering facility (onshore, offshore)
	GAS02: treatment plant
	GAS03: storage tank
	GAS04: station (compression; metering/pressure reduction; regulator; metering)
	GAS05: pipeline
	GAS06: SCADA
OIL: oil system	OIL01: production and gathering facility (onshore, offshore)
	OIL02: refinery
	OIL03: storage tank farm
	OIL04: pumping plant
	OIL05: pipeline
	OIL06: SCADA
WSS: water-supply network	WSS01: source (springs, rivers, natural lakes, impounding reservoirs, shallow or deep wells)
	WSS02: treatment plant
	WSS03: pumping station
	WSS04: storage tank
	WSS05: pipe
	WSS06: tunnel
	WSS07: canal
	WSS08: SCADA system
WWN: wastewater network	WWN01: wastewater treatment plant
	WWN02: pumping (lift) station
	WWN03: pipe
	WWN04: tunnel
	WWN05: SCADA system

S

(*continued*)

Seismic Vulnerability Assessment: Lifelines, Table 1 (continued)

System	Component (and subcomponents)
RDN: road network	RDN01: bridge (material, type of deck, deck structural system, pier to deck connection, type of pier to deck connection; type of section of the pier, spans, type of connection to the abutments, skew, bridge configuration, foundation type, seismic design level)
	RDN02: tunnel
	RDN03: embankment (road on)
	RDN04: trench (road in)
	RDN05: unstable slope (road on or running along)
	RDN06: road pavement (ground failure)
	RDN07: bridge abutment
RWN: railway network	RWN01: bridge
	RWN02: tunnel
	RWN03: embankment (track on)
	RWN04: trench (track in a)
	RWN05: unstable slope (track on or running along)
	RWN06: track
	RDN07: bridge abutment
	RWN08: station
HBR: harbor	HBR01: waterfront components (gravity retaining structures; sheet pile wharves; piers; breakwaters mooring and breasting dolphins)
	HBR02: earthen embankments (hydraulic fills and native soil material)
	HBR03: cargo handling and storage components (cranes, tanks, etc.)
	HBR04: buildings (sheds, warehouse, offices, etc.)
	HBR05: liquid fuel system (as per the OIL system)
FFS: fire-fighting system	FFS01: fire-fighter station
	FFS02: pumping station
	FFS03: storage tank
	FFS04: fire hydrant
	FFS05: pipe

deformation, bending and buckling of reinforcing bars, obstruction of the opening, or pavement cracks. In subway stations damages to columns may also occur. Damages to quay walls are related to excessive lateral pressures or decrease of shear strength of the foundation soil that can cause sliding, deformation, and tilting of the walls. Settlements and lateral movement of the backfill materials and cracking of apron pavements can be also induced. Damage to electric power substations includes failures of the various subcomponents (e.g., disconnect switches, circuit breakers, transformers) or damage of the building. Earth structures such as highway and railway embankments can spread laterally and settle, resulting in opening of cracks in the road pavement or displacement of the railway tracks. The list of possible damage patterns is unlimited.

Therefore, classification of damage and the subsequent definition of specific damage states are important in the vulnerability assessment as the seismic intensity is correlated to the expected damage level through the fragility or vulnerability functions. Again, the form of the fragility functions depends on the typology of the element at risk. For common structures (e.g., buildings, bridges) and other not extended elements (e.g., cranes, tanks, substations), the fragility curves describe the response and damage level of particular subcomponents (e.g., columns, transformers) or of the entire structure. For linear elements of extended networks such as gas pipelines, the fragility functions describe the number of expected damages along a certain length (i.e., per km). Examples and further details are given in the next sections.

Seismic Vulnerability Assessment: Lifelines, Table 2 Common types of materials and connections for buried pipelines

Material type	Connection type
Asbestos cement (AC)	Arc welded
Cast iron (CI)	Bell and spigot
Ductile iron (DI)	Cemented
Concrete (C)	Riveted
Polyvinyl chloride (PVC)	Rubber gasket
Welded steel (WS)	Gas welded
Medium-density polyethylene (MDPE)	
High-density polyethylene (HDPE)	

Seismic Vulnerability Assessment: Lifelines, Table 3 Typology of tanks (ALA 2001)

Unanchored redwood tank (5×10^4–5×10^5 gal)
Unanchored post-tensioned circular concrete tank ($>1 \times 10^6$ gal)
Unanchored steel tank with integral shell roof (1×10^5–2×10^6 gal)
Unanchored steel tank with wood roof (1×10^5–2×10^6 gal)
Anchored steel tank with integral steel roof (1×10^5–2×10^6 gal)
Unanchored steel tank with integral steel roof ($>2 \times 10^6$ gal)
Anchored steel tank with wood roof ($>2 \times 10^6$ gal)
Anchored reinforced (or prestressed) concrete tank (5×10^4–1×10^6 gal)
Elevated steel tank with no seismic design
Elevated steel tank with nominal seismic design
Open-cut reservoir
Fiberglass tanks

Damage States

In seismic risk assessment, the performance levels of a structure, for example, a reinforced concrete building, belonging to a specific class (Pitilakis et al. 2014a) can be defined through damage thresholds called limit states. A limit state defines a boundary between two different damage conditions often referred to as *damage states*. The thresholds are related to functionality and serviceability levels while they are usually defined based on engineering judgment and common sense. They are also strongly depending on the model applied for the analysis and derivation of the fragility functions. Uncertainties related to this stage of the analysis are normally referred as epistemic uncertainties (see Pinto 2014). Different damage criteria have been proposed depending on the typologies of elements at risk and the approach used for the derivation of fragility curves. The most common way to define earthquake consequences is a classification in terms of the following damage states: *no damage, slight/minor, moderate, extensive, and complete*. This qualitative approach requires an agreement on the meaning and the content of each damage state. The number of damage states is variable and is related with the functionality of the components and/or the repair duration and cost. In this way the total losses of the system (economic and functional) can be estimated. In particular, physical damages are related to the expected serviceability level of the component (i.e., fully or partial operational or inoperative) and the

corresponding functionality (e.g., power availability for electric power substations, number of available traffic lanes for roads, flow or pressure level for water system). These correlations provide quantitative measures of the component's performance and can be applied for the definition of specific *performance indicators (PIs)*, which are introduced in the systemic analysis of each network. Therefore, the comparison of a demand with a capacity quantity, the consequence of a mitigation action, or the accumulated consequences of all damages (usually referred as the "impact") can be evaluated.

Methods for deriving *fragility curves* generally describe damages on a discrete damage scale. In the empirical procedures, the scale is used in survey efforts to produce post-earthquake damage statistics, and sometimes it is rather subjective (i.e., there is often a discrepancy between the damage levels that any two different inspectors would assign for the same incident). In analytical procedures the scale is related to limit state of selected mechanical properties that are described by appropriate indices, such as the displacement capacity or the storey drift in the case of buildings or simple drift in pier bridges. For other elements at risk, the definition of the performance levels or the limit states may be more vague and follows

S

other criteria related, for example, in the case of pipelines, to the limit strength characteristics of the material used in each typology. The definition and consequently the selection of the damage thresholds, i.e., limit states, are among the main, yet, unavoidable sources of uncertainties.

Intensity Measures

An important issue related to the fragility curve construction and implicitly to the risk assessment is the selection of an appropriate earthquake *intensity measure (IM)* that characterizes the strong ground motion that best correlates with the response of each element, for example, building, pipeline, or harbor facilities like cranes. Several measures of the intensity of ground motion (IMs) have been developed. Each intensity measure may describe different characteristics of the motion, some of which may be more adverse for the structure or the system under consideration. The use of a particular IM in seismic risk analysis should be guided by the extent to which the measure corresponds to damage to the components of a system. Optimum intensity measures are defined in terms of practicality, effectiveness, efficiency, sufficiency, robustness, and computability (Mackie and Stojadinovic 2003).

In general, IMs are grouped in two general classes: empirical intensity measures and instrumental intensity measures. With regard to the empirical IMs, different macroseismic intensity scales could be used to identify the observed effects of ground shaking over a limited area. Instrumental IMs are, by far, more accurate and representative of the seismic intensity characteristics and the severity of ground shaking. For example, for bridges the best descriptor is a spectral response value at a specific period (i.e., T = 1.0 s). For other lifeline components, it may be the peak ground acceleration (e.g., buildings, tanks, electric power substations), peak ground velocity (e.g., pipelines), or even the permanent ground deformations (e.g., pipes, embankments, roadways, railways). The correlation between damages of specific elements at risk and intensity measures is not simple and never unique. Several other descriptors like peak ground strain, Arias intensity, cumulative absolute velocity, and other

parameters of the ground motion have been also used for different structures composing a lifeline system. More recently, it is proposed to use two descriptors instead of one, and hence the fragility curve is transformed in fragility surfaces (Seyedi et al. 2010; Douglas et al. 2014).

The selection of the adequate intensity parameter is also related to the approach that is followed for the derivation of fragility curves and the typology of elements at risk. The identification of the proper IM is determined from different constraints, which are first of all related to the adopted hazard model, but also to the element at risk under consideration and the availability of data and fragility functions for all different exposed assets. Empirical fragility functions are usually expressed in terms of the macroseismic intensity defined according to different macroseismic scales (EMS, MCS, and MM). Analytical or hybrid fragility functions are, on the contrary, related to instrumental IMs, which are related to parameters of the ground motion (PGA, PGV, PGD) or of the structural response (spectral acceleration Sa or spectral displacement Sd, for a given value of the period of vibration T). When the vulnerability of elements due to ground failure is examined (i.e., liquefaction, fault rupture, landslides), permanent ground deformation (PGD) is the most appropriate IM.

Vulnerability Assessment and Fragility Curves

The fundamental tool in seismic risk assessment of lifeline components is the *fragility curves* which describe the probability that a structure will reach or exceed a certain damage state for a given ground motion intensity. An extensive review of available fragility functions and state-of-the-art methods for vulnerability assessment of buildings and lifeline components can be found in Pitilakis et al. (2014a). Fragility curves are usually represented by two-parameter (median and log-standard deviation) cumulative lognormal distributions. Several approaches are used to establish fragility functions. They are grouped in the following five categories:

– Empirical fragility curves, based on post-earthquake surveys and observations of actual

Seismic Vulnerability Assessment: Lifelines, Fig. 9 Empirical fragility curves for power grids made up of substations of different voltages based on data from US west coast earthquakes (Dueñas-Osorio et al. 2007)
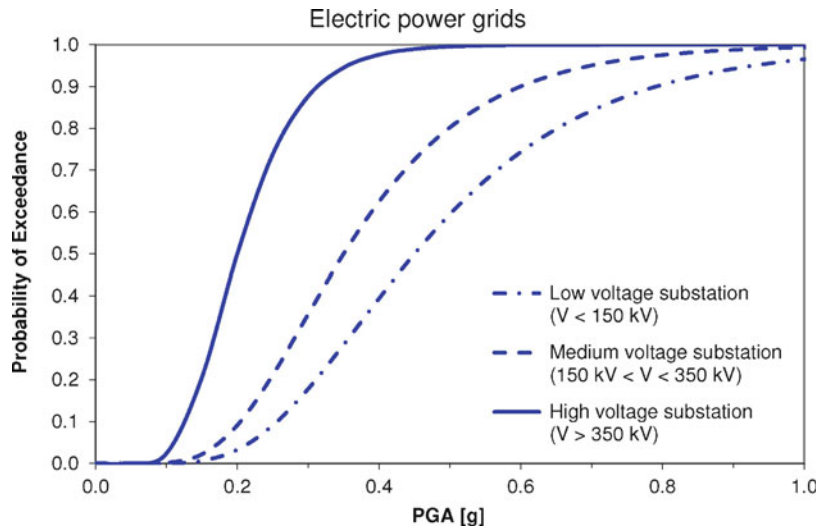
damage. They are specific to particular sites and seismotectonic, geological, and geotechnical conditions, as well as the properties of the damaged structures. Consequently, the use of these functions in different regions is always questionable (Figs. 9 and 10).

In case of pipelines, the empirical fragility functions relate the repair rates (RR) expressed as repairs/km with the peak ground velocity (PGV) or permanent ground deformation (PGD) (Fig. 11). The curves may be further adapted to the material properties and geometry of the pipelines and soil conditions. In the last two decades, the increased density of high-quality strong ground motion records in different soil conditions, in combination with new technologies such as geographical information systems (GIS) and remote sensing technologies (e.g., LiDAR) capable to measure more accurately ground movements, contributed significantly to the development and verification of such relationships (O'Rourke et al. 2012).

– Analytical fragility curves, based on numerical simulations of structural models under increasing earthquake loads (Figs. 12 and 13). Analytical methods, validated with large-scale experimental data and observations from recent strong earthquakes, have become more popular in recent years.

The main reason is the considerable improvement of computational tools, methods, and skills, which allows comprehensive parametric studies covering most common typologies to be undertaken. Moreover, several of the associated uncertainties, e.g., material properties, are better controlled.

– Judgmental or expert elicitation fragility curves, using questionnaires by which the experts are queried on the probability of a component being in a certain damage state for a given intensity (Fig. 14). They are versatile and relatively fast to establish, but their reliability is questionable because of their dependence on the experiences of the experts consulted.

– Hybrid fragility curves, which combine any of the above-mentioned techniques in order to compensate for their respective drawbacks.

– Fragility curves based on a fault-tree analysis, where complex components (e.g., substations, pumping plants, hospitals) are broken down into subcomponents and the global fragility is obtained based on the relationships between the subcomponents and their individual fragilities (Fig. 15).

An example of damage estimation for "on-ground unanchored steel tanks" due to ground shaking and "pipelines" due to ground failure is given in Table 4. In the first case, the exceedance probabilities of each damage state are estimated

Seismic Vulnerability Assessment: Lifelines, Fig. 10 Empirical fragility curves for on-ground steel tanks subjected to ground shaking (ALA 2001)
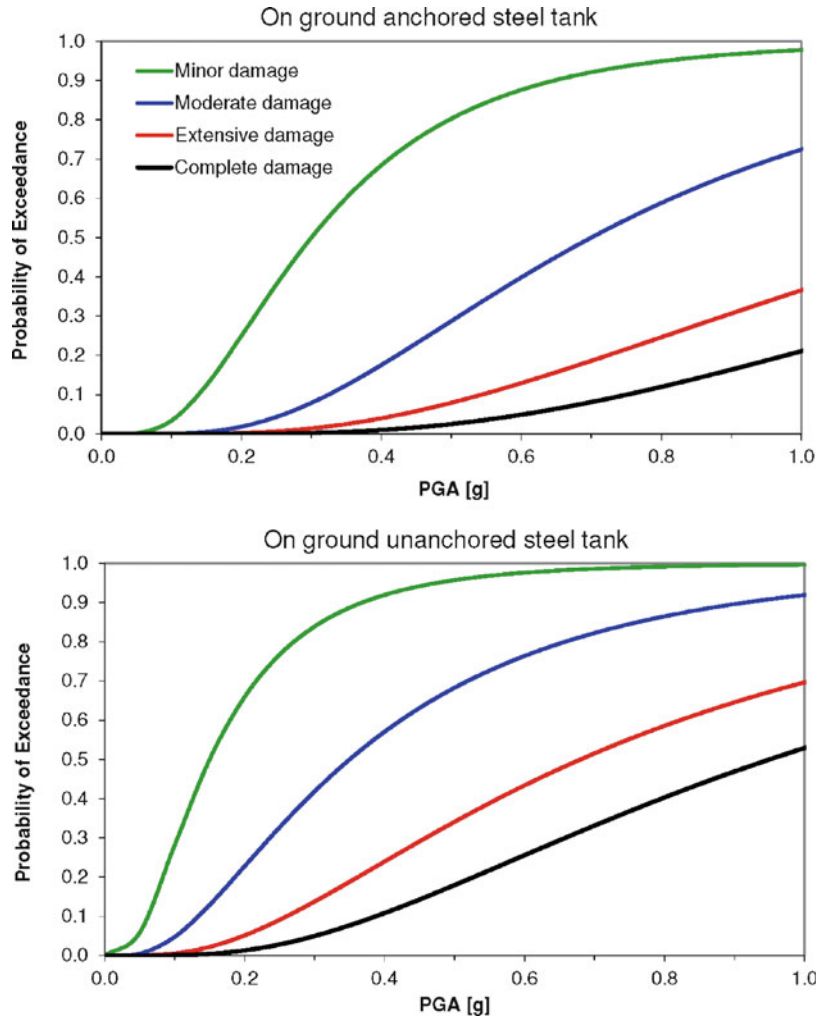
based on the fragility curves in Fig. 10, and then the punctual probabilities of each damage state are obtained. In the second case, the repairs per kilometer are estimated for a given value of PGD based on the curves in Fig. 11. The estimated repairs should be modified according to the length of the pipe segment under study.

Systemic Analysis

The majority of the methodologies that have been developed worldwide for the seismic risk assessment of lifelines and infrastructures refer to single systems, without considering interactions, cascading failures, and complex impacts. The study of interdependent infrastructures is challenging due to heterogeneous quality and

insufficient data availability and the need to account for their spatial and temporal aspects of complex supply–demand operation (Satumtira and Duenas-Osorio 2010). The various approaches available in the literature depend on the simulation method, modeling objectives, scale of analysis, availability of input data, and end-user type or needs. A comprehensive categorization scheme describing different levels is provided in Modaressi et al. (2014):

– Vulnerability analysis: This level considers only the potential physical damages of the components of the systems, with no consideration of functionality of either the elements or the whole system.
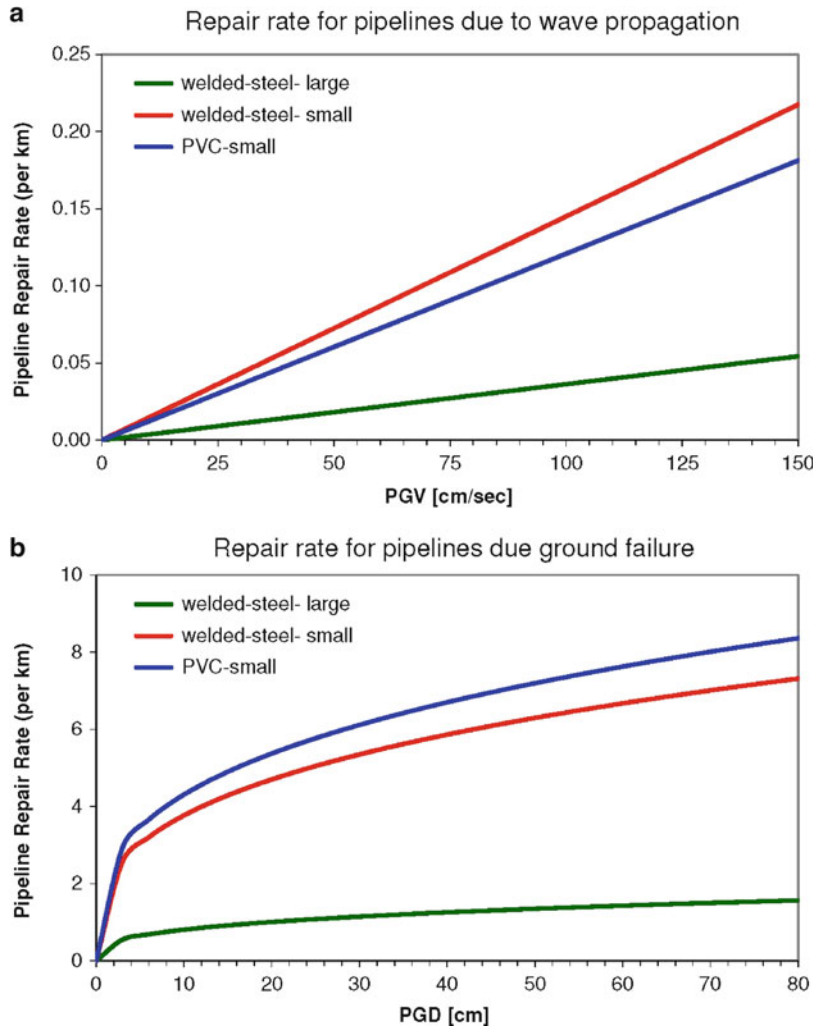
– Connectivity analysis: This level describes the probability of the demand nodes to be connected to functioning supply nodes through undamaged paths. In this approach the damaged components are removed from the network and the adjacency matrix is updated accordingly, thus pointing out the nodes or areas that are disconnected from the rest of the system. This qualitative approach is used for all utility networks (water, electricity, gas) and the road transportation system.

– Capacitive analysis: This level describes the ability of the system to provide to the users the required functionality through a quantitative approach. For utility networks, graph algorithms and flow equations can be used to estimate capacitive flows from sources (e.g., generators, reservoirs) to sinks (i.e., distribution nodes), based on the damages sustained by the network components (from total destruction to slight damages reducing the capacity).

– Fault-tree analysis: This level of analysis concerns critical infrastructures, where multiple conditions are necessary for the systems to ensure its task. This type of approach aims to evaluate the remaining operating capacity (residual operation capacity) of objects such as health-care facilities. The system is broken down into structural, nonstructural, or human components, each one of them being connected with logic operators.

Fig. 12 Analytical fragility curves for alluvial bored tunnels due to ground shaking, classified to ground type *B*, *C*, and *D* according to Eurocode 8 (Argyroudis and Pitilakis 2012)
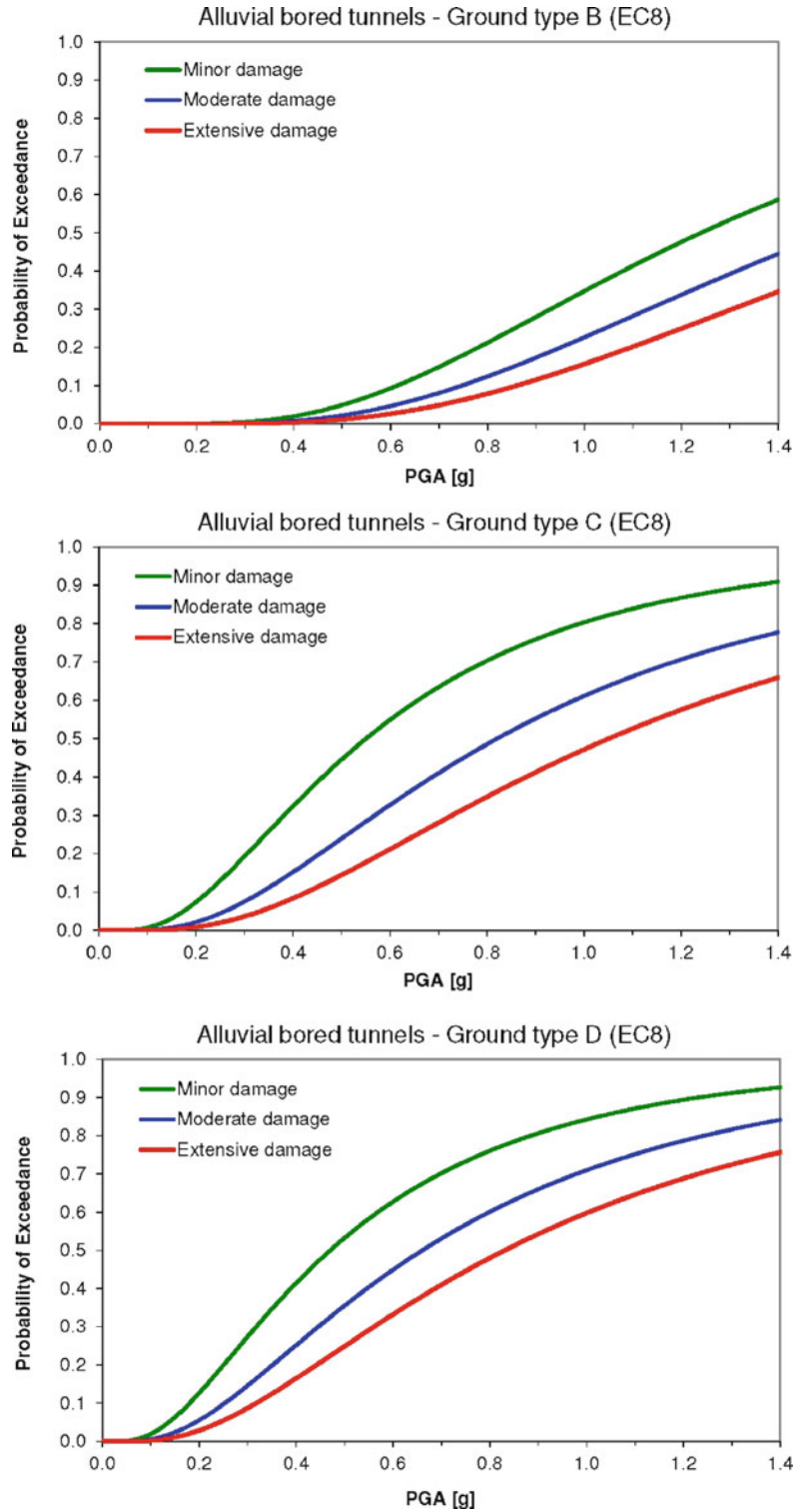
Seismic Vulnerability Assessment: Lifelines, Fig. 13 Analytical fragility curves for gravity waterfront structures due to ground shaking, classified according to the wall height (H) and the soil foundation conditions (Vs values) (Kakderi and Pitilakis 2010)
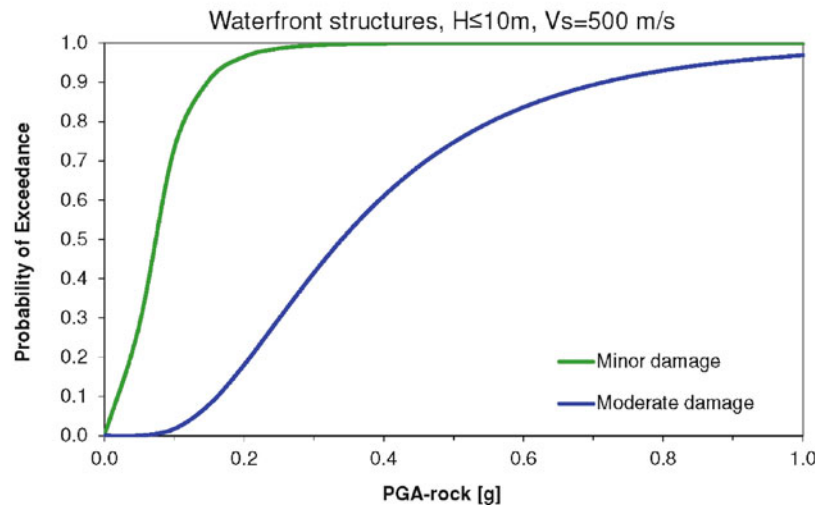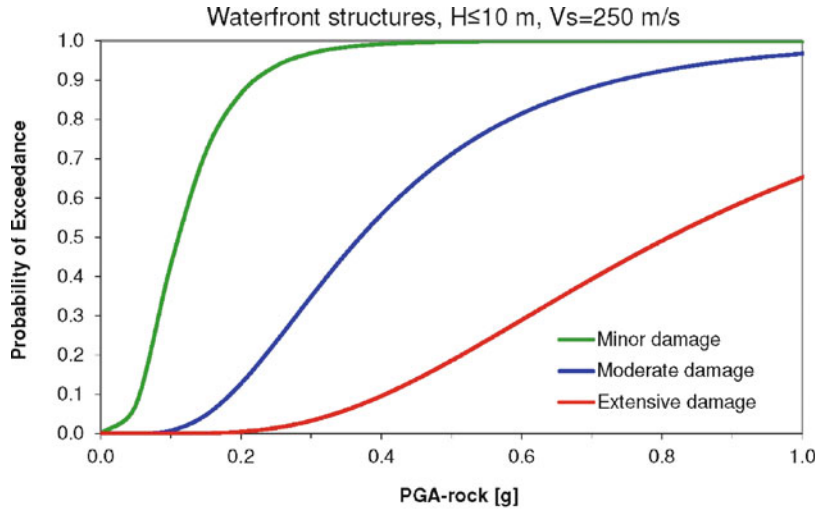
Seismic Vulnerability Assessment: Lifelines, Fig. 14 Expert judgment fragility curves for railway tracks subjected to permanent ground deformation (Argyroudis and Kaynia 2014)

S
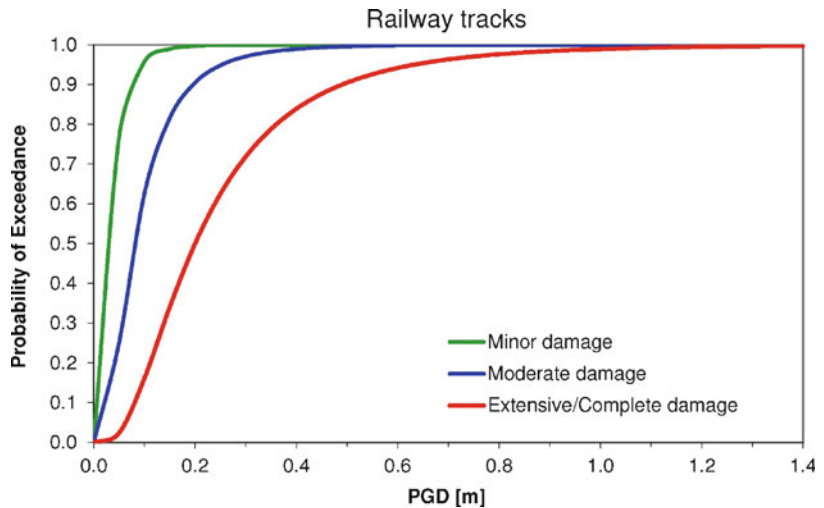
Seismic Vulnerability Assessment: Lifelines, Fig. 15 Fault-tree based fragility curves for water treatment plant with anchored components subjected to ground shaking (Pitilakis et al. 2014a)
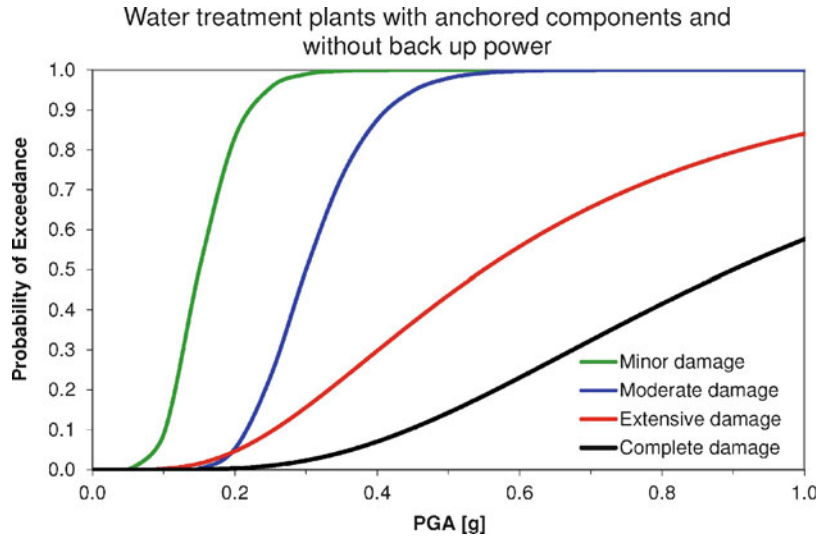
Seismic Vulnerability Assessment: Lifelines, Table 4 Example of damage estimation for steel tanks and pipelines

	On-ground unanchored steel tanks subjected to ground shaking (PGA = 0.4 g)				
Damage state	No damage	Minor	Moderate	Extensive	Complete
Probability of exceedance	–	0.919	0.571	0.240	0.108
Probability of occurrence	0.081	0.348	0.331	0.132	0.108
	(=1–0.919)	(=0.919–0.571)	(=0.571–0.240)	(=0.240–0.108)	
	Pipelines subjected to ground failure (PGD = 40 cm)				
Material diameter	Welded steel – large		Welded steel – small		PVC – small
Repairs/km	1.26		5.86		6.70

The performance of each network (e.g., utility or accessibility losses) is commonly measured through appropriate *performance indicators (PIs)*, which, if combined with direct losses from physical damages, can yield a first partial estimate of the overall socioeconomic impact of an earthquake. Performance indicators, at the component or the system level, depend on the type of analysis that is performed. Connectivity analysis gives access to indices such as the connectivity loss (measure of the reduction of the number of possible paths from sources to sinks). Capacitive modeling yields more elaborate performance indicators at the distribution nodes (e.g., head ratio for water system, voltage ratio for electric buses) or for the whole system (e.g., system serviceability index comparing the customer demand satisfaction before and after the seismic event). The fault-tree analysis method is generally used for the derivation of fragility

curves for specific components that comprise a set of subcomponents (e.g., health-care facilities, water treatment plants).

The importance of the interconnection between different systems is a more recent acquisition that targets two or, rarely, more systems (Pitilakis et al. 2014b). Several classifications have been proposed to categorize the types of interactions. The most common are physical, demand, and geographic interactions (Rinaldi et al. 2001). Physical interaction describes physical reliance on material flow from one infrastructural system to another as, for example, the supply of power to various network facilities by electric generators. Demand interactions correspond to a supply–demand from a given component to another system. An example is the number of casualties that should be treated by health-care facilities after an earthquake. Finally, geographic interactions describe the way that a local

environmental event affects components across multiple infrastructural systems due to physical proximity. For instance, the collapse of buildings in city centers can induce the blockage of adjacent roads due the debris accumulation.

A comprehensive methodological framework for the assessment of physical as well as socio-economic seismic vulnerability and risk of buildings, lifelines, and infrastructures at urban and regional level considering inter-element and intra-systems dependencies has been developed in SYNER-G project. The reader is referred to Pitilakis et al. (2014b) for more details and applications of the proposed framework.

Seismic Risk Assessment of Lifelines: Examples

Representative examples of seismic risk assessment studies for lifelines are given in the following, for different scales of analysis.

Medium- to High-Voltage Electric Power Network of the Sicily Region, Southern Italy (Regional Scale)

A power flow analysis is performed for the electric power network or Sicily, which is composed of 181 nodes and 220 transmission lines (Cavalieri et al. 2014). The nodes, i.e., the buses, are subdivided into 175 demand or load nodes and six supply nodes, five of which are power plants and one is the balance node (or slack bus), which is coinciding with the generation node providing the highest power. The load nodes (two for transmission/distribution and one for distribution substations) deliver power to users. In total, 390 municipalities are served by the network. All transmission lines are overhead lines and considered as non-vulnerable elements. They are classified into high- (HV), medium- (MV), and low-voltage (LV) lines (Fig. 16a). The vulnerable elements are the components within substations, called micro-components. A probabilistic evaluation of the performance of network is carried out by means of Monte Carlo simulation by sampling seismic events for 18 faults taken from the Italian DISS database employing the truncated Gutenberg and Richter recurrence model for the source activity.

The distribution of performance losses is shown in Fig. 16c as the mean annual frequency (MAF) of exceedance for the system serviceability index (SSI). SSI provides a global scalar measure of the system performance. It is defined as the ratio of the sum of the real power delivered from load buses after an earthquake to that before the earthquake. For each bus inside the substations, voltage ratio (VR) is defined as the ratio of the voltage magnitude in the seismically damaged network to the reference value for non-seismic, normal conditions. Figure 16b displays a contour map of the expected values of VR, averaged on the whole simulation for each demand node. It can be seen that the reduction in voltage due to induced damage is less than the tolerated threshold of 10 %, allowing the power demand delivery everywhere in the island.

Water Network of Thessaloniki, Greece (City Level)

A connectivity analysis is performed for the main water system (WSS) of Thessaloniki in Greece considering the interaction of electric power network (EPN) with pumping stations (Pitilakis et al. 2014b). If a pump serving a source node is not fed by the reference EPN node due to damage in EPN substation, then the pump itself is considered out of service and the relative WSS node is removed from the system. A probabilistic evaluation of the performance of networks is carried out by means of Monte Carlo simulation by sampling seismic events for five seismic zones with $M_{min} = 5.5$ and $M_{max} = 7.5$. Pipeline damage is evaluated for each simulation considering both wave propagation and ground failure due to liquefaction. The network is analyzed for each sampled event and the results are aggregated all over the sampled events, in order to numerically obtain the marginal distribution of performance losses (Fig. 17). The interaction can be important; as an example the water connectivity loss is increased from 1 to 1.8 % for $\lambda = 0.001$ (corresponding to mean return period T = 1,000 years) when the connections of water pumping stations to EPN are included in the analysis.

Figure 18 shows the level of correlation between the water connectivity loss and damages in pipes as well as the nonfunctional EPN

S

Seismic Vulnerability Assessment: Lifelines, Fig. 16 Seismic risk assessment of electric power network in Sicily, Italy (Source: Cavalieri et al. 2013). (**a**) Transmission lines, classified by voltage. (**b**) Contour map of expected values of voltage ratio (VR). (**c**) Mean annual frequency (λ) curve for electric power system serviceability index (SSI)
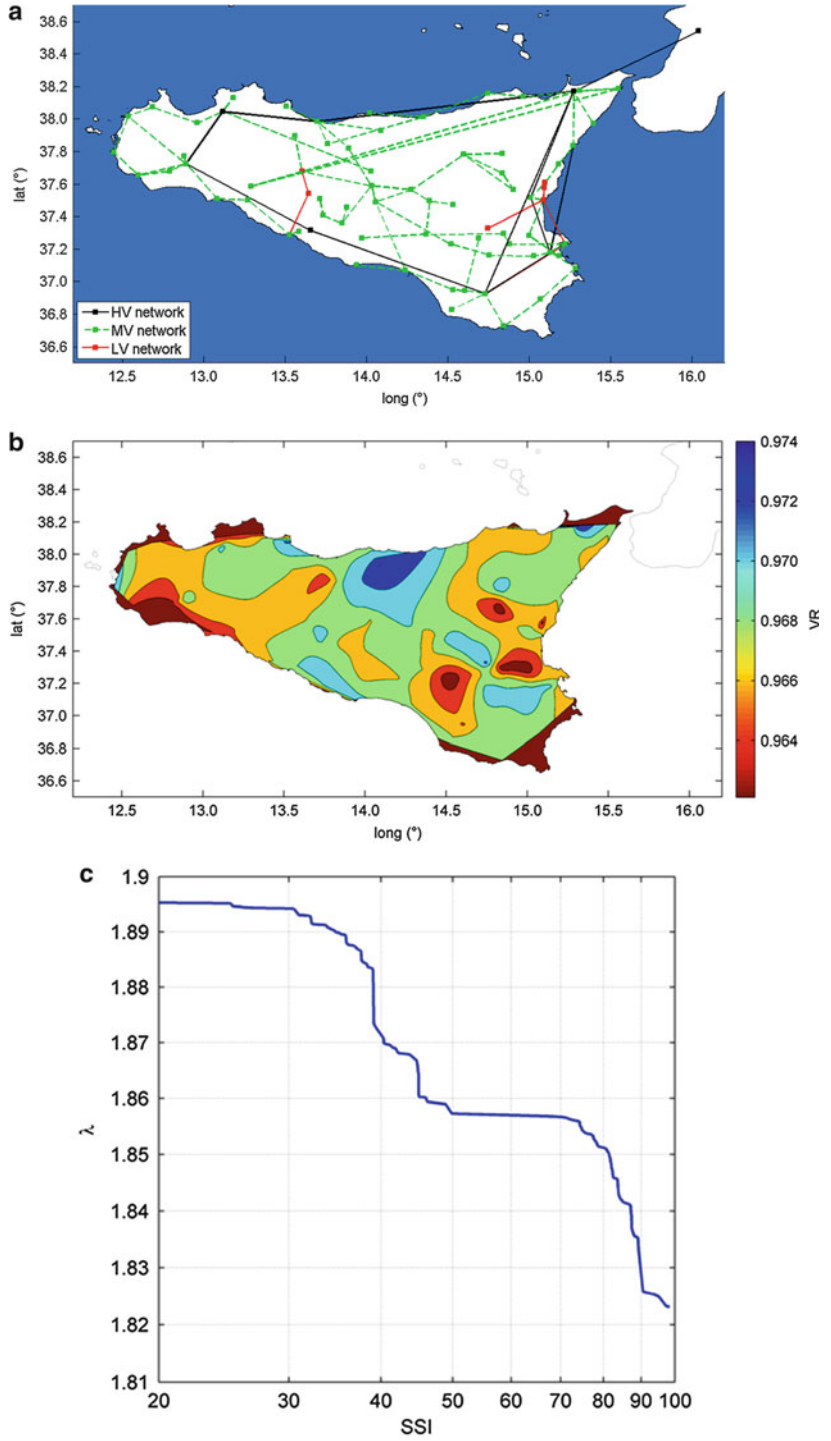
**Seismic Vulnerability
Assessment: Lifelines,
Fig. 17** Mean annual
frequency (λ) curve for
water connectivity loss
with and without
interaction with electric
power network (EPN)
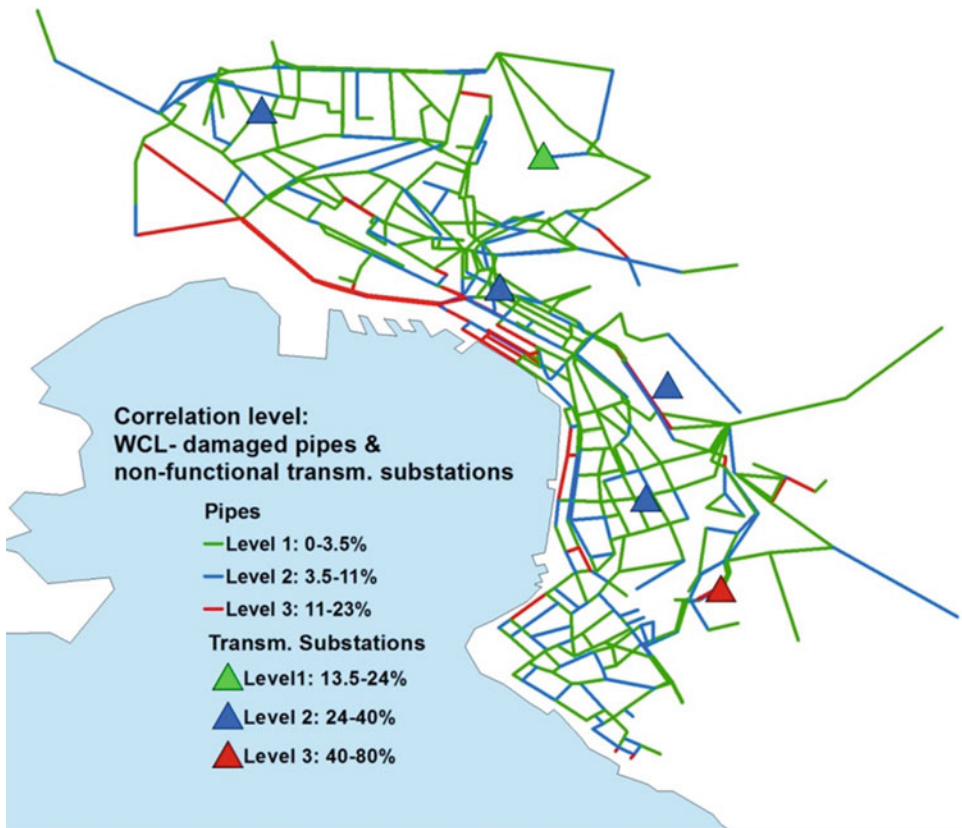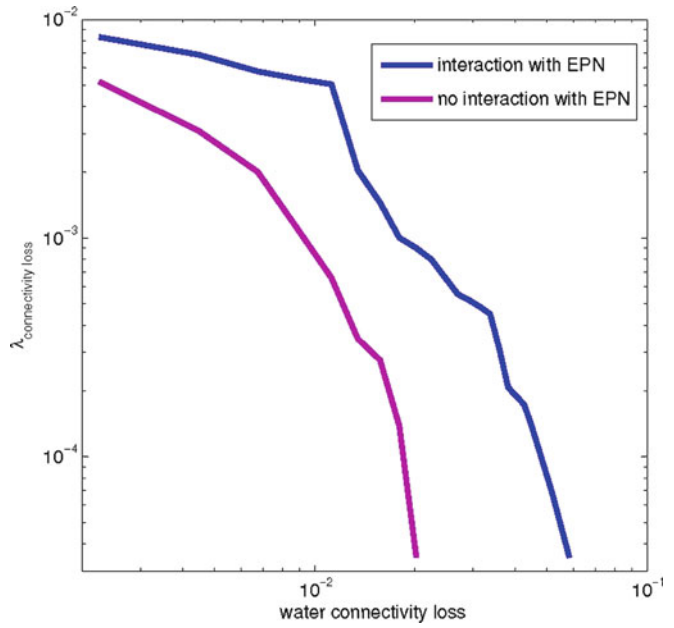
Seismic Vulnerability Assessment: Lifelines, Fig. 18 Correlation of damaged pipes and nonfunctional EPN transmission stations to water network connectivity

substations supplying the water pumping stations. The most correlated pipes are concentrated along the coastline where liquefaction susceptibility is high and therefore damages due to permanent ground displacement are expected. A higher level of correlation is obtained for the EPN transmission substations. The highest value of 80 % is attributed to a component in the southeast part of the city, where several pumping stations (connected to EPN) are located.

Road Network of Thessaloniki, Greece (City Level)
The main network of the urban area is considered in this case study, together with the ring road and the main exits of the city where the majority of bridges and overpasses are located. In particular, 594 nodes and 674 edges are included in the simulation. The nodes are subdivided into 15 external nodes, 127 traffic analysis zone (TAZ) centroids, and 452 simple intersections. Edges are assumed to be the only vulnerable components in the network. They are classified into road pavements and bridges, with fragility models expressed in terms of permanent ground deformation (PGD) due to liquefaction and peak ground acceleration (PGA) for ground shaking, respectively (Pitilakis et al. 2014b).

Road closures are estimated due to soil liquefaction and bridge damages. Moreover, the interaction with collapsed buildings that can induce road blockages is considered (Pitilakis et al. 2014b). A probabilistic evaluation of the network's performance is carried similarly to the one described in the previous example.

The interaction with building collapses can be important especially for return periods higher than 500 years ($\lambda = 0.002$). As an example the connectivity loss is increased from 20 % to 33 % for $\lambda = 0.001$ (T = 1,000 years) when the building collapses are included in the analysis (Fig. 19). Figures 20 and 21 show the level of correlation between the connectivity loss and the distribution of blockages due to building collapses and damage in bridges and road pavements, respectively. Relatively higher correlation factors are found for edges blocked by building collapse, demonstrating the importance of this failure mechanism in the analysis. A few road segments near the coastline which are subjected to ground failure due to liquefaction are also highly correlated to the network connectivity. The high risk of failure for bridges is attributed to their typology characteristics (old, simple span bridges) and the high values of PGA.
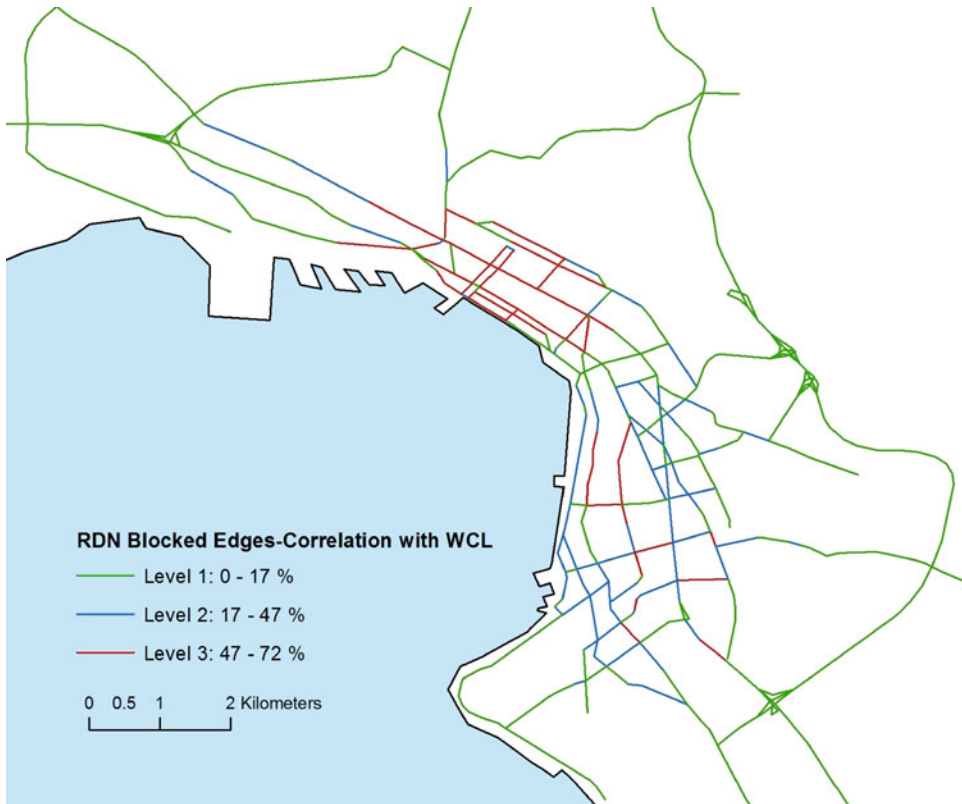
Seismic Vulnerability Assessment: Lifelines, Fig. 20 Correlation of edges blocked by buildings' collapse to road network connectivity

Port System of Thessaloniki, Greece (Infrastructure Level)

The port covers an area of 1,550,000 m^2 and trades approximately 16,000,000 t of cargo annually, having a capacity of 370,000 containers and six piers with 6,500 m length. In this case study, waterfront structures, cargo handling equipment, power supply system, roadway system, and buildings are examined. In particular, waterfront structures of a total 6.5 km length, 48 crane nodes, and two terminals (one container and one bulk cargo) are considered. The interactions accounted for in the analysis are the supply of EPN to cranes and the road closures due to building collapses. A probabilistic evaluation of the performance of networks is carried out by means of Monte Carlo simulation by sampling seismic events for five seismic zones affecting the city of Thessaloniki and the harbor with $M_{min} = 5.5$ and

$M_{max} = 7.5$. The performance of the port is described through the total cargo or containers handled in a predefined time frame per terminal and for the whole port system (Pitilakis et al. 2014b).

Figures 22 and 23 show the level of correlation between the total cargo handled per day (TCaH) and the distribution of damages in cranes and non-functionality of electric power distribution substations, respectively. In this way the most critical components can be defined in relation with their contribution to the performance loss of the system. All cranes have medium (40–70 %) to high (over 70 %) levels of correlation, indicating their great importance to the functionality of the overall port system. A higher level of correlation is estimated for the EPN distribution substations, with 40 % of the components having values greater than 70 %.
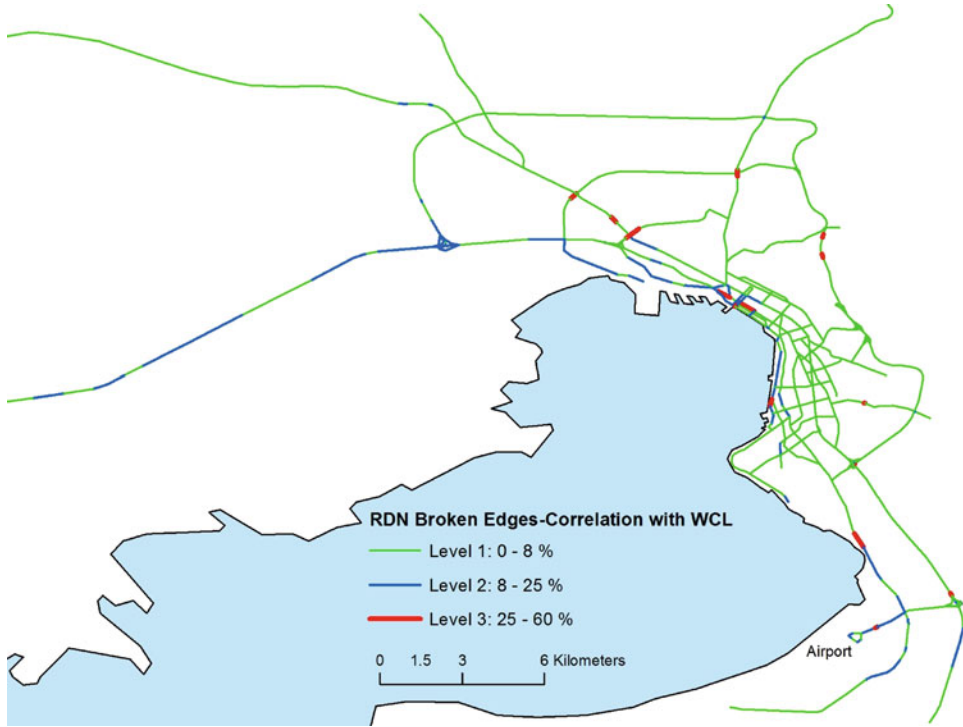
Seismic Vulnerability Assessment: Lifelines, Fig. 21 Correlation of broken edges (bridges due to ground shaking or road segments due to liquefaction) to road network connectivity
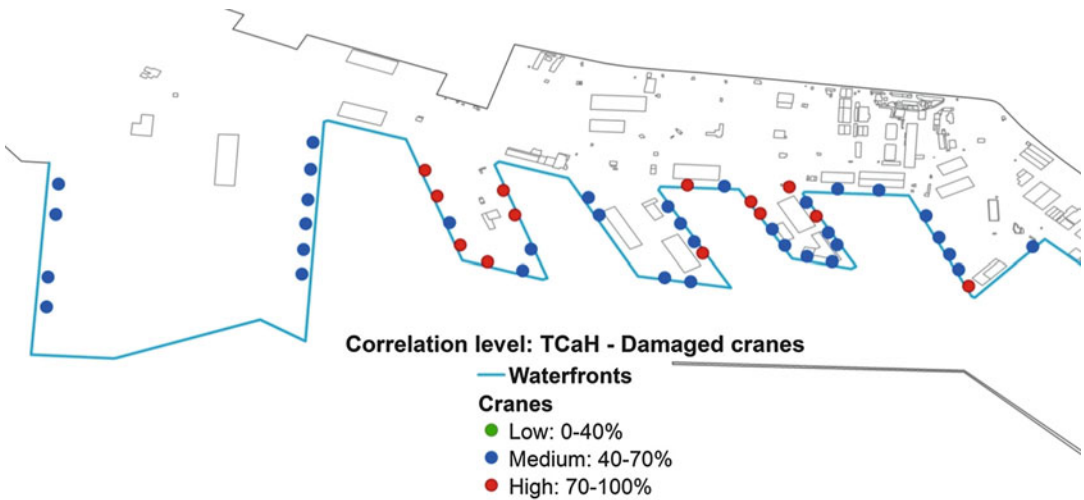
Seismic Vulnerability Assessment: Lifelines, Fig. 22 Correlation of damaged cranes to port performance (PI = TCaH)
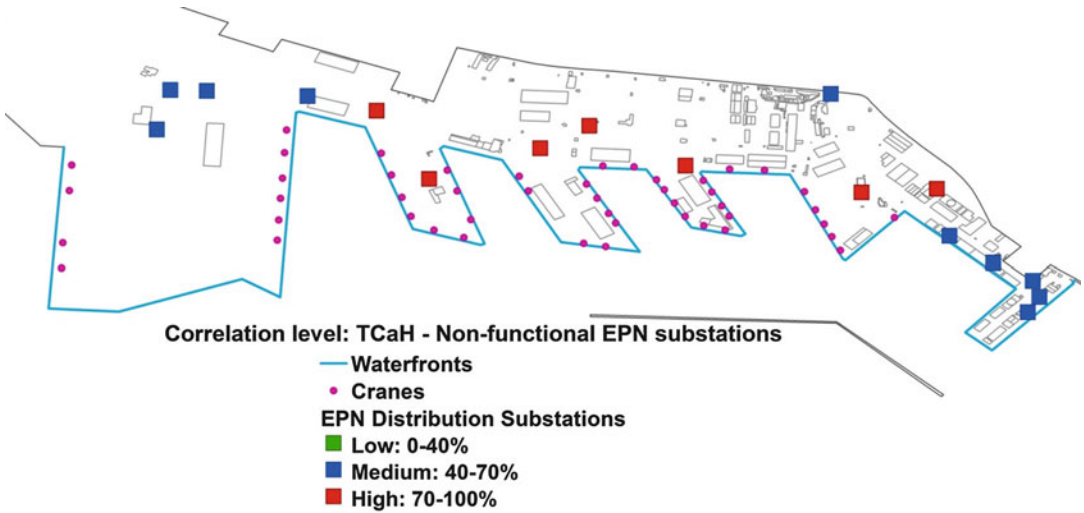
Seismic Vulnerability Assessment: Lifelines, Fig. 23 Correlation of nonfunctional electric power distribution substations to port performance (PI = TCaH)
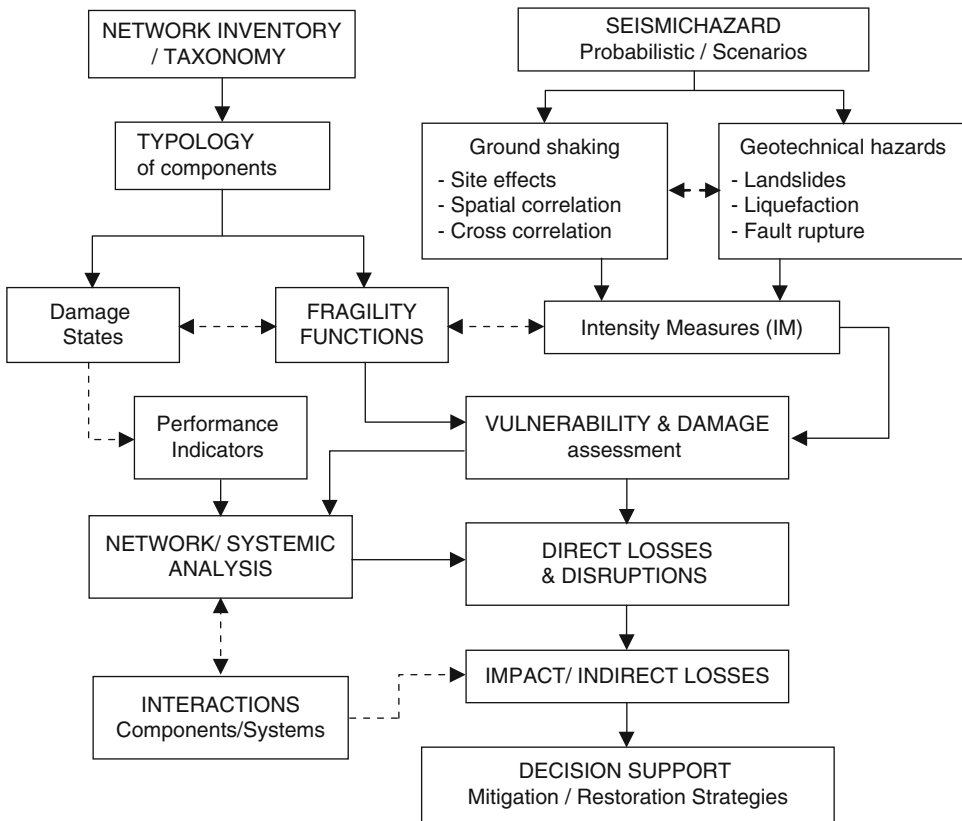
Seismic Vulnerability Assessment: Lifelines, Fig. 24 General layout for the seismic risk assessment of lifelines

Summary

Lifelines are spatially distributed systems that provide essential services to any modern society. They play also an important role for emergency response and restoration in the aftermath of disastrous earthquakes; in general they are vital for the resilience of society. They are often grouped to transportation and utility systems and comprise multiple components, which are exposed to different ground shaking effects and geotechnical hazards. Research efforts and studies undertaken the last 20 years, and after numerous devastating earthquakes that caused extensive losses and disruptions to lifelines and infrastructures, contributed to important improvement of knowledge and expertise in the seismic vulnerability and risk assessment of lifelines. The main objective of these efforts has been the development of methodologies and tools for the estimation of probable direct and indirect losses (physical, performance, economic, social) due to future earthquakes in order to develop efficient emergency response and mitigation strategies. A general layout for the seismic risk assessment of lifelines is outlined in Fig. 24. Inventory and taxonomy/typology definitions of all elements at risk and systems are the necessary initial steps to describe the elements and networks exposed to seismic risk. The estimation of site-specific seismic ground motion and other geotechnical hazards and the selection–estimation of appropriate seismic intensity measures (IM) is the basis of the seismic hazard analysis. Vulnerability is the expected response of an element or a network to a given seismic intensity. It is commonly assessed based on fragility models, which estimate the probability of exceeding certain damage states for given seismic intensity. Fragility functions are adapted with respect to the typological characteristics of each element at risk. The expected damage and loss of the entire network are estimated through appropriate systemic analysis, which takes into account the physical damage and the relations and interactions between the different components and systems. The interdependencies among different systems may considerably increase the overall impact. The results, which are commonly provided in terms of performance indicators that describe the network functional losses or the expected economic or socioeconomic losses, provide the means for an efficient mitigation or recovery planning. Obviously, several sources of uncertainties are inherent in the various definitions, models, tools, and methods for estimation of losses, which make the risk assessment of lifelines a complex and challenging topic.

Cross-References

▶ Damage to Infrastructure: Modeling
▶ Earthquake Risk Mitigation of Lifelines and Critical Facilities
▶ Empirical Fragility
▶ Liquefaction: Performance of Building Foundation Systems
▶ Probabilistic Seismic Hazard Models
▶ Seismic Actions Due to Near-Fault Ground Motion
▶ Seismic Fragility Analysis
▶ Seismic Loss Assessment
▶ Seismic Risk Assessment, Cascading Effects
▶ Seismic Vulnerability Assessment: Reinforced Concrete Structures
▶ Spatial Variability of Ground Motion: Seismic Analysis
▶ Strengthening Techniques: Bridges
▶ Uncertainty Theories: Overview

References

American Lifelines Alliance [ALA] (2001) Seismic fragility formulations for water systems. Part 1 – guideline. ASCE-FEMA, Washington, DC, 104 pp

Argyroudis S, Kaynia AM (2014) Fragility functions of highway and railway infrastructure. In: Pitilakis K, Crowley H, Kaynia AM (eds) SYNER-G: typology definition and fragility functions for physical elements at seismic risk, vol 27, Geotechnical, geological and earthquake engineering. Springer, Dordrecht. doi:10.1007/978-94-007-7872-6_10

Argyroudis S, Pitilakis K (2012) Seismic fragility curves of shallow tunnels in alluvial deposits. Soil Dyn Earthq Eng 35:1–12

Basöz N, Kiremidjian AS (1998) Evaluation of bridge damage data from the Loma Prieta and Northridge, California earthquake. Technical report MCEER-98-0004. State University of New York, Buffalo

Cardona OD, Ordaz MG, Reinoso E, Yamín LE, Barbar AH (2012) CAPRA – comprehensive approach to probabilistic risk assessment: international initiative for risk management effectiveness. In: Proceedings of the 15th world conference of earthquake engineering, Lisbon, 24–28 Sept 2012

Cavalieri F et al (2013) Application and validation study to an electric power network in Italy. In: Pitilakis K, Argyroudis S (eds) Systemic seismic vulnerability and loss assessment: validation studies SYNER-G. Reference report 6, Publications Office of the European Union, doi: 10.2788/16706

Cavalieri F, Franchin P, Pinto PE (2014) Application to selected transportation and electric networks in Italy. In: Pitilakis K et al (eds) SYNER-G: systemic seismic vulnerability and risk assessment of complex urban, utility, lifeline systems and critical facilities. Methodology and applications. Springer, Dordrecht. ISBN 978-94-017-8834-2

Daniell J, Vervaeck A (2013) CATDAT Damaging earthquakes in 2012 – the year in review. http://earthquake-report.com/2013/01/07/damaging-earthquakes-2012-database-report-the-year-in-review/

Douglas J, Seyedi DM, Ulrich T, Modaressi H, Foerster E, Pitilakis K, Pitilakis D, Karatzetzou A, Gazetas G, Garini E, Loli M (2014) Evaluation of seismic hazard for the assessment of historical elements at risk: description of input and selection of intensity measures. Bull Earthquake Eng. doi:10.1007/s10518-014-9606-0

Dueñas-Osorio L, Craig JI, Goodno BJ (2007) Seismic response of critical interdependent net-works. Earthq Eng Struct 36(2):285–306

Elnashai A, Hampton S, Lee JS, McLaren T, Myers JD, Navarro C, Spencer B, Tolbert N (2008) Architectural overview of MAEviz-HAZTURK. J Earthq Eng 12(S2):92–99. doi:10.1080/13632460802013610

Faccioli E (ed) (2007) Prediction of ground motion and loss scenarios for selected infrastructure systems in European urban environments: LESSLOSS report No. 2007/08, ISBN: 978-88-6198-012-9. IUSS Press, Pavia

Federal Emergency Management Agency [FEMA] (2003) Multi hazard loss estimation methodology: earthquake model – HAZUS-MH MR3 technical manual. FEMA, Washington, DC

Kakderi K, Pitilakis K (2010) Seismic analysis and fragility curves of gravity waterfront structures. In: Fifth international conference on recent advances in geotechnical. Earthquak Engineering and Soil Dynamics and Symposium in Honour of Prof. I. M. Idriss, 6.04a, San Diego, CA, May 24–29

Mackie K, Stojadinovic B (2003) Seismic demands for performance-based design of bridges. PEER Report 2003/16. Pacific Earthquake Engineering Research Center, University of California, Berkeley

Modaressi H, Desramaut N, Gehl P (2014) Specification of the vulnerability of physical systems. In: Pitilakis K et al (eds) SYNER-G: systemic seismic vulnerability and risk assessment of complex urban, utility, lifeline systems and critical facilities. Methodology and applications. Springer, Dordrecht. ISBN 978-94-017-8834-2

Mouroux P, Le Brun B (2006) Risk-UE project: an advanced approach to earthquake risk scenarios with application to different European towns. In: Oliveira CS, Roca A, Goula X (eds) Assessing and managing earthquake risk. Springer, Netherlands, pp 479–508. doi:10.1007/978-1-4020-3608-8_23

NCEER (1995) The Hanshin-Awaji earthquake of January 17, 1995: performance of lifelines, technical report NCEER-95-0015 (ed: Shinozuka M), State University of New York, Buffalo

O'Rourke TD, Jeon SS, Toprak S, Cubrinovski M, Jung JK (2012) Underground lifeline system performance during the Canterbury earthquake sequence. In: Proceedings of the 15th world conference on earthquake engineering, Lisbon

Pinto P (2014) Modeling and propagation of uncertainties. In: Pitilakis K, Crowley H, Kaynia AM (eds) SYNER-G: typology definition and fragility functions for physical elements at seismic risk, vol 27, Geotechnical, geological and earthquake engineering. Springer, Dordrecht. doi:10.1007/978-94-007-7872-6_2

Pitilakis K, Crowley E, Kaynia A (eds) (2014a) SYNER-G: typology definition and fragility functions for physical elements at seismic risk, vol 27, Geotechnical, geological and earthquake engineering. Springer, Heidelberg. ISBN 978-94-007-7872-6

Pitilakis K, Franchin P, Khazai B, Wenzel H (eds) (2014b) SYNER-G: systemic seismic vulnerability and risk assessment of complex urban, utility, lifeline systems and critical facilities. Methodology and applications, Geotechnical, geological and earthquake engineering. Springer, Heidelberg. ISBN 978-94-017-8834-2

Rinaldi SM, Peerenboom JP, Kelly TK (2001) Identifying, understanding, and analyzing critical infrastructure interdependencies. IEEE Contr Syst Mag 21(6):11–25

Satumtira G, Duenas-Osorio L (2010) Synthesis of modeling and simulation methods on critical infrastructure interdependencies research. In: Gopalakrishnan K, Peeta S (eds) Sustainable and resilient critical infrastructure systems. Springer, Berlin/Heidelberg. doi:10.1007/978-3-642-11405-2

Seyedi DM, Gehl P, Douglas J, Davenne L, Mezher N, Ghavamian S (2010) Development of seismic fragility surfaces for reinforced concrete buildings by means of nonlinear time-history analysis. Earthq Eng Struct 39:91–108

Tang C, Zhu J, Qi X (2011) Landslide hazard assessment of the 2008 Wenchuan earthquake: a case study in Beichuan. Can Geotech J 48:128–145

Weatherill G, Esposito S, Iervolino I, Franchin P, Cavalieri F (2014) Framework for seismic hazard analysis of spatially distributed systems. In: Pitilakis K et al (eds) SYNER-G: systemic seismic vulnerability and risk assessment of complex urban, utility, lifeline systems and critical facilities. Methodology and applications. Springer, Dordrecht. ISBN 978-94-017-8834-2

S

Seismic Vulnerability Assessment: Masonry Structures

Dina D'Ayala and Viviana Novelli
Department of Civil, Environmental and
Geomatic Engineering, University College
London, London, UK

Introduction

The assessment of the vulnerability of the building stock of an urban center is an essential prerequisite to its seismic risk assessment. The other two ingredients are the expected hazard over given return periods and the distribution and values of the assets constituting the building stock. All three elements of the seismic risk assessment are affected by uncertainties of aleatory nature, related to the spatial variability of the parameters involved in the assessment, and epistemic nature, related to the limited capacity of the models used to capture all aspects of the seismic behavior of buildings and to describe them in simple terms, suitable for this type of analysis. Hence it should always be kept in mind that the computation of a risk level is highly probabilistic and that to accurately represent the risk, the expected values should always be accompanied by a measure of the associated dispersion. A very preliminary estimate of the seismic capacity of the local building stock can be obtained by consulting requirements included in seismic standards and code of practices in force at the time of construction of such buildings. This information together with a temporal and spatial record of the growth of the urban center can provide a definition of classes of buildings assumed to have different capacities. This information can be obtained by looking at past and present cadastral maps with ages of buildings and knowing the historical development and enforcement of codes at the site. In general however for a correct assessment of the seismic risk, a more detailed inventory and classification should be considered, the extent of which is a function of the economic and

technical resources available and of the dimensions and building density of the area under investigation, as well as the diversity within the building stock.

Seismic vulnerability analysis of masonry structures has traditionally been carried out using "empirical" (or statistical) methods. These can be divided in categorization methods, which classify buildings into typologies characterized by propensity to damage, or inspection and rating methods, wherein scores are attributed to each significant characteristic of the buildings thought to have a bearing on its overall vulnerability (UNDP/UNIDO 1985). Several of these methods have been developed in the past 30 years, and the more robust and commonly used are reviewed in section "Empirical Methods."

Recent dramatic increase in computational powers and consequent improvement and refinement in numerical modelling of relatively complex structures, using either static or dynamic approaches (e.g., Fajfar and Gašperšič 1996; Vamvatsikos and Cornell 2005), has facilitated greater exploitation of analytical approaches for vulnerability assessment purposes. Analytical methods need experimental validation of the various parameters used to define the vulnerability, and so far relatively little experimental work on the behavior of masonry structures has been undertaken by the international community, albeit a wide variety of existing unreinforced masonry typologies, both historic and modern, exists. The framework within which analytical methods can be applied and their limitation are presented in section "Analytical Methods."

An attempt to exploit positive aspects of the two classes of methods above is made with the so-called hybrid methods. These combine numerical input/output from analytical models with statistical and probabilistic data to define exposure and vulnerability distribution. This allows to reduce the analytical burden while grounding results in a geographical context. Applicability and reliability of such approaches are constrained by the ability to define certain mechanical and structural characteristics in numerical terms and by the need to define a common approach for treating the various sources of uncertainty

associated with vulnerability, exposure, and hazard. This is further discussed in section "Hybrid Methods."

The review and discussion of existing methods sets out the rationale for a combined method, called FaMIVE (Failure Mechanism Identification and Vulnerability Evaluation), and developed by the first author and her co-researchers, in the past decade. The hypotheses and analytical approach to derive capacity curves and fragility functions are presented in detail in section "Development of Analytical Methods: FaMIVE." Results obtained from the FaMIVE method are presented in section "Damage Thresholds and Performance Indicators" in terms of damage limit states and drift and compared with the recommendation of Eurocode 8 (EC8) (CEN 2005) and experimental evidence to discuss issues of validation and calibration.

Vulnerability Approaches

Empirical Methods

The damage probability matrix (DPM) method has been the most common vulnerability assessment method used since the 1970s. Based on field observations, it expresses in a discrete form the conditional probability of obtaining a damage level D_j, due to a ground motion of macro intensity I_i,

$$P = \left[D = \frac{D_j}{I_i} \right].$$

The method relies on the wealth of observed damage data available from past earthquakes and the correlation of these, on the one hand, with assigned shaking macroseismic intensity at the site of observation and, on the other, with construction materials and methods in different geographical and seismic regions. First proposed by Whitman et al. (1973) after the San Fernando earthquake of 1971, specific applications of this method to masonry structures are numerous (Braga et al. 1982; Corsanego and Petrini 1990; Di Pasquale et al. 2005) and still widely used in developing countries and regions with extensive

historic seismicity records (Askan and Yucemen 2010; Zobin et al. 2010). Notwithstanding its popularity, the DPM has major limitations: discrete definition of damage levels/states and dependence on a specific seismic and architectural context. Hence, it may not be applicable to different geographic locations, in the absence of direct damage data.

Some of these limitations are overcome by the Vulnerability Index Method (VIM) (Benedetti et al. 1988), based on the vulnerability index I_v. The vulnerability index is a summation of weighted parameters, each associated with a constructional or structural feature of the typology, which affects its vulnerability. In this way the definition of vulnerability relates only to the mechanical characteristics of the building, while damage data from past earthquakes are used to calibrate the vulnerability functions, by relating I_v to observed global damage levels for buildings of the same typology and hence extending applicability in regions having experienced the same level of macroseismic intensity or peak ground acceleration (PGA).

The most substantial improvement of the VIM over the DPM is that the former provides a continuous vulnerability function, while the latter uses discrete vulnerability classes expressed in terms of expected damage. This means that possible interventions to shift the vulnerability level of a structural typology are readily quantifiable with I_v, while they are only recorded in the DPM if the intervention results in a shift of damage class. Moreover, the correlation of VIM with an instrumental measure of seismic action allows for application to and comparison between different seismic zones, even though in the last decade the reliability of correlating PGA and damage has been put into question by various authors (Elenas 2000; Wu et al. 2003). Giovinazzi and Lagomarsino (2004) have proposed a direct correlation of the two methods by introducing a definition of damage states and DPM as function of the I_v using the definition of the EMS-98 macroseismic scale (Grünthal 1998). This approach has been further developed in the European RISK-UE project for larger number of building typologies and vulnerability classes (Lagomarsino and Giovinazzi 2006).

The EMS-98 scale (Grünthal 1998) is in itself a DPM approach, where vulnerability classes are defined across different construction typologies. The uncertainty inherent in the attribution of a typology to a class and of its response to a seismic event in terms of damage level is dealt with by assuming a central value and a fuzzy set membership, so that both vulnerability and damage can be expressed in descriptive (few, many, most) rather than numerical terms. A method to compute total probability of damage of a given typology spreading over more vulnerability classes is proposed in Jaiswal et al. (2011).

A first attempt at deriving generic fragility curves was made by Spence et al. (1992) with the Parameterless Scale Intensity (PSI), aimed at providing fragility curves for different building typologies and at quantifying benefits from upgrading and strengthening interventions of the building stock. The advantage of this approach is that damage grades and fragility curves are independent of macroseismic intensity scales, as they are defined relative to each other, on the assumption that for each building type, it is possible to define the level of the scale corresponding to the median of the fragility curve for level of damage D3 (structural damage). Limitations of applicability are however related to the extensive need for observations to calibrate each fragility curve and to the difficulty to correlate observed or expected damage into construction and structural characteristics in a deterministic way.

Analytical Methods

Analytical methods present the advantage of framing the problem of seismic vulnerability of masonry structures in structural engineering terms, defining a direct relationship among construction characteristics, structural response to seismic action, and damage effects.

The development of attenuation equations for specific seismic regions and corresponding derivation of seismic hazard maps in terms of spectral ordinates, as opposed to macroseismic intensity or PGA, has given impetuous to the development of analytical methods. These methods tend to feature more detailed and transparent vulnerability assessment algorithms with direct physical meaning. In the development of analytical fragility functions for masonry structures, two methods can be identified: (i) correlation between damage index and damage thresholds and (ii) correlation between acceleration/displacement capacity curves and spectral demand curves, following the HAZUS99 (FEMA 2001) or N2 method (Fajfar and Gašperšič 1996). These two procedures, in line with the performance-based design for new built, define the seismic assessment of existing building stock and prediction of losses. The required analysis steps are (i) classification of buildings by typology and seismic design, (ii) definition of damage states, (iii) assignment of capacity curves, (iv) definition of demand spectra (associated with return periods and performance targets), and (v) evaluation of building response in terms of performance points. The two procedures, applied to large sets of buildings, allow to derive fragility curves and damage scenarios for given sites and as a function of ground motion parameters which are linked with hazard levels. The damage levels are directly related to the demand parameter, which is expressed in terms of displacement or drift. As the fragility curves are developed on the basis of lognormal distribution, once the typologies have been defined, it is not necessary to have a detailed knowledge of the building stock. All required are the parameters defining the capacity curve for each typology, various damage thresholds, and the number of buildings belonging to each typology. Once capacity curves are developed, then the range of behavior and variability of fragility within a building class or stock can be analyzed by parametric analysis, providing important insight for retrofitting (see D'Ayala 2005). Analytical approaches to define seismic vulnerability of masonry buildings are becoming more and more popular, as improved engineering knowledge on the behavior of masonry structures increases confidence on the reliability of such models. In the past decade, a relatively significant number of procedures aimed at defining reliable analytical vulnerability functions for masonry structures in urban context have been proposed (Lang and Bachmann 2004; Erberik 2008; Borzi

et al. 2008). Although they share similar conceptual hypotheses, they differ substantially in modelling/numerical complexity and treatment of uncertainties. Two main limitations are common to all: limited geographic applicability and limited number of failure mechanisms. Specifically, for the latter, most of the analytical models only consider in-plane or frame-like behavior, disregarding overturning and out-of-plane mechanisms, often occurring at lower levels of shaking and hence substantially affecting the seismic vulnerability in an urban context. Analytical vulnerability assessment methods for existing masonry buildings feature in many risk assessment integrated systems; however, their choices of capacity curves are predetermined. Among nonlinear analysis methods, the collapse-mechanism or limit-equilibrium method presents the advantage of requiring few input parameters and allowing consideration of different failure modes. Procedures using this approach are based on collapse load multipliers which identify the occurrence of different possible mechanisms for given typology and structural characteristics. Bernardini et al. (1990) developed a numerical routine (VULNUS), combined with a fuzzy set theory, in recognition of the limited knowledge associated with many of the parameters. Only one in-plane and one simple out-of-plane mechanism were initially considered as possible failures. D'Ayala and Speranza (2003) further enhanced this approach by developing the FaMIVE method, a mechanical approach based on a suite of 12 possible failure mechanisms directly correlated with in situ observed damage. The FaMIVE algorithm produces vulnerability functions for different building typologies and quantifies the effect of strengthening and repair intervention on reduction of vulnerability. The method has been applied in several locations worldwide and recently in Abruzzo, Italy, following the 2009 L'Aquila earthquake (D'Ayala and Paganoni 2011).

Hybrid Methods

Hybrid methods combine post-earthquake damage statistics with analytically derived nonlinear behavior using pushover analysis (e.g., Kappos et al. 1998; Barbat et al. 1996). Recently, Kappos

et al. (2008) have applied it for estimating direct losses for masonry buildings in Thessaloniki (Greece). This procedure is very useful when the damage data collection of an area of interest with a specific intensity is only partially available. However, the use of different data sources derived from different procedures, for which a direct cross-correlation is not readily available, might not necessarily result in reduced uncertainty of the output.

An application of this approach to masonry structure within the Risk-UE project shows the importance of knowledge of the building stock for reliable seismic vulnerability assessment, as increased level of uncertainties is compounded both on capacity and demand curves (Lagomarsino and Giovinazzi 2006). Application made using code prescription in terms of capacity curves provides a good first approximation but needs extensive in situ calibration as shown by Erberick (2008).

A summary of the methods reviewed above is presented in Table 1. While this is by no means an exhaustive list of the many applications of seismic vulnerability assessment for masonry structures available in literature, it concentrates on procedures that have been specifically developed for vulnerability assessment at territorial scale. For this reason procedures aimed at the assessment of single buildings are not included. The choice of the most suitable procedure is highly dependent on the resources available for the data collection, the computational expertise available, and ultimately the scale and aim of the study. Empirical procedures can be used for fairly large-scale studies to define damage scenarios; however, if the purpose of the study is to identify within a district or urban center specific buildings in need of strengthening, so as to increase their seismic capacity, then a suitable analytical procedure should be preferred.

Development of Analytical Methods: FaMIVE
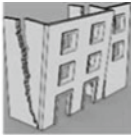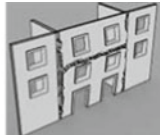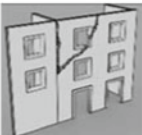
The seismic vulnerability assessment of unreinforced masonry historic buildings can be

Seismic Vulnerability Assessment: Masonry Structures, Table 1 Summary of procedures for the seismic vulnerability assessment of masonry structures

Method	References	Data requirement	Data collection	Assessment type	Approach	Demand input	Output
VULNUS	Bernardini et al. (1990)	Geometry, material parameters, structural details	On-site observation or systematic survey	Analytical	Mechanism method	PGA, response spectrum	Lateral force, lateral drift, vulnerability index
FaMIVE	D'Ayala and Speranza (2003); D'Ayala (2005)	Geometry, material parameters, structural details	On-site observation or systematic survey	Analytical	Capacity spectrum-based method, mechanism	PGA, response spectrum	Lateral force, lateral drift, damage scale, vulnerability index
SP-BELA	Borzi et al. (2008)	Structural description	Random generation	Analytical	Pushover-based method	PGA, response spectrum	Lateral force, lateral drift
CSBM	RISK-UE Kappos et al. (2008)	Structural description	Systematic survey and random generation	Analytical	Capacity spectrum-based method	PGA, response spectrum	Lateral force, lateral drift
VIM	Benedetti et al. (1988); Lagomarsino and Giovinazzi (2006)	Typological description	On-site observation	Empirical	Vulnerability index method	Macroseismic intensity, PGA	Phenomenological damage scale
DPM	Whitman et al. (1973)	Typological description	On-site observation	Empirical	Damage probability matrix	Macroseismic intensity	Phenomenological damage scale
AeDES	Baggio et al. (2009)	Typological description	On-site observation	Empirical	Vulnerability index method	Macroseismic intensity	Phenomenological damage scale
PSI	Spence et al. (1992)	Typological description	Exposure database	Empirical, statistical	Damage probability function	Macroseismic intensity	Phenomenological damage scale
Hybrid Methods	Kappos et al. (1998); Barbat et al. (1996)	Geometry, material parameters	On-site observation	Empirical/ analytical, statistical basis	Hybrid methods	Macroseismic intensity	Damage scale, lateral forces

Seismic Vulnerability Assessment: Masonry Structures,
Fig. 1 Mechanisms for computation of limit lateral capacity of masonry façades

Combined Mechanisms			
B1: façade overturning with one side walls	B2: façade overturning with two side walls	C: overturning with diagonal cracks involving corners	F: overturning constrained by rings beams or ties
In plane Mechanisms			
H1: diagonal cracks mainly in piers	H2: diagonal cracks mainly in spandrel	M1: soft storey due to shear	M2: soft storey due to bending
Out of Plane Mechanism			
A: façade overturning with vertical cracks	D: façade overturning with diagonal crack	E: façade overturning with crack at spandrels	G: façade overturning with diagonal cracks

performed with the Failure Mechanisms Identification and Vulnerability Evaluation (FaMIVE) analytical method (D'Ayala and Speranza 2003; D'Ayala 2005). The FaMIVE method uses a nonlinear pseudo-static structural analysis with a degrading pushover curve to estimate the performance points by way of a variant of the N2 method (Fajfar and Gašperšič 1996), included in EC8 part 3 (CEN 2005). It yields as output collapse load multipliers which identify the occurrence of possible different mechanisms for a given masonry construction typology, given certain structural characteristics.

It is based on computing the collapse load factor for 12 different failure modes, shown in Fig. 1. Each mode of failure corresponds to different constraints and conditions between the façade and the rest of the structure; hence a collapse mechanism can be univocally defined and its collapse load factor computed for each mechanism developing over either part or the whole façade. As shown in the flowchart of Fig. 2, the program FaMIVE first calculates the collapse load factor for each façade in a building and then, taking into account geometric and structural characteristics and constraints, identifies the one which is most likely to occur considering the combination of the largest portion mobilized with the lowest collapse load factor at building level.

The FaMIVE algorithm produces vulnerability functions in terms of ultimate lateral capacity for different building typologies and quantifies the effect of strengthening and repair intervention on reduction of vulnerability. In its current version, it also computes capacity curves and performance points and outputs fragility curves for different seismic scenarios in terms of intermediate and ultimate displacements or ultimate acceleration. Within the FaMIVE database, capacity curves and fragility functions are available for
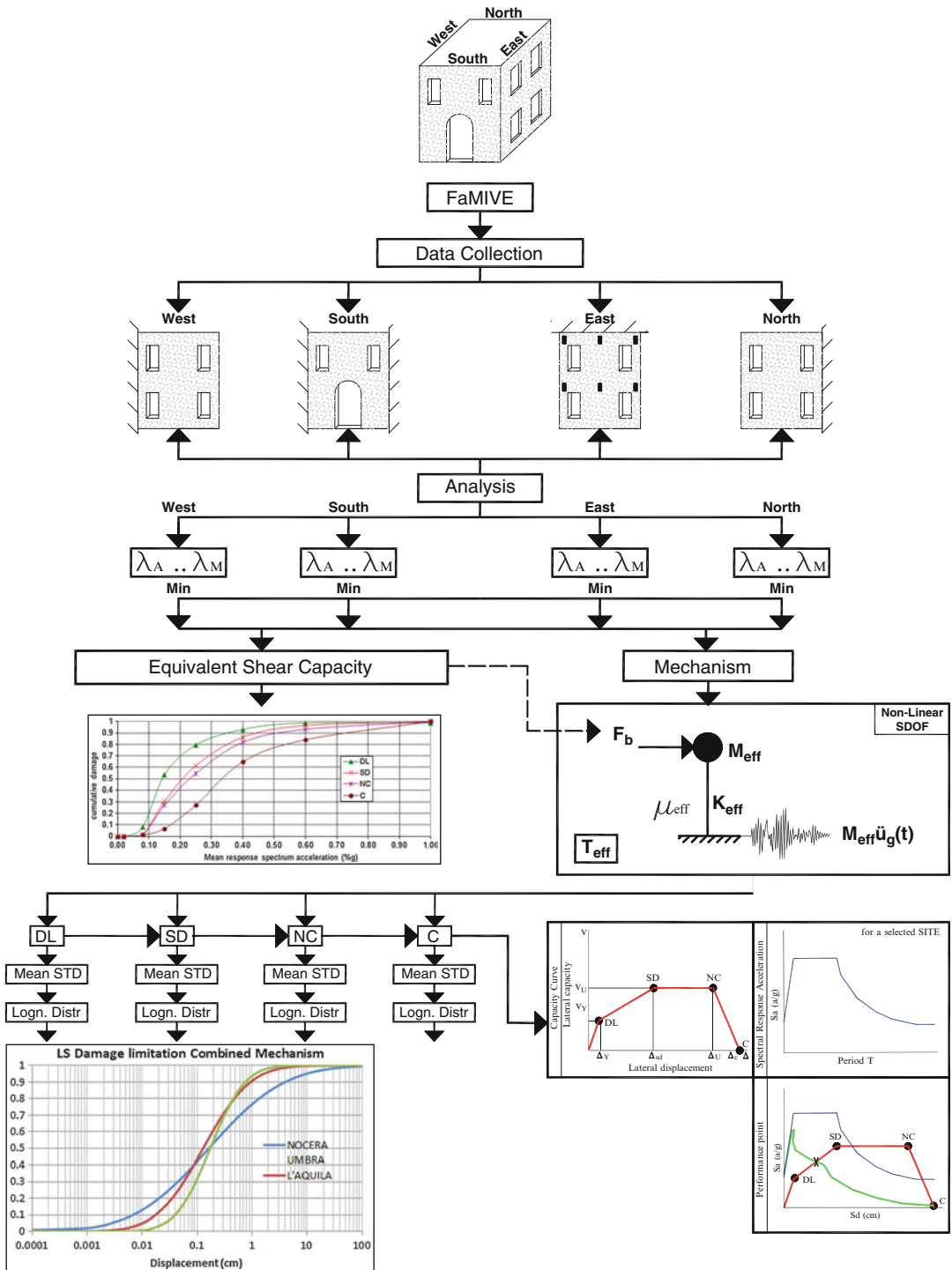
Seismic Vulnerability Assessment: Masonry Structures, Fig. 2 Flowchart setting out the rationale of the FaMIVE procedure

various unreinforced masonry typologies, from adobe to concrete blocks, for a number of reference typologies studied at sites in Italy, Spain, Slovenia, Turkey, Nepal, India, Iran, and Iraq. The procedure has been validated against the EMS-98 vulnerability classes and recently used to produce capacity curves and fragility curves for use in the USGS PAGER environment. The mechanism's characteristics are used to derive an equivalent nonlinear single degree-of-freedom capacity curve to be compared to a spectrum demand curve and eventually define performance points as illustrated in the flow-chart in Fig. 2.

In order to derive fragility curves, limit state performance criteria correlated to damage states need to be defined. This step is fraught with uncertainties, as very limited consolidated evidence exists to perform such correlation over a wide range of building typologies and shaking levels. While robust database of damage states exist in literature, no attempt has been so far made to record permanent drift and corresponding ground shaking in a consistent way so as to provide empirical evidence for capacity curves. As an alternative, a number of authors have worked on correlating performance indicators and damage indicators on experimentally obtained capacity curves, by way of shaking table tests or pushover tests.

Definition of Capacity Curve

Capacity curves can be derived for each façade on the basis of the following steps. The first step is to calculate the lateral effective stiffness for each wall and its tributary mass. As the floor and roof structures in traditional masonry buildings have relatively modest in-plane stiffness, it is assumed that redistribution of inertia forces among walls will be ineffective and torsional effects are not considered. The effective stiffness of each external wall is calculated on the basis of the type of mechanism attained, the geometry of the wall and layout of opening, the constraints to other walls and floors, and the portion of other walls involved in the mechanism:

$$K_{\text{eff}} = k_1 \frac{E_t I_{\text{eff}}}{H_{\text{eff}}^3} + k_2 \frac{E_t A_{\text{eff}}}{H_{\text{eff}}} \qquad (1)$$

where H_{eff} is the height of the portion involved in the mechanism; E_t is the estimated modulus of the masonry as it can be obtained from experimental literature for different masonry typologies; I_{eff} and A_{eff} are the second moment of area and the cross-sectional area, respectively, calculated taking into account extent and position of openings and variation of thickness over height; and k_1 and k_2 are constants which can assume values equal or greater than 0, depending on edge constraints and whether shear and/or flexural stiffness are relevant for the specific mechanism. For instance, for in-plane mechanisms, the inclusion of shear stiffness in Eq. 1 will depend on the slenderness of the pier and the relative stiffness between piers and spandrels of the façade being assessed.

The tributary mass Ω_{eff} is calculated following the same approach, and it includes the portion of the elevation activated by the mechanisms plus the mass of the horizontal structures involved in the mechanism:

$$\Omega_{\text{eff}} = V_{\text{eff}} \delta_m + \Omega_f + \Omega_r \qquad (2)$$

where V_{eff} is the solid volume of the portion of wall involved in the mechanism, δ_m is the density of the masonry Ω_f, Ω_r are the masses of the horizontal structures involved in the mechanism. Effective mass and effective stiffness are used to calculate a natural period T_{eff}, which characterizes an equivalent single degree-of-freedom (SDoF) oscillator:

$$T_{\text{eff}} = 2\pi \sqrt{\frac{\Omega_{\text{eff}}}{K_{\text{eff}}}} \qquad (3)$$

The mass is applied at the height of the center of gravity of the collapsing portion with respect to the ground, and a linear acceleration distribution over the wall height is assumed. The elastic limit acceleration S_{ay} is also computed depending on the failure mechanism identified; for failure

mechanisms involving flexural strain limit (assuming no tensile capacity in the material), for instance, S_{ay} will be the value of lateral acceleration that, combined with gravitational load resultant, will cause a triangular distribution of compression stresses at the base of the overturning portion, just before the onset of cracking:

$$S_{ay} = \frac{t_b}{6h_0}g \text{ with corresponding displacement}$$
$$S_{de} = \frac{S_{ay}}{4\pi^2}T_{\text{eff}}^2$$

$$(4)$$

where t_b is the effective thickness of the wall at the base of the overturning portion, h_0 is the height to the ground of the center of mass of the overturning portion, and T_{eff} is the natural period of the equivalent single degree-of-freedom (SDoF) oscillator. The maximum lateral capacity S_{au} is defined as

$$S_{au} = \frac{\lambda_c}{\alpha_1}g \qquad (5)$$

where λ_c is the collapse load factor of the mechanism chosen, calculated by FaMIVE, and α_1 is the proportion of total mass participating to the mechanism. This is calculated as the ratio of the mass of the façade and sides or internal walls and floor involved in the mechanism Ω_{eff} to the total mass of the involved macroelements (walls, portion of floors, and portion of roof supported by the wall). The displacement corresponding to the damage threshold of structural damage, S_{ds} is

$$3S_{dy} \leq S_{ds} \leq 6S_{dy} \qquad (6)$$

as suggested by (Tomazevic 2007). The range in Eq. 6 is useful to characterize masonry fabric of diverse levels of regularity and its integrity at ultimate conditions, with the lower bound better describing the behavior of adobe, rubblestone, and brickwork in mud mortar, while the upper bound can be used for massive stone, brickwork set in lime or cement mortar, and concrete blockwork. The near collapse condition is determined by the displacement S_{du} identified by the condition of loss of vertical equilibrium which, for overturning mechanisms, can be computed as a lateral displacement at the top or for in-plane mechanism by the loss of overlap of two units in successive courses as quantifiable with one of the two alternative expressions in Eq. 7:

$$S_{du} \geq t_b/3 \quad \text{or} \quad S_{du} \geq l_u/2 \qquad (7)$$

where t_b is the thickness at the base of the overturning portion and l_u is the typical length of units forming the wall. In the case of in-plane mechanism, the geometric parameter used for the elastic limit is, rather than the wall thickness, the width of the slender pier.

The threshold points identified by Eqs. 4, 5, 6, and 7 can be associated to corresponding states of damage. Specifically **DL**, *damage limitation*, corresponds to the elastic lateral capacity threshold (S_{de}, S_{ay}) defined by Eq. 4; **SD**, *significant damage*, corresponds to the peak capacity threshold (S_{ds}, S_{au}) defined by Eqs. 5 and 6; and **NC**, *near collapse*, corresponds to incipient or partial collapse threshold (S_{du}, S_{au}) defined by either conditions of Eq. 7.

The procedure's approach allows a direct analysis of the influence of different parameters on the resulting capacity curves, whether these are geometrical, mechanical, or structural. By way of example, Figs. 3 and 4 show a comparison of average capacity curves obtained by grouping the results using different criteria for the same sample of buildings in L'Aquila. In Fig. 3 the average curves represent the behavior according to the classes of mechanisms as per Fig. 1. In Fig. 4 the capacity curves are obtained by considering different structural typologies, as classified by the WHE-PAGER project (Jaiswal et al. 2011) and shown in Table 2. It can be seen

Seismic Vulnerability Assessment: Masonry Structures,
Fig. 3 Average capacity curves for sample grouped by collapse-mechanism classes
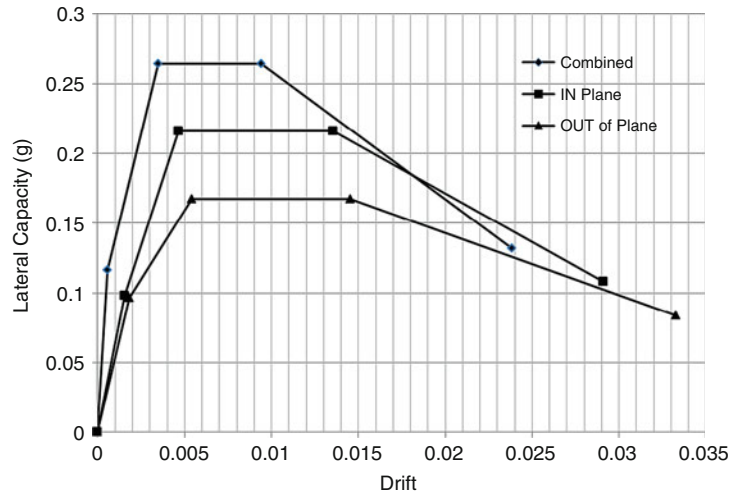
Seismic Vulnerability Assessment: Masonry Structures,
Fig. 4 Average capacity curves for sample grouped by structural typology
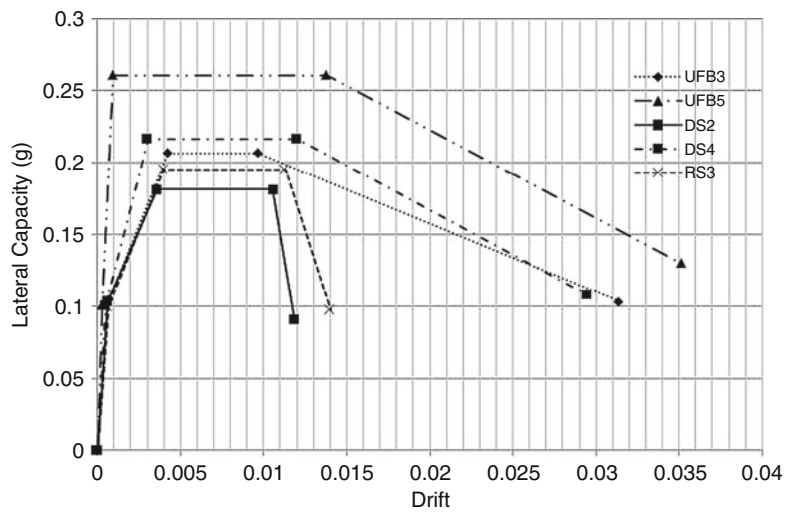
that the correlation between mode of failure and structural typology is qualitatively good but not univocal, and the grouping affects both ultimate lateral capacity and drift at all damage states.

Such results bring in sharper focus the limitation and inaccuracy of using idealized models and average curves without adequately considering the inherent aleatoric variation associated with any given site where the assessment is conducted and the importance of a detailed knowledge of the local construction characteristics when sampling the building representative of the building stock. The lack of such knowledge determines a non-negligible epistemic error.

In Figs. 3 and 4 the capacity curves show a fourth point, which represents the collapse of the façade corresponding to a lateral displacement S_{dc} and a lateral acceleration S_{ac} as computed in Eq. 8:

$$S_{dc} \geq t_b/2 \text{ or } S_{dc} \geq l_u, \quad S_{ac} = \frac{S_{au}}{2} \quad (8)$$

Such relationship are estimated on the basis of experimental work available in literature.

Seismic Vulnerability Assessment: Masonry Structures, Table 2 Structural typologies classification according to PAGER (Jaiswal et al. 2011)

Load-bearing material	PAGER code	Description
Stone masonry	RS3	Local fieldstones with lime mortar
	RS4	Local field stones with cement mortar, vaulted brick roof, and floors
	DS2	Rectangular cut stone masonry block with lime mortar
	DS3	Rectangular cut stone masonry block with cement mortar
	DS4	Rectangular cut stone masonry block with reinforced concrete floors and roof
	MS	Massive stone masonry in lime or cement mortar
Brickwork or blockwork	UFB1	Unreinforced brick masonry in mud mortar without timber posts
	UFB3	Unreinforced brick masonry in lime mortar. Timber flooring
	UFB4	Unreinforced fired brick masonry, cement mortar. Timber flooring
	UFB5	Unreinforced fired brick masonry, cement mortar, but with reinforced concrete floor and roof slabs
	UCB	Unreinforced concrete block masonry with lime or cement mortar

Damage Thresholds and Performance Indicators

In the previous section capacity curves have been determined as polylines characterized by four notable points which represent as many damage thresholds. A method for assessing the overall behavior by use of a global performance indicator is the computation of the performance point. In order to calculate the performance point, it is necessary to intersect the capacity curve derived above with the demand spectra for different return periods in relation to the performance criteria considered. Two broadly equivalent approaches for the derivation of the nonlinear demand spectra exist: the N2 method (Fajfar 1999) included in the EC8 and the capacity spectrum method (CSP) (FEMA 2001). The two methods differ essentially in the way the nonlinear demand spectrum is arrived at: the N2 method uses a reduction factor R, function of the structure expected ductility μ, while the CSP uses a fictitious damping factor derived from the hysteresis loop of the structure. In the following the N2 method will be used to illustrate the derivation of performance points.

The original version of the N2 method was derived for elastic-perfectly plastic SDoF oscillator, and the relationship between the reduction factor R and the ductility μ, used to determine the nonlinear response spectrum, is dependent on the bilinear shape of the capacity curve. Alternative relationships between the reduction factor, the ductility, and the natural period of the system, R-μ-T, have been developed to determine the inelastic displacement ratios and the inelastic response spectra for structures characterized by quadri-linear backbone curves, characterized by a plateau of reduced capacity and further ductility, after a negative slope branch (see Miranda 2001; Dolsek and Fajfar 2004; Vamvatsikos and Cornell 2005).

In the case of the capacity curves derived with the FaMIVE procedure however, as mentioned above, the behavior beyond the near collapse performance point is fraught with uncertainty and hence the descending branch of the curve is ignored, and the simplified relationship for bilinear curves is used.

To calculate the coordinates of the performance point in the displacement-acceleration space, the intersection of the capacity curve with the nonlinear demand spectrum for an appropriate level of ductility μ can be determined as shown in Eq. 8, given the value of ultimate lateral capacity of the equivalent SDoF oscillator, S_{au}:

$$\text{if } T^* < T_c$$
$$\text{if } S_{au} \geq S_a(T^*) \quad \Rightarrow S_d^*(\mu) = \frac{T_c^2(S_{ae}(T^*) - S_a(T^*))^2}{(\mu - 1)^2} * \frac{g\mu}{4\pi^2 S_a(T^*)}$$
$$\text{if } S_a(T_c) < S_{au} < S_a(T^*) \Rightarrow S_d^*(\mu) = \frac{T_c^2(S_{ae}(T^*) - S_{au})^2}{(\mu - 1)^2} * \frac{g\mu}{4\pi^2 S_{au}}$$
$$\text{if } S_{au} \leq S_a(T_c) \qquad \Rightarrow S_d^*(\mu) = \frac{gT_c^2(S_{ae}(T^*))^2}{4\pi^2 \mu S_{au}}$$
$$\text{if } T^* \geq T_c$$
$$\text{if } S_{au} \geq S_a(T^*) \qquad \Rightarrow S_d^*(\mu) = \frac{gT_c^2(S_{ae}(T^*))^2}{4\pi^2 \mu S_a(T^*)}$$
$$\text{if } S_{au} < S_a(T^*) \qquad \Rightarrow S_d^*(\mu) = \frac{gT_c^2(S_{ae}(T^*))^2}{4\pi^2 \mu S_{au}}$$

$$(9)$$

where two different formulations are provided for values of ultimate lateral capacity S_{au} greater or smaller than the nonlinear spectral acceleration $S_a(T_C)$ associated with the corner period T_C marking the transition from constant acceleration to constant velocity section of the parent elastic spectrum. Figure 5 shows an example of calculation of performance points for median capacity curves for some different types of mechanisms (see Fig. 1).

Derivation of Fragility Functions

Once capacity curves and limit state points are identified on them, fragility curves for different limit states are obtained by computing the median and standard deviation values of the performance point displacements for each index building in a given set of buildings surveyed on-site or created through randomization of the input parameters. The median and standard deviation are used to derive equivalent lognormal distributions. To this end the median displacement for each limit state can be calculated as:

$$\hat{S}_{ds_i} = e^{\overline{\mu}_i} \text{ with } \overline{\mu}_i = \frac{1}{n}\sum_{j=1}^{n} \ln S_{ds_i}^j \qquad (10)$$

and the corresponding standard deviation as:

$$\beta_{ds_i} = e^{\overline{\mu}_i + \frac{1}{2}\sigma_i^2}\sqrt{e^{\sigma_i^2} - 1} \text{ with } \sigma_i$$
$$= \sqrt{\frac{\sum_{j=1}^{n}\left(\ln S_{ds_i}^j - \ln \overline{S}_{ds_i}^j\right)^2}{n}} \qquad (11)$$

Figures 6 and 7 show the set of fragility curves obtained for the damage limit states of damage limitation and significant damage as computed for the two Italian sites of Nocera Umbra and Serravalle considering separately the three types of structural behavior. Once a structural typology has been assigned, the values of the mechanical characteristics are the same across the two samples, while the structural details are accounted for directly in the three classes of mechanisms. Hence the variability observed between samples in each chart can be related directly to geometric differences and masonry fabric, i.e., to the local aspects of the construction practice and architectural layout. Curves on the left of the diagrams relate to corresponding capacity curves with a stiffer elastic branch in the case of damage limitation and with lower ductility in the case of significant damage. However the distribution does not bare consistency across the three classes of mechanism for the two sites.

Conducting Seismic Vulnerability Assessment

As seen in the previous section, an analytical method based on limit state analysis and mechanism approach has the advantage of employing specific structural models to assess the seismic behavior of masonry buildings of different typologies, represented through capacity curves and performance points, while at the same time

Fig. 5 Example of
calculation of performance
points for median capacity
curves for types of
mechanisms triggered
(*COMB* combined
mechanism, *IP* in-plane
mechanism, *OOP* out-of-
plane mechanism; see also
Fig. 1)
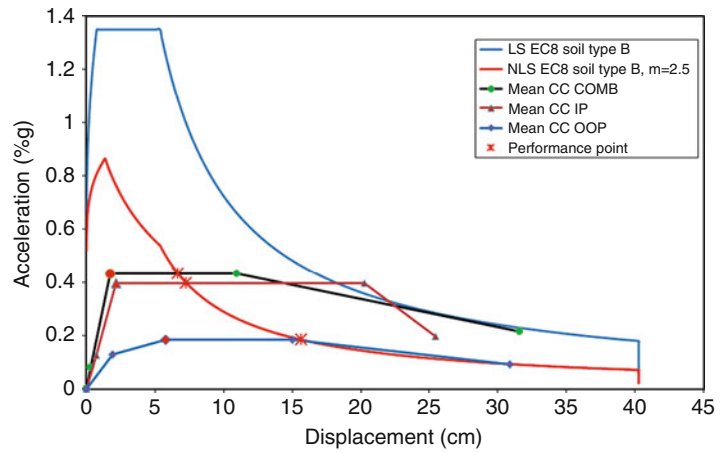
Fig. 6 Fragility curves for
limit state of damage
limitation for the three
classes of collapse
mechanisms for two
different Italian sites
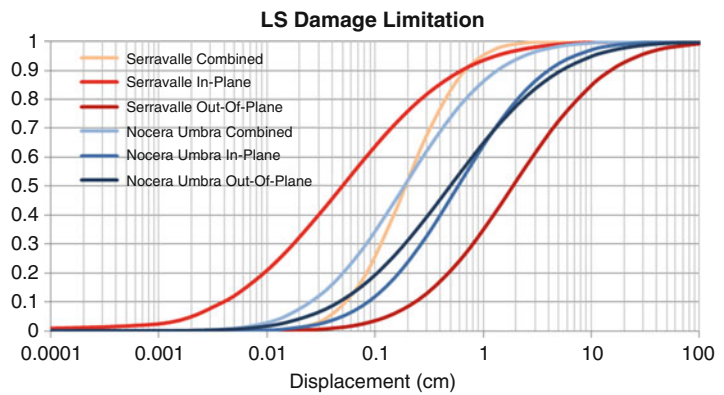
Fig. 7 Fragility curves for
limit state of significant
damage for the three classes
of collapse mechanisms for
two different Italian sites
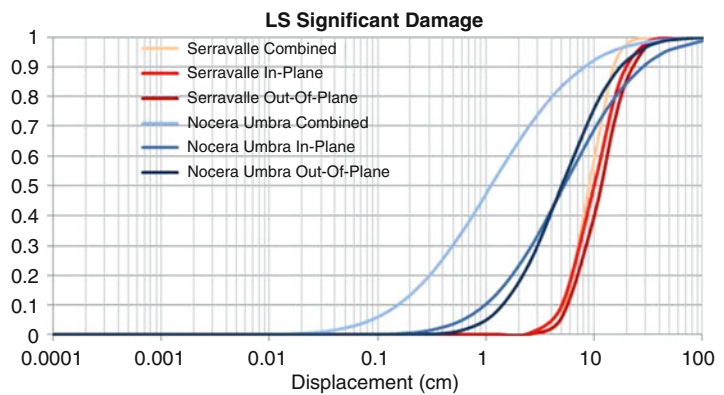
being applicable to relatively large numbers of
buildings. Hence procedures of this type, if prop-
erly developed, are particularly suitable for the
seismic assessment of building aggregates in his-
torical centers (Novelli and D'Ayala 2014) or

when a group of heritage structures at territorial
scale are considered (D'Ayala and Ansal 2012).

A robust approach for seismic vulnerability
assessment at territorial scale needs to address a
number of issues, mainly related to a correct

Seismic Vulnerability Assessment: Masonry Structures, Fig. 8 Fragility curves for two different samples – (**a**) 30 buildings and (**b**) 200 buildings – represented as actual distribution and lognormal distribution (Eqs. 10 and 11)
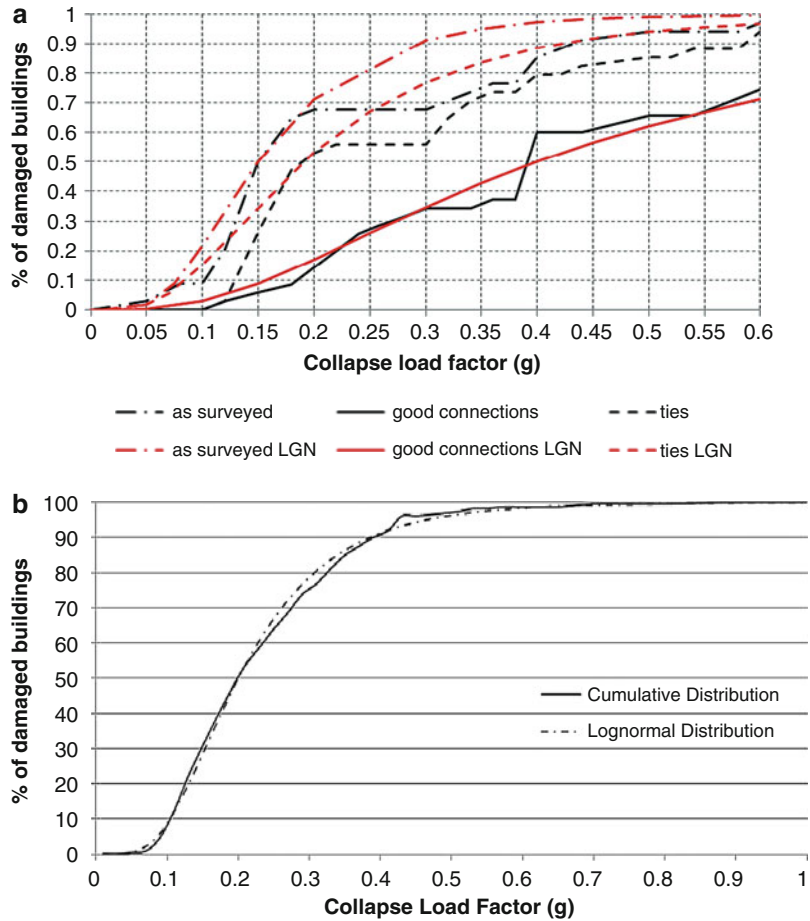
representation of the existing exposure so as to obtain meaningful fragility and vulnerability functions, while explicitly accounting for the uncertainties related to modelling limitation, the inherent randomness of the sample, and the randomness of the response. Given a territorial distribution, uncertainties are also present about the seismic action, relating to record to record, attenuation, soil characteristics, etc. Hence specific attention and due consideration should be given to the following aspects of a process leading to a seismic vulnerability assessment at territorial scale and identification of rehabilitation strategies.

Such process can be articulated in 10 steps:

• Consideration on the urban fabric and study area

• Identification of typologies and construction details
• Size and heterogeneity of sample to be surveyed
• Data collection
• Data analysis
• Safety and conservation requirements
• Seismic hazard
• Definition of performance points
• Definition of fragility functions
• Rehabilitation decisions

A detailed discussion of each step is beyond the scope of this document. However some observations in relation to specific applications of FaMIVE can help in highlighting some of the issues.

In relation to the sample size and heterogeneity, for instance, Fig. 8a, b shows the resulting

distributions for two different cases. The first sample included 30 buildings, which were considered relatively homogenous in terms of structural typology and materials but differing in geometry and quality of connections. The second one is made up of more than 200 buildings of varied characteristics. As it can be seen, the assumption of a lognormal model to describe the fragility cumulative distribution is more reliable as the size and heterogeneity of the sample increases.

Figure 8a also shows that results are highly affected by data collection and data analysis: the three cumulative distributions are obtained for the same sample of buildings when the presence of anchors or good wall connection is in turn ignored or considered. Because the collection and interpretation of this type of data is very sensitive to the expertise of the surveyor and the analyst, results can be highly affected by either incorrect interpretation of this information on the basis of the survey or use of a modelling tool that cannot properly simulate the behavior of the buildings when these constraints are in place.

The requirement of the codes and the limitations imposed by conservation principles also affect the way in which the fragility functions can be derived and used.

The reliability of the results obtained in the previous section can be considered within the framework set out in the EC 8 (CEN 2005), whereby the reliability associated to the results of a seismic assessment of a structure is expressed as a function of the level of knowledge and quantified by means of the confidence factor. Hence this can be considered a measure of the epistemic uncertainty. EC 8 (CEN 2005) recognizes three levels of knowledge (limited, normal, and full) and three fields of knowledge (geometry, construction details, and materials). As data used in the FaMIVE approach are collected by on-site visual inspection with some measurement and in situ accurate observation of construction details, while only very limited in situ nondestructive test on materials are performed and material characteristics are otherwise assigned based on literature or surveyor experience, then the level of knowledge is superior to KL1, *limited*, but not quite equal to KL2, *normal*. For this level of knowledge, a static nonlinear analysis, such as the limit state mechanism approach, leading to a capacity curve is deemed appropriate. Hence according to the recommended values, the confidence factor CF should be in the range 1.2 to 1.35 depending on how closer the actual knowledge can be considered to the reference level identified by KL2. The confidence factor is then used in EC8 to reduce the capacity values as obtained from the assessment.

The FaMIVE procedure uses a measure of uncertainty of the input data to determine the uncertainty of the output. Depending on whether data, in each section of the data collection form, has been collected and measured directly on-site, or collected on-site and confirmed by existing drawings or photograph, or collected from photographic evidence only, three levels of uncertainty are considered, as high, medium, and low, respectively, to which three confidence ranges of the value given for a parameter can be considered corresponding to 10 % variation, 20 % variation, and 30 % variation. The parameter value attributed during the survey is considered central to the confidence range so that the interval of existence of each parameter is defined as $\mu \pm 5\,\%$, $\mu \pm 10\,\%$, and $\mu \pm 15\,\%$, depending on highest or lowest uncertainty. The uncertainty applied to the output parameters, specifically lateral acceleration and limit states' displacement, is calculated as a weighted average of the uncertainty of each section of the data form, with minimum 5 % confidence range to maximum 15 % confidence range.

In Fig. 9 the cumulative distribution for drift at the structural damage state threshold is compared with the curve obtained assuming a confidence factor of 1.25 in agreement with EC8 knowledge level and with the confidence boundaries computed by FaMIVE according to the procedure outlined above.

The major limitation of using confidence factors in the way suggested by EC8 is the possible underestimation of the overall capacity of the sample and hence an appropriate representation of epistemic uncertainty.

Fig. 9 Cumulative
distribution for drift at
structural damage
threshold: comparison with
curve obtained for an EC8
confidence factor of 1.25
and with FaMIVE
confidence boundary
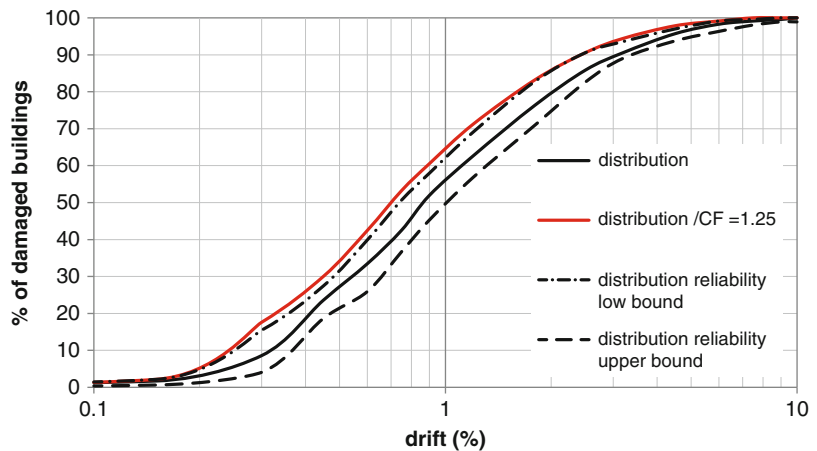
FaMIVE can also be used to foresee the shift in seismic performance of the building stock of a historic center if certain interventions are implemented to improve the seismic behavior. Among possible interventions, the improvement of cohesion of the masonry through grouting of multi-leaf walls with lime-based grout and the introduction of ties connecting walls together are seen as meeting the conservation principles of like for like, minimum intervention, and, in the case of ties, reversibility. Figure 10a, b shows the change in the median capacity curves when these interventions are considered: it can be seen that the response is different depending on the typology; in general both interventions cause an increase in lateral capacity. In some cases also a change in ductility and/or in the behavior up to structural damage as the intervention causes a change in the prevalent failure mechanisms.

Summary

Traditionally seismic vulnerability of historic buildings has been evaluated using empirical methods based either on damage observation or construction data. In the last decade, analytical methods for assessment of masonry structures have increased in number and quality. However, most of them assume only in-plane failure behavior reducing the walls to frame-like behavior. This approach has serious limitations and it is nonconservative, as it overlooks out-of-plane behavior which might occur for structures with insufficient out-of-plane wall capacity. Nevertheless, recommendations of the EC 8 (CEN 2005) only consider basic in-plane behavior, either controlled by flexure or controlled by shear, and set limit drifts for this performance criterion. The application of the FaMIVE procedure to a number of cases in regions of different seismicities shows the relevance of construction details and the importance of considering separately different failure mechanisms to identify realistic risks of failure. A procedure for the computation of the performance point given performance criteria and hazard return period is introduced to show the variability of performance points for different failure behaviors and how they affect probability of failure computed as cumulative. The random uncertainty associated with this type of analysis is presented in terms of fragility curves that can be modelled as lognormal distribution. In the same way, the epistemic uncertainty associated with the modelling procedure is also evaluated using the same approach. Results show that the reliability approach set out in FaMIVE correctly accounts for epistemic uncertainty and it is of the same order of magnitude as the confidence factor suggested by EC8, although this only accounts for negative effects.

Seismic Vulnerability Assessment: Masonry Structures,
Fig. 10 Comparison between the capacity curves for (*DS2*), (*DS4*), (*UFB3*), and (*RS4*) for (**a**) the condition of cohesion, c = 0 and c = 0.25 N/mm^2, and (**b**) with the addition of ties to the condition of cohesion c = 0.25 N/mm^2
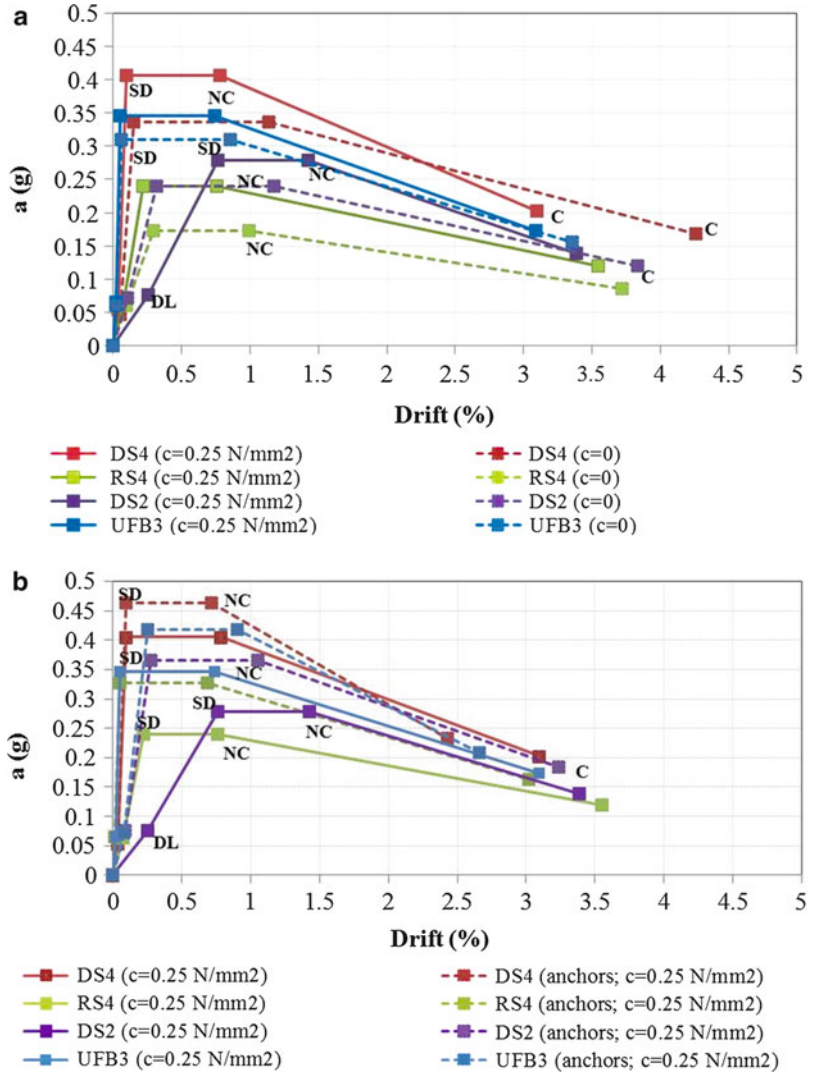
Cross-References

References

Askan A, Yucemen MS (2010) Probabilistic methods for the estimation of potential seismic damage: application to reinforced concrete buildings in Turkey. Struct Saf 32(4):262–271

Baggio C, Bernardini A, Colozza R, Corazza L, Della Bella M, Di Pasquale G, Dolce M, Goretti A, Martinelli A (2009) Manuale per la compilazione della scheda di 1° livello di rilevamento danno, pronto intervento e agibilità per edifici ordinari nell'emergenza post-sismica (AeDES). Editrice Italiani nel Mondo srl, Roma

Barbat AH, Yépez Moya F, Canas JA (1996) Damage scenarios simulation for seismic risk assessment in urban zones. Earthq Spectra 12(3):371–394

Benedetti D, Benzoni G, Parisi MA (1988) Seismic vulnerability and risk evaluation for old urban nuclei. Earthq Eng Struct Dyn 16:183–201

Bernardini A, Gori R, Modena C (1990) An application of coupled analytical models and experiential knowledge for seismic vulnerability analyses of masonry buildings. In: Koridze A (ed) Engineering aspects of earthquake phenomena, vol 3. Omega Scientific, Oxon, pp 161–180

Borzi B, Crowley H, Pinho R (2008) Simplified pushover-based earthquake loss assessment (SP-BELA) method for masonry buildings. Int J Architect Herit 2(4):353–376

Braga F, Dolce M, Liberatore D (1982) A statistical study on damaged buildings and ensuing review of the MSK −76 SCALE. In: Proceedings of the 7th European conference on earthquake engineering, Athens

CEN – EN 1998-3 (2005) Eurocode 8: design of structures for earthquake resistance, Part 3: strengthening and repair of buildings. Comitee' Europeen de Normalisation, Brussels

Corsanego A, Petrini V (1990) Seismic vulnerability of buildings – work in progress. In: Proceedings of the workshop II on seismic risk vulnerability and risk assessment, Trieste, pp 577–598

D'Ayala D (2005) Force and displacement based vulnerability assessment for traditional buildings. Bull Earthq Eng 3:235–265, Springer

D'Ayala D, Ansal A (2012) Nonlinear push over assessment of heritage buildings in Istanbul to define seismic risk. Bull Earthq Eng 10(1):285–306

D'Ayala D, Paganoni S (2011) Assessment and analysis of damage in L'Aquila historic city centre after 6th April 2009. Bull Earthq Eng 9(1):81–104

D'Ayala D, Speranza E (2003) Definition of collapse mechanisms and seismic vulnerability of historic masonry buildings. Earthq Spectra 19:479–509

Di Pasquale G, Orsini G, Romeo RW (2005) New developments in seismic risk assessment in Italy. Bull Earthq Eng 3(1):101–128

Dolsek M, Fajfar P (2004) Inelastic spectra for infilled reinforced concrete frames. Earthq Eng Struct Dyn 33:1395–1416

Elenas A (2000) Correlation between seismic acceleration parameters and overall structural damage indices of buildings. Soil Dyn Earthq Eng 20(1–4):93–100

Erberik MA (2008) Generation of fragility curves for Turkish masonry buildings considering in-plane failure modes. Earthq Eng Struct Dyn 37:387–405

Fajfar P (1999) Capacity spectrum method based on inelastic demand spectra. Earthq Eng Struct Dyn 28(9):979–993

Fajfar P, Gašperšič P (1996) The N2 method for the seismic damage analysis of R.C. buildings. Earthq Eng Struct Dyn 25(1):31–46

Federal Emergency Management Agency (FEMA) (2001) HAZUS99 technical manual. Service Release 2. FEMA, Washington, DC

Giovinazzi S, Lagomarsino S (2004) A macroseismic method for the vulnerability assessment of buildings. In: Proceedings of the 13th world conference on earthquake engineering, Vancouver, Paper No. 896

Grünthal G (ed) (1998) Cahiers du Centre Européen de Géodynamique et de Séismologie: volume 15 – European macroseismic scale 1998. European Center for Geodynamics and Seismology, Luxembourg

Jaiswal K, Wald D, D'Ayala D (2011) Developing empirical collapse fragility functions for global building types. Earthq Spectra 27:775

Kappos AJ, Stylianidis KC, Pitilakis K (1998) Development of seismic risk scenarios based on a hybrid method of vulnerability assessment. Nat Hazards 17(2):177–192

Kappos AJ, Panagopoulos G, Penelis GG (2008) Development of a seismic damage and loss scenario for contemporary and historical buildings in Thessaloniki, Greece. Soil Dyn Earthq Eng 28(10–11):836–850

Lagomarsino S, Giovinazzi S (2006) Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings. Bull Earthq Eng 4(4):415–443

Lang K, Bachmann H (2004) On the seismic vulnerability of existing buildings: a case study of the city of Basel. Earthq Spectra 20(1):43–66

Miranda E (2001) Estimation of inelastic deformation demands of SDOF systems. J Struct Eng (ASCE) 127(9):1005–1012

Novelli V, D'Ayala DF (2014) The PERPETUATE procedure for the seismic vulnerability assessment at territorial scale. Application to the Casbah of Algiers. Bull Earthq Eng 12:1–24

Spence R, Coburn AW, Pomonis A (1992) Correlation of ground motion with building damage: the definition of a new damage-based seismic intensity scale. In: Proceedings of the tenth world conference on earthquake engineering, Madrid, vol 1, pp 551–556

Tomaževič M (2007) Damage as a measure for earthquake-resistant design of masonry structures: Slovenian experience. Can J Civil Eng 34(11):1403–1412

UNDP/UNIDO Project RER/79/015 (1985) Post-earthquake damage evaluation and strength assessment of buildings under seismic condition, vol 4. UNDP, Vienna

Vamvatsikos D, Cornell C (2005) Direct estimation of seismic demand and capacity of multidegree-of-freedom systems through incremental dynamic

S

analysis of single degree of freedom approximation. J Struct Eng 131(4):589–599

Whitman RV, Reed JW, Hong ST (1973) Earthquake damage probability matrices. In: Proceedings of the 5th world conference on earthquake engineering, Rome, Italy, pp. 2531–2540

Wu Y-M, Teng T-l et al (2003) Relationship between peak ground acceleration, peak ground velocity, and intensity in Taiwan. Bull Seismol Soc Am 93(1):386–396

Zobin VM, Cruz-Bravo AA, Ventura-Ramırez F (2010) Microzonation of seismic risk in a low-rise Latin American city, based on the macroseismic evaluation of the vulnerability of residential buildings: Colima city, Mexico. Nat Hazard Earthq Syst Sci 10:1347–1358

Seismic Vulnerability Assessment: Reinforced Concrete Structures

Flavia De Luca[1] and Gerardo M. Verderame[2]
[1]Department of Civil Engineering, University of Bristol, Bristol, UK
[2]Department of Structures for Engineering and Architecture (DiSt), University of Naples Federico II, Naples, Italy

Synonyms

Axial load failure; Damage measure; Flexure; Nonlinear modeling; PBEE; RC structures; Shear

Introduction

Seismic vulnerability can be defined as the degree of loss to a given element at risk (e.g., buildings) resulting from the occurrence of an earthquake event (Coburn and Spence 2002). The development of methodologies for seismic vulnerability assessment is an essential tool for seismic risk management and for prioritizing pre-earthquake strengthening of the built environment.

Seismic risk can be approached through different methodologies. Most of them aim to deconstruct the problem into the typical elements defining any kind of risk: (i) hazard,

(ii) vulnerability, and (iii) exposure, and they are based on the total probability theorem.

Available seismic vulnerability methods, i.e., empirical, analytical, or hybrid (e.g., Calvi et al. 2006), developed in the last 30 years, differ because of the nature of tools and data employed. In the following only analytical approach to seismic vulnerability is considered, given the significance that such approach had in the last years. Analytical vulnerability methods stand at the basis of current worldwide codes and guidelines for seismic design and assessment of structures.

Therefore, next-generation codes (e.g., FEMA P-58-1 2012) are proceeding toward an explicit quantification of seismic risk through the Performance-Based Earthquake Engineering (PBEE) framework. Section "Performance Earthquake Engineering Framework" of this entry describes how in PBEE seismic risk assessment is deconstructed through the total probability theorem. PBEE's deconstruction of seismic risk in hazard, structural, damage, and loss analysis emphasizes how structural and damage analyses are strictly related to vulnerability assessment in strict sense.

Notwithstanding the fact that PBEE is a generalized methodology that can be applied for any kind of structure, the case of reinforced concrete (RC) structures is specifically addressed in the following. As far as the relationship between intensity measure and engineering demand parameter arises in PBEE, it is the phase in which it is necessary to focus on the nonlinear behavior of RC structural elements (section "Behavior of RC Elements and Non Structural Elements"), in order to approach analytical modeling (section "Analytical Modeling of RC Structures"), and qualify and quantify components damage measures (section "Damage Measures"). A review of experimental and analytical results collected in the last decades is provided herein. It aims at the quantification of the behavior of RC structures and, in turn, of RC elements. Most significant results on vulnerability assessment of RC structures are collected together, and they allow emphasizing future challenges and needs (section "Future Challenges and

Concluding Remarks") that scientific community and practitioners should take on in the next years.

The focus in the following is the quantification of nonlinear parameters of RC elements aimed at a reliable modeling of their behavior. Emphasis is also given to consolidated and more recent literature results that have been included in codes and guidelines.

Performance Earthquake Engineering Framework

"PBEE implies design, evaluation, construction, monitoring the function and the maintenance of engineered facilities whose performance under common and extreme loads respond to the diverse needs and objectives of owners-users and society. It is based on the premise that performance can be predicted and evaluated with quantifiable confidence to make, together with the client, intelligent and informed trade-offs based on life cycle considerations rather than construction costs alone" (Krawinkler and Miranda 2004). Guidelines and codes, since the 1990s, have partially implemented, in their general framework, the basic concepts of PBEE in various forms, resulting in the attempt to tie design criteria to a performance level, usually that of collapse prevention. The Pacific Earthquake Engineering Research (PEER) Center has focused for several years on the development of procedures, knowledge, and tools for a comprehensive seismic performance assessment of buildings and bridges. The efforts made have resulted in a general framework that is now shared by the earthquake engineering community all over the world. Different publications describe the basic steps of PBEE methodology (e.g., Cornell and Krawinkler 2000; Krawinkler 2002; Moehle 2003; Porter 2003; Deierlein et al. 2003; Krawinkler and Miranda 2004; Moehle and Deierlein 2004), and a comprehensive document for a punctual methodological description of last enhancements of PBEE is available in FEMA P-58-1 document (2012).

Sharing a common vision and approach allows placing further enhancements and progresses in this PBEE consolidated framework, thus aiming at the final challenge to contribute effectively to the reduction of losses and the improvement of safety (Cornell and Krawinkler 2000).

The basis of performance assessment can be summarized in a single equation (see Eq. 1). This equation suggests a generic structure for coordinating, combining, and assessing the many considerations implicit in performance-based seismic assessment.

$$\lambda(DV) = \iiint |G\langle DV|DM\rangle dG\, \langle DM|EDP\rangle dG \\ \langle EDP|IM\rangle d\lambda(IM)|$$

(1)

Based on total probability theorem, Eq. 1 allows deconstructing the problem in four steps: (i) hazard analysis, (ii) structural analysis, (iii) damage analysis, and (iv) loss analysis. Each step carries out a specific generalized variable: *intensity measure* (*IM*), *engineering demand parameter* (*EDP*), *damage measure* (*DM*), and *decision variable* (*DV*). The key issue of PBEE methodology is to identify and quantify *DV* of primary interest to the decision makers with consideration to all important uncertainties. *DV*s have been defined in terms of different quantities, such as repair costs, downtime, and casualty rates.

First member of Eq. 1 is a probabilistic description of the *DV*, such as the mean annual frequency of the *DV* exceeding a specified value. $\lambda(DV)$ might be the mean annual frequency of the direct dollar loss (repair cost) exceeding 50 % of the replacement cost of the facility. The terms *dG* (*EDP*|*IM*), *dG*(*DM*|*EDP*), and *G*(*DV*|*DM*) or their derivatives on the right side of Eq. 1 are conditional probabilities relating one quantity to another, while dλ(*IM*) is the derivative of the hazard curve, relating ground motion intensity measure to its mean annual frequency of exceedance.

Aimed at facilitating probability calculation, but mostly aimed at compartmentalizing discipline-specific knowledge, the above PBEE framework should choose intermediate variables (*EDP* and *DM*), so that conditional probabilities

S

are independent of one another (Deierlein et al. 2003). A brief intro to each step of PBEE is provided herein. It is a general methodological framework description that does not refer yet specifically to RC structures.

Hazard Analysis

Hazard analysis allows describing earthquake hazard in a probabilistic manner. The result of this phase is a hazard curve, which shows the variation of the selected *IM* versus mean annual frequency of exceedance (*MAF*). Probabilistic seismic hazard analysis, for a specific site, passes through four main steps: (i) characterization of the seismic sources (e.g., identified faults or geographical areas), in which it is generally assumed that the occurrence of an event at one source does not affect the occurrence of events at other sources; (ii) characterization of magnitude distribution, generally obtained through a truncated exponential distribution or through characteristic magnitude distribution estimated for specific faults; (iii) ground motion estimation evaluated through prediction equation of the considered *IM*; and (iv) probability analysis for the evaluation of the hazard curve (see also McGuire 2004). Hazard analysis can be performed with different intensity measures, provided the appropriate prediction equation. Notwithstanding the fact that advanced intensity measures have been considered in literature (e.g., Tothong and Luco 2007) and prediction equations for these IMs are available (e.g., Tothong and Cornell 2006; De Luca et al. 2014a), still the employment of peak ground acceleration (*PGA*) and, more commonly, of spectral acceleration at the fundamental period of the structure, Sa(T), represents the basic practice.

Hazard analysis is generally performed for the evaluation of a target spectrum finally aimed at ground motion selection, if the assessment is made through nonlinear dynamic analysis. It is worth to note that according to FEMA P-58-1 (2012), hazard analysis for vulnerability assessment and loss assessment can be performed with three different approaches according to the final aim of the study: intensity-based, scenario-based, and time-based.

Notwithstanding the fact that hazard analysis is a common step for each structural typology to which PBEE methodology is applied, it is worth noting that it is also the step in which most of the variability of the whole problem is considered. The so-called record-to-record variability is the source of the most significant variability in risk evaluation (e.g., Goulet et al. 2007). On the other hand, such source of variability is also a common aspect for each structural typology, and it is not an aspect that strictly characterizes RC more than other structural typologies (e.g., steel, masonry, etc.).

Structural Analysis

Structural analysis is the phase of PBEE in which structural simulations are performed for the evaluation of *EDP*s at given *IM*s. At this step, a structural model is built up for the estimation of *EDP*s, such as internal member forces and local or global deformations, including structural collapse. The computational model is built up taking into account uncertainties in model parameters. A proper nonlinear model of a structure is the result of a reliable analytical model of the elements of the structure considered and of the analysis of experimental behavior of the elements.

The choice of *EDP*s depends on the performance target and on the type of system of interest. The general approaches for the evaluation of *dG(EDP/IM)* is to perform incremental dynamic analyses or other kinds of nonlinear analyses; see also FEMA P-58-1 (2012).

Damage Analysis

Damage analysis is the third step of PBEE methodology. It relates the *EDP*s to *DM*s. The *DM*s include quantitative descriptions of damage to structural elements, nonstructural elements, and contents. This quantification must be in sufficient detail to enable subsequent quantification of the necessary repairs, disruption of function, and safety hazards (Moehle and Deierlein 2004). The *EDP*s considered in structural analysis are the input to a set of fragility functions that model the probability of various level of physical damage, conditioned on structural response. Physical damage is described at a detailed level, defined relative to particular repair efforts required to

restore the component to its undamaged state. The current approach is a component-based approach (e.g., Aslani 2005; FEMA P-58-1 2012). The study of experimental behavior of components is still a key issue for reliable evaluation of damage measures.

Loss Analysis

The last step of a PBEE loss assessment is the probabilistic estimation of performances through *DV*s. *DV*s represent the outcome parameters of PBEE and allow a transformation of engineering evaluations in terms of variables of interest for stakeholders (e.g., dollars, deaths, downtime); see Porter (2003). Different studies are available in literature for the evaluation of direct and indirect losses (e.g., Mitrani-Reiser 2007), and results of such studies are now the basis of the loss assessment framework provided by FEMA P-58-1 (2012). It is worth noting that loss analysis works on repair costs and downtime that, in turn, are strictly related to the social and economical environment in which they are evaluated.

Behavior of RC Elements and Nonstructural Elements

The basis of a reliable vulnerability analysis for any structural material passes through a solid modeling approach. PBEE assessment framework naturally points to nonlinear dynamic analysis as the reference structural analysis tool for RC structures. Thus, it is necessary to consider reliable nonlinear hysteretic models of all components of RC structures. Structural modeling of RC elements is based on the observation and interpretation of experimental tests and post-earthquake evidence from which it is possible to carry out nonlinear models of element behavior. It is important to state that a significant difference has to be made between elements with and without seismic detailing. In fact, existing elements are susceptible of various modes of failure, while capacity design in seismic detailed elements prevents brittle modes of failure.

Seismic vulnerability assessment has to account for all typical modes of failure that RC elements can show; so, in the following, behavior of RC elements is described considering the case of existing elements that represent the most general situation in vulnerability assessment problems. This section is primarily focused on the vulnerability of RC frame structures; thus the behavior of columns and beam–column joints is specifically addressed. At the end of this section, behavior of masonry infills is also considered. Masonry infills are considered because of the significant effect they can have on vulnerability of RC frame structures. In fact, they provide on one hand the strength and stiffness increase and on the other hand, the local interaction between infill and RC frame.

RC Members

The specification of nonlinear structural component models in the form of monotonic backbone curves and hysteretic rules for RC structures is one of the key targets for the earthquake engineering community worldwide. Monotonic backbone curves and hysteretic rules should be able to capture the experimental behavior of RC elements. Any RC element component model is defined as function of specific parameters, such as shear span ratio, materials' strengths, longitudinal and transversal reinforcement ratios, and axial load ratio (e.g., CEB 1996; Elwood 2004; Elwood and Moehle 2005a, b; Zhu et al. 2007; Haselton et al. 2008; Biskinis and Fardis 2010a, b). Specific combinations of the above ruling parameters end up to the different definitions of wall, beam, or column. Furthermore, the structural difference between walls, beams, and columns is also made by the structural typology in which the above prismatic RC members are included, and by the structural role they play with respect to loads they are subjected to (e.g., gravity loads, earthquake loads, etc.).

In particular a prismatic RC member that carries gravity loads through shear and bending and that is designed under the assumption of uniaxial load with zero axial force is referred as a beam in RC frame structures. Again, a prismatic RC member in which the ratio between the two nondominant dimensions (i.e., width and depth of the cross section) is higher than four is typically defined as wall. This latter geometrical

assumption implies that a "wall" resists lateral forces mainly in one direction, and it can be designed for such unidirectional resistance by assigning flexural resistance to the two far ends of the section and shear resistance to the web (Fardis 2009). The classification of an RC prismatic elements as "column" is the more general possible; it covers the widest range of parameters ruling RC component modeling (e.g., variable axial load ratio, biaxial load, variable reinforcement ratios). Therefore, in the following, specific attention is given to modeling issues related to RC columns rather than beams or walls.

Notwithstanding the fact that columns are characterized by biaxial loading and by the widest variety of ruling parameters, experimental data more commonly available address the case of typical columns in RC frame structures. More specifically, most test specimens have rectangular cross section, symmetric reinforcement, and nonzero axial load, and rarely they are characterized by biaxial loading. The latter is the reason why most reliable nonlinear structural component models are conceived for columns under uniaxial loading. It is worth noting that recent experimental and modeling efforts have been made for the characterization of biaxial component models (e.g., Bousias 1993; Bousias et al. 2002; Di Ludovico et al. 2013), and some guidelines now provide recommendations for their employments. On the other hand, biaxial modeling efforts are still in a preliminary phase, and most common PBEE nonlinear analysis applications are performed discarding biaxial nonlinear component models. Thus, in the following, the only uniaxial behavior of elements is described.

Failure Mode Classification of RC Elements

Columns are generally subjected to the contemporary presence of axial load, shear forces, and bending moment. The interaction between those internal forces strictly affects the response of structural elements and consequently their mode of failure. Interaction effects, i.e., the influence of shear forces on flexural deformations and that of normal stress resultants on shear deformations, are important. In general, they cannot be neglected,

especially for inelastic cyclic loading for non-capacity-designed elements. Flexural response can be affected by the interaction with shear also in linear elastic phase, when plastic shear capacity of the element is not attained. Conversely, in the case of inelastic response, shear capacity can decrease after yielding as a result of increasing ductility demand in the element. In fact, damage caused by cyclic degradation in flexural response decreases post-cracking shear capacity mechanisms (e.g., aggregate interlock, loss of anchorage of transverse reinforcement, etc.) resulting in a decrease of shear capacity.

In Fig. 1, two experimental tests on rectangular RC columns from PEER database by Berry et al. (2004) are shown. The first column, Fig. 1a, is characterized by flexural-dominated behavior, while the second column, Fig. 1b, is characterized by post-yielding interaction between flexure and shear.

The above observations allow indentifying three different failure modes for columns resulting from flexure–shear interaction, as shown, schematically, in Fig. 2. The analytical classification of different modes of failure can be made only considering degrading shear capacity models which account for shear capacity degradation after yielding.

1. Shear failure mode, S – initial shear capacity (V_0), i.e., not affected by cyclic degradation, is lower than the plastic shear capacity (V_p), and flexural response is modified by a preemptive shear failure in the elastic response phase of the element. Deformation capacity (i.e., drift, chord rotation) is very limited. Subsequent response of the element is characterized by a strictly degrading behavior with significant shear strength reduction and a consequently loss of axial load-carrying capacity; see Fig. 2a.

2. Flexure–shear failure mode, FS – V_0 is higher than V_p, and flexural response is characterized by an inelastic phase. However, if degraded shear strength (V_d), the lower red dotted branch in Fig. 1, is lower than V_p, the inelastic flexural response is modified by a shear failure that occurs at an intermediate shear strength
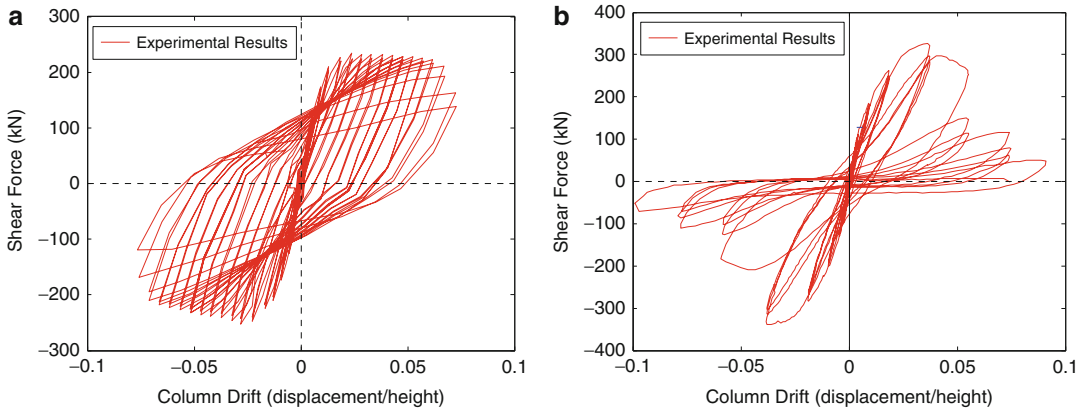
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 1 Experimental results for Berry et al. (2004) database. Experimental tests by (**a**) Legeron and Paultre (2000) and (**b**) Pujol (2002) (Adapted from Haselton et al. 2008)
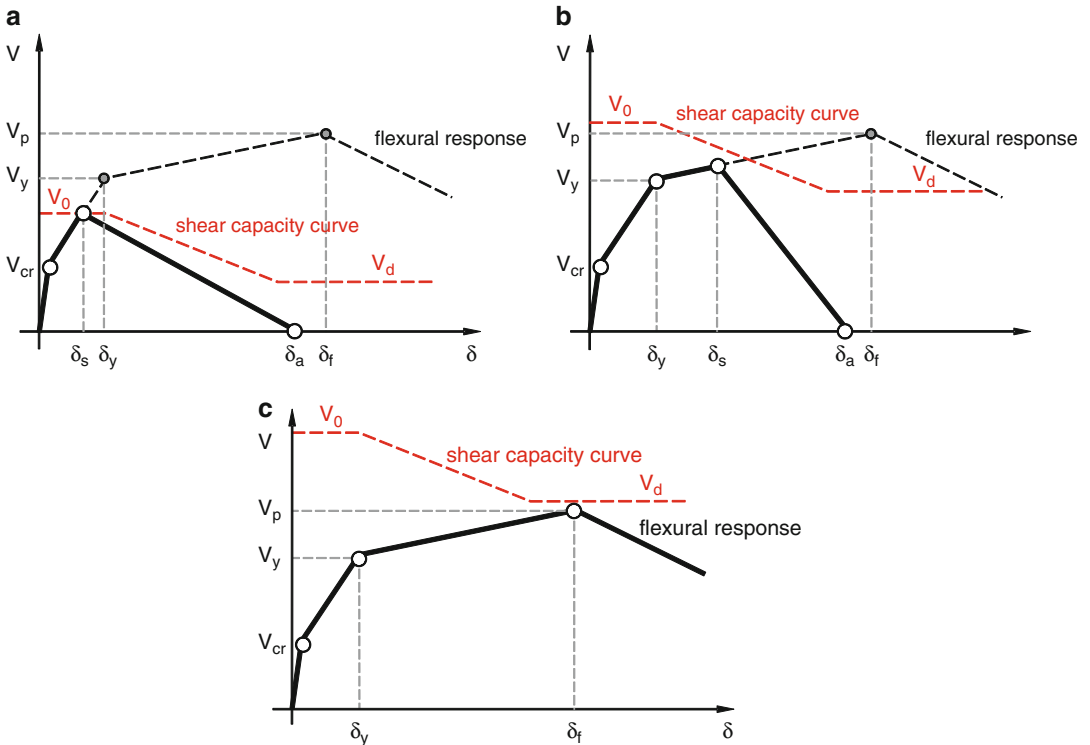
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 2 Collapse modes of RC elements subjected to axial load, shear, and bending moment: (**a**) shear failure mode, S; (**b**) flexure–shear failure mode, FS; (**c**) flexural failure mode, F

between V_0 and V_d. The corresponding deformation capacity is higher with respect to that of case (1). The behavior of the element succeeding this kind of shear failure is strictly

degrading up to the loss of axial load-carrying capacity; see Figs. 1b and 2b.

3. Flexural failure mode, $F - V_d$ is higher than V_p, and flexural response is not affected by

interaction with shear. Thus, the element is characterized by inelastic deformation dominated by flexure and its typical damage, such as local buckling of longitudinal reinforcement and concrete spalling. This damage typically corresponds to the ultimate flexural capacity (θ_f). The softening branch of shear force–displacement envelope ends up to the loss of vertical load-carrying capacity; see Figs. 1a and 2c. This is the failure mode typical of capacity-designed elements in which shear–flexure hierarchy is controlled during the design process (see also De Luca and Verderame 2013).

The classification of failure mode is made on the basis of the ratio between plastic shear capacity V_p and degrading shear strength V_n. Several models have been developed to represent the degradation of shear strength with increasing inelastic deformations (Watanabe and Ichinose 1992; Aschheim and Moehle 1992; Priestley et al. 1994; Biskinis et al. 2004; Sezen and Moehle 2004).

In particular, the shear-degrading model proposed by Sezen and Moehle (2004) estimates column shear strength as the summation of shear carried by concrete (V_c) and shear carried by transverse reinforcement through a 45° truss model (V_s). This shear strength model is calibrated on 51 experimental column tests, and it has the form shown in Eq. 2.

$$
\begin{aligned}
V_n &= k \cdot V_o = k \cdot (V_c + V_w) \\
&= k \cdot \left[\left(\frac{0.5\sqrt{f_c}}{L_V/d} \cdot \sqrt{1 + \frac{P}{0.5 \cdot \sqrt{f_c} \cdot A_g}} \right) \cdot \right. \\
&\qquad \left. 0.8 A_g + \frac{A_{sw} \cdot f_{yw} \cdot d}{s} \right]
\end{aligned}
$$
(2)

Shear strength contribution carried by concrete depends on concrete compressive strength, (f_c); shear span, (L_V); section depth, (d); axial load, (P); and gross area section, (A_g), while shear strength contribution carried by reinforcements depends on transverse reinforcement area (A_{sw}) within a spacing (s) in the loading direction,

yielding steel strength (f_y), and section depth (d), computed as distance from the extreme compression fiber to centroid of tension reinforcement. This model takes into account shear strength degradation after yielding through a degradation coefficient (k); k is defined to be equal to 1.0 for displacement ductility less than 2, to be equal to 0.7 for displacement ductility exceeding 6, and to vary linearly for intermediate displacement ductilities. The degradation coefficient k is applied to both V_c and V_s, under the assumption that concrete component degrades due to increased cracking and degradation of the aggregate interlock mechanism, while steel component likewise degrades due to hoop opening and bond degradation. This shear capacity model is adopted in ASCE/SEI 41 guidelines (2007).

It is worth noting that the classification of column failure modes based only on shear strength and V_p can be still not adequate. Other column parameters may also influence the observed failure mode (Zhu et al. 2007). These considerations are reflected by ASCE/SEI 41 – supplement 1 provisions (Elwood et al. 2007) as follows: the expected failure is first evaluated based on the ratio between V_p and the shear capacity, and then, when classifying a column as flexure-controlled, further parameters are considered, namely, the transverse reinforcement ratio, assuming 0.002 as the lower limit, as in Zhu et al. (2007), and the tie spacing-to-section depth ratio, assuming 0.5 as the upper limit.

Characteristic points of shear force–displacement envelope (i.e., moment–chord rotation), for each failure mode, can be identified schematically through the representation shown in Fig. 2. Monotonic and cyclic deformation capacities to be attributed to each characteristic point can be found in literature. It is worth to note that if in the elastic phase the distinction between monotonic and cyclic deformation capacity can be unnecessary (pre-yielding phase), such distinction becomes significant in the inelastic phase (post-yielding), especially if those deformation capacities are employed in the analytical modeling of the element.

Monotonic deformation capacities are employed for the definition of the force–deformation envelope
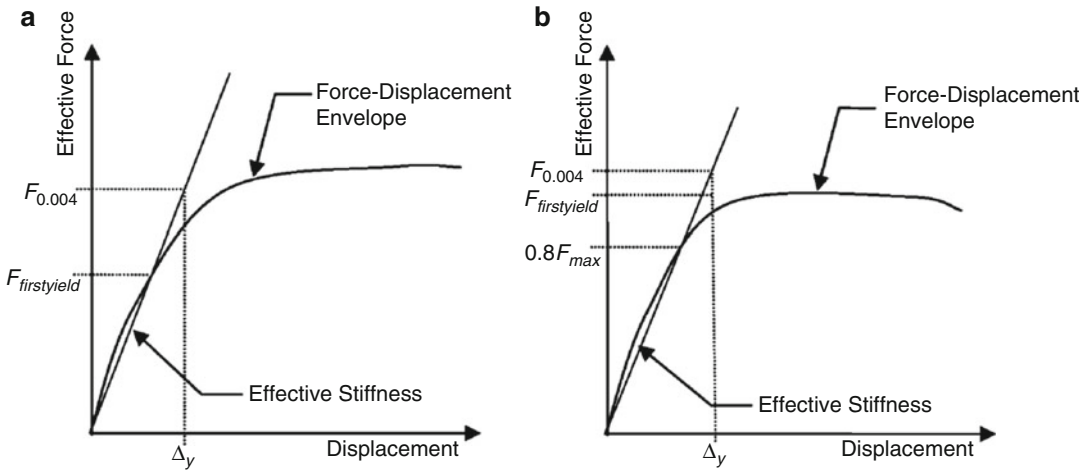
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 3 Definition of yield displacement and effective stiffness from test data for (**a**) yielding columns and (**b**) columns that did not yield (From Elwood and Eberhard 2009)

of the element. Successively, it is necessary to assign cyclic degrading rules. So, for example, a specific degradation percentage of peak force (conventional failure) is attained at a deformation capacity lower with respect to that considered on the monotonic envelope. When cyclic degrading rules are not modeled, the conventional failure has to be fixed at a displacement capacity that implicitly accounts for it (i.e., a value on the monotonic envelope that corresponds to the lower capacity degraded by cyclic effects). It is worth noting that the latter solution is obviously an approximated way for the estimation of a reliable deformation capacity. Cyclic degradation depends on the load path, and it cannot be accurately captured analytically with monotonic models. Still, for some specific failure modes, the monotonic degraded solution can be more straightforward with respect to the calibration of fulfilling cyclic models.

Yielding Displacement in RC Members

For the estimation of a reliable inelastic backbone of an RC element, the estimation of the yielding point is quite relevant. In literature, different empirical or hybrid models for the estimation of yielding drift capacity based on experimental observations are available. Direct estimations of yielding drift (Elwood and Eberhard 2009; Biskinis and Fardis 2010a) are generally based

on the assumption that yielding drift is the sum of three different components: a flexural component, a shear component, and a fixed end rotation component due to longitudinal bar slip; see Eq. 3. Alternatively yield drift can be evaluated indirectly from the estimation of yield stiffness (Elwood and Eberhard 2009; Biskinis and Fardis 2010a).

$$\delta_y = \delta_{y,\,\text{flex}} + \delta_{y,\,\text{shear}} + \delta_{y,\,\text{slip}} \qquad (3)$$

Yield drift obtained through Eq. 3 is generally evaluated as the corner for an ideal envelope approximating a bilinear force deformation response of the element and not as the drift corresponding to the first yielding in the reinforcement steel (or equivalently in the concrete) at the first section of the element. Figure 3 shows the procedure by Elwood and Eberhard (2009) for the definition of yielding displacement on the envelope of measured lateral load–displacement relationship corrected for P-delta. Elwood and Eberhard's procedure covers both the cases in which yield force is attained in the envelope (Fig. 3a) and cases in which the calculated yield force is not attained (Fig. 3b); in fact, it provides an estimate of effective stiffness also in the case of columns that do not yield (see Fig. 3b).

Eurocode 8 part 3 (CEN 2005) recommends a three-component model for the estimation of yielding drift, while ASCE/SEI 41 (2007) provides an effective stiffness to be computed as function of the axial load to which the element is subjected. Experimental databases on which the above formulations are calibrated are often characterized by modern code-conforming elements with proper seismic detailing. In the case of nonconforming elements, such formulations have still shown a fair agreement; emphasizing that seismic detailing does not affect strictly yielding capacity (Biskinis and Fardis 2010a). It is worth noting that nonconforming definition changes according to the specific code considered; on the other hand, this definition is generally referred to lack of detailing in transversal reinforcements, smooth bar presence, and lack of confinement (e.g., no 135° hooks, poor stirrup spacing, etc.). Analogously, database of nonconforming elements with smooth bars have shown a fair agreement with empirical formulations based on conforming elements with ribbed bars (e.g., Ricci et al. 2013).

Ultimate Drift in RC Members

Other characteristic point of RC member force–displacement response is the ultimate drift capacity for each failure mode considered. Ultimate drift capacity is generally evaluated at the drift characterizing a 20 % loss of the maximum shear strength (FIB 2003) attained in the element; see Fig. 4. According to the failure mode, different empirical ultimate drift capacity formulations are available in the literature. Most experimental databases are characterized by cyclic tests and few monotonic tests. Thus, most of those empirical formulations account implicitly for cyclic degradation.

Zhu et al. (2007) provide an ultimate drift capacity model for two collapse modes. They considere a database of 125 nonconforming columns. Tests are divided in two different sub-databases according to a binary classification approach: Zone S columns (shear-dominated failures) and Zone F columns (flexural-dominated failures). Zone F is composed of 85 tests, while Zone S is composed of 40 tests that include both
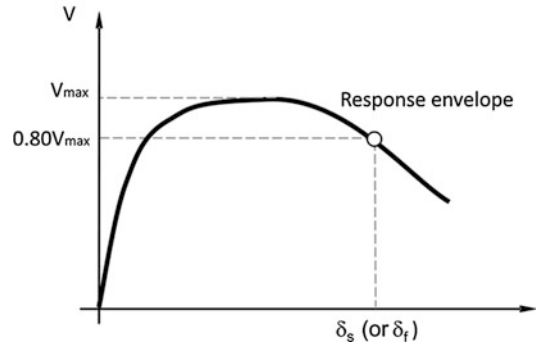
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 4 Conventional criterion for the evaluation of ultimate drift capacity independently of the failure mode observed

shear- and flexure–shear-dominated failures. Classification is made on the basis of three column parameters: ratio of plastic shear demand to strength ratio (V_p/V_n), the aspect ratio (a/d), and the transverse reinforcement ratio (ρ''). The two sub-databases are employed for the estimation of the median ultimate drift capacity (20 % degradation of maximum shear strength) in the two cases, with characterization of logarithmic standard deviations. Equations 4 and 5 show the median estimates for Zone S and Zone F ultimate drift capacity, respectively. Zone S ultimate drift is function of hoop spacing to depth ratio (s/d), aspect ratio (a/d), and normalized axial force (v). Zone F ultimate drift depends on longitudinal reinforcement ratio (ρ_l), mechanical transversal reinforcement ratio ($\rho'' f_{yt}/f_c$), s/d, and v.

$$\delta_s = 2.02\rho'' - 0.025\frac{s}{d} + 0.013\frac{a}{d} - 0.031v \quad (4)$$

$$\delta_f = 0.049 + 0.716\rho_l + 0.120\frac{\rho'' f_{yt}}{f_c} \\ - 0.042\frac{s}{d} - 0.070v \quad (5)$$

An alternative solution for the characterization of ultimate drift capacity for flexure–shear mode of failure (i.e., after yielding) is provided by Elwood and Moehle (2005a) based on a
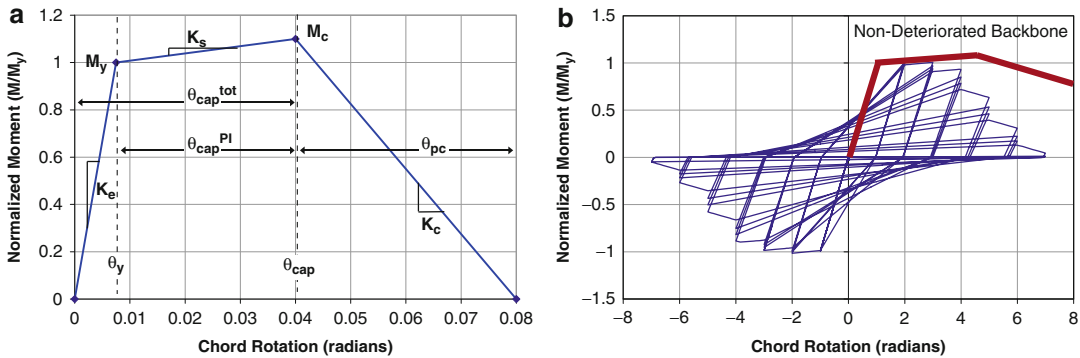
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 5 (**a**) Monotonic and (**b**) cyclic behavior of component model used in the calibration study by Haselton et al. (2008) (From Haselton et al. 2008)

database of 51 lightly transversal-reinforced columns (Sezen and Moehle 2004); see Eq. 6. This ultimate drift depends on the maximum nominal shear stress (v), ρ'', and v.

$$\delta_s = \frac{3}{100} + 4\rho'' - \frac{1}{40}\frac{v}{\sqrt{f_c}} - \frac{1}{40}v \geq \frac{1}{100} \quad (6)$$

For flexural-dominated behavior, an empirical ultimate drift capacity is provided by Biskinis and Fardis (2010b). It represents the evolution of the first formulation provided by Panagiotakos and Fardis (2001). This formulation is based on a database of 1,352 uniaxial columns with good detailing and continuous bars (299 monotonic and 1,053 cyclic tests). According to the loading of the test, Fardis and his coauthors provide two empirical formulations for ultimate drift capacity, one for monotonic loading (see Eq. 7) and one for cyclic loading (see Eq. 8). Biskinis and Fardis' formulations depend on v, f_c, shear span ratio (L_v/h), longitudinal mechanical reinforcement ratio in compression and tension (ω' and ω, respectively), confinement effectiveness factor (α), transversal reinforcement ratio (ρ_w or ρ''), transversal reinforcement yielding strength (f_{yw} or f_{yt}), and diagonal reinforcement ratio (ρ_d).

$$\delta_f = 0.028 \cdot (0.3^v)\left[\frac{\max(0.01;\omega')}{\max(0.01;\omega)}f_c\right]^{0.225}$$
$$\times \left[\min\left(9;\frac{L_v}{h}\right)\right]^{0.35} 25^{\frac{\alpha\rho_w f_{yw}}{f_c}} 1.25^{100\rho_d} \quad (7)$$

$$\delta_f = 0.016 \cdot (0.3^v)\left[\frac{\max(0.01;\omega')}{\max(0.01;\omega)}f_c\right]^{0.225}$$
$$\left[\min\left(9;\frac{L_v}{h}\right)\right]^{0.35} 25^{\frac{\alpha\rho_w f_{yw}}{f_c}} 1.25^{100\rho_d} \quad (8)$$

Considering a database of 48 cyclic tests, Biskinis and Fardis (2010b) provide a correction factor to Eq. 8 that accounts for poor detailing (i.e., older columns), and it is equal to (1/1.20). For substandard elements with smooth bars (based on 31 tests), the same authors provide an additional correction factor, equal to 0.95, to be applied to Eq. 8 (see also Verderame et al. 2010).

Flexural Component Model
For flexural-dominated behavior, Haselton et al. (2008) provide a complete calibration based on Berry et al. (2004) database for the specification of a complete nonlinear structural component model based on the beam–column element model developed by Ibarra et al. (2005). Figure 5 shows the monotonic and cyclic behavior of the component model employed by Haselton et al. (2008). Parameters to be considered for the employment of Haselton et al. (2008) model are (1) M_y, yielding moment; (2) θ_y, yielding rotation; (3) θ_{cap}, monotonic chord rotation at onset of strength loss (i.e., capping); (4) K_s, hardening stiffness; (5) θ_{pc} or K_c, post-capping stiffness; (6) λ, normalized hysteretic energy degradation capacity (i.e., cyclic degradation
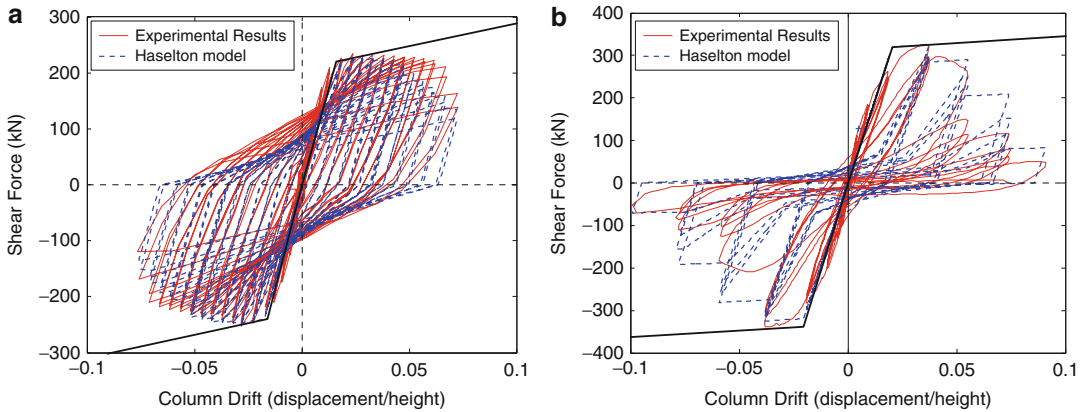
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 6 Example of Haselton et al.'s calibration for test specimens 193 and 212 (see Fig. 1) characterized by (**a**) flexural failure mode and (**b**) flexure–shear failure mode (Adapted from Haselton et al. 2008)

parameter); and (7) c, exponent term to model rate of deterioration (i.e., cyclic parameter). This model accounts for cyclic degradation directly through the modeling of hysteretic behavior. It is significantly different with respect to the other models considered and also more complete. On the other hand, drift capacity thresholds that implicitly account for cyclic degradation cannot be evaluated on the monotonic backbone provided by this model. In fact, the monotonic backbone by Haselton without the degrading parameters λ and c is incomplete and nonconservative for RC elements capacity evaluation through static methods.

It is worth to note that Haselton et al. make a calibration effort for most of columns of Berry et al. database considering both columns that failed in flexural mode and columns characterized by flexure–shear mode. In Fig. 6, an example of the calibration made by Haselton et al. (2008) is shown for the two experimental test specimens provided in Fig. 1. The comparison of the two examples of calibration emphasizes that this component model is very well fitted for columns characterized by flexural behavior and less accurate for flexure–shear-dominated columns.

Loss of Vertical Carrying Capacity

Some of the experimental tests characterized by shear-dominated failure have shown, after the

attainment of ultimate deformation, a loss of vertical carrying capacity. In particular, Elwood and Moehle (2005b), on the basis of 12 tests characterized by flexure–shear mode of failure, observed that the drift ratio at axial load failure (δ_a) of a shear-damaged column is inversely proportional to the magnitude of the axial load and directly proportional to the amount of transverse reinforcement. Based on the classical shear friction model, Elwood and Moehle (2005b) propose a model for the drift at axial failure for shear-damaged columns. The axial load on a shear-damaged column is assumed to be supported by a combination of compression of the longitudinal reinforcement and force transfer through shear friction on an idealized shear failure plane (see Fig. 7).

The effective coefficient of friction from the classical shear friction equation is related to the drift ratio at axial failure using the results from 12 full-scale pseudo-static column tests. The effective coefficient of friction along the critical shear failure plane can be calculated using equilibrium and subsequently related to the drift ratio at axial failure. Elwood and Moehle found that the effective coefficient of friction calculated by ignoring the contribution of longitudinal reinforcement provided good agreement with the experimental data. According to the above observations, the empirical formulation provided in Eq. 9 is proposed for the evaluation of drift at
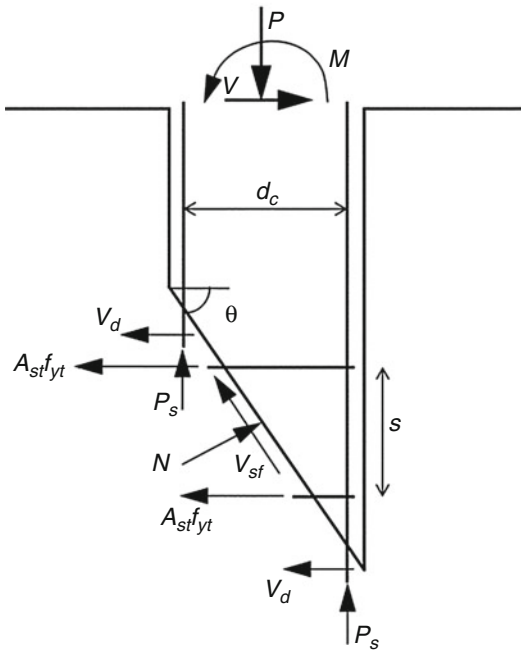
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 7 Free body diagram of column after shear failure (From Elwood and Moehle 2005b)

loss of axial load-carrying capacity, where θ is equal to $65°$, d_c is the depth of the column core (centerline to centerline of transverse reinforcement), and A_{st} and s are area and spacing of transverse reinforcement, respectively.

$$\delta_a = \frac{4}{100} \frac{1 + \tan(\theta)^2}{\tan(\theta) + P\left(\dfrac{s}{A_{st}f_{yt}d_c \tan(\theta)}\right)} \quad (9)$$

Elwood and Moehle's model is successively recalibrated by Zhu et al. (2007). Zhu et al. employed a database of 28 column specimens. All tests are unidirectional pseudo-static tests and were terminated after loss of axial load capacity. All the considered columns experienced flexural yielding prior to shear failure. Zhu et al. provide Eq. 10 for the evaluation of drift ratio at loss of vertical load-carrying capacity (axial failure), in which μ is the effective coefficient of friction evaluated according to Eq. 11. The same ruling parameters are assumed with respect to Eq. 9, and θ is equal, again, to $65°$. Both Eqs. 9 and 10 are

based on the same model, but the most recent study (i.e., Eq. 10 by Zhu et al.) is based on a larger database, and it also provides a coefficient of variation equal to 0.35 for the median estimate of Eq. 10.

$$\delta_a = 0.184 \exp(-1.45\mu) \quad (10)$$

$$\mu = \frac{\dfrac{P}{A_{st}f_{yt}d_c/s} - 1}{\dfrac{P}{A_{st}f_{yt}d_c/s}\dfrac{1}{\tan\theta} - \tan\theta} \quad (11)$$

Beam–Column Joints

The performance of beam–column joints has been identified as a critical issue in the seismic resistance of RC moment-resisting frames (RC MRF). In RC MRF structures under severe ground motions, beam–column connections are subjected to moment reversals across the joint due to the adjacent beams and columns. As a result, the joint regions undergo significant horizontal (and vertical) shear forces whose magnitudes are much larger than those in the adjacent members (CEB 1996); see Fig. 8.

As a result, beam–column joints are susceptible to shear failure which generally involves a brittle process. Such brittle shear failure must be avoided through appropriate design to ensure ductile response of the frame. Despite the importance of shear design of RC beam–column joints, the research community has not yet developed a commonly accepted approach for the determination of the shear strength of RC beam–column joints, probably due to the complexity of joint behavior.

Many reinforced concrete (RC) buildings constructed without transverse steel shear reinforcement in the beam–column joint region still widely exist in seismically active regions, since the transverse reinforcement requirements for the design of beam–column joints were not addressed in earlier code provisions. Such unreinforced joints are considered vulnerable to brittle shear failure under earthquake shaking due to insufficient shear reinforcement in the joint region. These problems have been highlighted, in recent past, by the damage observed after
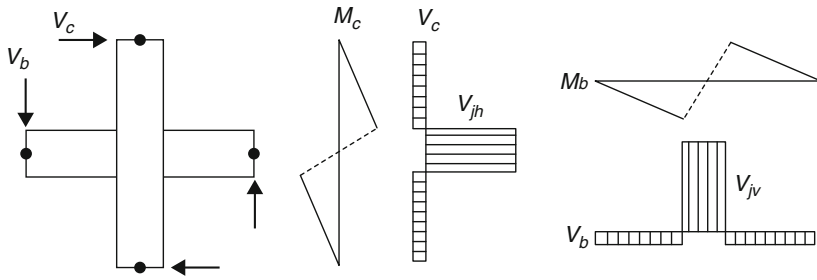
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 8 Shear forces within the joint. M_c, V_c, M_b, and V_b are column and beam moment and shear, respectively. V_{jh} and V_{jv} are horizontal and vertical joint shears (Adapted from Fardis 2009)

Seismic Vulnerability Assessment: Reinforced Concrete Structures, Table 1 γ_n for joint shear strength evaluation

$\gamma_n (MPa)^{0.5}$					
ρ''	Interior joint with transverse beams	Interior joint without transverse beams	Exterior joint with transverse beams	Exterior joint without transverse beams	Knee joint with or without transverse beams
<0.003	1.0	0.8	0.7	0.5	0.3
≥ 0.003	1.7	1.2	1.2	1.0	0.7

devastating earthquakes in different countries; moreover, many tests have proven the poor seismic performance of unreinforced joints, especially of exterior joints.

Design recommendations for beam–column joints in code provisions have been considerably revised since the ACI-ASCE Committee 352 (ACI 352R-02 2002) published its first seismic design guidelines in 1976. Still, nominal joint shear strengths proposed in current code provisions are only applicable to joints satisfying the minimum requirements of transverse reinforcement in the joint region. Some analytical models (Vollum 1998; Hwang and Lee 1999; Lowes and Altoontash 2003; Wong 2005) for predicting joint shear strength have been developed based on the strut-and-tie (SAT) idealization incorporating Mohr's strain compatibility and softening concrete behavior, e.g., the modified compression field theory (MCFT) proposed by Vecchio and Collins (1986). However, LaFave and Shin (2005) indicated that the strength prediction using the MCFT tends to underestimate the shear strength of lightly reinforced and that of unreinforced joints.

Therefore, current code provisions and available analytical models may be inappropriate to predict the shear strength of unreinforced joints; see also Park and Mosalam (2013) for details.

Alternatively, ASCE/SEI 41 (2007) provides recommendations for the shear strength and joint shear stress–strain relationship of unreinforced joints for seismic rehabilitation purposes based on the pre-standard developed in FEMA 273 (FEMA 1997) and FEMA 356 (FEMA 2000). According to ASCE 41, nominal joint shear strength (V_n) is defined according to Eq. 12; it is a function of a coefficient γ_n, joint width (b_j), column depth (h_c), and concrete tensile strength ($\sqrt{f_c}$)

$$V_n = \gamma_n \sqrt{f_c} b_j h_c \qquad (12)$$

The values of γ_n for joint shear strength calculation are presented in Table 1. To investigate the relevance of the strength recommendations in ASCE 41, Park and Mosalam (2012) collected 62 previous unreinforced exterior or corner joint test data. The evaluation of the joint shear
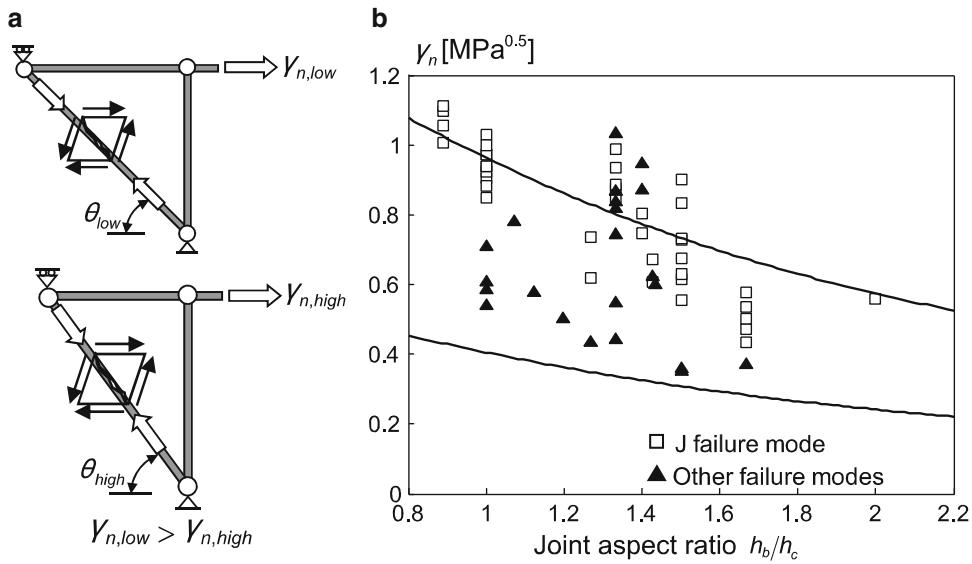
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 9 Effect of the joint aspect ratio: (**a**) SAT idealizations for two joint aspect ratios and (**b**) evaluation of the database against joint aspect ratio (Adapted from Park and Mosalam 2012)

strength for the collected database reveals that ASCE 41 may underestimate the shear strength of unreinforced exterior joints. This parametric study shows that the shear strength of unreinforced exterior joints is affected by the joint aspect ratio, defined as the ratio of beam to column cross-sectional heights. The effect of the joint aspect ratio can be explained by the SAT idealization where a steeper diagonal strut is developed in the unreinforced joint with a high aspect ratio; see Fig. 9. Consequently, this steeper diagonal strut results in less effective shear resistance to equilibrate the horizontal joint shear force. Hence, the shear strength of unreinforced exterior joints decreases with the increase of the joint aspect ratio. Similar results are reported by Kim and LaFave (2007), Vollum and Newman (1999), and Bakir and Boduroğlu (2002).

Another key parameter, joint shear demand calculated from beam longitudinal bars and frame geometry, has been suggested by Anderson et al. (2008) based on results of unreinforced interior joint tests by Walker (2001) and Alire (2002). Depending on the joint shear demand, two types of joint failure have been commonly

recognized by other researchers (Wong 2005; Kim and LaFave 2007):

1. Joint failure prior to beam longitudinal reinforcement yielding (referred to as J type failure)
2. Joint failure after beam longitudinal reinforcement yielding (referred to as BJ-type failure)

The joint shear strength in most code provisions is pertinent to joint failure independent of beam longitudinal bars yielding or not, whereas the maximum joint shear stress for the BJ-type failure is directly determined from the joint shear demand dictated by the yielding of the beam longitudinal reinforcement. Conversely, experimental results show that the joint shear strength of unreinforced joints increases with the beam longitudinal reinforcement ratio. This evidence can be explained as follows: (i) increasing the beam longitudinal reinforcement ratio leads to the increase of the horizontal joint shear force without the yielding of the beam longitudinal bars, i.e., larger horizontal joint shear force is imposed with less deterioration of bond resistance around the beam longitudinal bars in the
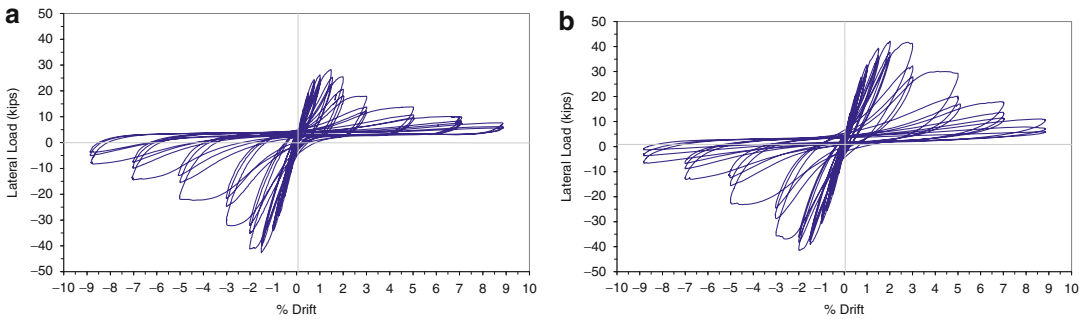
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 10 Experimental results of unreinforced exterior joint: (**a**) effect of beam–column joint failure on the sub-assemblage; (**b**) effect of anchorage failure (Adapted from Pantelides et al. 2002)

joint region; (ii) this more stable bond resistance produces a wider diagonal strut which can carry the larger horizontal joint shear force. Some strength models use either ductility factor (Park 2002; Hakuto et al. 2000) to predict the joint shear strength of BJ type of joint failure. A similar approach is adopted to predict the shear strength of columns (Sezen and Moehle 2004).

In general, experimental results on beam–column joints without transverse steel shear reinforcement emphasize a strictly degrading behavior caused, on one hand, by the shear behavior of the joint panels and, on the other hand, by the bond–slip behavior of longitudinal reinforcements anchored in it; see Fig. 10. Experimental research on the seismic performance of the beam–column joints (Walker 2001; Alire 2002; Pantelides et al. 2002) has revealed that the joint shear stress–strain response, typically, has a degrading envelope and a highly pinched hysteresis. Moreover, the common practice of terminating the beam's bottom reinforcement within the joints makes the bottom reinforcement prone to pullout under a seismic excitation. Insufficient beam bottom bar anchorage precludes the formation of bond stresses necessary to develop the yield stress in the beam's bottom reinforcement.

Finally, as described in section "Loss of Vertical Carrying Capacity," axial failure of a shear-damaged column can occur by sliding along an inclined crack plane, with resistance provided by transverse reinforcement clamping the crack and longitudinal reinforcement supporting axial force directly. Axial failure of a joint may be viewed similarly (Hassan 2011). After developing joint shear failure, the axial load will be supported by shear friction on the diagonal shear failure plane and the axial capacity of column reinforcing bars. Two axial capacity models designed for unconfined joints were developed. However, this result is obtained based on a small dataset.

Infills

The practice of realizing RC frames with masonry infill panels is very common in European countries. Masonry infill walls affect the strength and stiffness of infilled RC frame structures. The general approach to discard their contribution in design and assessment could be acceptable when proper capacity design rules are employed, since code recommendations account implicitly for their contribution through proper acceptance criteria and drift limitations (CEN 2004; ASCE/SEI 41 2007; DM14/01/2008 2008). In addition, their contribution to overall strength and stiffness is less significant in modern designed structures (e.g., Dolšek and Fajfar 2001). On the other hand, assessment of RC structures often involves substandard buildings in which the presence of masonry infill can strictly affect the structural behavior. Typical two-layer hollow clay brick infills employed in the Mediterranean area and very common in existing RC buildings cannot be neglected.
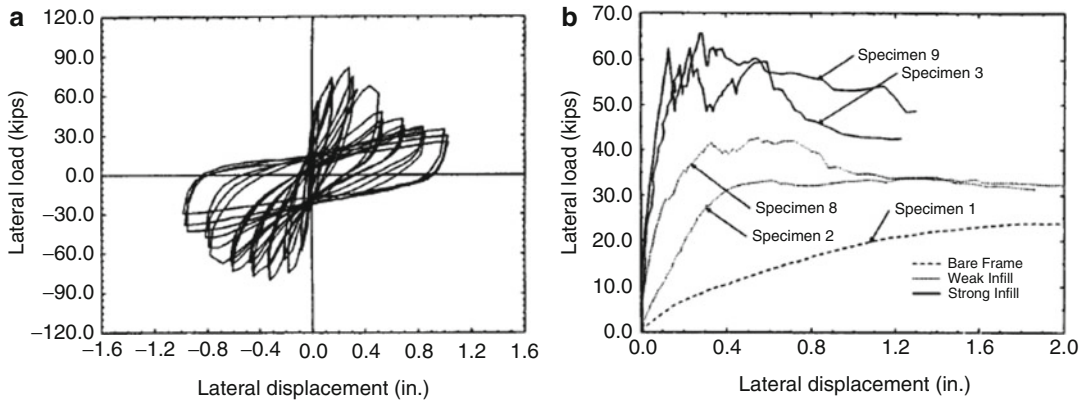
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 11 Example of (**a**) cyclic- and (**b**) monotonically loaded specimens of RC-infilled frames (From Mehrabi et al. 1996)

Their contribution represents the first capacity source to earthquake shaking in this kind of buildings (e.g., De Luca et al. 2014b; Manfredi et al. 2014), and, at the same time, they can lead to preemptive brittle failures due to local interaction between infill and RC frame (e.g., Verderame et al. 2011).

Notwithstanding the above considerations, it is worth noting that these elements are considered as *nonstructural* in the design process, and they are not subjected to the same acceptance and quality controls of RC elements. Thus their behavior can be significantly affected by specific executive practice and very different material properties.

Experimental investigations of the last decades (e.g., Mehrabi et al. 1996; Mosalam 1996; Colangelo 2003; Calvi et al. 2004, among others) emphasized a considerable reduction in the response of infilled frames under reversed cyclic loading. This behavior is due to the rapid degradation of stiffness, strength, and low energy dissipation capacity, resulting from the brittle and sudden damage of the unreinforced masonry (URM) infill walls. In Fig. 11 an example of cyclic (Fig. 11a) and monotonic (Fig. 11b) behavior of masonry infill frames from the experimental campaign by Mehrabi et al. (1996) is shown.

Based on both experimental and analytical results of the last decades, different failure mode classifications of masonry-infilled frames

have been proposed. Infill failure can be classified into five distinct modes (e.g., Shing and Mehrabi 2002; Calvi et al. 2004; Asteris et al. 2011). Infill failure graphical examples according to Shing and Mehrabi (2002) are shown in Fig. 12.

– *Sliding shear mode* represents horizontal sliding shear failure through bed joints of a masonry infill. This mode is associated with infill of weak mortar joints and a strong frame.
– *Diagonal cracking mode* is seen in the form of a crack across the compressed diagonal of the infill panel and often takes place with simultaneous initiation of the sliding shear mode. This mode is associated with a weak frame or a frame with weak joints and strong members infilled with a rather strong infill.
– *Diagonal compression mode* represents the crushing of the infill within its central region. This mode is associated with a relatively slender infill, where failure results from out-of-plane buckling of the infill.
– *Corner crushing mode* represents the crushing of the infill in at least one of its loaded corners. This mode is usually associated with infilled frames consisting of a weak masonry infill panel surrounded by a frame with weak joints and strong members.
– *Frame failure mode* is seen in the form of plastic hinges developing in the columns or
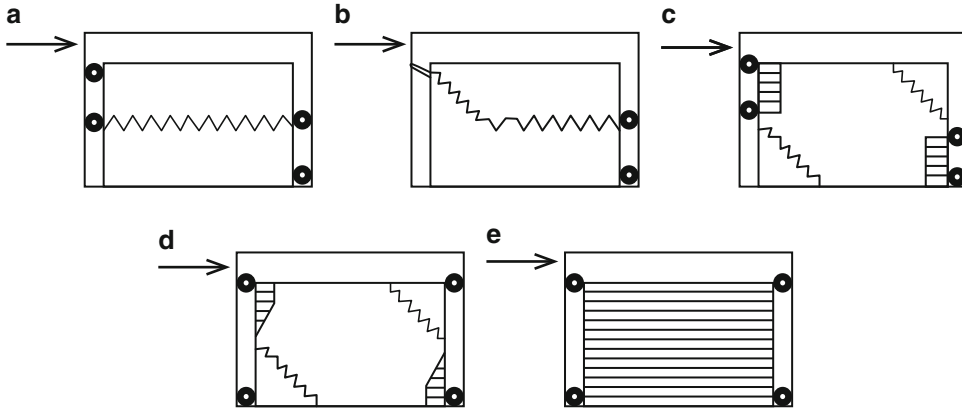
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 12 Failure mode according to the classification made by Shing and Mehrabi (2002), (**a**) midheight crack, (**b**) diagonal crack, (**c** and **d**) corner crushing, (**e**) horizontal slip (Adapted from Calvi et al. 2004)
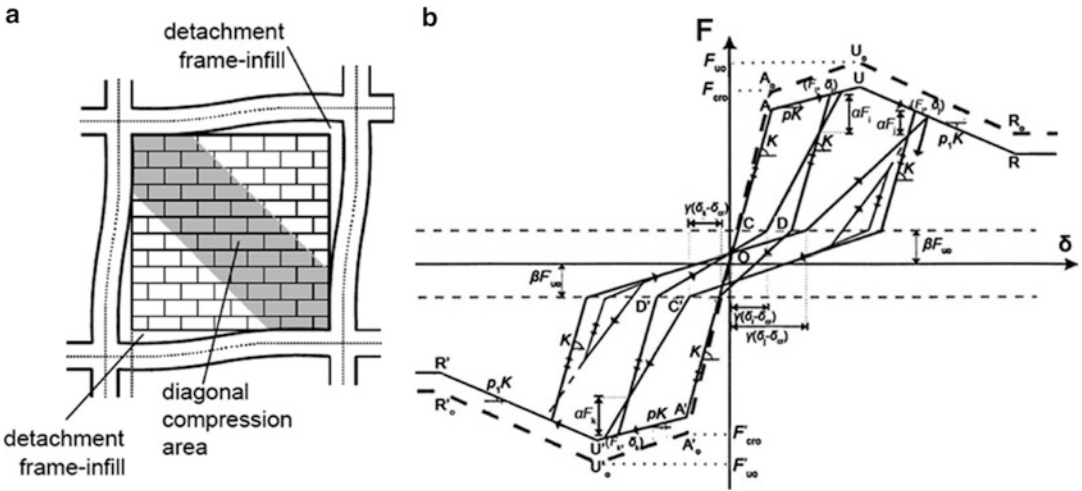
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 13 (**a**) Strut model analogy of infilled frame (From Asteris et al. 2011); (**b**) Fardis and Panagiotakos (1997) analytical model for infills (From Fardis 2009)

the beam–column connections. This mode is associated with a weak frame or a frame with weak joints and strong members infilled with a rather strong infill.

Analytical modeling of infills can be carried out according to different macro-modeling approaches. Equivalent strut macro-model approach is the simplest way to model the global interaction between infill and RC elements; see Fig. 13a; it is also explicitly suggested by codes (e.g., ASCE/SEI 41 2007). On the other hand, the characterization of the strut can be made according to different formulations. In literature are available different overviews of the different analytical approaches (e.g., Crisafulli 1997; Chrysostomou and Asteris 2012, among others). Some analytical approaches are able to account for all the failure modes considered above (e.g., Bertoldi et al. 1993), other macro-models does not account explicitly for all the five modes of failure but show a fair agreement with experimental results (e.g., Fardis and Panagiotakos 1997); see Fig. 13b.
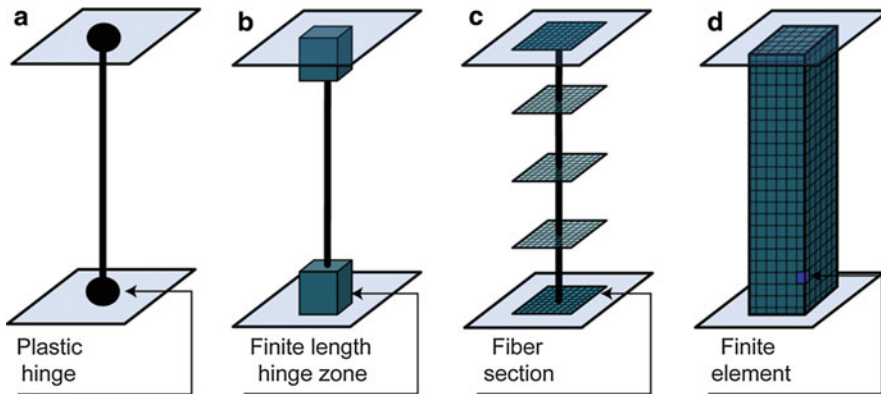
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 14 Idealized models of RC elements (Adapted from Deierlein et al. 2010)

Local interaction between infills and RC frame can be captured by multiple strut macro-models (e.g., Crisafulli 1997) or by single strut macro-models with ends located at the edge of the column or of the beam (e.g., ASCE/SEI 41 2007). Notwithstanding the fact that it can be quite hard to distribute strength among the different struts of the analytical model, still modeling local interaction allows to check if brittle failures are likely to occur in the RC elements (e.g., Verderame et al. 2011).

Analytical Modeling of RC Structures

Inelastic analytical structural component models can be differentiated by the way that plasticity is distributed through the member cross sections and along its length. Modeling approaches can be divided in three main groups: (i) lumped plasticity models, (ii) distributed plasticity models, and (iii) finite element micro-modeling. The main difference between lumped plasticity, distributed plasticity, and finite element models resides in the parameters employed for the definition of inelastic behavior. In the first case, moment–rotation relationships are employed, while, in the other two options, moment–curvature or stress–strain relationships are assumed, and it is necessary to pass through a numerical integration step.

The analysis of experimental behavior of structural elements in section "Behavior of RC Elements and Nonstructural Elements" allowed

the definition of nonlinear structural component models for RC buildings, in terms of M-θ. It was emphasized that component models are generally defined on the basis of experimental results, and often curvature is not the easiest parameter to be employed for the description of all the modes of failures characteristic of RC element. Thus, most commonly, PBEE applications employ lumped plasticity models. On the other hand, it is possible to find PBEE loss and vulnerability assessment applications on RC structures that employ fiber models for low-intensity levels (e.g., Mitrani-Reiser 2007).

In the following, a brief description of both analytical modeling solutions is provided. A more complete overview of modeling approaches is provided in Deierlein et al. (2010). In Fig. 14, the comparison of different idealized model types for simulating the inelastic response of elements is shown.

Lumped plasticity models concentrate the inelastic deformations at the end of the element, such as through a rigid-plastic hinge or an inelastic spring with hysteretic properties (Fig. 14a). By concentrating the plasticity in zero-length hinges with moment–rotation model parameters, these elements have relatively condensed numerically efficient formulations.

Distributed plasticity models are different according to the way the element is discretized, and numerical integration is carried out resulting in finite length hinge models (Fig. 14b) or fiber
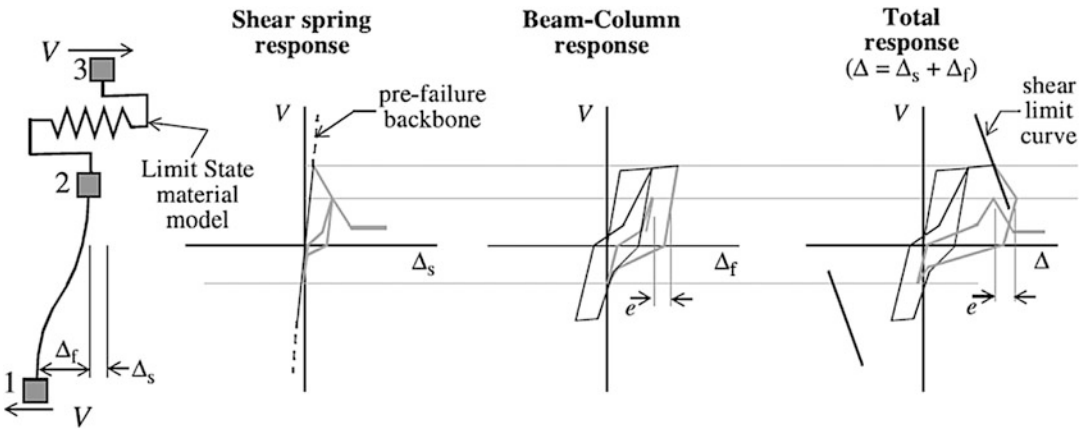
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 15 Shear spring in series model using limit state material model (From Elwood 2004)

model (Fig. 14c). In the finite length hinge model, cross sections in the inelastic hinge zones are characterized through either nonlinear moment–curvature relationships or explicit fiber-section integrations that enforce the assumption that plane sections remain plane. The inelastic hinge length may be fixed or variable, as determined from the moment–curvature characteristics of the section together with the concurrent moment gradient and axial force. The fiber formulation models distribute plasticity by numerical integrations through the member cross sections and along the member length. In this latter case, stress–strain relationship is defined and plane-sections-remain-plane assumption is enforced; see Deierlein et al. (2010). Distributed fiber formulations do not generally report plastic hinge rotations, but instead report strains in concrete cross-section fibers. The calculated strain demands can be quite sensitive to the moment gradient, element length, integration method, and strain hardening parameters on the calculated strain demands. These modeling approaches allow capturing the interaction between axial load and biaxial components of bending moment at section level, under cyclic loading.

Finally, finite element models (Fig. 14d) represent a detailed micro-modeling solution in which material constitutive relationships are assigned to each element. It is obvious that this latter approach represents the most detailed way to represent the elements; on the other hand, it is computationally demanding, and it asks for the definition of numerous input parameters that, in turn, need to be calibrated.

While distributed plasticity formulations capture variations of stress and strain through the section and along the member in more detail, important local behaviors, such as strength degradation due to local buckling of steel reinforcing bars or nonlinear interaction of flexural and shear, are difficult to capture without sophisticated and numerically intensive models. Phenomenological concentrated hinge/spring models may be better suited to capture the nonlinear degrading response of members through calibration (see Fig. 15), using member test data, on experimental moment rotations and hysteresis curves (Deierlein et al. 2010); see Fig. 6. It is worth noting that even when opting for lumped plasticity models, it is possible to refine the analytical modeling choice in different ways, and the accuracy of results is consequently affected; see, for example, Yavari et al. (2009).

Beam–column joint modeling is still at an early and less mature stage with respect to member modeling (i.e., beam and columns), and practical available solutions are to consider rotational springs capable to describe the beam–column joint behavior that is governed by shear and
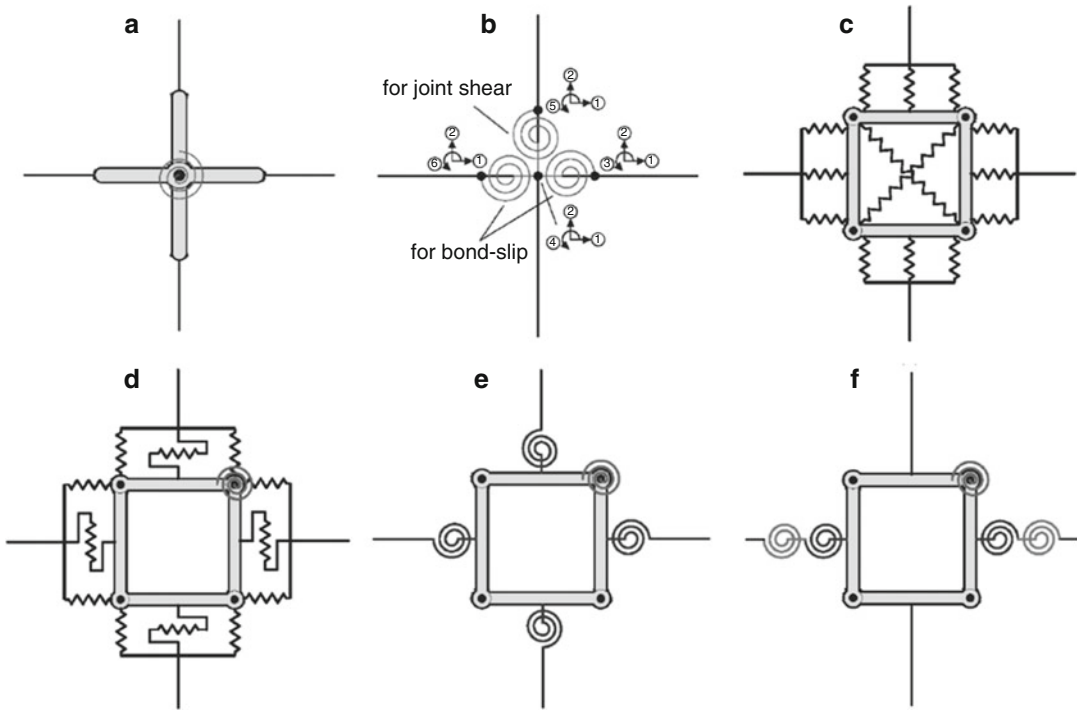
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 16 Existing beam–column joint models: (**a**) Alath and Kunnath (1995), (**b**) Biddah and Ghobarah (1999), (**c**) Youssef and Ghobarah (2001), (**d**) Lowes and Altoontash (2003), (**e**) Altoontash (2004), and (**f**) Shin and LaFave (2004) (From Celik and Ellingwood 2008)

bond–slip deformation; see some modeling scheme examples in Fig. 16 as collected by Celik and Ellingwood (2008). Alath and Kunnath (1995) modeled the joint shear deformation with a rotational spring model with degrading hysteresis (Fig. 16a). Biddah and Ghobarah (1999) modeled the joint shear and bond–slip deformations with separate rotational springs (Fig. 16b). Youssef and Ghobarah (2001) proposed a joint element in which two diagonal translational springs connecting the opposite corners of the panel zone simulate the joint shear deformation, and 12 translational springs at the panel zone interfaced simulate all other modes of inelastic behavior (Fig. 16c). Lowes and Altoontash (2003) proposed a 4-node 12° of freedom joint element that explicitly represents three types of inelastic mechanisms of beam–column joints under reversed cyclic loading (Fig. 16d). Successively Altoontash (2004) simplified the model by Lowes and Altoontash introducing a four

zero-length rotational springs located at beam–column joint interfaces (Fig. 16e). Shin and LaFave (2004) represented the joint by rigid elements located at the edges of the panel zone and rotational springs embedded in one of the four hinges linking adjacent rigid elements (Fig. 16f).

Damage Measures

The nature and amount of structural damage depends necessarily on the quality of the materials that compose structural and nonstructural elements, and on the configuration and type of structural systems. In the early years of modern earthquake engineering, damage definition was basically approached in qualitative terms (e.g., through the definition of probable localization of such damage in a structure). This kind of approach relied fundamentally on the observation of damaged structures after seismic events or it

was based on experimental tests. The first studies for damage quantification of RC structural members date back to late 1980s (e.g., Park and Ang 1985; Fardis et al. 1992). The availability of experimental data allowed more and more detailed approaches to damage analysis up to the most recent PBEE applications.

Most recent studies on damage analysis employ component-based approaches (e.g., Aslani 2005; Mitrani-Reiser 2007). The PBEE damage analysis overpasses the approach to global estimate of damage implicit in previous damage scales (e.g., ATC-13 1985; Grunthal 1998; NIBS 1999) and switch to a component-based damage analysis that allows a more accurate definition of losses. On the other hand, component-based approach requires a significant amount of data, such as experimental data, earthquake experience, expert opinion, and some combination of these.

For component-based approaches, it is necessary to define fragility functions for various damage states for each structural and nonstructural component in typical RC frames. In literature, there are available different benchmark studies for damage assessment of RC buildings (e.g., Aslani 2005; Mitrani-Reiser 2007). FEMA P-58-1 (2012) provides a list of fragility specifications for structural components, nonstructural components, and contents.

The functional form most commonly used for components' fragility functions is the lognormal distribution that asks for the definition of median and logarithmic standard deviation. Therefore, for each component, it is necessary to define median capacity and logarithmic standard deviation (σ_{log}) in terms of EDP for each damage state.

In section "Examples of Damage Analysis", there is a brief overview of component-based damage analyses for RC structural components and infills according to different authors. This overview is aimed at providing an example of damage evaluation made for the elements whose behavior was considered in section "Behavior of RC Elements and Non Structural Elements." Notwithstanding the fact that quantitative approaches to component damage analysis rely on experimental databases, they are still based on a limited number of data, and they cannot be considered fully reliable yet, since they need further numerical refinement.

Finally, in section "Code Limit States," a brief overview of code acceptance criteria for RC elements at different limit states is provided, and compared with the information provided in section "Examples of Damage Analysis." The two sections allow a comparison between next-generation performance-based codes (i.e., section "Examples of Damage Analysis") with respect to current performance-based codes (i.e., section "Code Limit States").

Examples of Damage Analysis

Various damage indices or EDP thresholds are used to quantify damage of RC structural members. As an example, Mitrani-Reiser (2007) employed for beam and columns' damage states the fragility curves developed by Beck et al. (2002) and, in turn, based on Williams and Sexsmith (1997) studies. Beck et al.'s fragility functions use a deformation damage index (DDI) that can be evaluated from chord rotations obtained in structural analyses. This approach allows the definition of four damage states for RC members. On the other hand, Aslani (2005) provides fragility functions for RC columns with light transversal reinforcement in terms of IDR, evaluated from experimental data, and based on the component modeling provided in section "Behavior of RC Elements and Non Structural Elements." For both approaches, four damage states are defined resulting in damage descriptions compared in Table 2. Notwithstanding the fact that Aslani refers specifically to lightly reinforced RC columns (i.e., non-ductile columns), similar damage descriptions are provided in Table 2. In both approaches, the typical lognormal shape is adopted for the fragility curve, so that for each damage state, a median EDP and a σ_{log} are estimated.

Aslani estimates median and σ_{log} for each damage states, based on a database of 92 lightly reinforced columns. The preliminary result of this estimation leads to a significant dispersion for all damage states; so an additional parameter (α) is considered in the regression to reduce the

Seismic Vulnerability Assessment: Reinforced Concrete Structures, Table 2 Comparison of damage states for columns

Williams and Sexsmith (1997) – ductile columns		Aslani (2005) – non-ductile columns	
Damage state	Damage description	Damage state	Damage description
None	None or small number of light cracks, either flexural (90°) or shear (45°)		
Light	Widespread light cracking or a few cracks >1 mm or light shear cracks tending to flatten toward 30°	DS1: light cracking	Visible cracks; crack widths smaller than 0.3 mm. Light repair to improve aesthetic appearance
Moderate	Significant cracking, e.g., 90° cracks >2 mm; 45° cracks >2 mm; 30° cracks >1 mm	DS2: severe cracking	Wider and deeper crack widths, more extensive compared to DS1
Severe	Very large flexure or shear cracks, usually accompanied by limited spalling of cover concrete	DS3: shear failure	Deterioration of shear capacity after yielding (that occurs at earlier stages of loading compared to ductile columns) leads to shear failure
Collapse	Very severe cracking and spalling of concrete; buckling, kinking, or fracture of rebar	DS4: axial failure	Loss of vertical carrying capacity. This damage state has possible disastrous consequences, if there is no possibility to redistribute vertical load to other members
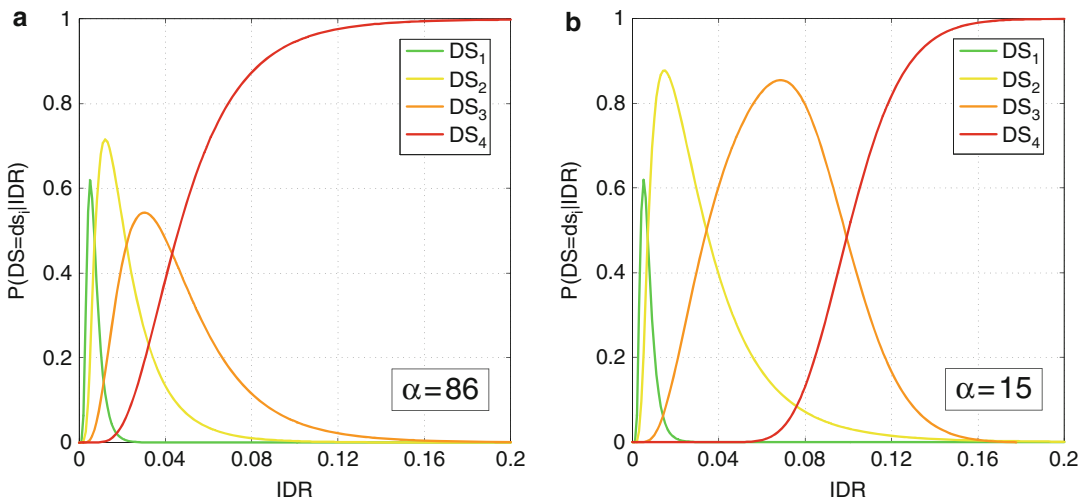
Seismic Vulnerability Assessment: Reinforced Concrete Structures, Fig. 17 Probability of being at each damage state for non-ductile RC columns characterized by different α parameter

resulting $\sigma_{\log}$ of the fragilities and to adjust the estimated median IDR at each damage states. The parameter $\alpha = P/(A_g\rho''f_c)$ is function of axial load (P), gross area section (A_g), concrete compressive strength (f_c), and transversal reinforcement ratio (ρ''). In Fig. 17, the probability of being in a specific damage state for lightly reinforced columns is shown, and the effect of α is emphasized through the comparison of the curves in Fig. 17a, b. It is worth noting that α affects strictly shear failure and axial failure as it can be expected by the observation of the parameters ruling equations shown in section "Behavior of RC Elements and Non Structural Elements."

Seismic Vulnerability Assessment: Reinforced Concrete Structures, Table 3 Damage state definition for interior and exterior beam–column joints according to Aslani (2005)

Damage state	Damage description
DS1: sever cracking in the beam	Wide and extensive cracking of the beam
DS2: severe cracking in the column	Wide and extensive cracking of the column
DS3: severe cracking in the joint	Severe cracking at the beam–column joint
DS4: joint spalling	Spalling of the concrete cover of the beam–column joint
DS5: loss of vertical carrying capacity	The joint collapses under its gravity load

For beam–column connections, Aslani considers different fragility curves for interior beam–column connections and exterior beam column connections. Damage state definitions are identical in the two cases, and they are shown in Table 3, while fragility functions, based on experimental data, can be really different. The numerical estimation for joints is again affected by high uncertainties.

Fragility functions for masonry infill panels are not available in the damage and loss studies cited up to this point. On the other hand, experimental tests (e.g., Mehrabi et al. 1996; Colangelo 2003) and post-earthquake damage observations (e.g., Ricci et al. 2011; De Luca et al. 2014b; Manfredi et al. 2014) have shown that this kind of elements can have a significant impact on both damage and loss analysis of RC structures, especially in the Mediterranean area, in which this kind of structural typology is quite common. Notwithstanding the fact that damage observations based on experimental tests are available in literature and macro-modeling approaches and component modeling backbones as well are available in literature, still characteristic parameters for infill fragilities are characterized by significant differences according to different authors (e.g., Gu and Lu 2005; Özcebe et al. 2012; Colangelo 2013). The point is that in the case of damage analysis of infills, authors do not even agree on damage states' qualitative definitions. The damage degrees are distinguished on the basis of the description of physical damage in terms of cracking, crushing, etc. and the feasibility of repair. As an example, Table 4 provides damage description at different damage states according to different authors; see Colangelo (2013). Significant differences can be found in damage descriptions and, as consequences, also in the fragility function parameters estimated from each description (see also Colangelo 2013).

Despite the significant differences that can be found in literature formulations for component fragilities of infills, it is worth noting that widely employed damage scales for RC structures that interpret global damage of buildings often refer to infill damage. An example is the EMS 98 scale (Grunthal 1998) in which the first three grades of global damage are mainly characterized by damage description of masonry infills; see, for example, De Luca et al. (2014b) and Manfredi et al. (2014), in which infill damage analysis is employed as analytical proxy for damage classification of the whole RC building. On the other hand, EMS 98 damage scale is not a component-based approach, but the single most damaged element represents the damage classification of the whole structure.

Code Limit States

Current PBEE approach is also the result of the huge effort made in ATC-58 project. The purpose of this project is to develop next-generation seismic design procedures that will provide a more reliable means of predicting and designing performances of structures. On the other hand, PBEE has been already implemented in codes providing target performance levels. As an example, FEMA 356 (FEMA 2000) and successively ASCE/SEI 41 (2007) do not attempt to quantify the probability of achieving a given performance level or to quantify losses; but, both FEMA 356 and ASCE/SEI 41 address component-level and system-level damage states, and they relate damage to life safety and post-earthquake operability. The damage states given in FEMA 356 tend to be qualitative and open to multiple interpretations. In Table 5, target performance levels according to FEMA 356 and ASCE/SEI 41 are shown.

ASCE/SEI 41 (2007) succeeded FEMA 356 as document for the seismic assessment of existing

Seismic Vulnerability Assessment: Reinforced Concrete Structures, Table 4 Damage state definition for masonry infills according to Colangelo (2013)

Damage state	Damage description (Colangelo 2013)	Damage description (Gu and Lu 2005)	Damage description (Özcebe et al. 2012)
DS1	Onset of cracking in the bricks, associated with the first noticeable reduction of stiffness	Minor cracking and falling of plaster; only local repair needed with function maintained; maximum strength	Negligible cracks; maximum strength (i.e., base shear)
DS2	Moderate cracks before attaining the maximum strength (i.e., base shear)	Continual diagonal cracking and flaking; repairable damage; 30 % reduction of the maximum strength	Appreciable damage; maximum stress in the infill
DS3	Extensive cracks with tensile splitting and falling of the outer layer of a few bricks; repairable damage	Loss of wall integrity; 70 % reduction of the maximum strength	Ultimate strain in the infill
DS4	So many broken bricks that repair is unreasonable; reconstruction needed		

buildings; it provides a performance-based engineering framework whereby deformation and force demands for different seismic hazards are compared against deformation and force capacities for various performance levels. Elwood et al. (2007) provided an update to concrete provisions in ASCE/SEI 41 given new data available in literature and considering that practitioners observed that previous concrete provisions tend to err on the conservative side. In their update to concrete provisions, Elwood et al. provide modeling parameters and acceptance criteria modifications for columns, beam–column joints, etc. Regarding columns, the classification described in section "RC Members" is considered. As an example, for columns, drift limits at the different acceptance criteria change according to the classification of the expected mode of failure.

European code for the assessment of RC elements (CEN 2005) provides, as collapse prevention acceptance criteria for RC members, the value of the ultimate rotation capacity according to Biskinis and Fardis (2010b); as life safety acceptance criteria, 75 % of the ultimate rotation capacity according to Biskinis and Fardis (2010b); and, as immediate occupancy acceptance criteria, the value of yielding rotation

according to Biskinis and Fardis (2010a). All these chord rotation capacities are intended as member capacity thresholds if shear capacity evaluated according to Biskinis et al. (2004) is not attained, and a preemptive brittle failure (computed through a force-based approach) does not occur before the attainment of such chord rotations.

Regarding acceptance criteria for infills, codes (e.g., CEN 2004), in general, do not provide specific drift limits for infill elements, since it is not taken for granted the fact that infills are explicitly modeled. On the other hand, immediate occupancy and operational acceptance criteria are evaluated to implicitly account for damage to masonry infills.

In FEMA 356 and ASCE/SEI 41, four nonstructural performance levels are classified (see Table 5). The limit for exterior walls differs from the limit for heavy partitions (light partitions may be drywall partitions with studs, for instance) for both occupancy and life safety performance level. For the occupancy level for heavy partitions, the limit is 0.5 %, the same as Eurocode 8 (CEN 2004), while for exterior walls, the limit is 1 %. For life safety level, the above occupancy limits double. The Italian code (DM14/1/2008 2008) also prescribes 0.5 % for

Seismic Vulnerability Assessment: Reinforced Concrete Structures, Table 5 Damage control and building performance level extracted from Table C1-2 in FEMA 356 (FEMA 2000) and ASCE/SEI 41 (2007)

	Collapse prevention level	Life safety level	Immediate occupancy level	Operational level
Overall damage	Severe	Moderate	Light	Very light
General	Little residual stiffness and strength, but load-bearing columns and walls function. Large permanent drifts. Some exits blocked. Infills and unbraced parapets failed or at incipient failure. Building is near collapse	Some residual strength and stiffness left in all stories. Gravity-load-bearing elements function. No out-of-plane failure of walls or tipping of parapets. Some permanent drift. Damage to partitions. Building may be beyond economical repair	No permanent drift. Structure substantially retains original strength and stiffness. Minor cracking of facades, partitions, and ceilings as well as structural elements. Elevators can be restarted. Fire protection operable	No permanent drift. Structure substantially retains original strength and stiffness. Minor cracking of facades, partitions, and ceilings as well as structural elements. All systems important to normal operation are functional
Nonstructural components	Extensive damage	Falling hazards mitigated but many architectural, mechanical, and electrical systems are damaged	Equipment and contents are generally secure but may not operate due to mechanical failure or lack of utilities	Negligible damage occurs. Power and other utilities are available, possibly from standby sources

occupancy, while 0.5 % times 2/3 is prescribed to ensure operation for civil protection. Moreover, with regard to existing buildings, in the commentary to the Italian code (CM 617 2009), both limits are reduced if the analysis model includes the infills. In such a case, the limits become the same as for the masonry buildings, which are equal to 0.3 % and 0.2 % to ensure occupancy and operation, respectively.

Future Challenges and Concluding Remarks

The overview of the seismic assessment framework for RC building provided in this entry has been organized according to the most recent state of the art and guidelines that provide a vision for next-generation codes for RC structures. On the other hand, there are still different aspects on which research and, in turn, code standards should focus.

In particular, behavior of RC elements and their modeling still requires significant experimental and numerical efforts. Behavior of non-ductile RC elements (e.g., characterization of axial load failure) should be better characterized on the basis of higher number of experimental tests in order to better characterize both modeling issues and damage analysis issues in terms of median and logarithmic standard deviation. Same considerations can be made on beam–column joints and masonry infills.

Finally, masonry infills represent a relevant issue for all countries in the Mediterranean area in which this kind of construction practice is very common. Despite the huge efforts made in the last decades, the characterization of nonstructural component modeling and damage is still challenging considering that they are not subjected to the same design and control process to which new RC elements are subjected to.

Performance evaluation of RC structure according to the methodological framework of FEMA P-58-1 (2012) provides a significant enhancement toward a more reliable performance evaluation for stakeholders and decision makers. Still, some procedures need to be codified in simpler and more user-friendly tools to be employed by practitioners and professional

engineers in any structural performance evaluations as current and consolidated practice.

Summary

The framework in which modern design and assessment of reinforced concrete structures are placed is the well-known Performance-Based Earthquake Engineering (PBEE) framework. The most recent methodological organization of PBEE tools is represented by FEMA P-58-1 document. The brief analysis of the main steps of PBEE provides the basis to introduce where, in such framework, the specific structural material – in this case RC – plays a role. In particular, when the relationship between intensity measure and engineering demand parameter arises in PBEE, it is the phase in which it is necessary to focus on the behavior of RC structural elements, in order to approach analytical modeling, and quantify components damage measures.

Seismic vulnerability assessment has to account for all the typical modes of failure that RC elements can show. Thus, behavior of RC elements (e.g., beams, columns, beam–column joints) is described considering the case of existing elements representing the most general situation for vulnerability assessment problems. The behavior of masonry infills is also considered, given the significant effect they can have on vulnerability in terms of strength and stiffness increase, and in terms of occurrence of brittle failures caused by local interaction of infill and RC frame. Critical review of behavior of RC structures and, in turn, of RC elements, allows emphasizing future challenges and needs that the scientific community and practitioners should take on in the next years.

Cross-References

References

ACI 352R-02 (2002) Recommendations for design of beam-column connections in monolithic reinforced concrete structures. American Concrete Institute, Farmington Hills

Alath S, Kunnath SK (1995) Modeling inelastic shear deformations in rc beam-column joints. In: Engineering mechanics proceedings of 10th conference, University of Colorado at Boulder, Boulder, 21–24 May, vol 2. ASCE, New York, pp 822–825

Alire DA (2002) Seismic evaluation of existing unconfined RC beam–column joints. MSc thesis, University of Washington

Altoontash A (2004) Simulation and damage models for performance assessment of reinforced concrete beam-column joints. PhD dissertation, Department of Civil and Environment Engineering, Stanford University, Stanford

American Society of Civil Engineers (ASCE) (2007) Seismic rehabilitation of existing buildings, ASCE/SEI 41-06. American Society of Civil Engineers, Reston, VA

Anderson M, Lehman D, Stanton J (2008) A cyclic shear stress-strain model for joints without transverse reinforcement. Eng Struct 30:941–954

ATC Applied Technology Council (1985) ATC-13, earthquake damage evaluation data for California. Applied Technology Council, Redwood City, CA, pp 492

Aschheim AM, Moehle JP (1992) Shear strength and deformability of RC bridge columns subjected to inelastic displacements. Technical Report No UCB/EERC 92/04, University of California at Berkeley, Berkeley

Aslani H (2005) Probabilistic earthquake loss estimation and loss disaggregation in buildings. PhD thesis, Stanford University

Asteris PG, Kakaletsis DJ, Chrysostomou CZ, Smyrou EE (2011) Failure modes in infilled frames. Electron J Struct Eng 11(1):11–20

Bakir PG, Boduroğlu HM (2002) A new design equation for predicting the joint shear strength of monotonically loaded exterior beam-column joints. Eng Struct 24:1105–1117

Beck JL, Porter KA, Shaikhutdinov R, Au SK, Moroi T, Tsukada Y, Masuda M (2002) Impact of seismic risk on lifetime property values, final report for CUREE-Kajima phase IV project. Consortium of Universities for Research in Earthquake Engineering, Richmond

S

Berry M, Parrish M, Eberhard M (2004) PEER structural performance database user's manual. Pacific Engineering Research Center, University of California, Berkeley, 38 pp. Available at http://nisee.berkeley.edu/spd/ and http://maximus.ce.washington.edu/~peera1/. Accessed 23 Feb 2014

Bertoldi SH, Decanini LD, Gavarini C (1993) Telai tamponati soggetti ad azioni sismiche, un modello semplificato: confronto sperimentale e numerico. In: Proceeding of 6° Convegno Nazionale ANIDIS, vol 2, 13–15 Perugia, Italy, 815–824 (in Italian)

Biddah A, Ghobarah A (1999) Modelling of shear deformation and bond slip in reinforced concrete joints. Struct Eng Mech 7(4):413–432

Biskinis D, Fardis MN (2010a) Deformations at flexural yielding of members with continuous or lap-spliced bars. Struct Concr 11(3):127–138

Biskinis D, Fardis MN (2010b) Flexure-controlled ultimate deformations of members with continuous or lap-spliced bars. Struct Concr 11(2):93–108

Biskinis DE, Roupakias GK, Fardis MN (2004) Degradation of shear strength of reinforced concrete members with inelastic cyclic displacement. ACI Struct J 101(6):773–783

Bousias SN (1993) Experimental and analytical study of RC columns in cyclic biaxial bending with axial force. Doctoral thesis, Civil Engineering Department, University of Patras, Patras

Bousias SN, Panagiotakos TB, Fardis MN (2002) Modelling of RC members under cyclic biaxial flexure and axial force. J Earthq Eng 6(3):213–238

Calvi GM, Bolognini D, Penna A (2004) Seismic performance of masonry-infilled RC frames–benefits of slight reinforcements. Invited lecture to "Sísmica 2004 – 6° Congresso Nacional de Sismologia e Engenharia Sísmica", Guimarães, 14–16 Apr

Calvi GM, Pinho R, Magenes G, Bommer JJ, Restrepo-Vélez LF, Crowley H (2006) Development of seismic vulnerability assessment methodologies over the past 30 years. ISET J Earthq Technol Paper No 472 43(3):75–104

CEB (1996) RC frames under cyclic loading. T. Telford, London, for Comite Euro-international du beton, Lausanne

Celik OC, Ellingwood BR (2008) Modeling beam–column joints in fragility assessment of gravity load designed reinforced concrete frames. J Earthq Eng 12(3):357–381

CEN (2004) EN 1998-1 Eurocode 8: design of structures for earthquake resistance-part 1: general rules, seismic actions and rules for buildings. European Committee for Standardization, Brussels

CEN (2005) Eurocode 8: design of structures for earthquake resistance – part 3: assessment and retrofitting of buildings. European Standard EN 1998-1:2005-Comité Européen de Normaliation, Brussels

Chrysostomou CZ, Asteris PG (2012) On the in-pane properties and capacities of infilled frames. Eng Struct 41:385–402

CM 617 (2009) Istruzioni per l'applicazione delle nuove norme tecniche per le costruzioni. Ministero delle Infrastrutture e dei Trasporti, Rome (in Italian)

Coburn A, Spence R (2002) Earthquake protection, 2nd edn. Wiley, Chichester

Colangelo F (2003) Experimental evaluation of member-by-member models and damage indices for infilled frames. J Earthq Eng 7(1):25–50

Colangelo F (2013) Drift-sensitive non-structural damage to masonry infilled reinforced concrete frames designed to Eurocode 8. Bull Earthq Eng. doi:10.1007/s10518-013-9503-y

Cornell A, Krawinkler H (2000) Progress and challenges in seismic performance assessment. PEER News, Apr

Crisafulli FG (1997) Seismic behaviour of reinforced concrete structures with masonry infills. PhD thesis, University of Canterbury, Christchurch

De Luca F, Verderame GM (2013) A practice-oriented approach for the assessment of brittle failures in existing RC elements. Eng Struct 48:373–388

De Luca F, Ameri G, Iervolino I, Bindi D (2014a) Toward validation of simulated accelerograms via prediction equations for nonlinear SDOF response. Boll Geofis Teor Appl 55(1):81–101

De Luca F, Verderame GM, Gómez-Martínez F, Pérez-García A (2014b) The structural role played by masonry infills on RC building performances after the 2011 Lorca, Spain, earthquake. Bull Earthq Eng 12(5):1999–2026

Deierlein GG, Krawinkler H, Cornell CA (2003) A framework for performance-based earthquake engineering. In: Proceedings of 2003 pacific earthquake conference on earthquake engineering, University of Canterbury, Christchurch

Deierlein GG, Reinhorn AM, Willford MR (2010) Nonlinear structural analysis for seismic design, NEHRP Seismic Design Technical Brief No. 4, produced by the NEHRP Consultants Joint Venture, a partnership of the Applied Technology Council and the Consortium of Universities for Research in Earthquake Engineering, for the National Institute of Standards and Technology, Gaithersburg, NIST GCR 10-917-5

Di Ludovico M, Verderame G, Prota A, Manfredi G, Cosenza E (2013) Experimental behavior of nonconforming RC columns with plain bars under constant axial load and biaxial bending. J Struct Eng 139(6):897–914

DM 14/1/2008(2008) Norme tecniche per le costruzioni. Ministero delle Infrastrutture e dei Trasporti, Rome (in Italian)

Dolšek M, Fajfar P (2001) Soft storey effects in uniformly infilled reinforced concrete frames. J Earthq Eng 5(1):1–12

Elwood K (2004) Modelling failures in existing reinforced concrete columns. Can J Civ Eng 31:846–859

Elwood KJ, Eberhard M (2009) Effective stiffness of reinforced concrete columns. ACI Struct J 106(4):476–484

Elwood K, Moehle JP (2005a) Drift capacity of reinforced concrete columns with light transverse reinforcement. Earthq Spectra 21(1):71–89

Elwood K, Moehle JP (2005b) Axial capacity model for shear-damaged columns. ACI Struct J 106(S45):578–587

Elwood KJ, Matamoros AB, Wallace JW, Lehman DE, Heintz JA, Mitchell AD, Moore MA, Valley MT, Lowes LN, Comartin CD, Moehle JP (2007) Update to ASCE/SEI 41 concrete provisions. Earthq Spectra 23(3):493–523

Fardis MN (2009) Seismic design, assessment and retrofitting of concrete buildings based on EN-Eurocode 8. Springer, Dordrecht

Fardis MN, Panagiotakos TB (1997) Seismic design and response of bare and infilled reinforced concrete buildings – part II: infilled structures. J Earthq Eng 1(3):473–503

Fardis MN et al (1992) Damage measures and failure criteria during seismic response. In: Cooperative research program on the seismic response of reinforced concrete structures. First phase, Final report, II, National Laboratory for Civil Engineering, Lisbon

Federal Emergency Management Agency (FEMA) (1997) FEMA 273, guidelines for the seismic rehabilitation. Federal Emergency Management Agency, Washington, DC

Federal Emergency Management Agency (FEMA) (2000) FEMA 356, prestandard and commentary for the seismic rehabilitation of buildings. Federal Emergency Management Agency, Washington, DC

Fédération internationale du béton (FIB). Task Group 7.1 (2003) Seismic assessment and retrofit of reinforced concrete buildings: state-of-art report (vol 24). International Federation for Structural Concrete (ed). FIB-Féd. Int. du Béton

Federal Emergency Management Agency (FEMA) P-58-1 (2012) Seismic performance assessment of buildings volume 1 – 558 methodology. Federal Emergency Management Agency, Washington, DC

Goulet CA, Haselton CB, Mitrani-Reiser J, Beck JL, Deierlein GG, Porter KA, Stewart JP (2007) Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building – from seismic hazard to collapse safety and economic losses. Earthq Eng Struct Dyn 36(13):1973–1997

Grunthal G (ed) (1998) European Macroseismic Scale 1998 (EMS-98), Cahiers du Centre Européen de Géodynamique et de Séismologie, 15, Luxembourg

Gu X, Lu Y (2005) A fuzzy–random analysis model for seismic performance of framed structures incorporating structural and non-structural damage. Earthq Eng Struct Dyn 34(10):1305–1321

Hakuto S, Park R, Tanaka H (2000) Seismic load tests on interior and exterior beam-column joints with substandard reinforcing details. ACI Struct J 97(1):11–25

Haselton CB, Liel AB, Taylor Lange S, Deierlein GG (2008) Beam-column element model calibrated for predicting flexural response leading to global collapse of RC frame buildings, PEER report 2007/03

Hassan WM (2011) Analytical and experimental assessment of seismic vulnerability of beam-column joints without transverse reinforcement in concrete buildings. PhD dissertation, University of California, Berkeley

Hwang S, Lee H (1999) Analytical model for predicting shear strengths of exterior reinforced concrete beam-column joints for seismic resistance. ACI Struct J 96(5):846–858

Ibarra LF, Medina RA, Krawinkler H (2005) Hysteretic models that incorporate strength and stiffness deterioration. Earthq Eng Struct Dyn 34:1489–1511

Kim J, LaFave JM (2007) Key influence parameters for the joint shear behavior of reinforced concrete (RC) beam-column connections. Eng Struct 29:2523–2539

Krawinkler H (2002) A general approach to seismic performance assessment. In: Proceedings of international conference on advances and new challenges in earthquake engineering research, ICANCEER, Hong Kong, 19–20 Aug

Krawinkler H, Miranda E (2004) Performance-based earthquake engineering, chapter 9. In: Bozorgnia Y, Bertero VV (eds) Earthquake engineering: from engineering seismology to performance-based engineering. CRC Press, Boca Raton

LaFave M, Shin M (2005) Discussion of 'Modeling reinforced-concrete beam-column joints subjected to cyclic loading', by Lowes, L. N. and Altoontash, A. ASCE J Struct Eng 131(6):992–993

Legeron F, Paultre P (2000) Behavior of high-strength concrete columns under cyclic flexure and constant axial load. ACI Struct J 97(4):591–601

Lowes LN, Altoontash A (2003) Modeling reinforced-concrete beam-column joints subjected to cyclic loading. ASCE J Struct Eng 129(12):1686–1697

Manfredi G, Prota A, Verderame GM, Ricci P, De Luca F (2014) 2012 Emilia earthquake, Italy: reinforced concrete buildings response. Bull Earthq Eng 12(5):2275–2298

McGuire RK (2004) Seismic hazard and risk analysis. Report MNO-10. Earthquake Engineering Research Institute Publication, Oakland

Mehrabi AB, Shing PB, Schuller MP, Noland JL (1996) Experimental evaluation of masonry-infilled RC frames. J Struct Eng ASCE 122(3):228–237

Mitrani-Reiser J (2007) An ounce of prevention: probabilistic loss estimation for performance-based earthquake engineering. PhD dissertation, California Institute of Technology (Caltech), Pasadena

Moehle JP (2003) A framework for performance-based earthquake engineering. In: Proceedings of ATC-15-9 workshop on the improvement of building structural design and construction practices, Maui, June

Moehle JP, Deierlein GG (2004) A framework for performance-based earthquake engineering.

S

In: Proceedings of 13th world conference on earthquake engineering, Paper No 679, Vancouver

Mosalam KM (1996) Experimental and computational strategies for the seismic behavior evaluation of frames with infill walls. PhD dissertations, Cornell University, Ithaca

National Institute of Building Science (NIBS) (1997, 1999 and 2002) Earthquake loss estimation methodology. HAZUS. Technical manuals, vols 1, 2, 3. Federal Emergency Management Agency (FEMA), Washington, DC. http://www.fema.gov/hazus/

Özcebe S, Crowley H, Bal IE (2012) Distinction between no and slight damage states for existing RC buildings using a displacement-based approach. In: Proceedings of the 15th world conference on earthquake engineering, paper 5126, 24–28, Lisbon, Portugal

Panagiotakos TB, Fardis MN (2001) Deformations of RC members at yielding and ultimate. ACI Struct J 98(2):135–148

Pantelides CP, Hansen J, Nadauld J, Reaveley LD (2002) Assessment of reinforced concrete building exterior joints with substandard details, PEER 2002/18, May 2002

Park R (2002) A summary of results of simulated seismic load tests on RC beam–column joints, beam and columns with substandard reinforcing details. J Earthq Eng 6(2):147–174

Park YJ, Ang AMS (1985) Mechanistic seismic damage model of reinforced concrete. ASCE J Struct Eng 111:722–739

Park S, Mosalam KM (2012) Parameters for shear strength prediction of exterior beam–column joints without transverse reinforcement. Eng Struct 36:198–209

Park S, Mosalam KM (2013) Experimental investigation of nonductile RC corner beam-column joints with floor slabs. ASCE J Struct Eng 139(1):1–14

Porter KA (2003) An overview of PEER's performance-based earthquake engineering methodology. In: Conference on Applications of Statistics and Probability in Civil Engineering (ICASP9), Civil Engineering Risk and Reliability Association (CERRA), San Francisco, 6–9 July

Priestley MJN, Verma R, Xiao Y (1994) Seismic shear strength of reinforced concrete columns. J Struct Eng 120(8):2310–2329

Pujol S (2002) Drift capacity of reinforced concrete columns subjected to displacement reversals. PhD thesis, Purdue University

Ricci P, De Luca F, Verderame GM (2011) 6th April 2009 L'Aquila earthquake, Italy: reinforced concrete building performance. Bull Earthq Eng 9(1):285–305

Ricci P, Verderame G, Manfredi G (2013) ASCE/SEI 41 provisions on deformation capacity of older-type reinforced concrete columns with plain bars. ASCE J Struct Eng 139(12):04013014

Sezen H, Moehle JP (2004) Shear strength model for lightly reinforced concrete columns. ASCE J Struct Eng 130(11):1692–1703

Shin M, LaFave JM (2004) Testing and modeling for cyclic joint shear deformations in rc beam-column connections. In: Proceedings of the thirteenth world conference on earthquake engineering, Vancouver, 1–6 Aug 2004, Paper No 0301

Shing PB, Mehrabi AB (2002) Behaviour and analysis of masonry-infilled frames. Prog Struct Eng Mater 4:320–331

Tothong P, Cornell CA (2006) An empirical ground motion attenuation relation for inelastic spectral displacement. Bull Seismol Soc Am 96:2146–2164

Tothong P, Luco N (2007) Probabilistic seismic demand analysis using advanced ground motion intensity measures. Earthq Eng Struct Dyn 36:1837–1860

Vecchio FJ, Collins MP (1986) The modified compression-field theory of reinforced concrete elements subjected to shear. ACI Struct J 83(2):219–231

Verderame GM, Ricci P, Manfredi G, Cosenza E (2010) Ultimate chord rotation of RC columns with smooth bars: some considerations about EC8 prescriptions. Bull Earthq Eng 8(6):1351–1373

Verderame GM, De Luca F, Ricci P, Manfredi G (2011) Preliminary analysis of a soft-storey mechanism after the 2009 L'Aquila earthquake. Earthq Eng Struct Dyn 40(8):925–944

Vollum RL (1998) Design and analysis of exterior beam column connections. PhD dissertation, Imperial College of Science Technology and Medicine-University of London

Vollum RL, Newman JB (1999) Strut and tie models for the analysis/design of external beam-column joints. Mag Concr Res 51(6):415–425

Walker SG (2001) Seismic performance of existing RC beam–column joints. MSc thesis, University of Washington

Watanabe F, Ichinose T (1992) Strength and ductility of RC members subjected to combined bending and shear. In: Concrete shear in earthquake. Elsevier Applied Science, New York, pp 429–438

Williams MS, Sexsmith RG (1997) Seismic assessment of concrete bridges using inelastic damage analysis. Eng Struct 19(3):208–216

Wong HF (2005) Shear strength and seismic performance of non-seismically designed reinforced concrete beam-column joints. PhD dissertation, Department of Civil Engineering, The Hong Kong University of Science and Technology, August

Yavari S, Elwood KJ, Wu C (2009) Collapse of a nonductile concrete frame: evaluation of analytical models. Earthq Eng Struct Dyn 38(2):225–241

Youssef M, Ghobarah A (2001) Modelling of RC beam-column joints and structural walls. J Earthq Eng 5(1):93–111

Zhu L, Elwood K, Haukaas T (2007) Classification and seismic safety evaluation of existing reinforced concrete columns. J Struct Eng 133(9):1316–1330

Seismometer Arrays

Johannes Schweitzer
NORSAR, Kjeller, Norway

Synonyms

Arrays; Beamforming; fk-analysis; Seismic arrays

Introduction

Definition and Purpose

In seismology, the term "array" (of seismometers) has been known for more than 50 years. Today this phrase is often used for any group of seismic stations in the sense as defined in Macmillan's online dictionary, where an array is described as "a number of pieces of equipment of the same type, connected together to do a particular job" (Macmillan 2013). In this chapter, the term "seismometer arrays" is used with a more restricted definition, as it was originally introduced to seismology. A seismometer array can be defined as (Schweitzer et al. 2012): "A seismic array is a set of seismometers deployed so that characteristics of the seismic wavefield at a specified reference point, within or close to the array, can be inferred by analyzing the waveforms recorded at the different sites. A seismic array differs from a local network of seismic stations mainly by the techniques used for data analysis. Thus, in principle, a network of seismic stations can be used as an array, and data from an array can be analyzed as data from a network. The size of an array is defined by its aperture, which is the largest horizontal distance between two sensors of the array. In practice, the geometry and the number of seismometer sites of an array are determined by the intended scientific purpose and economic limits." The main purpose to install a seismic array is threefold. At first, an array can be steered as an antenna to amplify the signals of interest by stacking (summing) the seismic wave energy recorded at the different array sites after applying appropriate phase delays. With these so-called "beamforming" techniques, arrays show superior signal detection capabilities with respect to single 3-component (3C) seismic stations. The second main purpose to install a seismic array is the capability to estimate the station-to-seismic source azimuth (backazimuth, BAZ) and the apparent velocity of unknown seismic signals crossing the array. Thirdly, a single seismic array often provides enough information about the incoming seismic signals that an automatic algorithm can be used to estimate a first (preliminary) solution of the seismic source location.

History of Seismometer Arrays

The history of seismometer array installations is going back to the 1950s. At that time, the idea of installing arrays of sensors to improve the signal-to-noise ratio (SNR) of a seismic onset was adopted from radio astronomy, radar, acoustics, and sonar. Since then, classified arrays deployed to monitor nuclear test activities at teleseismic distances have been built worldwide. In 1958, "The Conference of Experts to study the methods of detecting violations of a possible agreement on the suspension of nuclear tests" was held under the auspices of the United Nations in Geneva. This conference was followed up by several initiatives for improving the quality of seismic stations worldwide. Many of these classified arrays became known in the 1990s and are today part of the International Monitoring System (IMS) for the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) (see e.g., Douglas 2002; Dahlman et al. 2009) as primary or auxiliary stations, e.g., AKASG (Malin, Ukraine), ASAR (Alice Springs, Australia), BRTR (Keskin, Turkey), CMAR (Chiang Mai, Thailand), ESDC (Sonseca, Spain), ILAR (Eielson, Alaska), and KURK (Kurchatov, Kazakhstan). Many of these arrays have quite diverse geometries (Fig. 1) and in some cases comprise different installations for short-period (frequencies above 0.5 Hz) and long-period (frequencies below 0.1 Hz) signals as, e.g., the Belbaşi array nearby the Keskin array. Similar to arrays of seismometers, the
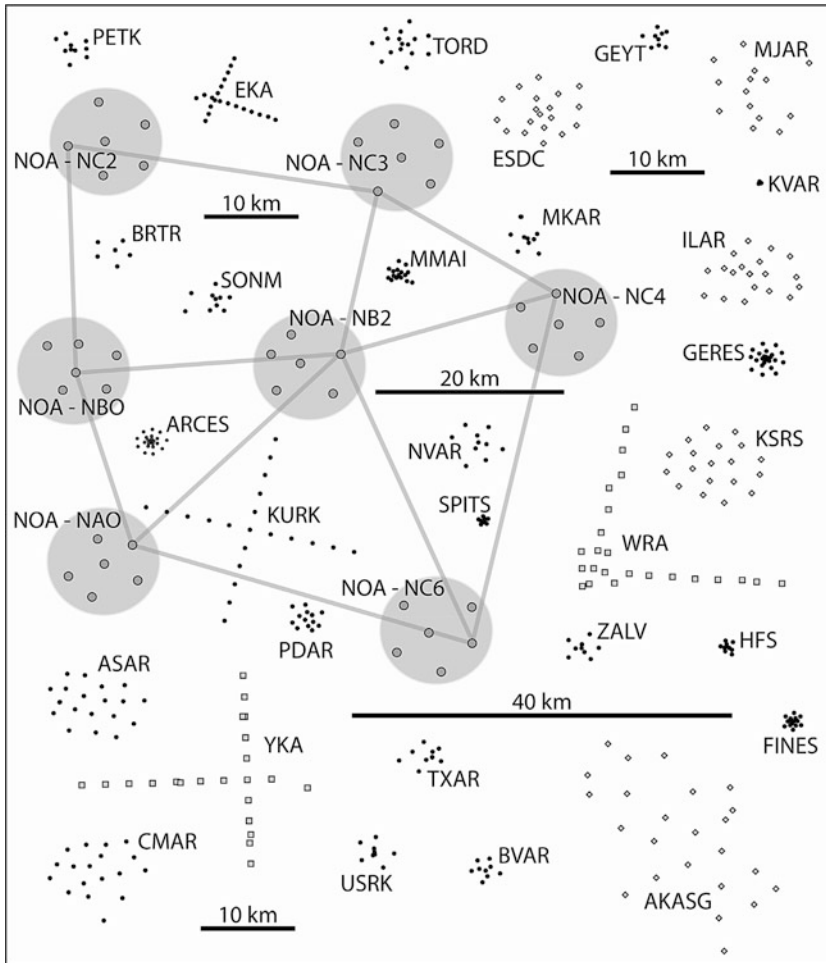
Seismometer Arrays, Fig. 1 Seismic arrays equipped with short-period or broadband seismic sensors in operation in December 2013 as part of the International Monitoring System (IMS) for the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO). The code name with which it is registered in the international registry of seismic stations at the International Seismological Centre (http://www.isc.ac.uk/registries/) is provided for each array. All array maps are plotted in the same scale (Courtesy of S. J. Gibbons, NORSAR)

IMS includes also arrays of infrasound sensors and hydrophones, and similar data analysis techniques are applied as for data from seismic arrays.

To the best of knowledge, the first experimental seismic array with more than four elements and openly available data was established in February 1961 by the United Kingdom Atomic Energy Agency (UKAEA) on the Salisbury Plain (UK), followed in December 1961 by the Pole Mountain array (PMA, Wyoming, USA), in June 1962 by the Eskdalemuir array (EKA, Scotland, UK), and in December 1963 by the Yellow-knife array (YKA, Canada). This array type

(known as UK array) is orthogonal linear or L-shaped, with apertures between 10 and 25 km. Later, arrays of the same type were built in Australia (Warramunga), Brazil (Brasilia), and India (Gauribidanur).

In the 1960s, arrays with very different apertures and geometries were tested, from small circular ones with apertures of a few kilometers to huge arrays with apertures of up to 200 km. The largest arrays were the Large Aperture Seismic Array (LASA) in Montana (USA), opened in 1965 and in operation until 1978 with 525 seismometer sites, and the original Norwegian

Seismic Array (NORSAR) in southern Norway, consisting of 132 sites over an aperture of approximately 100 km with altogether 198 seismometers, which became fully operational in spring 1971. The original NORSAR array was reduced in 1976 to seven subarrays and was assigned the new code name NOA.

LASA, NORSAR, and the UKAEA arrays had narrowband short-period seismometers (for signal frequencies around 1 Hz) at all sites and additional long-period seismometers (for signal periods around 20 s) at selected sites in their original configurations. In Germany, a new array type was planned and installed in the 1970s with an aperture of about 110 km. The Gräfenberg-Array (GRF) was installed on the limestone plateau of the Franconian Jura as the world's first seismometer array equipped entirely with broadband sensors (for frequencies between 0.01 and 8 Hz). Since then, the short-period and long-period sensors of many arrays were or will be in the near future exchanged with broadband seismometers.

In the 1980s, the geometry of the so-called regional arrays was developed. This is often called a NORES-type array design and has array sites located on concentric rings (each with an odd number of sites) spaced at log-periodic intervals. It is now used for the design of most modern, small aperture arrays; only the number of rings and the aperture (diameter of the outermost ring of sites) differ from installation to installation.

Another approach to seismometer arrays was developed in the 1990s. In parts of the world, the networks of seismometer stations became so dense that data from these single station networks could be combined and analyzed as data from a seismic array. Examples are the J-array in Japan, the German Regional Seismic Network, the Californian array, the Kyrgyz Network (K-Net), and the ongoing USArray project with semi-temporary stations. Most of the known array techniques could be applied to analyze data from these and other networks.

Array installations demonstrated during the last decades of the twentieth century that seismic arrays could facilitate detection and characterization of seismic signals that was superior to that of single three-component (3C) stations. Today, many of the seismic stations of the IMS for the CTBTO are arrays (e.g., Dahlman et al. 2009).

During the last two decades, temporary, very small aperture arrays (apertures usually below 1 km) were used to investigate the distribution of (mostly) S-wave velocities below and nearby the array installation. Knowing S-wave velocities of the uppermost layers is essential for seismic hazard mitigation; for details see, (e.g., Schweitzer et al. 2012).

Further details about different array configurations can be found in the literature (e.g., Barber 1958; Haubrich 1968; Harjes and Henger 1973; Mykkeltveit et al. 1983). As examples of seismic arrays, Fig. 1 shows maps of all primary and auxiliary IMS arrays in operation in December 2013; note the huge variety in geometries, number of sites, and apertures.

Basics of Array Seismology

Basic Requirements for Seismometer Arrays

The observed apparent-velocity range and the dominant frequency content of seismic signals are quite different for different types of signals (i.e., local, regional, or teleseismic observations, body waves, or surface waves). Thus, the geometry, aperture, and instrumentation of seismic arrays have to be adjusted with respect to the scientific requirements of an array installation.

Proper analysis of array data depends on a stable, high-precision relative timing of all array elements. This is required because most of the parametric information calculated using an array involves the measurement of (usually very small) time differences (phase shifts) between the seismic signals recorded at the different sensors. Most array-data analysis algorithms assume that seismic energy crosses an array as a plane wave. From the theory of signal processing, it is known that signals can be constructively summed up as long as the time shift between the signals is not larger than about ¼ of the dominant signal period T. This rule applies also to the case of

S

seismic signals observed with an array. Many different effects can influence the arrival time of seismic onsets and thereby disturb the plane-wave approximation and should be taken in account when analyzing array data:

- Seismic waves usually propagate with spherical (in the case of body waves) or circular geometry (in the case of surface waves). Therefore, the aperture of an array should be small enough that the plane-wave approximation is still valid, i.e., the (theoretical) arrival-time differences with respect to the plane-wave approximation should be so small that they do not influence the analysis results.
- Since the fundamental work of Hans Benndorf (1870–1953) and Emil Wiechert (1861–1921) during the first decade of the twentieth century, it is known that in the case of a body wave, there is a direct relation between the observed (apparent) propagation velocity of a seismic wave, the ray parameter (the derivative of the travel-time curve), and the seismic velocity at the turning point of the observed seismic wave within the Earth. With an array of seismometers, this (apparent) velocity can be directly measured as long as all sensors are located in one horizontal plane.

  In the case that the sensors are not located on one horizontal plane, the theoretical time shifts between the array sensors do not only depend on the propagation direction and (apparent) velocity of the plane wave but also on the wavelength-dependent seismic velocities below the array sites. If travel-time effects due to array topography become larger than about one fourth of the signal period, it becomes important for the accuracy of data analysis results to take this effect in account. However, analysis algorithms then become rather complicated, and if one plans for a new array, one should try to avoid this additional complexity and locate all seismometer sites on one horizontal plane.
- The seismic wavefield is always influenced by lateral heterogeneities in the Earth, which can disturb the plane-wave approximation due to velocity inhomogeneities. In particular local

lateral heterogeneities can have large effects. To avoid this, many arrays are built on homogeneous geological units, but this is not everywhere possible, and uniform geology at the Earth's surface does not guarantee that deeper structures are similarly homogeneous. Therefore, it is necessary to investigate recorded data from each seismic array for systematic effects which may disturb the plane-wave approximation.

- There are many examples of how timing-system errors at single array sites can produce difficulties for array-data analysis. However, stable and correct relative timing can be achieved by installing at the array a central clock, which distributes a common time signal to all digitizers at all array sites. Even in the case that this central timing system fails and has some offset to absolute time, all array-analysis algorithms can still be applied. So, whenever it is possible, a central timing system should be considered for new array installations.

The previously mentioned "beamforming" techniques delay (phase shift) and stack the signals from the different sensors. Due to constructive interference of the signals, the SNR is then enhanced, whereas the (random and uncorrelated) background noise is suppressed. One can show that a seismic array can theoretically improve the SNR by a factor of $\sqrt{M}$, where M is the number of array elements. Figure 2 shows an example from the GERES array in Germany. The tiny onset of a P phase, which was reflected at the Earth's core, was recorded with vertical sensors at 20 sites of the array. Figure 2 shows these data in blue and on top the array beam in red. On many single traces the signal is hardly discernible, but it is clearly visible on the beam trace due to the drastically improved SNR. To achieve the theoretical factor of $\sqrt{M}$, most array-data processing techniques require not only high clock stability of the recording systems at the different array sites and best knowledge about possible deviations from the plane-wave approach but also high signal coherency across the array.

Waveforms of seismic signals can be influenced by interference with wave energy
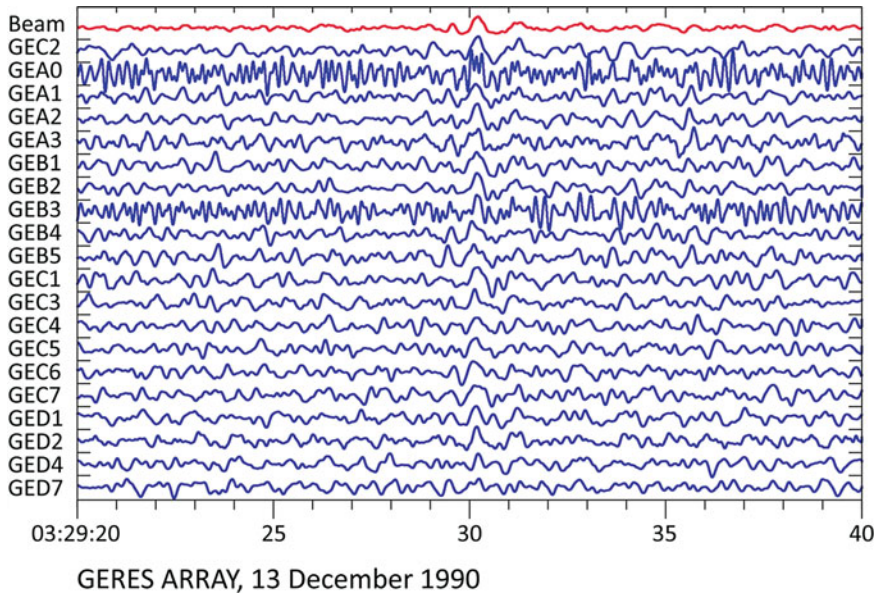
Seismometer Arrays, Fig. 2 Vertical seismograms (*blue*) with the onset of a tiny P onset reflected from the Earth's core (PcP) and recorded at 03:29:29.5 on 13 December 1990 with the GERES array after an earthquake in the Tyrrhenian Sea at an epicentral distance of about 1,070 km. The top trace (*red*) shows the array beam of the shown vertical records after applying "delay-and-sum" processing. All data are Butterworth band-pass filtered between 0.8 and 4 Hz and equally scaled

scattered at heterogeneities along the entire ray path, and as a result, waveforms observed at different seismometer sites may significantly differ. Since high signal coherency is required for seismic arrays, such waveform-altering effects should be minimized by avoiding array locations close to known lateral inhomogeneous structures. This constitutes an additional reason to prefer installing seismic arrays at one relatively homogeneous geological unit, almost transparent to seismic waves. Since such local site effects are signal frequency dependent, signal coherency over the array also becomes frequency dependent, and this may impose constraints on array geometry, spatial extent, and instrumentation.

As mentioned earlier, seismometer arrays can provide estimates of the station-to-event azimuth (backazimuth, BAZ) and the apparent velocity of seismic signals. These estimates are important both for event location purposes and for signal identification and classification, e.g., as P, S, local, regional, or teleseismic phases. The slowness resolution of an array – i.e., how accurately

the propagation direction and the apparent velocity of a wave front can be measured – improves with increasing array aperture. However, the apparent velocity of a seismic onset changes with epicentral distance and the shape of a seismic waveform can change drastically due to interference with other seismic phases; in addition, signal coherency diminishes with increasing sensor separation. Therefore, when building an array, one has to find a balance between coherency and theoretical slowness resolution.

Finally, some remarks about the instrumentation of arrays. Traditionally, arrays were equipped with vertical sensors at all sites and only a few sites with additional horizontal sensors. This was mostly related to the additional costs for more sensors, data transmission, and data storage. However, many studies have shown that arrays equipped with more 3C sensors are superior in analyzing S-type onsets. Since costs for data transmission and storage have drastically dropped, one should always consider installing 3C instruments at all sites for new array installations.

S

The Array-Transfer Function

In signal-processing theory, the process of steering a seismic array to a specific target, i.e., optimizing the array beam for a given BAZ and apparent velocity of a seismic wavefield, can be described as a linear filter, which allows only signals with these characteristics to pass. In the case of an array, it is a two-dimensional filter, which is defined in the frequency-wavenumber (fk) space. The wavenumber k of a signal is defined as $k = 2\pi f/v$ or $k = 2\pi/\lambda$, where f is the frequency, λ is the wavelength, and v is the apparent velocity of the signal. Filter-transfer functions are mathematical descriptions of the filter characteristics (bandpass filter). The two-dimensional array filter is described by the sharpness of its main lobe (maximum of passing signals) and the position and relative height of eventually existing side maxima (side lobes). The narrower the main lobe and the smaller are eventual side lobes, the better the array's performance as a filter for signals with very specific BAZ and apparent velocity. For more details on the theory of seismic arrays, one may look in, e.g., Capon (1973), Aki and Richards (1980), Buttkus (2000), Johnson and Dudgeon (2002), Schweitzer et al. (2012), and the citations therein. By comparing array-transfer functions, the quality of an array of seismometers as a frequency-wavenumber filter can be discussed. The amount of literature about general criteria used to evaluate array-transfer functions for a given array geometry is huge, and some suggestions for more details are, e.g., in Harjes and Henger (1973) or Schweitzer et al. (2012) and the references therein. However, some general rules about transfer-function characteristics of seismometer arrays can be formulated as follows (from Harjes and Henger 1973; and Schweitzer et al. 2012):

1. The aperture a of an array defines the resolution of the array for small wavenumbers k. The larger the aperture is, the smaller the wavenumbers that can be measured with the array. The upper limit for the longest wavelength λ that can be resolved by array techniques is approximately similar to the aperture of the array. The array behaves like a single station for signals with $\lambda \gg a$.

2. The number of sites controls the quality of the array as a wavenumber filter, i.e., its ability to suppress seismic energy crossing simultaneously the array with a different slowness than the one on which the array is steered.

3. The distances between the seismometers define the position of the side lobes of the array-transfer function and the largest resolvable wavenumber: the smaller the mean distance, the smaller the wavelength of a resolvable seismic phase will be (for a given apparent velocity).

4. The geometry of the array defines the azimuth dependence of the aforementioned points.

Figure 3 shows two examples of such array-transfer functions, the ARCES array in northern Norway and the Yellowknife array in northern Canada for a 1 Hz signal. The geometry and aperture of these two arrays is included in Fig. 1. The ARCES shows no differences for signals from different azimuths, and the side lobes of the transfer function are far away (outside the plot) from the main lobe. However, because of the small aperture of only 3 km, this array cannot distinguish between waves with small wavenumber differences, as can be seen in the relatively wide main lobe of the transfer function. In contrast, in the case of Yellowknife, the main lobe is very narrow because of the much larger aperture of the array of about 25 km. This results in a much higher resolution in measuring small apparent-velocity differences. However, the array shows resolution differences for the different azimuths: the many side lobes of the transfer function along north-south and east-west oriented lines are caused by its cross-shaped geometry and the relative large distances between the single array sites.

For further reading, reviews on array theory, together with quite comprehensive citation lists, see, e.g., Douglas (2002), Rost and Thomas (2002, 2009), and Schweitzer et al. (2012).

Data Analysis Algorithms for Seismometer Arrays

One standard analysis technique for array data is the earlier described beamforming algorithm,
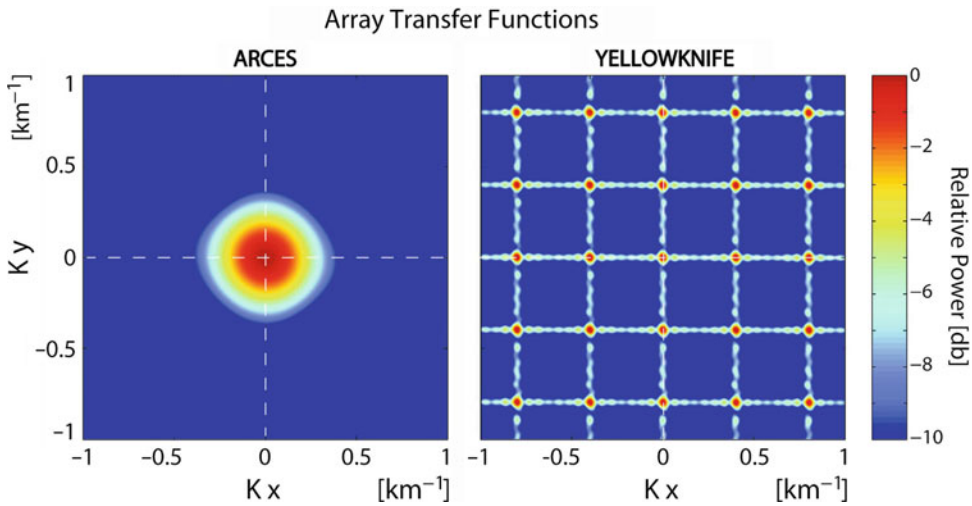
Seismometer Arrays, Fig. 3 The array-transfer functions of the circular ARCES array (*left*) and of the cross-shaped Yellowknife array (*right*) for a 1 Hz signal as relative power (*color coded*) of the array response normalized with its maximum in [db] (Modified from Schweitzer et al. 2012)

also referred to as "delay-and-sum" processing. In seismic prospecting, "beamforming" is called "stacking." An extension of this concept is the so-called "double-beam" technique, in which not only data of seismometer arrays are stacked but also the array beams of different seismic sources closely located to each other (Krüger et al. 1993).

Another standard array-data analysis technique is frequency-wavenumber analysis (fk-analysis). This method, originally developed for narrowband filtered one-component data, determines the slowness vector (i.e., BAZ and apparent velocity) for a single signal at a single frequency. Fk-analysis can be performed either in the frequency domain or the time domain (see Schweitzer et al. 2012). Fk-analysis applies systematically different phase shifts (or delay times) to a selected time window of the different array-site data and calculates the power of the seismic wave. By searching for the maximum power, the algorithm estimates the corresponding BAZ and apparent velocity of the signal. A variety of fk-analysis algorithms have been developed, which mostly differ in the method applied to find this maximum: e.g., maximum-likelihood method (Levin 1964; Capon et al. 1967, 1973), maximum-entropy method (Burg 1964), multiple

signal classification (MUSIC) algorithm (Schmidt 1986), and various adaptive algorithms for the estimation of spectral power density (e.g., Goldstein and Archuleta 1991). Later, fk-analysis was expanded to wider frequency bands and 3C data (Kværna and Doornbos 1986), the inversion for spherical waveforms (Almendros et al. 1999), and arrays with incoherent data (Gibbons et al. 2008, 2012).

One can also directly measure signal-arrival times at the different array sites and invert the observed travel-time differences for the best fitting plane wave. The (relative) arrival times can be either determined by an analyst or by automatic arrival-time-picking algorithms, as, e.g., by correlation analysis (Cansi 1995).

One widely applied technique to analyze data observed with an array is the VElocity SPectrum Analysis (VESPA) process (Davies et al. 1971), also called vespagram. This algorithm is very similar to the time-domain version of the fk-analysis, but it measures the power of seismic signals from a constant BAZ with different apparent velocities as function of time to investigate apparent-velocity changes for the different onsets in a seismogram. Later, the vespagram concept was expanded by calculating the observed power from different azimuths for a constant, specific

apparent velocity. In this case, the vespagram is a very useful tool to investigate the BAZ of the seismic phases, their precursors, and in particular the scattered energy of seismic waves arriving after seismic onsets.

Seismological Research with Seismometer Arrays

There are numerous publications from authors using data of seismometer arrays, which cannot be discussed in detail herein. In the following, only some general topics are named to illustrate the diversity of array-seismology applications.

Seismic array data have been used to detect and investigate all kinds of small amplitude phases: seismic onsets observed in front of seismic phases, which had traversed the Earth's core (PKP), were often interpreted as regular seismic phases, which led to many, quite controversial velocity models of the lower part of the outer core of the Earth. Using array observations it could be shown that these precursors are scattered energy from lateral heterogeneities in the Earth's interior and not regular seismic phases. There exist numerous array-data studies of the global and local structure of the lower mantle, the laterally heterogeneous lowermost part of the mantle directly above the core-mantle boundary (known as D" region), the core-mantle boundary, and the inner-core boundary.

Also strong-motion instruments have been installed in array configurations at the Earth's surface, as well as in boreholes to study near-field effects of earthquakes, and arrays were used to track the aftershock activity after larger earthquakes or to monitor earthquake swarms. In combination with cross-correlation techniques, this became a very powerful tool to lower the detection thresholds for seismic events in monitored areas. Similar is the quite recent approach to track the earthquake slip process itself with array data observed at teleseismic distances. For such studies data from seismic networks are combined to virtual arrays and analyzed.

Seismic arrays have been used to measure dispersion curves of surface-wave velocities.

This technique became recently quite important in context with seismic risk and hazard studies. Since ambient noise mostly consists of surface-wave energy, one can use local dispersion-curve data, measured with temporary, very small aperture arrays, to invert for the near-surface S-velocity structure; for further details see, e.g., Schweitzer et al. (2012).

Seismic arrays are also used to investigate the nature and source regions of microseisms and to locate and track volcanic tremor for analyzing complex seismic wavefield properties in volcanic areas.

The capability of seismometer arrays to locate seismic events was already mentioned at the beginning of this chapter. The data analysis of arrays can be automated for signal detection and fk-analysis of the detected onsets. Then, BAZ observations can be used together with additional travel-time constrains to automatically group the onsets to events. The observed apparent velocities can be used to classify the type of seismic onset, together with other signal characteristics, as, e.g., the dominant frequency. This way, P- and S-type onsets from local and regional events can be automatically identified with the use of small aperture, regional arrays and the associated seismic event can be located. A detailed description of such an automatic regional event location algorithm can be found in (Schweitzer et al. 2012).

Apparent-velocity observations from seismic arrays of at least 10 km aperture can be directly inverted to an epicentral distance for the first arriving P-type onsets from seismic events at teleseismic distances (i.e., from about 22° to about 100° epicentral distance). Knowing the epicentral distance, the observed BAZ can then be used to define the epicentral coordinates. This algorithm cannot be applied to local or regional events because measurable apparent velocities of the P-wave onsets can no longer directly be inverted to epicentral distance. At distances beyond ~90°, these derivatives become again very small for P waves, and in the Earth's shadow zone (distance >100°), the interpretation of the different core-phase onsets is also quite difficult and limits the location capabilities

of a seismic array. The described event location technique has been in use at least since the 1960s, and a quick look in the bulletins of the International Seismological Centre shows the huge amount of reported teleseismic event locations made with, e.g., the Large Aperture Seismic Array (LASA) in Montana, USA, the Yellowknife Array (YKA) in Northern Canada, the Gräfenberg Array (GRF) in Bavaria, Germany, or the large Norwegian Seismic Array (NOA) in southern Norway.

For further reading, more details on the research topics mentioned herein and for many more array applications in seismology, together with quite comprehensive citation lists, see, e.g., Gibbons et al. (2008, 2012), Rost and Thomas (2002, 2009), and Schweitzer et al. (2012).

Summary

A seismometer array is a set of seismic sensors deployed so that characteristics of the seismic wavefield at a specified reference point can be inferred by analyzing together the waveforms recorded at the different array sites. Seismometer arrays have been shown to be superior to single, 3C stations: they are able to lower the detection threshold for seismic onsets due to beamforming; they are able to measure the apparent velocity of seismic onsets and their propagation direction; they can be used to locate seismic events in an automated data processing scheme; and they are very successful tools to investigate the Earth's interior on local, regional, or global scale.

Cross-References

▶ Earthquake Location
▶ Earthquake Swarms
▶ Principles of Broadband Seismometry
▶ Recording Seismic Signals
▶ Seismic Event Detection
▶ Seismic Network and Data Quality
▶ Seismic Noise
▶ Time History Seismic Analysis

References

Aki K, Richards PG (1980) Quantitative seismology – theory and methods, vol II. Freeman and Company, San Francisco, pp 609–625. ISBN 0-7167-1059-5

Almendros J, Ibáñez JM, Alguacil G, Del Pezzo E (1999) Array analysis using circular-wave-front geometry: an application to locate the nearby seismo-volcanic source. Geophys J Int 136:159–170

Barber NS (1958) Optimum arrays for direction finding. N Z J Sci 1:35–51

Burg JP (1964) Three-dimensional filtering with an array of seismometers. Geophysics 29:693–713

Buttkus B (2000) Spectral analysis and filter theory in applied geophysics. Springer, Berlin, xv + 667 pp. ISBN 3-540-62674-3

Cansi Y (1995) An automatic seismic event processing for detection and location: P.M.C.C. method. Geophys Res Lett 22:1021–1024

Capon J (1973) Signal processing and frequency-wavenumber spectrum analysis for a large aperture seismic array. In: Bolt BA (ed) Methods in computational physics, 13 geophysics. Academic, New York, xiii + 473 pp. ISBN 0-12-460813-2, 1–59

Capon J, Greenfield RJ, Kolker KJ (1967) Multidimensional maximum-likelihood processing of a Large Aperture Seismic Array. Proc IEEE 55(2):192–211

Dahlman O, Mykkeltveit S, Haak H (2009) Nuclear test ban – converting political visions to reality. Springer, Dordrecht. doi:10.1007/978-1-4020-6885-0, xviii + 277 pp. ISBN 978-1-4020-6883-6

Davies D, Kelly EJ, Filson JR (1971) The VESPA process for the analysis of seismic signals. Nature 232:8–13

Douglas A (2002) Seismometer arrays – their use in earthquake and test ban seismology. In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) Handbook of earthquake and engineering seismology Part A. Academic, Amsterdam, xxiii + 933 pp. ISBN 0-12-440652-1, 357–367

Gibbons SJ (2012) The applicability of incoherent array processing to IMS seismic arrays. Pure Appl Geophys. doi:10.1007/s00024-012-0613-2

Gibbons SJ, Ringdal F, Kværna T (2008) Detection and characterization of seismic phases using continuous spectral estimation on incoherent and partially coherent arrays. Geophys J Int 172:405–421. doi:10.1111/j.1365-246X.2007.03650.x

Goldstein P, Archuleta RJ (1991) Deterministic frequency-wavenumber methods and direct measurements of rupture propagation during earthquakes using a dense array: theory and methods. J Geophys Res 96:6173–6185

Harjes H-P, Henger M (1973) Array-Seismologie. Z Geophys 39:865–905

Haubrich RA (1968) Array design. Bull Seismol Soc Am 58:977–991

S

Johnson DH, Dudgeon DE (2002) Array signal processing: concepts and techniques, Prentice Hall signal processing series. PTR Prentice Hall, Upper Saddle River, XIII + 533 pp. ISBN 0-13-048513-6

Krüger F, Weber M, Scherbaum F, Schlittenhardt J (1993) Double beam analysis of anomalies in the core-mantle boundary region. Geophys Res Lett 20: 1475–1478

Kværna T, Doornbos DJ (1986) An integrated approach to slowness analysis with arrays and three-component stations. NORSAR Sci Rep, 2-85/86, 60–69

Levin MJ (1964) Maximum-likelihood array processing. Lincoln Laboratory report, 31 Dec 1964

Macmillan (2013) Array. http://www.macmillan-dictionary.com/dictionary/british/array. Accessed 08 Dec 2013

Mykkeltveit S, Åstebøl K, Doornbos DJ, Husebye ES (1983) Seismic array configuration optimization. Bull Seismol Soc Am 73:173–186

Rost S, Thomas C (2002) Array seismology: methods and applications. Rev Geophys 40(3):1008. doi:10.1029/2000RG000100

Rost S, Thomas C (2009) Improving seismic resolution through seismic arrays. Surv Geophys 30:271–299. doi:10.1007/s10712-009-9070-6

Schmidt RO (1986) Multiple emitter location and signal parameter estimation. IEEE Trans Ant Prop AP-34:276–280

Schweitzer J, Fyen F, Mykkeltveit S, Gibbons SJ, Pirli M, Kühn D, Kværna T (2012) Seismic arrays. http://ebooks.gfz-potsdam.de/pubman/item/escidoc 43213:7. doi:10.2312/GFZ.NMSOP-2_ch9; 80 pp. In: Bormann P (ed) New manual of seismological observatory practice (NMSOP-2), 2nd (rev) edn. Deutsches GeoForschungsZentrum GFZ, Potsdam. doi:10.2312/GFZ.NMSOP-2. Accessed 08 Dec 2013

Seismometer Self-Noise and Measuring Methods

Adam T. Ringler[1], Reinoud Sleeman[2],
Charles Robert Hutt[1] and Lind S. Gee[1]
[1]Albuquerque Seismological Laboratory,
U.S. Geological Survey, Albuquerque, NM, USA
[2]Seismology Division, Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands

Synonyms

Coherence analysis; Instrument noise; Seismometer testing

Introduction

Seismometer self-noise is usually not considered when selecting and using seismic waveform data in scientific research as it is typically assumed that the self-noise is negligibly small compared to seismic signals. However, instrumental noise is part of the noise in any seismic record, and in particular, at frequencies below a few mHz, the instrumental noise has a frequency-dependent character and may dominate the noise. When seismic noise itself is considered as a carrier of information, as in seismic interferometry (e.g., Chaput et al. 2012), it becomes extremely important to estimate the contribution of instrumental noise to the recordings.

Noise in seismic recordings, commonly called seismic background noise or ambient Earth noise, usually refers to the sum of the individual noise sources in a seismic recording in the absence of any earthquake signal. Site noise (e.g., cultural sources, nearby tilt signals, etc.) and noise introduced by the sensitivity of an instrument to non-seismic signals (e.g., temperature and pressure variations, magnetic field changes, etc.) both contribute to the ambient seismic noise levels. The background noise ultimately defines a lower limit for the ability to detect and characterize various seismic signals of interest. Background noise levels have also been found to introduce a systematic bias in arrival times because the amplitude of the seismic phase must rise above the station's noise levels (Röhm et al. 1999). The upper limit of useful signals is governed by the clip level of the recording system (the point at which a recording system's output is no longer a linearly time-invariant representation of the input).

Site noise can be reduced by careful site selection (e.g., hard rock far from strong noise sources) and by emplacing instruments in good vaults or boreholes. It is also possible to reduce sensitivity to non-seismic signals by thermal insulation and appropriate shielding such as pressure chambers (Hanka 2000). At quiet sites with well-installed instrumentation, instrument noise may be the dominant noise source

(Berger et al. 2004); this is especially true for long-period seismic data (>100 s period) on very broadband instruments (e.g., Streckeisen STS-1 seismometer). The interpretation of such data only makes sense if the instrumental noise level is known. Also, research on noise levels in seismic recordings, the effect of noise reduction by the installation technique, and the nature and contribution of different noise sources to the recordings require knowledge of instrumental self-noise to rule out bias from the instrumentation self-noise.

A number of tests have been developed, under the assumption that instrument self-noise is approximately constant as a function of time (e.g., Evans et al. 2010; Holcomb 1989, 1990; Sleeman et al. 2006). Only recent studies have started to look at the potential time dependence of instrument self-noise (e.g., Sleeman and Melichar 2012). Understanding the self-noise of a given piece of recording equipment helps station operators to better identify sites that can take advantage of low-self-noise instruments. This knowledge also allows a network operator to provide higher-quality data with limited resources by making better use of their high-quality instruments. Having a rough understanding of an instrument's self-noise also gives a first-order diagnostic for determining if the recording system, as installed, is performing satisfactorily.

In order to estimate the self-noise of a seismometer, it is necessary to remove non-instrumental noise signals (e.g., earthquakes and ambient Earth noise) from the data. This is often accomplished by using coherence analysis techniques. In its simplest form, one can select a quiet time period at a low-noise site and attribute all recorded noise to the sensor (one-sensor method). However, it can be difficult to find stations with sufficiently low site noise to evaluate high-quality broadband seismometers. This is especially true at periods between the primary and secondary microseism (approximately 4–22 s period). In such cases, more sophisticated techniques, where one removes coherent signals using colocated instruments, are required. This can be done using a second instrument or two additional instruments (two- and three-instrument methods). For passive sensors, it

is possible to obtain an estimate of the self-noise by locking the mass of the instrument.

This is a general overview on the various methods currently used for estimating the instrument self-noise when using one, two, or three sensors. As there is no universally "best" method for all types of instruments and test conditions, some of the advantages and disadvantages of each method are discussed. Variants of these methods, such as rotating horizontal components of Earth motions to maximize the coherence and correcting for misalignments, are also included. Finally, the test setup used for estimating self-noise for the various methods is discussed and examples of corresponding test results are given.

Methods

To discuss the various methods currently in use for estimating the self-noise of a seismometer, a mathematical framework common to all the methods is developed. The system under test is assumed to be a linear time-invariant system (LTI) making the system completely determined by its impulse response (Scherbaum 2007).

Basic Assumptions and Conventions
Let x_i denote an input seismic signal, h_i denote the seismometer's impulse response, and n_i denote the self-noise of instrument i. Then the instrument's output signal y_i can be modeled, in the time domain, as

$$y_i = h_i * (x_i + n_i) \tag{1}$$

where "*" denotes convolution (Holcomb 1990; Sleeman et al. 2006). It is possible to write this in the frequency domain as

$$Y_i = H_i \cdot (X_i + N_i), \tag{2}$$

where capital letters denote the Fourier transforms of the corresponding lowercase letter time domain terms of index i. With the assumption that the self-noise of two different instruments is incoherent, the term $N_{ij} = N_i N_j$ is zero for $i \neq j$. It is
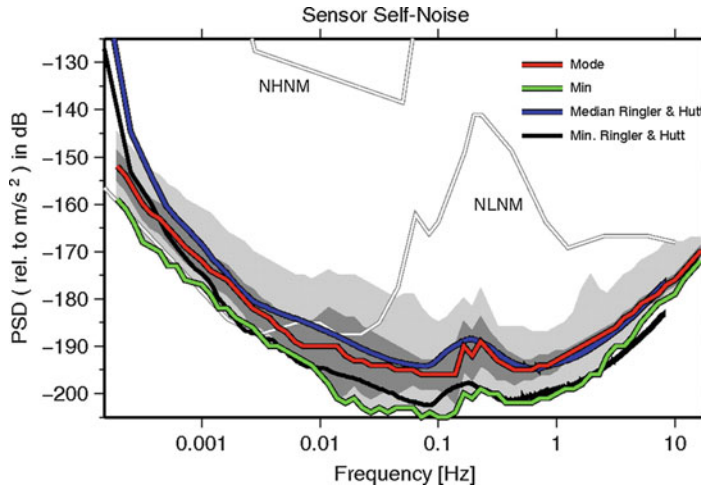
Seismometer Self-Noise and Measuring Methods, Fig. 1 Self-noise estimates of the vertical component of the STS-2 seismometer for the minimum noise (*green*) and the mode (*red*) of sensor noise measured at the Conrad Observatory, Austria. The *light gray* and *dark gray* bands depict the 95 % and 68 % percentile power spectral density (PSD) estimates. Finally, median (*blue*) and minimum (*black*) self-noise estimates for the STS-2 are taken from the data at the Albuquerque Seismological Laboratory (ASL)

also assumed that the self-noise and the input signal are incoherent, $X_i N_j$ is zero for $i \neq j$. The cross power between instrument i and j is denoted by P_{ij} so for $i \neq j$ we can write the cross power as

$$P_{ij} = H_i X_i \overline{H_j X_j} \qquad (3)$$

where the bar denotes the complex conjugate. The coherence between instruments i and j is given by

$$\gamma^2 = \frac{\left| P_{ij} \right|^2}{P_{ii} P_{jj}} \qquad (4)$$

Finally, it is assumed that all of the instruments in the test have common output units so that they have similar ground motion units (e.g., m/s or m/s^2) after removing the instrument response. This assumption can generally be made true by using an "omega correction" (multiplying or dividing by $\omega = 2\pi f$, where f is the frequency) in the frequency domain (Stearns 1975).

Noise levels are often compared to absolute Earth noise models such as Peterson's New Low-Noise Model (NLNM) and New High-Noise Model (NHNM), which are global seismic noise models derived from the 74 Global Seismographic Network (GSN) stations along with one additional station (Peterson 1993). This NLNM represents the lower envelope of 2,000 data records and represents an approximation of the lowest observed seismic noise levels, whereas the NHNM represents the upper envelope of the same 2,000 data records and represents an approximation of the highest observed seismic noise levels (Fig. 1).

Single-Sensor Method

When testing instruments at locations where the site noise is well below the self-noise of the instrument, it is possible to attribute the power in a given frequency band entirely to the instrument's self-noise (Fig. 2). This often occurs when testing strong-motion accelerometers in a quiet vault or lower-grade sensor in almost any good site (Evans et al. 2010). In such cases the simple relation is obtained:

$$N_{ii} = \frac{P_{ii}}{H_i \overline{H_i}} \qquad (5)$$

assuming that $X_i \ll N_{ii}$. Even when this assumption is not satisfied across the entire frequency

Seismometer Self-Noise and Measuring Methods, Fig. 2 Self-noise estimates for a strong-motion accelerometer using the single-sensor method (*red*), the Sleeman (Sleeman et al. 2006) three-sensor method (*green*), and the Holcomb (Holcomb 1989) two-sensor method (*blue*). For reference the New Low-Noise Model (NLNM) is included (*black*)
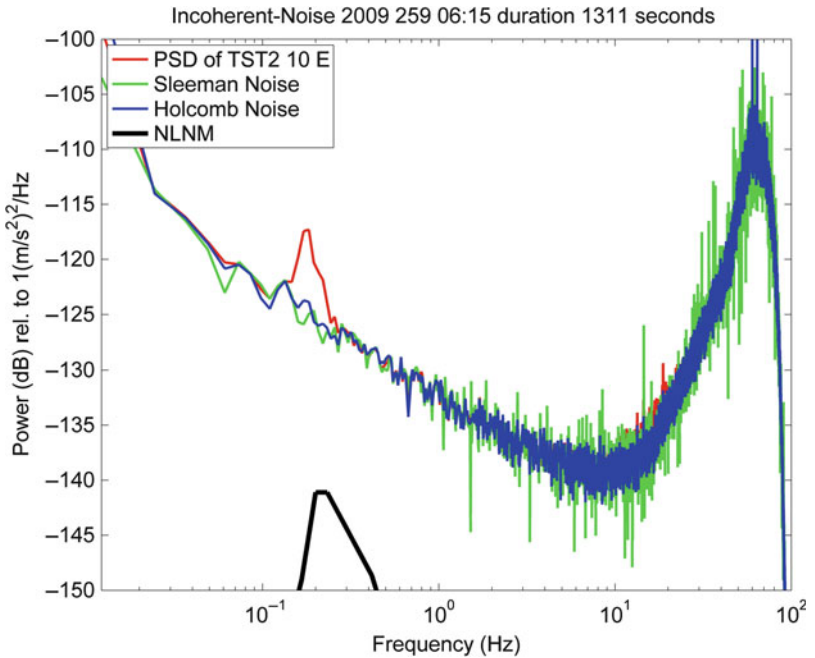
band, it is often possible to get an initial estimate of the instrument's self-noise in a specific frequency band (e.g., outside the primary and secondary microseism bands for state-of-the-art accelerometers). The single-sensor method can also be used to get an upper bound on the self-noise of an instrument. If this upper bound is obtained at a relatively quiet location, it might characterize the self-noise sufficiently well to identify the suitability of an instrument at a station with potentially much higher site noise. This method gives a first approximation of an instrument's self-noise, which can be valuable when testing time is limited or there are insufficient resources to use multiple sensors.

Two-Sensor Methods

Estimating self-noise using two sensors has been the traditional method for noise testing broadband seismometers for many years (Holcomb 1989). As very few locations have background noise levels below that of high-quality broadband sensors across a wide frequency range, it becomes important to remove the local background noise (Fig. 3). Assuming that the colocated sensors are recording similar seismic

signals (same ground motion), $X_i = X_j$, it is possible to derive the self-noise of instrument i from Eq. 2. Recalling the assumption that $N_{ij} = 0$ for $i \neq j$ we have

$$N_{ii} = \frac{P_{ii}}{|H_i|^2} - \frac{P_{ij}}{H_i \overline{H_j}}. \tag{6}$$

Under the assumption that the sensors both have well-known responses, the instrument corrected output, from the two sensors, should only differ in instrumental self-noise. By estimating the coherent signal between the two records, the coherence signal is removed resulting in the incoherent signal which is attributed to the self-noise. It can be seen from this estimate of the self-noise that it is critical to have well-described transfer functions for both instruments i and j. Since the transfer functions are used in the calculation, errors in the transfer function will produce errors in the self-noise estimates of the instrument. This method was originally proposed by Holcomb (1989), who later characterized error sources (Holcomb 1990). In the latter work, Holcomb also suggested alternative two-sensor methods, under the assumptions that the two sensors
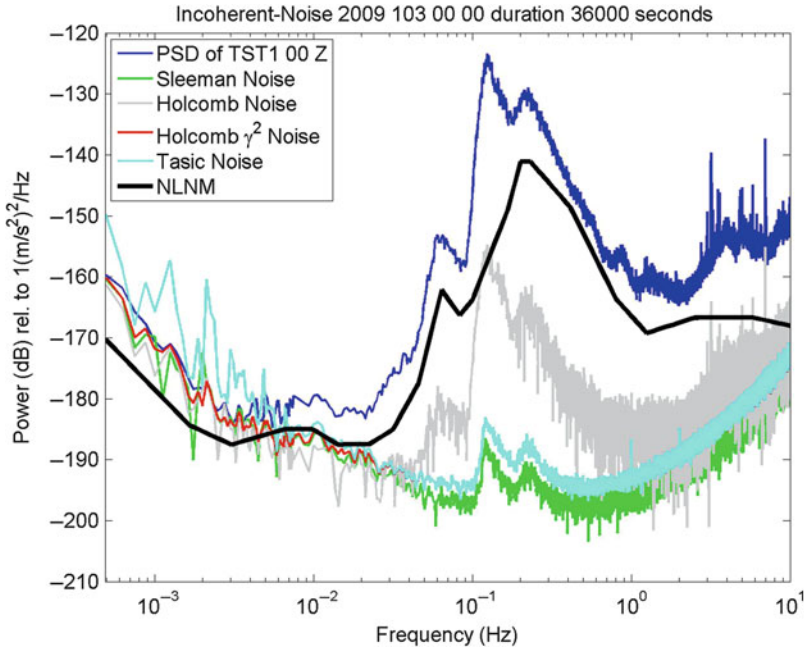
Seismometer Self-Noise and Measuring Methods, Fig. 3 Comparison of five different methods for estimating the self-noise using vertical broadband sensor data: the one-sensor (direct power spectral density (PSD)) method (*blue*), the three-sensor Sleeman (Sleeman et al. 2006) method (*green*), the two-sensor Holcomb (Holcomb 1989) method (*light gray*), the two-sensor γ^2 Holcomb method (*red*), and the two-sensor Tasič (Tasič and Runovc 2012) method (*cyan*) (for the different methods, see text). The cyan and red lines overlay at frequencies less than approximately 0.01 Hz. For reference the New Low-Noise Model (NLNM) is the solid black line. The elevated noise in the microseism band is caused by misalignment

under test have equal noise and a high signal-to-noise ratio (ratio of input signal to instrument self-noise). In this case, it is possible to derive the following estimate for the self-noise of instrument i:

$$N_{ii} = \frac{\left(P_{ii} \dfrac{(1 - \gamma^2)}{\gamma^2} \right)}{H_i \overline{H_j}}. \tag{7}$$

Using a related approach, Tasič and Runovc (2012) develop a different method for estimating the self-noise using two instruments. In their approach the self-noise of instrument i is given by

$$N_{ii} = \frac{P_{ii}(1 - \gamma^2)}{\left(H_i \overline{H_J} \right)} \tag{8}$$

where again they assume the signal-to-noise ratio is large.

Three-Sensor Method

Using three colocated sensors, it is possible to estimate the self-noise of each instrument while minimizing errors in the estimate due to uncertainty in the transfer functions (Sleeman et al. 2006). This method has become the preferred approach for estimating self-noise for broadband sensors, even though it requires additional resources and setup; it typically "sees through" site noise more deeply to extract lower estimates of instrument noise.

Using Eq. 2 and the assumption that $X_i = X_j = X_k$, one can estimate the self-noise of instrument i as

$$N_{ii} = \frac{\left(P_{ii} - P_{ij} \dfrac{P_{ik}}{P_{jk}} \right)}{H_i \overline{H_l}}. \tag{9}$$

The term P_{ik}/P_{jk} can be viewed as a "relative transfer function" H_{ij} between instruments i and j. The three-sensor method is related to the

Seismometer Self-Noise and Measuring Methods, Fig. 4 Self-noise estimate of the Geotech GS-13 geophone using a theoretical self-noise model (*blue*) versus the estimated low self-noise model of the Geotech GS-13 (*red*) from 11 self-noise measurements; the median self-noise from these 11 tests is also shown (*green*). The self-noise of the digitizers used, Quanterra Q330HRs, is also shown after correction with the response of the Geotech GS-13 (*cyan*). For reference the New Low-Noise Model (NLNM) is included (*black*)
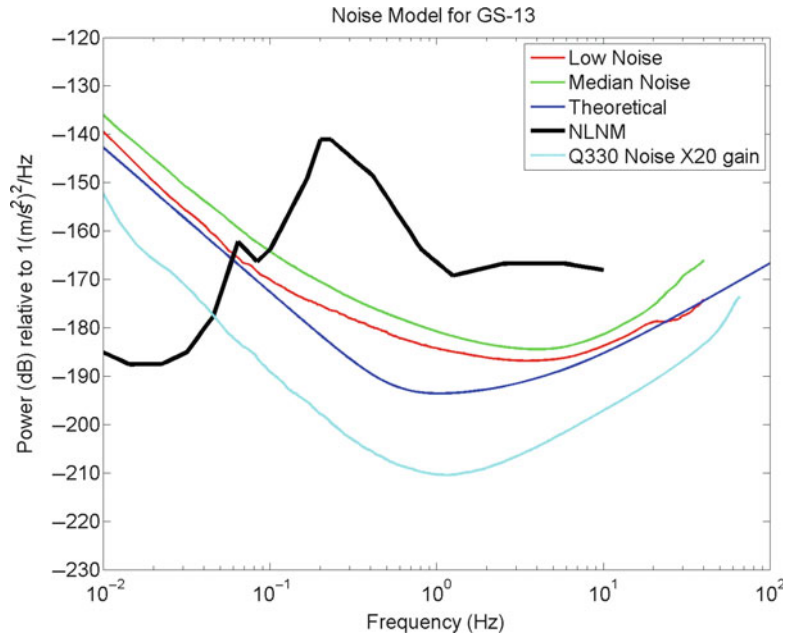
two-sensor method developed by Tasič and Runovc (2012) when one compares relative transfer functions. It is important to note that the transfer function of instrument *i* is only used to convert the noise estimate to units of ground motion; thus, errors in the transfer functions do not propagate to the self-noise estimate before deconvolution with the individual instrument (Ringler et al. 2011). Because the three-sensor method is not as sensitive to errors in the transfer functions, it extracts the self-noise with a potentially higher accuracy than the two-sensor method (Fig. 3).

Other Techniques

When instrument designers are selecting components for building a seismometer, they often have noise estimates for each electronic component as well as estimates of fundamental noise contributions (e.g., Brownian noise). Through modeling these individual noise sources, it is possible to estimate the total self-noise of a seismometer without having a working prototype. A modeling example for Brownian noise is described by Aki and Richards (2002), using the physical parameters of a simple gravitational pendulum. Such methods often give a first approximation of the instrumental noise to be aimed for during design (Fig. 4).

Such theoretical methods for estimating the self-noise of an instrument have been applied to a number of seismometers in conjunction with their digitizers (Rodgers 1994). However, when such methods are used, a number of assumptions must be made (e.g., that the pendulum obeys the small-angle approximation and all electronic components perfectly match their specified noise levels).

For seismometers that are not controlled by feedback but having self-noise well below any available test site's ambient noise, such as geophones at high frequency (e.g., GeoTech GS-13), it is possible to estimate the self-noise using a "locked mass" test (Havskov and Alguacil 2004). This test is performed by locking the mass of the instrument and recording the output of a single instrument. This test provides a lower noise-bound estimate but may not fully characterize the self-noise of the instrument because it lacks noise contributions from moving hinges and other mechanical sources.

Testing and Analysis

Characterizing the self-noise of a seismometer requires attention to detail in the test setup,

Seismometer Self-Noise and Measuring Methods, Fig. 5 Three colocated Geotech GS-13 short-period seismometers setup for test to be sensitive to vertical ground motion. All three are sitting on a granite slab supported at three points and are in the Albuquerque Seismological Laboratory (ASL) underground vault. On the back right is a reference Streckeisen STS-2 seismometer in a steel bell jar (not used in the test)

careful selection of data windows to avoid transients, identification of potentially non-seismic noise sources, and methods for selecting data to analyze. In both the two- and three-sensor methods, it is necessary to make sure all instruments truly are recording the same ground motion.

Test Setup

Both the two- and three-sensor methods assume that all instruments are recording the same ground motion. In order for this to be true, it is critical that the instruments be colocated and co-aligned. Furthermore, since local tilt signals can vary widely over even a few inches, one must take care to insure that all three instruments are measuring similar tilt signals so that they can be removed using coherence analysis. Locating all instruments on a stiff slab (e.g., granite or gabbro) with a three-point suspension is usually sufficient for this purpose (Fig. 5).

It is also important to isolate the instruments from non-seismic noise sources (e.g., temperature variations, pressure variations, locally induced tilt). These noise sources can increase the self-noise estimate of the individual instruments as they generally do not respond coherently to such sources (Anderson et al. 2012). Misalignment of the horizontal components can

also contribute to a decrease in signal coherency and increase the estimate of the self-noise. Such relative errors in misalignment can be seen by elevated incoherent noise-level estimates in the microseism band (Fig. 6), so it is possible to reduce such misalignment errors by rotating the instrument outputs numerically to maximize coherence (Tasič and Runovc 2012). Recent work by Gerner and Bokelmann (2013) shows that the leakage of microseism signals into the self-noise can be eliminated successfully by numerical rotation of the three-component traces and using the three-sensor method. Their study optimally aligns two RefTek 151-60A sensors with a third one and shows that self-noise estimates outside the microseism bands are not compromised by misalignment errors.

Finally, if one wants to isolate the self-noise of a seismometer, it is critical that the digitizer and other recording equipment have self-noise levels well below that of the sensor. This can be a problem on low-gain instruments, such as strong-motion accelerometers. Verifying that the digitizer's self-noise levels are below that of the seismometer requires a terminated-input test of the digitizer, using a terminating resistor similar to the output impedance of the sensors to be tested. When the digitizer's noise level is not below the seismometer noise level for any of the
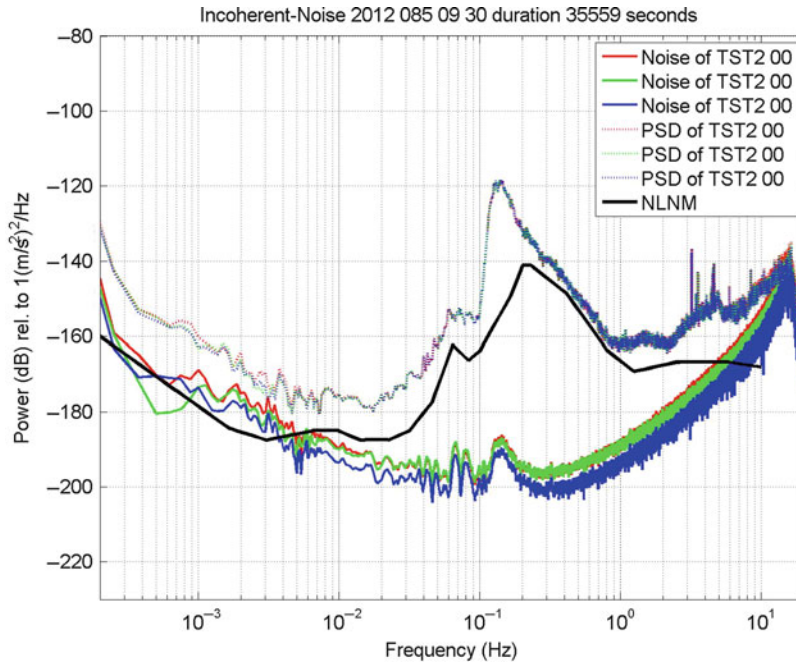
Seismometer Self-Noise and Measuring Methods, Fig. 6 Self-noise estimates from three horizontal very broadband Metrozet M-2166 seismometers (flat to velocity from 0.0028 to10 Hz) using Sleeman (Sleeman et al. 2006) three-sensor self-noise estimate (solid lines). Single-sensor estimates are in dotted lines. In order to minimize incoherent tilt noise, all three instruments were installed with the same orientation on the same baseplate and the baseplate is supported at three points only. The relatively small increase in the self-noise in the secondary microseism frequency band (at approximately 0.125–0.25 Hz) suggests all three instruments are well aligned to one another

frequency range of interest, digitizer preamps generally can be used for low-noise amplification of the sensor signal to a level above the digitizer noise level.

Data Selection

Data selection is currently an important item in the development of algorithms and techniques to estimate self-noise. This requirement makes it necessary to estimate self-noise noise during only "quiet" time periods (e.g., during nighttime and periods with low pressure-induced tilt noise). No current agreement exists as to whether data selection should be applied or not, or which criteria should be used for data selection. In particular, when features in the data are not fully understood, the debate on data selection is ongoing and not yet entirely resolved, though a best-case result seems to require such minimization (Ringler et al. 2011). For example, the occurrences of periodic pulsing in the data are not yet understood (though many have been named "popcorn noise," "spherics," and so forth and are tentatively explained in various ways). Some of these noise sources could properly be considered either as part of the self-noise of the instrument or as transient external signals that would artificially elevate the self-noise if included (Sleeman and Melichar 2012). The discussion on data selection of time segments becomes even more controversial when limited amounts of data are available. By using multiple quiet time segments, it is possible to understand the self-noise of an instrument in at least a best-case scenario (Ringler and Hutt 2010). One such example of a best-case scenario, along with a self-noise estimate using the median of multiple tests, is shown in Fig. 4. This example also suggests that even when using coherency analysis techniques, it is important to do such tests in a

S

low-noise environment. The selection of time segments is additionally controversial when limited amounts of test data exist, because the assumption that the self-noise of a seismometer is time invariant is only approximately true, as seen in the relatively large gray bands in Fig. 1, which depict the 95 % and 68 % of the self-noise of the Streckeisen STS-2 seismometer as recorded under very stable conditions at the Conrad Observatory (Vienna, Austria) over an entire year.

A synthetic-data experiment by Sleeman and Melichar (2012) shows that the three-sensor technique can reliably extract instrumental noise even for high "seismic-signal-to-instrumental-noise (SNR)" ratios, but is in practice limited to lower ratios due to misalignments or mis-leveling between the sensors. As discussed earlier, small alignment or leveling differences between the sensors decrease the coherency between the recordings and thus affect the noise estimate. It was found that an alignment error of 0.2° in any axis allows the technique to extract self-noise for SNR values up to 60 dB, which was confirmed in their associated experiment using real data. As this order of misalignment between corresponding axes in similar colocated sensors is in agreement with the precision of best-practices manufacturing, one should only use data taken during "quiet" time periods (SNR below 60 dB, e.g., during night and periods with minimal pressure-induced tilt). Even during these times, microseism noise can make it difficult to estimate the self-noise of broadband sensors in the period bands of 4–8 s and 18–22 s for instruments (Ringler et al. 2011).

Analysis Parameters

As in any time series analysis, appropriate time window lengths, number, and overlapping must be used to resolve the low frequencies of interest while minimizing variance. For very broadband seismometers, it is often of interest to understand the self-noise for periods up to several thousands of seconds, which corresponds to a total window length of at least 8 h divided into overlapping sub-windows, each of about an hour (e.g., Table 1 of Evans et al. 2010).

To directly compare results between noise tests, it is important that the tests be conducted using similar methods (e.g., settling time, installation methods, data selection, and spectral processing parameters).

Expected Results and Caveats

A discussion on the expected self-noise results for various types of seismometers now follows. This discussion is in no way exhaustive but simply discusses some of the more common situations and what is generally expected for various test scenarios.

Strong-Motion Accelerometers

The noise levels of strong-motion accelerometers make estimating the self-noise of such sensors easier than typical low-noise broadband sensors (Fig. 2). Since the user community is generally only interested in strong motions with frequencies from a few Hz to a few tens of seconds period, it is often possible to resolve the self-noise using relatively short testing intervals.

As noted earlier, one must be careful to make sure that the digitizer noise is below that of the sensor. Since strong-motion accelerometers are generally sampled at higher rates than broadband sensors ($\geq$100 sps), it is necessary to identify the self-noise to much higher frequencies than for broadband sensors (Cauzzi and Clinton 2013). At these higher frequencies, it becomes more difficult to isolate the instrument from potentially non-seismic noise sources such as electrical power grid frequencies (e.g., 50 Hz in Europe and 60 Hz in North America) noise. At these higher frequencies, the coherency between recordings from colocated sensors presumably decreases, but this decrease in coherency is typically compensated by decreasing ambient noise and increasing instrument self-noise.

Broadband Seismometers

After installing a broadband seismometer, it is important to let the instrument fully settle down from mechanical stresses that can build up during transport and adjust to temperature before
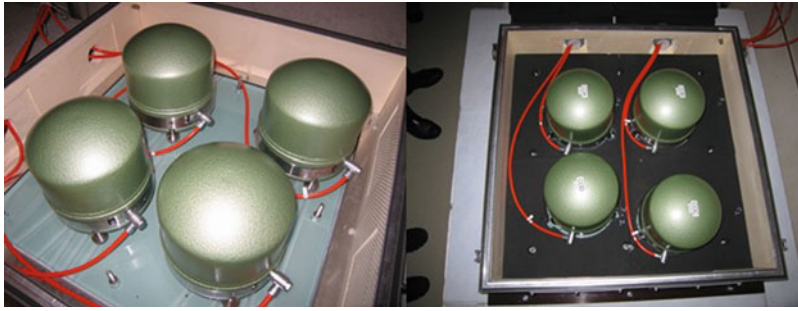
Seismometer Self-Noise and Measuring Methods, Fig. 7 Colocated Streckeisen STS-2 sensors at the Conrad Observatory, Austria. Sensor casings are co-aligned by parallel grooves in the underlying glass plate. Thermal insulation consists of thin layers of neoprene around the sensors, which also reduces the noise contribution due to air convection around the sensors

attempting to estimate the self-noise. What constitutes an instrument that is "fully settled" is not well understood and currently a topic of some controversy. After the instrument has been allowed to settle, one can estimate the self-noise by using long-running quiet time periods (Hutt et al. 2010). Local changes in wind and pressure can introduce incoherent elevated noise levels on horizontal channels at long periods (e.g., >100 s period) as well as short (e.g., <1 s period).

Estimates Using Other Data Selection Techniques

Although the process of self-noise measuring is not yet mature, the methods and techniques are rapidly improving. As manufacturers' abilities to reduce sensor noise levels improves, methods for measuring the self-noise will also need to improve. A better understanding of the sensitivity of seismometers to non-seismic phenomena and techniques to thermally isolate the sensors (e.g., Fig. 7) or correct the data from these sources will be needed. Although there has been initial work on this subject (e.g., Forbriger et al. 2010; Zürn et al. 2007) with promising results, there is still much to be done before these methods become widely used by the seismological community.

Other Geophysical Instruments

The above methods for estimating seismometer and accelerometer self-noise can be applied more generally to other instruments and recording systems. As the only assumptions are that all instruments will have a common input signal and output units, it is also possible to estimate the self-noise of digitizers (a "terminated noise" test), infrasound sensors (Hart et al. 2007), tilt meters, gravimeters, and rotational seismic sensors.

In all cases, the general approach for estimating the self-noise will be similar, though the details may differ (e.g., test setup logistics, frequency band of interest, etc.). For a discussion on applying the three-sensor self-noise technique to digitizers, the reader is referred to Kromer (2006) and Sleeman et al. (2006). Details on applying self-noise techniques to other geophysical instruments are beyond the scope of this entry.

Summary

With the suite of various methods for testing seismometer self-noise, it is possible to get an initial estimate of the self-noise of a seismometer even in less-than-ideal testing situations. For instruments with relatively high self-noise, simple (one-sensor) self-noise methods may prove sufficient.

Using additional high-quality sensors and more sophisticated techniques (two- or three-sensor methods) can help to eliminate common signals when there is concern of the local background noise being above the sensor's self-noise.

This will almost always be the case when testing broadband seismometers, but will not necessarily be the case when testing accelerometers though at least microseisms are above the noise floor of the state-of-the-art accelerometers even at an inland continental site.

Equally important to the two- and three-sensor methods is that the common seismic signals be coherent among all instruments in the test. As discussed, this can be done by installing the instruments in a quiet vault on a granite slab or other stiff material to increase the coherence among sensors. Care should also be taken to reduce the influence on instruments under test from local non-seismic noise sources such as pressure or temperature variations and that adequate thermal settling time be provided before testing.

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US government.

Cross-References

▶ Seismic Network and Data Quality
▶ Seismic Noise
▶ Sensors, Calibration of

References

Aki K, Richards PG (2002) Quantitative seismology. University Science, Sausalito

Anderson KE, Anderson JF, Anthony RE, Chaput J, McMahon ND, Morton EA, Aster R (2012) A site comparison between shallow vault-deployed and direct burial broadband seismometers, IRIS 2012 Workshop Abstracts. http://www.iris.edu/hq/iris_workshop2012/scihi/WebPages/0070.html

Berger J, Davis P, Ekström G (2004) Ambient earth noise: a survey of the global seismographic network. J Geophys Res 109, B11307

Cauzzi C, Clinton J (2013) A high- and low-noise model for high-quality strong-motion accelerometer stations. Earthq Spectra 29:85–102

Chaput JA, Zandomeneghi D, Aster RC, Hunter K, Kyle PR (2012) Imaging of Erebus volcano using body wave seismic interferometry of Strombolian eruption coda. Geophys Res Lett 39, L07304

Evans JR, Followill F, Hutt CR, Kromer RP, Nigbor RL, Ringler AT, Steim JM, Wielandt E (2010) Method for calculating self-noise spectra and operating ranges for seismographic inertial sensors and recorders. Seismol Res Lett 81:640–646

Forbriger T, Sidmer-Schnidrig R, Wielandt E, Hayman M, Ackerley N (2010) Magnetic field background variations can limit the resolution of seismic broad-band sensors. Geophys J Int 183:303–312

Gerner A, Bokelmann G (2013) Instrument self-noise and sensor misalignment. Adv Geosci 36:17–20

Hanka W (2000). Which parameters influence the very long period performance of a seismological station? GEOFON Program. http://geofon.gfz-potsdam.de/geofon/manual/welcome.html

Hart D, Merchant B, Chael E (2007) Seismic and infrasound sensor testing using three-channel coherence analysis. In: 29th monitoring research review: ground-based nuclear explosion monitoring technologies. LA-UR-07-5613, 935–944

Havskov J, Alguacil G (2004) Instrumentation in earthquake seismology. Springer, Dordrecht

Holcomb LG (1989) A direct method for calculating instrument noise levels in side-by-side seismometer evaluations. U.S. Geological Survey Open File Report, 89–214, 34 pp

Holcomb LG (1990) A numerical study of some potential sources of error in side-by-side seismometer evaluations. U.S. Geological Survey Open File Report, 90–406, 41 pp

Hutt CR, Evans JR, Followill F, Nigbor RL, Wielandt E (2010) Guidelines for standardized testing of broadband seismometers and accelerometers. U.S. Geological Survey Open File Report, 2009–1295, 62 pp

Kromer RP (2006) Technology report evaluation of the Kinemetrics/Quanterra Q330HR Remote Seismic System for IRIS/GSN Q330HR/GainX1 Configuration, 32286. http://www.iris.edu/hq/files/programs/gsn/documents/IRIS_Progress_Report2_Q330HR.pdf

Peterson J (1993) Observations and modeling of seismic background noise. U.S. Geological Survey Open File Report, 93–322, 95 pp

Ringler AT, Hutt CR (2010) Self-noise models of seismic instruments. Seismol Res Lett 81:972–983

Ringler AT, Hutt CR, Evans JR, Sandoval LD (2011) A comparison of seismic instrument noise coherence analysis techniques. Bull Seismol Soc Am 101:558–567

Rodgers PW (1994) Self-noise spectra for 34 common electromagnetic seismometer/preamplifier Pairs. Bull Seismol Soc Am 84:222–228

Röhm AHE, Trampert J, Paulssen H, Snieder RK (1999) Bias in reported seismic arrival times deduced from the ISC Bulletin. Geophys J Int 137:163–174

Scherbaum F (2007) Of poles and zeros: fundamentals of digital seismology, 2nd edn. Springer, Dordrecht

Sleeman R, Melichar P (2012) A PDF representation of the STS-2 self-noise obtained from one year of data recorded in the Conrad Observatory, Austria. Bull Seismol Soc Am 102:587–597

Sleeman R, van Wettum A, Trampert J (2006) Three-channel correlation analysis: a new technique to measure instrumental noise of digitizers and seismic sensors. Bull Seismol Soc Am 96:258–271

Stearns SD (1975) Digital signal analysis. Hayden Book Company, Rochelle Park

Tasič I, Runovc F (2012) Seismometer self-noise estimation using a single reference instrument. J Seismol 16:183–194

Zürn W, Exß J, Steffen H, Kroner C, Jahr T, Westerhaus M (2007) On reduction of long-period horizontal seismic noise using local barometric pressure. Geophys J Int 171:780–796

Seismometer, Extended Response

Dieter Stoll
Lennartz Electronic GmbH, Tübingen, Germany

Synonyms

Feedback system; Inverse filtering; Negative impedance converter

Introduction

Until the advent of the broadband seismometer (Wielandt and Streckeisen 1982), electrodynamic geophones were used for picking up short-period (1 s and less) earth movement, and LP (long period) seismometers were used for anything with lower frequency than one second. Owing to their bulk and weight, LP seismometers are not practical for use in the field; they typically require a controlled observatory-type environment. Most types of geophones, on the other hand, are by their very nature field worthy, compact, and cheap. This is because they are mass-produced for use in seismic exploration where they are typically deployed in large numbers (hundreds or thousands). The downside is that their natural frequency is usually 4.5 Hz or more. Some geophones designed for scientific use have a natural frequency of 1 Hz, but this low-frequency limit comes at the expense of

significant bulk and weight and an unwelcome increase in mechanical complexity.

Thus, combining the ruggedness of the short-period geophone with the extended frequency response of an LP seismometer has long been a goal for researchers and instrument makers alike.

Geophone Frequency Response

In its simplest form, a geophone consists of a mass suspended by a spring. To pick up the relative motion between the mass (which, due to inertia, tends to remain steady) and the frame (which will follow the motion of the Earth's surface), a moving coil system is typically used. A coil that is attached to the suspended mass moves in a magnetic field provided by a permanent magnet attached to the frame, thus providing an electrical output voltage that is proportional to the relative velocity between mass and frame.

If the suspended mass were an ideal oscillator (i.e., without any damping forces at play), a geophone would not pick up Earth motion very faithfully. In fact, a single impulse would cause it to oscillate forever at its natural frequency. A certain amount of damping is always provided by mechanical damping (friction) and electromagnetic damping, with the latter being responsible for the major part of damping. Electromagnetic damping is a consequence of Lenz's law (in a closed circuit moving inside a magnetic field, a current is induced which in turn generates a magnetic field opposed to the original magnetic field, thus resulting in a force counteracting the movement of the conductor), and is proportional to the sum of the coil resistance and the external load on the coil (the input resistance of the amplifier or recorder).

In a plain geophone (without any active electronics involved), the minimum amount of damping is given by the open circuit damping (typically in the range 0.25–0.5, but some commercially available geophones go down to 0.02 (Rodgers 1993)), meaning there is no external load connected. With no external load connected,

S

Fig. 1 Geophone
frequency response for
different damping factors
(normalized to be 1.0 for f
>> f0)
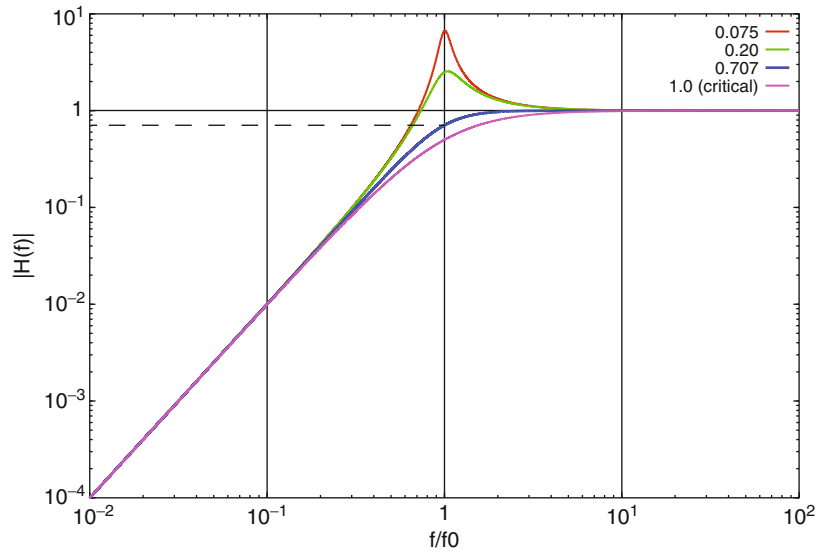
the circuit is open, no current flows through the coil, and no additional damping is contributed by the coil. The other extreme is a shorted coil – in this case, the external load is 0 Ω, and only the intrinsic resistance (inductive reactance, to be more precise) limits the amount of current flowing. Consequently, this setup provides the maximum possible amount of damping (this is why classical geophones should be transported only with the output coil(s) shorted – the high damping will help inhibit extraneous mass movement).

The modulus of the complex transfer function of an electrodynamic geophone with respect to ground velocity depends on damping as follows (Bormann 2012, Chapter 5.2):

$$H_V(\omega) = \frac{\omega^2}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4h^2\omega^2\omega_0^2}} \quad (1)$$

where h is the damping factor expressed as the ratio of actual to critical damping (when h = 1), ω is angular frequency, and ω_0 denotes the natural angular frequency.

Figure 1 shows the modulus of a geophone's transfer function to ground velocity over a wide range of frequencies for different damping

factors, ranging from open circuit (very low damping, strong resonance peak) to 1.0. Usually, the desired value is 0.707 of critical damping (thick line). At this setting, the modulus of the transfer function at the natural frequency is 0.707 ($0.5*\sqrt{2}$).

Since open circuit damping is lower than the desired value of 0.707, an external damping resistor (often referred to as a "shunt") is used to achieve higher damping. An inevitable consequence of shunting the geophone coil is that its output decreases (because the external resistor and the geophone coil resistance form a voltage divider). In the flat part of the transfer function, the geophone's output is proportional to

$$\left(\frac{R_d}{R_d + R_{coil}}\right) \quad (2)$$

where R_d is the damping (shunt) resistance, and R_{coil} is the coil resistance.

This fact has been deliberately omitted from Fig. 1; all curves have been normalized to 1.0 for frequencies much higher than the natural frequency.

For frequencies $<< f_0$, the geophone's output rolls off at -12 dB/octave (proportional to f^2). For frequencies $>> f_0$, the transfer function is flat.

Seismometer, Extended Response, Fig. 2 Schematic operation of inverse filter, new f0 = 0.1* original f0
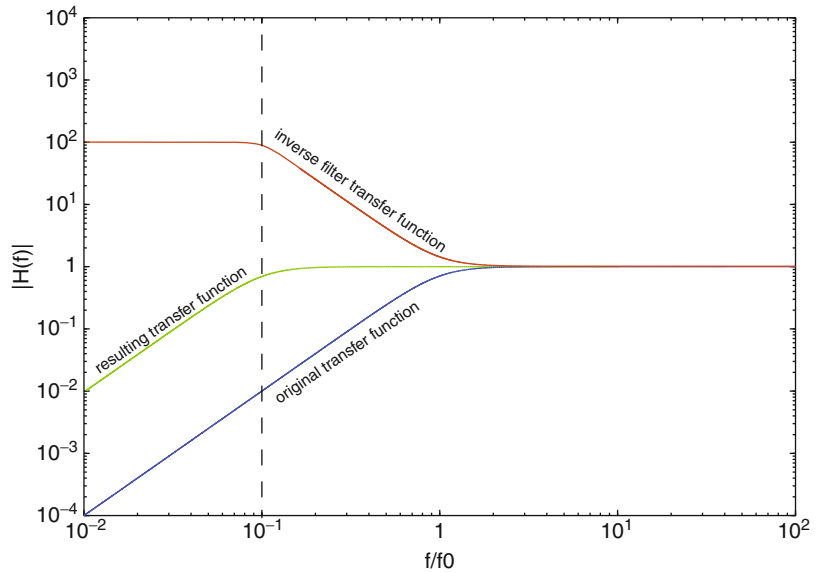
Methods of Frequency Response Extension

Extending the low-frequency response of a geophone without increasing its bulk and weight can be achieved with several methods. The most common ones will be briefly outlined here.

Using an Inverse Filter

An electronic filter can be implemented to compensate the geophone's roll-off at low frequencies (Havskov and Alguacil 2004). In real life, the filter's transfer function cannot increase indefinitely towards lower frequencies. A new fictitious natural frequency is defined by the point where the filter's transfer function returns to being flat. Figure 2 shows the schematic function of an inverse filter (Lippmann 1982).

Intriguing as it may seem at first sight, this approach is plagued with many practical problems. Noise appearing in electronic circuits is not necessarily "white," i.e., equally distributed over a wide frequency range. There is one particularly annoying type of noise called 1/f noise, which, as its name implies, means it gets increasingly stronger the lower the frequency gets. This phenomenon occurs below a certain corner frequency

which, unfortunately, typically is found to be not less than a few Hz – a typical operation amplifier such as the OP-27 has a 1/f corner frequency of 2.7 Hz (Analog Devices 2006). Since the inverse filter method requires an increasing amount of amplification towards lower frequencies, intrinsic 1/f noise will be amplified overproportionally. Further, maintaining thermal stability of low-frequency circuits is not trivial. Thus, signal-to-noise ratio for low frequencies will be problematic. Also, the signal chain may take a prohibitively long time to recover from even a brief overload. Inverse filtering was applied in the 1960s–1980s but, owing to its drawbacks and the advent of superior methods, is no longer in practical use today.

Positive Feedback

As outlined previously, the bare-bones version of a geophone consists of a coil suspended by one or more springs, and moving in a magnetic field. When motion occurs, a voltage is induced across the coil's electrical terminals. However, fancier versions exist where a second coil is present. This coil is not a movable coil but is present so as to be able to feed an external current through it which, again through induction, will exert a force on the

Fig. 3 Geophone
frequency response for
different(over) damping
factors
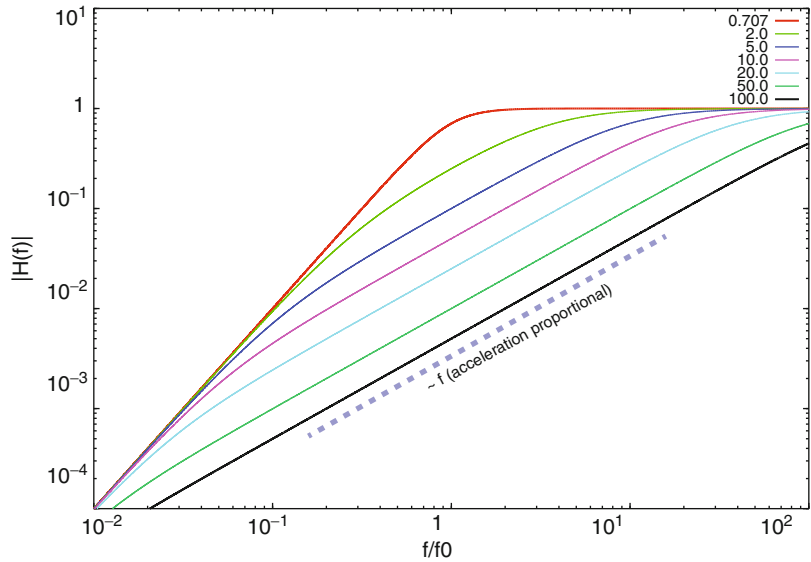
suspended mass. This procedure is usually applied just for calibration (hence the name "calibration coil") but can also be applied for positive feedback.

Hypothesizing the presence of such a coil, positive feedback could be applied by integrating the output signal (thus rendering it proportional to displacement) and feeding the integrated voltage back through the feedback coil (Lippmann 1982). When applied with the correct polarity, the feedback signal will "nudge" the mass so that the original movement will be amplified, thus counteracting the decay of sensitivity. The effect is similar to using a softer spring, leading to a smaller restoring force and, consequently, to a lower natural frequency. Like the inverse filter method, positive feedback has practical stability problems and is not in practical use nowadays.

Negative Feedback

Rather than integrating the output signal, this method differentiates it (thus rendering it proportional to acceleration of the suspended mass) and feeds it back through the second coil choosing the polarity so that it will not amplify but attenuate the mass movement. Ideally, the geophone mass would be damped so strongly that it does not perform any movement relative to the frame at

all; in other words, the mass is accelerated in exactly the same way as the Earth, and the current required to keep it steady with respect to the frame is proportional to acceleration for frequencies higher than the new corner frequency (it is not feasible to extend the bandwidth to arbitrarily low frequencies). Practical limitations of this method are two coils required (thus excluding cheap mass-produced exploration geophones) and stability problems of feedback electronics due to inductive coupling of signal and feedback coils (Lippmann 1982).

Broadband sensors also use negative feedback, but they use a capacitive transducer to pick up mass movement, thus avoiding the stability issues. However, it is a nontrivial undertaking to fit a capacitive pickup in an off-the-shelf exploration geophone, so this method, while promising, does not lend itself well to commercializing.

Response Extension Using NIC (Negative Impedance Converter)

Recalling Eq. 1, let us now extend the series of damping factors in Fig. 1 towards higher values, and we get this (Fig. 3).

It is evident that the higher the damping factor gets, the wider is the frequency range in which the output is proportional to acceleration. If a way can be found to achieve such high damping factors, the geophone will effectively be transformed into an accelerometer with a wide frequency range, albeit at the expense of output. As can be seen from Fig. 3, output decreases by several orders of magnitude in a certain frequency range.

Achieving Overdamping

Neglecting the mechanical contribution to damping (which is a reasonable simplification for typical geophones if a very high damping is required), we can limit our consideration to electrical damping which is proportional to the combined resistances of the generator coil and the external shunt resistor (Havskov and Alguacil 2004):

$$h = \frac{CDR}{R_{coil} + R_{shunt}} \qquad (3)$$

where CDR is the total resistance needed to get $h = 1$ (critical damping resistor).

As we can see, the maximum amount of damping that can be achieved in a bare-bones geophone is when the signal coil is shorted ($R_{shunt} = 0$). In this case, only the intrinsic coil resistance is responsible for damping. Loading the coil with any external resistance at all will only decrease damping. In order to arrive at damping factors higher than short-circuit damping, a negative resistance would have to be used – clearly a feat that cannot be achieved using passive electronics. An NIC (negative impedance converter) provides exactly that functionality (Lippmann 1982; Ulmann 2005). In its simplest form, it is just an operational amplifier and two resistors.

A practical value for the negative resistance is 80 % of the coil resistance (Ulmann 2005). Completely compensating the coil resistance with close to 100 %, negative resistance is not practical for reasons of stability.

Practical Considerations

Linearity

Conventional geophones are plagued with linearity issues on account of the inherent nonlinearity of the suspension spring(s). Positive feedback would magnify this type of problems, whereas negative feedback (by overdamping) will actually alleviate them. Since, in a first-order approximation, the mass will remain steady, the suspension springs will be flexed only very little. Since spring nonlinearity increases with the amount of flexing or elongation, this is a very efficient means of suppressing spring-induced nonlinearity.

Amplification

Since, as seen in Fig. 3, output all across the acceleration proportional passband is significantly lower than the velocity proportional signal of the non-overdamped geophone, additional amplification is called for. However, it is not convenient to amplify the acceleration proportional signal since this would create dynamic range problems. Applying linear amplification by, say, a factor of 1,000 would lead to impractically large amplitudes for the high frequencies. It is a much more practical approach to apply amplification only after the acceleration proportional signal has been converted back to velocity proportional (see next section).

Converting the Output Signal Back to Velocity

The obvious solution of simply integrating the acceleration proportional signal back to velocity is not a very practical one. As can be seen in Fig. 3, the low-frequency roll-off of the standard damped seismometer is no longer seen in the overdamped transfer function, but there is still quite an amplitude gap to cover. The difference in amplitude is largest around the original corner frequency and decreases towards higher and lower frequencies, but it is especially towards lower frequencies that trouble is to be expected if one were to attempt a simple integration without any frequency band limitation. As an

Seismometer, Extended Response, Fig. 4 Design of integrating/differentiating filter
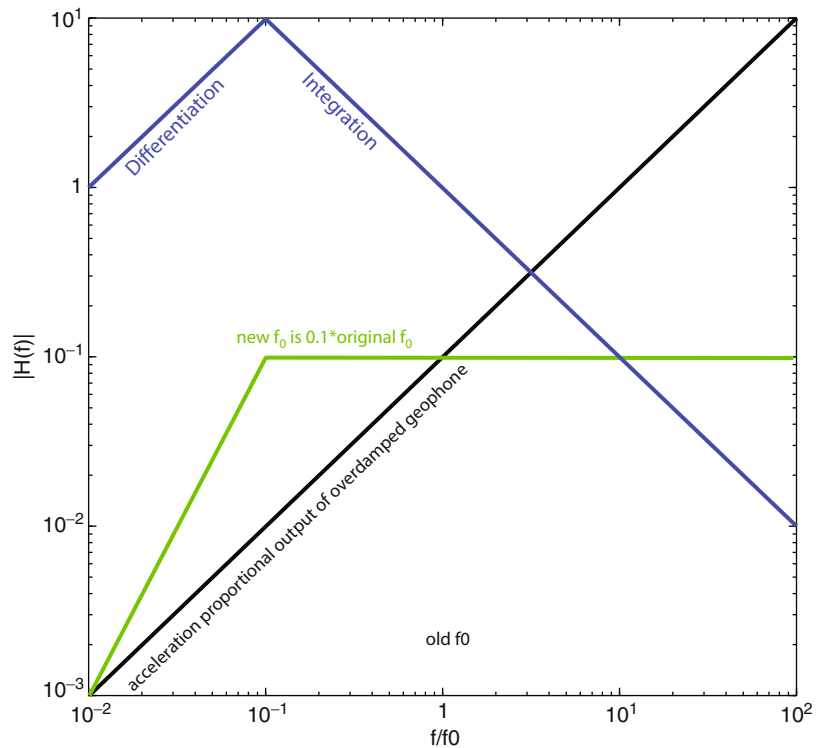
example, let's take a look at 0.01 f/f$_0$. The output of the standard damping seismometer is down to 10^{-4} of its plateau output, and the overdamped (acceleration proportional) output is down even further, already off scale. Bringing these low frequencies back to "flat" would lead to extremely high amplification factors, leading to the same type of problems that were described above in "Using an Inverse Filter." Rather than simulating an extremely broadband seismometer, a band-limited approach is a lot more promising. To this end, a filter is needed that integrates frequencies above its corner frequency and differentiates those below. Figure 4 (after Lippmann 1982) schematically shows its functional principle. With a frequency proportional rise for f < f$_{0NEW}$, and a frequency proportional decay for f > f$_{0NEW}$, the result of multiplying the frequency proportional acceleration signal with the filter's transfer function will be exactly the transfer function of a geophone with a virtual natural frequency of f$_{0NEW}$. However, in order to arrive at an output

level comparable to the non-overdamped geophone, the filter must provide more amplification than shown in our schematic example. The exact amount of peak amplification depends on the ratio between the original and the new virtual natural frequency.

A filter with the desired characteristics (6 dB/octave roll-off on either side of a given frequency) is not difficult to implement – it is basically a combination of a first-order high-pass and a first-order low-pass filter, with both corner frequencies being set to the desired new f$_0$ and of course an amplification stage.

Environmental Influences

Except for a few broadband models, seismometers will typically be used in field scenarios, i.e., they have to work under adverse environmental conditions, and will have to maintain their characteristics over a wide range of temperatures.

The negative feedback principle outlined above basically amounts to measuring one small

physical entity (in our case, the current flowing through the generator coil when the coil starts moving), then adding a counteracting entity of similar magnitude (in our case, a current that will create a force to keep the coil steady). Effectively two similar entities are subtracted from each other. Consequently, if one or both of the two entities are susceptible to changes in environmental parameters, the result of the subtraction will also be susceptible to changes. The coil resistance is a "real" resistor consisting of many windings of copper wire whose value will increase with temperature. The negative impedance is implemented by active circuitry which will not be immune to temperature changes but will certainly react differently from the purely ohmic coil resistance. The following example from Lippmann (1982) illustrates the problem, focusing only on the coil.

A coil resistance of 600 Ω at 20 °C will change to 640 Ω at 40 °C. If our negative resistance is 520 Ω, the effective damping resistance will change from 80 Ω (20 °C) to 120 Ω at 40 °C – a 50 % increase, leading to drastic changes in the sensor's characteristics.

Clearly, temperature compensation needs to be applied. Lippmann (1982) describes the measures and results.

Electronic Noise

Electronic noise is the prime limiting factor for how much the response can be extended since it defines the smallest detectable ground velocity amplitude. Electronic noise is generally frequency dependent. As outlined previously in the "Inverse Filter" entry, below a certain threshold frequency (which depends on the particular active electronic element, e.g., an operational amplifier), noise tends to be proportional to 1/f. In other words, lower frequencies tend to exhibit more noise.

In an extended response seismometer using negative feedback and a subsequent integrating/ differentiating filter, electronic noise can be considered separately for the acceleration proportional part (i.e., the combination of geophone and NIC) and for the filter/amplifier part. Lippmann (1982) reveals that the contribution of the

latter part is close to negligible, compared to the NIC's contribution.

The only thing of relevance to the practitioner, though, is the composite noise, expressed not in electrical but in seismological units, i.e., in equivalent ground velocity or acceleration. The following slightly edited diagram from Wielandt (1991) shows the equivalent RMS velocity for two different types of Lennartz (2013) seismometers – the LE-3Dlite type based on physical 4.5 Hz geophones and converted to 1 Hz seismometers and the LE-3D/5 s type based on physical 2 Hz geophones and converted to 0.2 Hz (5 s) seismometers. For reference, the NLNM (New Low Noise Model, Peterson (1993)) and the noise level at the BFO (Black Forest Observatory) site are also given. Note that the noise levels have been calculated for frequency bins of constant relative bandwidth. For the 1/3 octave bandwidth shown here, the relative bandwidth is 23.2 %. For example, the 10 Hz bin is 2.32 Hz wide, whereas the 50 Hz bin is 11.6 Hz wide. Using constant relative bandwidth bins is best suited for comparing and converting noise measurements (Bormann 2012, Chapter 4.1) (Fig. 5).

It turns out that both seismometer types are able to resolve ground noise at a very quiet site like BFO over a wide range of frequencies. The vast majority of sites will exhibit much more pronounced noise than BFO, thus rendering the extended response seismometers useable over their full frequency range.

A Practical Implementation

The photo shows a classical, purely mechanical seismometer next to an electronically enhanced seismometer built around a 4.5 Hz geophone. In both cases, natural frequency is one hertz, and both are single-component instruments (Fig. 6).

Summary

Range-extended seismometers based on robust, readily available, and comparatively cheap
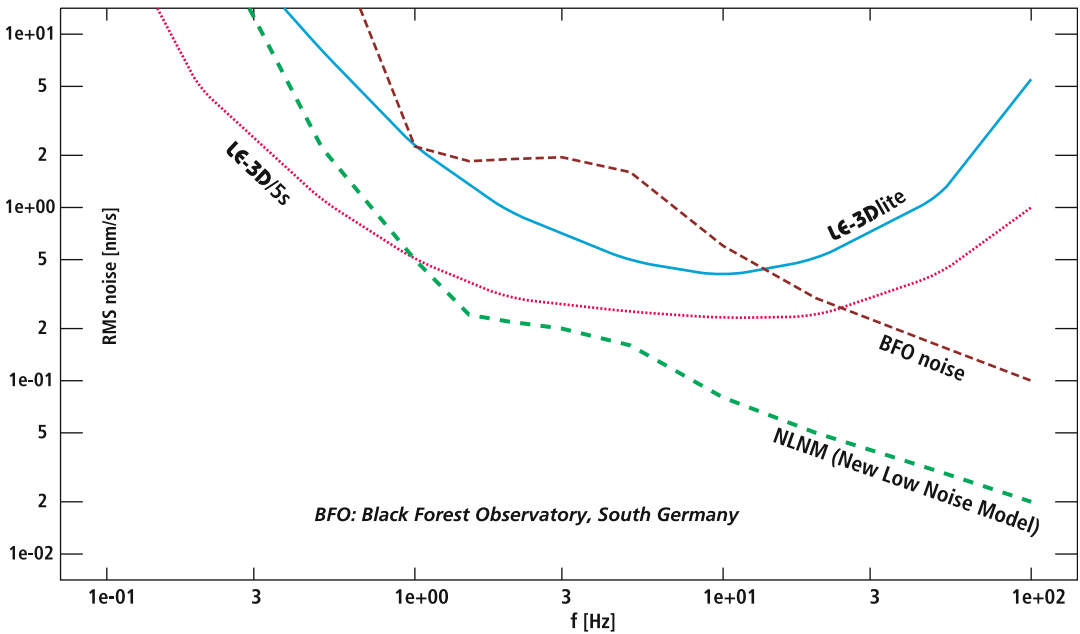
Seismometer, Extended Response, Fig. 5 L∈-xD Self Noise in 1/10 decade (1/3 octave)

Seismometer, Extended Response, Fig. 6 Classical Geophone (left) vs. Range-Extended Seismometer (right), both of them single-component, 1 Hz units

exploration-grade geophones provide a viable alternative to clumsy and delicate mechanical seismometers for short- and intermediate-period applications. Unlike typical broadband sensors, range-extended seismometers do not contain physical elements tuned to very low frequencies and do not respond to very low frequencies. Consequently they usually require no shielding from

pressure and temperature changes and provide extended high-frequency response, typically up to and above 100 Hz. Inherent limitations such as electronic noise can be mitigated by proper design and implementation.

Cross-References

References

Analog Devices (2006) OP27 datasheet, downloadable from http://www.analog.com/static/imported-files/data_sheets/OP27.pdf

Bormann P (ed) (2012) New manual of seismological observatory practice (NMSOP-2), IASPEI. GFZ German Research Centre for Geosciences, Potsdam. http://nmsop.gfz-potsdam.de

Havskov J, Alguacil G (2004) Instrumentation in earthquake seismology. Springer, Dordrecht

Lennartz electronic (2013) LE-xD seismometer family (Document 990-0003), PDF document downloadable from http://www.lennartz-electronic.de

Lippmann E (1982) Umwandlung eines elektrodynamischen Seismometers in einen Beschleunigungs-Aufnehmer und Entwicklung eines Filters zur Simulation eines niederfrequenten Seismometers. Diploma thesis, Ludwig-Maximilians-Universität München

Peterson J (1993) Observations and modeling of seismic background noise. US geological survey open file report 93-322

Rodgers P (1993) Maximizing the signal-to-noise ratio of the electromagnetic seismometer: the optimum coil resistance, amplifier characteristics, and circuit. B Seismol Soc A 83:561–582

Ulmann B (2005) Overdamping geophones using negative impedances. http://www.vaxman.de/publications/teach_gp.pdf

Wielandt E (1991) Messungen der kurzperiodischen Bodenunruhe in Schiltach (BFO). Internal report, published in German only. PDF of original text and English translation available from the author of present article

Wielandt E, Streckeisen G (1982) The leaf-spring seismometer: design and performance. B Seismol Soc A 72:2349–2367

Selection of Ground Motions for Response History Analysis

Anastasios G. Sextos
Division of Structural Engineering, Department of Civil Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece

Synonyms

Earthquake records; Ground motion; Record scaling; Response history analysis

Introduction

The evolution in computational power and the parallel processing capabilities of modern engineering software make nowadays the use of complicated structural analysis methods an attractive alternative for the design and assessment of structures. In contrast to the past, when the elastic static analysis was almost exclusively used for the seismic design of structures, the state of practice has progressively moved toward dynamic-elastic, nonlinear-static (i.e., single mode or multi-modal "pushover"), and even nonlinear response history analysis. The latter, capturing more efficiently the hierarchy of failure mechanisms, the energy dissipation, the force redistribution among the structural members, and contact issues (such as gap, impact, sliding, and uplift), is deemed preferable in cases of significant material or geometrical nonlinearities and, as such, is used for the design of seismically isolated buildings and bridges or the assessment of existing structures with various degrees of damage. Elastic response history analysis is also extensively used, primarily for structures whose response is dominated by higher modes (mostly tall and irregular buildings and towers) or structures of high importance that are typically designed to remain elastic even for long return-period earthquake intensities (i.e., industrial facilities, power plants, dams, critical administrative buildings, etc).

S

In all cases, the main task of the design procedures is to achieve more predictable and reliable levels of safety and operability against different levels of seismic intensity, a framework known as performance-based design and assessment. Despite the above major advances made in terms of structural analysis, the reliability of the analysis output and the subsequent structural performance prediction strongly depend on the decisions made for the selection of the seismic input which is used as ground excitation. Research has shown that among all possible sources of uncertainty stemming from structural and soil material properties, the modeling approximations, and the design and analysis assumptions as well as the earthquake-induced ground motion, the latter has by far the highest effect on the variability observed in the structural response (Elnashai and McClure 1996; Padgett and Desroches 2007; Shome et al. 1998). Therefore, the selection of a "reliable" suite of earthquake ground motions constitutes an important prerequisite for the reliability of the structural analysis procedure as a whole.

Along these lines, numerous computational methods and tools have been developed for (a) selecting suites of earthquake records from available strong ground motion record databases (b) generating synthetic and artificial ground motions or (c) modifying existing ground motions until they present desirable target characteristics.

Objectives of Ground Motion Selection

Currently more than 40 methods exist for selecting, modifying, and scaling earthquake ground motions. These methods can be grouped conceptually by their objective as follows based on the Ground Motion Selection and Modification (GMSM) Program of the Pacific Earthquake Engineering Research (PEER) Center (Haselton 2009) in the following broad four categories:

- *Predict* the *Probability Distribution* (mean and dispersion) of *Structural Response* (i.e., engineering demand parameter, EDP) from ground motions that comply with an

earthquake scenario of *given magnitude, M, source-to-site distance, R*, and, in some cases, faulting type or soil class (i.e., $V_{s,30}$) at the site of interest.
- *Predict* the *Median or Mean of Structural Response* from ground motions that are selected (or generated) to *match a median target response spectrum* for a given M-R pair, obtained from a ground motion prediction equation (or attenuation relationship). Dispersion of structural response might also be of interest; however, minimizing the standard error of response quantities is not explicitly envisaged.
- *Predict* the *Probability Distribution* (mean and dispersion) of *Structural Response* using ground motions that satisfy a given *spectral acceleration at the fundamental period* of the structure, $S_a(T_1)$, that has resulted from an associated M and R (as well as fault mechanism and soil class or $V_{s,30}$) scenario. Notably, a priori knowledge of the structural properties (i.e., its fundamental period T1) is a prerequisite.
- *Predict* the *Median or Mean of Structural Response* for a given *set of $S_a(T_1)$, M, and R*. In this case, the scenario refers to a Maximum Considered Earthquake $S_a(T_1)$ and not a MCE response spectrum.

A fifth objective can also be distinguished *to minimize the structural response discrepancy* by considering a threshold confidence level for the standard error of the response quantities, additional to conventional spectral matching (Katsanos and Sextos 2013). This involves consideration of structural analysis results within the GMSM process.

Methodologies for Ground Motion Selection and Scaling

Independently of the ultimate objective of the GMSM process, the various methodologies for developing earthquake ground motion sets for (linear or nonlinear) dynamic analysis of structures can also be classified based on the concept

of selection and the procedure of modulating or scaling ground motions. Naturally, method objectives and selection and scaling procedures are not fully correlated. The following classes of methods are listed below. A detailed review of most important methods is also performed elsewhere (Katsanos et al. 2010):

Selection by M and R and Scaling to a Target Intensity Measure

The simplest GMSM method involves the formation of set (bin) of motions that satisfy, as close as possible but without explicit constraints, preliminary magnitude, M, source-to-site distance, R, and often seismotectonic or soil class criteria. It is recalled that the total number of sets N_{tot} that can be formed, from m, potentially eligible, records out of a larger group of s records, can be calculated by the following factorial formula of the binomial coefficient:

$$N_{tot} = \binom{s}{m} = \frac{s!}{m!(s-m)!} \qquad (1)$$

Once eligible ground motions are selected, their accelerations are multiplied by a scaling factor to match a target intensity measure (IM), typically being the peak ground acceleration (PGA) or the spectral acceleration at the fundamental period of the structure, $S_a(T_1)$. In the latter case, it is evident that all ground motion spectra will have identical ordinates at period T_1 and different spectral accelerations in all other periods. More elaborate IMs have also been proposed, involving the spectral shape and the structural characteristics. Such IMs are expected to result into a relatively more accurate prediction of the seismic demand (Baker and Cornell 2005; Luco and Cornell 2007; Tothong and Luco 2007); however, the approximate definition of the seismic scenario is a significant limitation.

Selection and Scaling to a Target Uniform Hazard Spectrum

In the light of the above criteria, it is also common to envision matching of the response spectra of the selected acceleration time series with a target, Uniform Hazard Spectrum (UHS) (Kramer 1996; American Society of Civil Engineers 2005; McGuire 2004), which is determined from (a) a ground motion prediction relationship, (b) a seismic hazard assessment for the site of interest, or (c) the seismic code provisions. In this case, ground motions are scaled with a single or individual scaling factors so that their individual spectrum, or the mean of their response spectra closely match and in fact exceed, in terms of "shape," the ordinates of the target UHS, typically within a given range of periods around the fundamental period of the structure. When a uniform scaling factor is sought, it may be determined through the following expression:

$$sf_{avg} = \left\{ \min\left(\frac{Sa_{avg}(T_i)}{Sa_{target}(T_i)} \right) \right\}^{-1}, \quad i = 1 \text{ to } N \qquad (2)$$

where T_i is the sample period and N is the size of the sample into which the prescribed period range is discretized.

Quite commonly, all eligible suites are ranked according to their "goodness-of-fit" to the target spectrum, as quantified by the normalized root-mean-square-error, δ, between the scaled average $Sa_{,avg}$, and the target spectrum $Sa_{,trt}$ (Iervolino et al. 2010b):

$$\delta = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \left(\frac{Sa_{avg}(T_i) - Sa_{trt}(T_i)}{Sa_{trt}(T_i)} \right)^2} \qquad (3)$$

A wide variety of similar expressions have also been used in the literature. Further discussion on the efficiency of various spectral matching indicators can be found elsewhere (Beyer and Bommer 2007; Jayaram et al. 2011; Kottke and Rathje 2008; Naeim et al. 2004).

In addition to spectral shape, these methods may consider other earthquake, site, or ground motion parameters in selecting ground motions. Finally, the mean or the maximum (depending on the number of ground motions within the set) of the response quantities is used as the design value.

Seismic Codes of Practice

Most contemporary seismic codes and design recommendations, such as Eurocode 8 for buildings (CEN 2004) and bridges (CEN 2005), ASCE standards 7-10 (ASCE/SEI 2010) and FEMA P-750 (FEMA 2009a) as well as various national norms (New Zealand Standards, Italian Code, and Greek Seismic Code), describe relatively similar procedures that are based on spectral matching for the selection of earthquake ground motions in the framework of dynamic analysis of structures. In most cases, seismic motions can be represented by real, artificial, or simulated records, typically complying with the aforementioned preliminary criteria of earthquake magnitude, distance, the seismotectonic environment, and the local soil conditions. Differences among the codes can be summarized as follows:

Duration Strong-motion duration is not explicitly considered as an additional selection parameter in most documents inclusive of ASCE standards 7-10 (ASCE/SEI 2010) and the Eurocodes.

Simulated and Artificial Ground Motions Use of simulated ground motions is permitted in a number of seismic codes; however, it is restricted only to cases of inadequate number of real accelerograms in others, i.e., FEMA P-750 and Eurocode 8 Part 2 for bridge design (CEN 2005).

Near-Field Considerations Near-field effects are typically either ignored or considered in a quantitative manner in terms of direction of fault rupture and velocity pulses, an example being FEMA P-750 (FEMA 2009a). More detailed distinction is made in FEMA P695 (FEMA 2009b) by forming distinct far-field and near-field record sets. ASCE standards 7-10 prescribe different design quantities for sites within 3 miles (5 km) of the active fault that controls the hazard. In this case, each pair of ground motion components shall be rotated to the fault-normal and fault-parallel directions of the causative fault. Next, it shall be appropriately scaled so that the

average of the fault-normal components is not less than the MCE response spectrum for the period range of spectral matching (as described below).

Spectral Matching A distinction is typically made between two-dimensional and three-dimensional analysis, the latter involving selection and application of pairs of records for linear or nonlinear dynamic analysis. In principle, the procedure is similar among most important seismic codes and recommendations and involves the following steps: (a) the 5 %-damped elastic spectra are derived for each component of the eligible horizontal motions selected; (b) the mean of the individual (EC8-Part 1) or the square root of the sum of squares (SRSS spectra) of the individual elastic spectra ordinates is determined; (c) the mean spectrum of the ground motions is compared with the target code-prescribed spectrum; and (d) records are scaled so that the spectral ordinates of the mean spectrum (either from individual or SRSS spectra) exceed a lower bound of the target spectral acceleration (Table 1) within a prescribed period range (also depicted in Table 1).

Scaling Scaling of the individual records toward spectral matching can be performed either with the use of a uniform scaling factor (i.e., EC8–Part2, ASCE 7-10) or a record-dependent, individual scaling factor (FEMA P-750). Other codes (i.e., EC8-Part1) do not provide specific guidelines regarding the scaling of seismic records in order to establish the required compliance with the target design spectrum. New Zealand Standards (Standards New Zealand (SNZ) 2004) on the contrary prescribe two distinct scale factors k1 and k2, the first minimizing the difference between recorded and target response spectrum in a least mean square scheme over the period range of interest, while the latter ensures that the energy in the spectrum of the strongest ground motion exceeds the energy of the target spectrum.

Design Quantities Structural analysis results are processed in a statistical manner, and the

Selection of Ground Motions for Response History Analysis, Table 1 Earthquake records selection and spectral matching criteria prescribed in the seismic codes and guidelines studied here

Seismic codes and guidelines	Selection criteria	Ensemble spectrum	Spectral matching period range	Lower bound of mean spectral acceleration	Scaling factor
EC8 Part 1 (buildings)	Seismotectonic features, soil type	Mean of individual spectra	$0.2–2.0\ T_1$	$0.90 \cdot Sa(T)_{design}$	Not specified
EC8 Part 2 (bridges)	Source mechanism, M, R	Mean of SRSS spectra	$0.2–1.5\ T_1$	$1.30 \cdot Sa(T)_{design}$	Uniform
FEMA P-750	Source mechanism, M, R	Mean of SRSS spectra	$0.2–1.5\ T_1$	$1.17 \cdot Sa(T)_{design}$	Individual
ASCE/SEI 7-10	Source mechanism, M, R	Mean of SRSS spectra	$0.2–1.5\ T_1$	$1.00 \cdot Sa(T)_{design}$	Uniform
NZS 1170.5	Seismotectonic features, soil type	Individual record scaling and matching	$0.4–1.3\ T_1$	Matching as nearly as practicable	k_1 and k_2 factors
Taiwan seismic code	Source mechanism, M, R	Mean of individual spectra	$0.2–1.5\ T_1$	$0.90 \cdot Sa(T)_{design}$	Not specified

design quantities are defined as either the mean (in case of seven records or more) or the maximum (in case of three records) of the engineering demand parameters of interest (EC8, FEMA P-750). Certain codes solely provide the option of obtaining the maximum response out of the three response history analyses (CPA 2011; Standards New Zealand (SNZ) 2004). The most refined to date provision to derive seismic design values is prescribed in ASCE 7-10 clearly distinguishing between (a) *force response parameters* that shall be multiplied for each ground motion analyzed by a factor I_e / R, where I_e is the importance factor and R is the Response Modification Coefficient, and (b) *drift response quantities*, which shall be multiplied by C_d / R, C_d being the deflection amplification factor. It is noted herein that the above methods refer to the design of new buildings according to modern seismic codes and do not apply to the probabilistic assessment of the nonlinear response of existing buildings (Jalayer and Cornell 2009). It is also noted that despite the simplification and applicability of the above code-based procedures, significant limitations also exist (Sextos et al. 2010).

More details on comparative seismic code provisions on ground motion selection and scaling can be found elsewhere (Hachem et al. 2010).

Selection and Scaling to a Target Conditional Mean Spectrum

An alternative group of methods for ground motion selection and scaling is similar to the aforementioned spectral matching, but the Uniform Hazard target Spectrum, the Conditional Mean Spectrum (CMS) (Baker 2011), is used instead. The main reason behind this substitution (Kottke and Rathje 2008; Lin et al. 2013a; Wang 2010) is that the UHS is shown to be unsuitable to serve the main purpose of dynamic analysis, which is to excite the structure under consideration with ground motions having a specified spectral acceleration at a given period. In fact, UHS compatible ground motions are often associated with large-amplitude spectral values in a wide range of periods. Furthermore, UHS can hardly be considered as a spectrum of a single earthquake event as it rather represents an envelope of spectra corresponding to different seismic events and sources. On the contrary, a Conditional Mean Spectrum (CMS) represents the expected (i.e., mean) response spectrum, *conditioned* on the occurrence of a target spectral acceleration at the period of interest, typically, though not exclusively, the fundamental period of the structure T_1.

Similarly to the UHS, the target CMS is calculated for the scenario $S_a(T_1)$, magnitude, M, and source-to-site distance, R, together with

S

other preliminary criteria such as fault type or soil, if desirable. In simpler words, the CMS represents the expected response spectrum for the defined ground motion scenario, which is based on a target $S_a(T_1)$ at a single structural period, in contrast to the UHS, which represents equally rare $S_a(T)$ values at many periods (including T_1) simultaneously (Haselton 2009). To develop this new target spectrum, Probabilistic Seismic Hazard Assessment can be used to determine the spectral acceleration $S_a(T_1)$ that corresponds to the target probability of exceedance at the site of interest, denoted as $S_a(T_1)^*$. De-aggregation can then be used to estimate the mean values of magnitude, source-to-site distance, and "epsilon" ("ε") $\left(\overline{M}, \overline{R}, \overline{\varepsilon}\right)$ that lead to an acceleration equal to $S_a(T_1)^*$. The definition of parameter "epsilon" is provided in the following section. The way in which spectral matching to the target CMS is achieved is similar to the procedures described above.

Selection and Scaling to a Proxy ("ε") of the Target Conditional Mean Spectrum

A more advanced approach category of methods for ground motion selection and scaling to a target Conditional Mean Spectrum utilizes the parameter "epsilon" ("ε") as a proxy of the desirable CMS spectral shape (Baker and Cornell 2006). It is recalled that "epsilon" was first defined by engineering seismologists as the number of standard deviations by which a given spectral acceleration, expressed in logarithmic terms, differs from the mean logarithmic spectral acceleration provided by a ground motion prediction (attenuation) equation. In other words, "ε" is derived by subtracting the predicted mean logarithmic spectral acceleration at given period T_1 ($\ln\{S_a(T_1)\}$) from the corresponding value ($\ln\{S_a(T_1)\}$) of the record under examination and then dividing by the logarithmic standard deviation estimated by the attenuation relationship. In practical terms, this implies that a record with "ε of 1.5 at 0.5 s" has a spectral acceleration at a period $T1 = 0.5$ s that is 1.5 standard deviations higher than the predicted mean spectral value.

An advantage of the "ε" parameter is that it is determined with respect to the unscaled record

and does not change in case of record scaling. On the other hand, it is also noticeable that for a given ground motion record, "ε" is clearly a function of the period of interest and depends on the particular ground motion prediction model used, since different attenuation relationships lead to different mean and standard deviation of $\ln\{S_a(T)\}$. Therefore, it is important to ensure that the ground motion prediction model used to compute ε is the same model used in the ground motion hazard assessment. This dependence of ε to attenuation relationships is perhaps the most important drawback in the use of this parameter.

Ground motion selection methods (Tothong and Luco 2007) that use "ε" as a proxy of the CMS spectral shape ensure that the value of "ε" at the fundamental period of the building, denoted as $\varepsilon(T_1)$, is as close as possible to the target $\varepsilon(T_1)$ of the ground motion scenario. This practically implies that the record-to-record variability at the fundamental period of the structure is reduced, and in turn, the discrepancy in the structural response of the corresponding SDOF system is also lower. Some concerns have been expressed regarding the appropriate choice of CMS and the conditioning period, primarily related to the importance of higher modes of vibration of MDOF systems and the anticipated period elongation under strong ground motions (Katsanos et al. 2014). However, it has been shown that risk-based assessments are relatively insensitive to the choice of conditioning period provided that the ground motions are carefully selected to ensure hazard consistency (Lin et al. 2013a).

Available Computational Tools and Databases

Strong Ground Motion Databases

Among numerous strong ground motion databases in Japan, Taiwan, and Europe (European Strong Ground Motion database, www.isesd.hi. is), the PEER-NGA Next Generation Attenuation strong-motion database is a continuously developing project currently consisting of 3,551 publicly available, three-component seismic records (i.e., about 10,650 individual earthquake

acceleration time series) that have been recorded during 173 shallow crustal earthquakes from active tectonic regions worldwide. The corresponding seismic events, which have been recorded primarily in California, range in magnitude from 4.2 to 7.9 and cover epicentral distances in the range 0.2–600 km. Apart from the magnitude and the distance, the earthquake database contains basic information about the seismic source including date and time of the event, hypocenter location, faulting mechanism, seismotectonic environment, and others. Detailed data about 1,600 strong-motion stations are also provided (i.e., site characterizations, surface geology, shallow subsurface conditions, the location of the instrument inside the structure's installation place). Furthermore, each acceleration time history has been corrected for the response of the strong-motion instrument itself and filtered out the noise included while it can also be automatically scaled online.

Software and Tools for Ground Motion Selection

Given the above extensive repository of earthquake records and the fact that the most common earthquake record selection procedures involve spectral matching, recent work evolved to develop computational tools for quantifying and/or optimizing spectrum compatibility to a code prescribed or CMS (Youngs et al. 2007). REXEL (Iervolino et al. 2010a, 2011, 2012) was the first, all-in-one, software introduced for this purpose and facilitates the search for suites of waveforms compatible to target spectra that are either user defined or automatically generated according to Eurocode 8 and the recently issued Italian seismic code. An alternative web-based software for earthquake record selection is SelEQ (Dias 2010), offering various filtering options. More recently, the Integrated System for Structural Analysis and Record Selection (ISSARS) software has been developed (Katsanos and Sextos 2013), retrieving dynamically ground motions from the PEER-NGA database, to form suites of records that not only comply with specific criteria but also ensure, through numerical analyses of the structure that

run at the background, a target level of dispersion of structural response quantities.

Accepted Knowledge on Ground Motion Selection for Response History Analysis Purposes

– Structural response is inherently probabilistic in nature due to the variability among ground motions.
– A major challenge in choosing the most appropriate method for ground motion selection and scaling is to understand the purpose of the dynamic analysis for which the ground motions are sought. A major distinction is between design and assessment purposes. Careful selection of methods, intensity measures, and engineering demand parameters is necessary.
– An additional challenge, at least for a design viewpoint, is to keep structural response discrepancy low.
– Numerous GMSM methods exist in the literature with distinct advantages and drawbacks. The direct comparison of these methods is not always feasible.
– Most GMSM methods adopted in contemporary seismic codes are rather simplified compared to the breadth of the existing methods available in the literature. Still, they require significant effort, deep understanding of the physical problems and the parameters involved as well as specialized computational tools in order to overpass their inherent limitations.

Summary

Linear and nonlinear dynamic analysis of structures is becoming increasingly popular in structural design and assessment practice. Along these lines, the selection and scaling of the appropriate set of earthquake ground motions, required in the framework of dynamic analysis, has become of paramount importance due to the significant sensitivity of structural response to the assumptions made in forming the necessary set of earthquake

records. A large number of ground motion selection and modification (GMSM) methods have been proposed in the literature; still though, the major progress made during the last decade is not yet reflected in modern seismic codes. Along these lines, the present entry aims at briefly presenting, in the simplest possible terms but not simpler than necessary, the objectives and fundamental concepts of GMSM methods along with the current seismic code framework and the computational tools developed to facilitate code-prescribed procedures.

Cross-References

▶ Assessment of Existing Structures Using Response History Analysis
▶ Conditional Spectra
▶ Engineering Characterization of Earthquake Ground Motions
▶ Nonlinear Dynamic Seismic Analysis
▶ Performance-Based Design Procedure for Structures with Magneto-Rheological Dampers
▶ Probabilistic Seismic Hazard Models
▶ Response-Spectrum-Compatible Ground Motion Processes
▶ Seismic Collapse Assessment
▶ Spatial Variability of Ground Motion: Seismic Analysis
▶ Stochastic Ground Motion Simulation
▶ Time History Seismic Analysis

References

American Society of Civil Engineers (2005) Seismic design criteria for structures, systems, and components in nuclear facilities, Structural Engineering Institute, Working Group for Seismic Design Criteria for Nuclear Facilities, ASCE/SEI 43-05, Reston, VA, 81 pp

ASCE/SEI (2010) Minimum design loads for buildings and other structures: ASCE standard 7-10. American Society of Civil Engineers/Structural Engineering Institute, Reston

Baker JW (2011) The conditional mean spectrum: a tool for ground motion selection. J Struct Eng 137(3):322–331. doi:10.1061/(ASCE)ST.1943-541X.0000215

Baker JW, Cornell CA (2005) A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon. Earthq Eng Struct Dyn 34:1193–1217. doi:10.1002/eqe.474

Baker JW, Cornell CA (2006) Spectral shape, epsilon and record selection. Earthq Eng Struct Dyn 35(9):1077–1095. doi:10.1002/eqe.571

Beyer K, Bommer JJ (2007) Selection and scaling of real accelerograms for bi-directional loading: a review of current practice and code provisions. J Earthq Eng 11:13–45. doi:10.1080/13632460701280013

CEN (2004) European Standard EN 1998-1. Eurocode 8: design of structures for earthquake resistance, part 1: general rules, seismic actions and rules for buildings, committee for standardization. Design, vol 3. European Committee for Standardization, Brussels

CEN (2005) European Standard EN 1998-2. Eurocode 8: design of structures for earthquake resistance – part 2 bridges, committee for standardization, vol 3. European Committee for Standardization, Brussels

CPA (2011) Seismic design code and commentary for buildings, construction and planning agency. Ministry of Interior Affair, Taipei (in Chinese), pp 4–51

Dias J (2010) SelEQ⃝: a web-based application for the selection of earthquake ground motions for structural analysis. In: The 14th European conference on earthquake engineering, Ohrid, 30 Aug–3 Sept

Elnashai AS, McClure DC (1996) Effect of modelling assumptions and input motion characteristics on seismic design parameters of RC bridge piers. Earthq Eng Struct Dyn 25(5):435–463

FEMA (2009a) NEHRP recommended seismic provisions for new buildings and other structures, FHMA 750. Building Seismic Safety Council, Washington, DC

FEMA (2009b) Quantification of building seismic performance factors, FEMA P695. Federal Emergency Management Agency, Washington, DC

Hachem MM, Mathias NJ, Wang YY, Fajfar P, Tsai K-C, Ingham JM, ... Francisco S (2010) An international comparison of ground motion selection criteria for seismic design. In: Codes in structural engineering, developments and needs for international practice, joint IABSE – fib conference. Dubrovnik, pp 237–250

Haselton CB (2009) Evaluation of ground motion selection and modification methods: predicting median interstory drift response of buildings, PEER report, 2009/01

Iervolino I, Galasso C, Cosenza E (2010a) New features of REXEL 2. 61 beta, a tool for automated record selection. In: Bulletin of earthquake engineering, vol 8. Ohrid, 30 Aug–3 Sept, pp 339–362. doi:10.1007/s10518-009-9146-1

Iervolino I, Galasso C, Cosenza E (2010b) REXEL: computer aided record selection for code-based seismic structural analysis. Bull Earthq Eng 8:339–362. doi:10.1007/s10518-009-9146-1

Iervolino I, Galasso C, Paolucci R, Pacor F (2011) Engineering ground motion record selection in the ITalian

ACcelerometric Archive. Bull Earthq Eng 9(6):1761–1778. doi:10.1007/s10518-011-9300-4

Iervolino I, Galasso C, Chioccarelli E (2012) REXEL 3.3: closing the loop of computer aided record selection. In: The 15th world conference on earthquake engineering, vol 10, Lisbon

Jalayer F, Cornell CA (2009) Alternative nonlinear demand estimation methods for probability-based seismic assessments. Earthq Eng Struct Dyn 38(8):951–972

Jayaram N, Lin T, Baker JW (2011) A computationally efficient ground-motion selection algorithm for matching a target response spectrum mean and variance. Earthq Spectra 27(3):797–815. doi:10.1193/1.3608002

Katsanos EI, Sextos AG (2013) ISSARS: an integrated software environment for structure-specific earthquake ground motion selection. Adv Eng Softw 58:70–85. doi:10.1016/j.advengsoft.2013.01.003

Katsanos EI, Sextos AG, Manolis GD (2010) Selection of earthquake ground motion records: a state-of-the-art review from a structural engineering perspective. Soil Dyn Earthq Eng 30(4):157–169. doi:10.1016/j.soildyn.2009.10.005

Katsanos EI, Sextos AG, Elnashai AS (2014) Prediction of inelastic response periods of buildings based on intensity measures and analytical model parameters. Eng Struct 71:161–177. doi:10.1016/j.engstruct.2014.04.007

Kottke AR, Rathje EM (2008) A semi-automated procedure for selecting and scaling recorded earthquake motions for dynamic analysis. Earthq Spectra 24(4):911–932

Kramer SL (1996) Geotechnical earthquake engineering. Prentice Hall, Upper Saddle River, N.J., 653 pp

Lin T, Haselton CB, Baker JW (2013a) Conditional-spectrum-based ground motion selection. Part II: intensity-based assessments and evaluation of alternative target spectra. Earthq Eng Struct Dyn. doi:10.1002/eqe

Lin T, Haselton CB, Baker JW (2013b) Conditional-spectrum-based ground motion selection. Part I: hazard consistency for risk-based assessments. Earthq Eng Struct Dyn. doi:10.1002/eqe

Luco N, Cornell CA (2007) Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. Earthq Spectra 23(2):357–392. doi:10.1193/1.2723158

McGuire RK (2004) Seismic hazard and risk analysis, Earthquake Engineering Research Institute, Berkeley, 240 pp

Naeim F, Alimoradi A, Pezeshk S (2004) Selection and scaling of ground motion time histories for structural design using genetic algorithms. Earthq Spectra 20(2):413–426

Padgett JE, Desroches R (2007) Sensitivity of seismic response and fragility to parameter uncertainty. J Struct Eng 133(12):1710–1718

Sextos AG, Katsanos EI, Manolis GD (2010) EC8-based earthquake record selection procedure evaluation: validation study based on observed damage of an irregular R/C building. Soil Dyn Earthq Eng 1–15. doi:10.1016/j.soildyn.2010.10.009

Shome N, Cornell CA, Bazzurro P, Carballo EJ (1998) Earthquakes, records and nonlinear responses. Earthq Spectra 14(3):469–500

Standards New Zealand (SNZ) (2004) NZS 1170.5:2004 – Structural design actions. Earthquake actions. Standards New Zealand, Wellington

Tothong P, Luco N (2007) Probabilistic seismic demand analysis using advanced ground motion intensity measures. Earthq Eng Struct Dyn 36:1813–1835. doi:10.1002/eqe

Wang G (2010) A ground motion selection and modification method capturing response spectrum characteristics and variability of scenario earthquakes. Soil Dyn Earthq Eng 1–15. doi:10.1016/j.soildyn.2010.11.007

Youngs RR, Power MS, Wang G, Makdisi F, Chin CC (2007) Design ground motion library (DGML) – tool for selecting time history records for specific engineering applications. In: Proceedings of SMIP07 seminar on utilization of strong-motion data, Sacramento

Sensitivity of First-Excursion Probabilities for Nonlinear Stochastic Dynamical Systems

Marcos Valdebenito and Hector A. Jensen
Department of Civil Engineering, Santa Maria University, Casilla, Valparaiso, Chile

Synonyms

First Excursion; Stochastic Dynamics; Sensitivity; Simulation Methods

Introduction

Quantification of the performance of structural systems subject to dynamic loading is of paramount interest in several fields of engineering and particularly in the case of earthquake engineering. Knowledge on the performance of a structure during seismic events allows taking design decisions that ensure its serviceability and safety throughout its life. Nonetheless, quantification

S

of performance is a challenging task as there is always uncertainty on future loadings that affect a structure during its lifetime. Structural reliability has emerged as a discipline that allows accounting for the unavoidable effects of uncertainty over performance. Thus, probability theory is used to describe the uncertainty associated with different relevant parameters that affect performance by means of random variables, random fields, and/or stochastic processes. In this manner, uncertainty is *propagated* from these *input* parameters to the responses of interest such as displacements, accelerations, forces, etc.

A particularly useful way to measure the effects of uncertainty in the dynamic response of structural systems is the so-called first-excursion probability. This probability is widely used in stochastic structural dynamics and measures the chances that one or more structural responses exceed a prescribed threshold level within the duration of a dynamical excitation (Soong and Grigoriu 1993). First-excursion probability estimation is particularly challenging as characterization of uncertain loading usually comprises stochastic processes whose discrete representation can involve hundreds or even thousands of random variables. Similarly, the number of possible failure criteria involved can be extremely large as well, i.e., there can be several responses of interest that must be controlled at a large number of discrete time instants. Hence, several different techniques have been proposed in order to estimate first-excursion probabilities. Among these, methods based on simulation (such as the Monte Carlo method and its more advanced variants) have been shown to be the most appropriate approach to compute these probabilities (Schuëller et al. 2004).

Although first-excursion probability provides a most useful way to rationally account for the effects of uncertainty on structures subject to stochastic loading, it is certainly not the only metric that should be taken into account when designing a system. In fact, it is also of interest analyzing the *sensitivity* of the probability with respect to variations in the properties of the structural system. For example, determining the variation in probability due to a change in the size of a structural member can provide useful information to increase the safety level or to identify the most influential design parameters. Nonetheless, estimation of the sensitivity of first-excursion probabilities for dynamical systems is a challenging task as it comprises not only taking into account the uncertainty in input parameters but also assessing how performance is affected due to variations in properties of a structure.

This contribution presents an approach for assessing probability sensitivity of systems subject to stochastic excitation with emphasis on structures whose response is nonlinear. The approach combines state-of-the-art simulation strategies with a series of approximation concepts. Salient features of the approach are the capability of considering problems involving a large number of random variables (in the order of thousands), the possibility of estimating sensitivity with respect to several variables simultaneously (scalability), and a high numerical efficiency achieved by integrating an advanced simulation algorithm with local approximations of the functions modeling the structural performance.

Formulation of the Problem

Structural Model

Let $f(t)$ be a scalar representing loading acting over a structure during a time span $t \in [0, T]$. This load is modeled at discrete time instants of analysis $t_k = (k - 1)\Delta t$, $k = 1, \ldots, n_T$ where Δt is the time step and n_T is the number of time points considered (clearly, $\Delta t = T/(n_T - 1)$). The loading $f(t)$ is uncertain and is characterized by means of a stochastic process using an appropriate representation (Schuëller 1997), e.g., Karhunen-Loève (KL) expansion, polynomial chaos (PC) expansion, etc. Thus, the loading can be represented as $f(t_k, z)$ where z is a vector of random variables of dimension n_z whose associated probability density function $f_z(z)$ depends on the characteristics of the stochastic process. Note that depending on the specific situation under

analysis, load could also be represented as a -vector-valued stochastic process (i.e., a vector whose entries are stochastic processes) instead of a scalar. However, in order to simplify the presentation of this contribution and with no loss of generality, the vector-valued case is not explored further.

In addition to the stochastic loading $f(t, z)$, consider a vector $y \in \Omega_y \subset \mathbb{R}^{n_y}$ of dimension n_y grouping the *design* variables of the structural system. This vector can group those variables that can be altered during the design process (such as cross section of structural members, material properties, etc.). Moreover, consider a classically damped structural system represented by an appropriate model (e.g., a finite element model (Bathe 1996)) comprising a total of n degrees of freedom. Then, the differential equation describing the response of the structure subject to the stochastic excitation $f(t, z)$ is (see, e.g., Chopra, 1995)

$$
\begin{aligned}
&M(y)\ddot{x}(t, y, z) + C(y)\dot{x}(t, y, z) + K(y)x(t, y, z) \\
&+ f_{NL}(y, x(t, z), \dot{x}(t, y, z)) = gf(t, z)
\end{aligned}
\tag{1}
$$

where x is the displacement response vector of dimension n; M, C, and K are the mass, damping, and stiffness matrices of dimension $n \times n$; $f_{NL}(x, y(t, y, z), \dot{x}(t, y, z))$ is a vector of dimension n representing the effect of nonlinear forces of the structure due to, e.g., special devices for energy dissipation, yielding, nonlinear behavior, etc.; and g is the vector of dimension n that couples the stochastic excitation $f(t, z)$ with the degrees of freedom of the structure. Note that mass, damping, stiffness, and nonlinear forces are a function of the vector of design variables y.

First-Excursion Probability

The performance of the structural system in view of the stochastic excitation is characterized by means of n_r responses of interest $r_i(t, y, z)$, $i = 1, \ldots, n_r, t \in [0, T]$ measuring, e.g., displacements, accelerations, stresses, etc. For example, in applications associated with earthquake

engineering, a response of interest r_i could be the interstory drift displacement, which can be calculated as a linear combination of the displacement vector, i.e., $r_i(t, y, z) = \gamma_i^T x_i(t, y, z)$, where γ_i is a vector of size n whose entries are 0 and 1. For design purposes, the responses of interest $r_i, i = 1, \ldots, n_r$ are checked against allowable threshold levels $r_i^*, i = 1, \ldots, n_r$. In a deterministic design framework, the objective is verifying that these responses do not exceed their prescribed thresholds in order to avoid undesirable situations (such as loss of serviceability or collapse). However, when uncertainties are explicitly taken into account, the aforementioned condition cannot be always satisfied, i.e., there is the chance that the responses surpass their prescribed thresholds thus leading to an undesirable situation.

In order to account for the effects of uncertainties and potential undesirable performance of a structure during its lifetime, reliability offers the means for quantifying the level of safety associated with a structural system. A criterion widely used for characterizing safety of a structure is the first-excursion probability (see, e.g., Soong and Grigoriu, 1993). This probability measures the chances that uncertain structural responses exceed in magnitude prescribed thresholds within a specified time interval. That is, first-excursion probability measures the chances of occurrence of the following event F (which is termed in the sequence as *failure* event):

$$
F = D_N(y, z) \geq 1
\tag{2}
$$

where $D_N(y, z)$ is the so-called normalized demand (Au and Beck 2001) and is defined as

$$
D_N(y, z) = \max_{i=1,\ldots,n_r} \left(\max_{t \in [0, T]} \left(\frac{|r_i(t, y, z)|}{r_i^*} \right) \right)
\tag{3}
$$

The normalized demand represents the maximum of the quotient between the structural responses of interest and their corresponding threshold levels. Clearly, whenever $D_N(y, z)$ exceeds 1, there is failure as the response exceeds its maximum acceptable value. In this context, note

the word *failure* is not a synonym of collapse. It is intended to denote unacceptable system performance which can range from partial damage states (e.g., loss of serviceability) to collapse depending on the specific application under study.

The probability of occurrence of the failure event can be defined by means of the following classical probability integral:

$$
\begin{aligned}
P_F(\mathbf{y}) = P[D_N(\mathbf{y}, \mathbf{z}) \geq 1] &= \int_{D_N(\mathbf{y}, \mathbf{z}) \geq 1} f_z(\mathbf{z}) d\mathbf{z} \\
&= \int_{\mathbf{z} \in \Omega_z} I_F(\mathbf{y}, \mathbf{z}) f_z(\mathbf{z}) d\mathbf{z}
\end{aligned}
\tag{4}
$$

In the above equation, $P[\cdot]$ denotes probability of occurrence of the argument between brackets, $P_F(\mathbf{y})$ represents the probability of failure (i.e., probability of occurrence of the event F), and $I_F(\mathbf{y}, \mathbf{z})$ denotes the indicator function which is equal to 1 in case the normalized demand is equal or larger than 1 and 0 otherwise.

The evaluation of the failure probability integral is a challenging task as it usually involves a high number of dimensions (for applications of practical interest, n_z can be in the order of thousands) and the normalized demand function $D_N(\mathbf{y}, \mathbf{z})$ can be evaluated point-wise only through (numerically demanding) FE analyses. A possible means for evaluating probability integrals is applying simulation methods. Among different available *simulation methods*, Monte Carlo simulation (MCS) (Metropolis and Ulam 1949) is the most general technique. However, MCS is numerically demanding for estimating low failure probabilities (which are typical in engineering applications). In order to circumvent this issue, *advanced simulation techniques* have been developed, allowing to estimate small failure probabilities at affordable numerical costs (Schuëller et al. 2004).

Sensitivity of First-Excursion Probability

As shown in Eq. 4, the probability of failure is a function of the design variable vector $\mathbf{y}$. Such dependence can be understood as follows: modifications on the design vector do affect the response of the structure and, consequently, also affect the probability of exceeding the prescribed thresholds. In consequence, for decision making and risk analysis, it is important to evaluate the probability of failure $P_F(\mathbf{y})$ and its sensitivity with respect to changes in the design vector $\mathbf{y}$. That is, besides computing the value of the failure probability, it is also of relevance estimating how much the probability changes due to a modification of the design variable vector. A classical measure for sensitivity is calculating the gradient of the quantity of interest. However, within the context of nonlinear dynamics, the estimation of such quantity may not be feasible as the gradient may not exist due to the nonsmooth normalized demand defined in Eq. 3 (see, e.g., Kang et al. 2006). In order to circumvent this difficulty, an approximate representation of the failure probabilities that is differentiable is constructed, and then, the gradient of this approximation is estimated. Details on these approximations are discussed below.

Estimation of First-Excursion Probability Sensitivity

From a mathematical viewpoint, estimating the gradient of the first-excursion probability implies solving the following limit:

$$
\frac{\partial P_F(\mathbf{y})}{\partial y_q} = \lim_{\Delta y_q \to 0} \frac{P_F(\mathbf{y} + \mathbf{v}(q)\Delta y_q) - P_F(\mathbf{y})}{\Delta y_q}, \quad q = 1, \ldots, n_y
\tag{5}
$$

In Eq. 5, $\mathbf{v}(q)$ is a vector of dimension n_y with all entries equal to zero, except by the q-th entry, which is equal to one. Introducing the definition of failure probability (see Eq. 4) in Eq. 5 yields

$$\frac{\partial P_F(\mathbf{y})}{\partial y_q} = \lim_{\Delta y_q \to 0} \frac{P_F\big[D_N\big(\mathbf{y} + \mathbf{v}(q)\Delta y_q, \mathbf{z}\big) \geq 1\big] - P[D_N(\mathbf{y}, \mathbf{z}) \geq 1]}{\Delta y_q}, \, q = 1, \ldots, ny \qquad (6)$$

Note the limit (as well as the partial derivative) in the above equation may not exist as the normalized demand function $D_N(\mathbf{y}, \mathbf{z})$ may be nonsmooth with respect to both the design parameter vector $\mathbf{y}$ and the uncertain variable vector $\mathbf{z}$. In order to avoid this issue and still obtain a sensitivity measure of the first-excursion probability, approximate representations for the normalized demand function and the excursion probability are introduced. These approximations – which were proposed in Jensen et al. (2009) and Valdebenito and Schuëller (2011) – are discussed in the following. The first approximation comprises an approximate representation of the normalized demand function, i.e.,

$$D_N(\mathbf{y} + \Delta\mathbf{y}, \mathbf{z}) \approx \tilde{D}_N(\mathbf{y} + \Delta\mathbf{y}, \mathbf{z})$$
$$= D_N(\mathbf{y}, \mathbf{z}) + \sum_{q=1}^{n_y} a_q \Delta y_q \qquad (7)$$

where $\tilde{D}_N(\mathbf{y}, \mathbf{z})$ is the approximate normalized demand function, $\Delta\mathbf{y}$ is a certain perturbation of

the design variable vector, and a_q, $q = 1, \ldots, n_y$ are real, constant coefficients. The procedure to determine these coefficients is described in section "Numerical Implementation." The second approximation involves an approximate representation of the probability that the normalized demand exceeds a threshold level b, i.e.,

$$P[D_N(\mathbf{y}, \mathbf{z}) \geq b] \approx e^{\psi_0 + \psi_1(b-1)}, \quad b \in [1 - \epsilon, 1 + \epsilon] \qquad (8)$$

where ϵ is a small constant and where ψ_0, ψ_1 are real coefficients. The issue of how to calculate them is analyzed in section "Numerical Implementation." A thorough discussion on the applicability of these two approximations can be found in Jensen et al. (2009) and Valdebenito and Schuëller (2011).

Using the approximations introduced in Eqs. 7 and 8, it can be shown (see Valdebenito and Schuëller (2011)) that the sought gradient can be approximated as follows:

$$\frac{\partial \tilde{P}_F(\mathbf{y})}{\partial y_q} = \lim_{\Delta y_q \to 0} \frac{P\big[\tilde{D}_N\big(\mathbf{y} + \mathbf{v}(q)\Delta y_q, \mathbf{z}\big) \geq 1\big] - P(D_N(\mathbf{y}, \mathbf{z}) \geq 1)}{\Delta y_q} \qquad (9)$$

$$= \lim_{\Delta y_q \to 0} \frac{P[D_N(\mathbf{y}, \mathbf{z}) \geq 1] - a_q \Delta y_q] - P[D_N(\mathbf{y}, \mathbf{z}) \geq 1]}{\Delta y_q} \qquad (10)$$

$$= \lim_{\Delta y_q \to 0} \frac{e^{\psi_o - \psi_1 a_q \Delta y_q} - e^{\psi_o}}{\Delta y_q} \qquad (11)$$

$$= -\psi_1 a_q e^{\psi_o} \qquad (12)$$

$$= -\psi_1 a_q \tilde{P}_F(\mathbf{y}) \qquad (13)$$

In the above equations, $\partial \tilde{P}_F(\mathbf{y})/\partial y_q$ represents the partial derivative of the approximate

representation of the failure probability with respect to y_q, $q = 1, \ldots, n_y$ and $\hat{P}_F(\mathbf{y})$ is the estimate for the failure probability in Eq. 4 which is calculated using an advanced simulation method. Thus, for estimating the gradient of the probability, it is necessary to determine the probability $\hat{P}_F(\mathbf{y})$ and the coefficients ψ_1 and a_q, $q = 1, \ldots, n_y$. A procedure for determining this probability and these coefficients is discussed in the following section.

Numerical Implementation

General Remarks

This section provides details on the procedure to calculate the different coefficients and probability involved in the estimation of the first-excursion probability sensitivity according to the approximate formula proposed in Eq. 13. First, it is explained how excursion probabilities are estimated. Then, the sensitivity of the first-excursion probability with respect to a normalized threshold level is analyzed, and finally, an approximate representation of the normalized demand function is discussed.

First-Excursion Probability Estimations

As already pointed out, estimating a first-excursion probability implies solving the integral in Eq. 4. In this contribution, this probability is evaluated by means of the so-called advanced simulation methods (Schuëller et al. 2004). In particular, in this contribution, subset simulation (SS) (Au and Beck 2001) is applied to estimate first-excursion failure probabilities.

In SS, the failure domain F (see Eq. 2) is defined as a sequence of subsets (or *intermediate* failure events) F_i, $i = 1, \ldots, m$ such that $F_1 \supset F_2 \supset \ldots \supset F_m = F$. Thus, the failure probability is cast as a product of conditional failure probabilities, i.e.,

$$P_F = P[F_m] = P[F_1] \prod_{i=1}^{m-1} P[F_{i+1}|F_i] \qquad (14)$$

where $P[F_{i+1}|F_i]$ is the probability of occurrence of the event F_{i+1} conditioned on the event F_i. In this way, a small failure probability is expressed as the product of larger, conditional probabilities which can be calculated using, e.g., Monte Carlo simulation (MCS). The practical implementation of SS requires an efficient algorithm for generating samples of the uncertain parameters (z) conditioned on an intermediate failure event, such as the modified Metropolis algorithm (Au and Beck 2001). Details on the algorithmic implementation of SS can be found in Au and Beck (2001).

A salient feature of SS is that it *populates* the space of uncertain parameters by means of successive subsets F_i, $i = 1, \ldots, m$, each of which is of

a rarer occurrence than the previous one, i.e., $P[F_i] < P[F_{i-1}]$. Hence, a full run of SS does not only provide the probability of occurrence associated with a normalized threshold $b = 1$ (see Eq. 4), but actually for a range of the normalized threshold levels (Au and Beck 2001).

Sensitivity of First-Excursion Probability with Respect to Threshold

The approximation introduced in Eq. 8 allows estimating the probability that the normalized demand function $D_N(y, z)$ exceeds a prescribed threshold b, i.e., it is sought to estimate the curve relating normalized thresholds b and their corresponding first-excursion probabilities. Note that this curve can be calculated as a byproduct of reliability analysis applying advanced simulation methods such as SS. Thus, the coefficients ψ_0 and ψ_1 can be estimated with no additional numerical efforts once a reliability analysis has been carried out. In fact, when applying SS, the sought coefficients can be calculated with the samples of the normalized demand generated at the last stage of SS. The main idea is generating the curve probability versus normalized threshold in a discrete manner using samples already available. Then, the sought coefficients are calculated in a least square sense considering the analytic approximation of Eq. 8. Details on how to implement this procedure are discussed in Valdebenito and Schuëller (2011).

Approximate Representation of Normalized Demand Function

The approximate representation in Eq. 7 suggests that changes in the normalized demand function $D_N(y, z)$ due to changes in the design variable vector y can be explained through a linear relation. Such relation is clearly not captured exactly by such a simple expression due to a number of reasons: higher-order terms involving y are ignored, no interaction between y and z is considered, the nonlinear nature of the normalized demand function (see Eq. 3) is not captured appropriately, etc. However, in case the coefficients a_q, $q = 1, \ldots, n_y$ are chosen appropriately, it could be expected that $\tilde{D}_N(y + \Delta y, z)$ approximates $D_N(y + \Delta y, z)$ sufficiently well. That is, the coefficients a_q represent an *average sensitivity* on how

Fig. 1 Example – 2-story RC frame structure including nonlinear devices (NLDs)
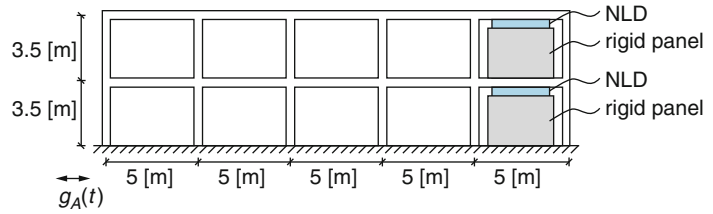
the design variables affect the normalized demand. Numerical validation (see Jensen et al. 2009; Valdebenito and Schuëller 2011) has shown such an assumption is appropriate within the scope of the problems studied in this contribution.

The procedure applied to calculate the coefficients a_q, $q = 1, n_y$ is quite straightforward. Samples of the uncertain variable vector z are taken from the last stage of SS. It is expected that for these samples the normalized demand function is close to 1. Then, for each of these samples, the value of the normalized demand is reevaluated considering perturbed values of the design variable vector. Then, the sought coefficients are estimated in a least square sense considering the sampled data and the analytic model of Eq. 7. For details on the implementation of this procedure, the reader is referred to Jensen et al. (2009) and Valdebenito and Schuëller (2011).

Example

Description of the Problem

In order to illustrate the application of the procedure for failure probability sensitivity estimation, the following example is considered. It involves a two-story reinforced concrete (RC) frame which includes nonlinear hysteretic devices (NLDs). Figure 1 illustrates the elevation of the model. The frame is excited by a horizontal ground acceleration of 15 [s] duration, which is modeled as a stochastic process. The failure event

takes place whenever displacement of each floor of the building exceeds a prescribed threshold within the duration of the stochastic ground acceleration. The design variables of the problem refer to the dimensions of the columns of the RC frame.

The RC frame possesses a Young's modulus equal to 2×10^{10} [N/m^2]. Each of its floors (of mass 1.5×10^5 [kg]) is supported by six columns of square cross section (side length equal to 0.5 [m]) and a height of 3.5 [m]. The beams of the frame are rigid in the axial direction, so each floor can be described by a single horizontal degree of freedom (DOF); thus, the model involves a total of two DOFs. It is assumed that the columns and beams remain linear within the duration of the stochastic ground acceleration and classical modal damping of 5 % for all modes is assumed.

In order to improve the safety of the frame, two hysteretic NLDs are included. The restoring force $(F_{R,i}(\cdot))$ associated with the i-th NLD is described by the following model (Pradlwarter and Schuëller 1993):

$$F_{R,i}(t) = k_d\big(\Delta_i(t) - q_i^1(t) + q_i^2(t)\big), \ i = 1, 2 \tag{15}$$

where $\Delta_i(\cdot)$ is the relative displacement between the $(i, i - 1)$-th floors, k_d is the stiffness of the NLD, and $q_i^1(\cdot)$ and $q_i^2(\cdot)$ denote the plastic elongations of the NLD, which are governed by the following equations:

$$\dot{q}_i^1(t) = \dot{\Delta}_i(t)H\big(\dot{\Delta}_i(t)\big)\left[H\big(\tau_i(t) - \Delta_p\big) + H\big(\tau_i(t) - \Delta_y\big)\frac{\tau_i(t) - \Delta_y}{\Delta_p - \Delta_y}H\big(\Delta_p - \tau_i(t)\big)\right], \ i = 1, 2 \tag{16}$$

$$\dot{q}_i^2(t) = -\dot{\Delta}_i(t)H\big(-\dot{\Delta}_i(t)\big)\left[H\big(-\tau_i(t) - \Delta_p\big) + H\big(-\tau_i(t) - \Delta_y\big)\frac{-\tau_i(t) - \Delta_y}{\Delta_p - \Delta_y}H\big(\Delta_p + \tau_i(t)\big)\right], \ i = 1, 2 \tag{17}$$

where $\tau_i(t)$ is an auxiliary variable defined as $\tau_i(t) = \Delta_i(t) - q_i^1(t) + q_i^2(t)$, $i = 1, 2$, and Δ_y and Δ_p are the yielding and plastic displacements, respectively. The numerical values considered for the parameters of the NLD are $k_d = 10^8$ [N/m], $\Delta_p = 6 \times 10^{-3}$ [m], and $\Delta_y = 0.8\Delta_p$.

The ground acceleration $(g_A(t))$ is modeled as a filtered white noise of 15 [s] duration. The ground acceleration is calculated as $g_A(t) = \boldsymbol{\alpha}^T \boldsymbol{p}(t)$; the vectors $\boldsymbol{\alpha}^T$ and $\boldsymbol{p}(t)$ are defined as

$$\boldsymbol{\alpha}^T =< \Omega_1^2, 2\xi_1\Omega_1, -\Omega_2^2, -2\xi_2\Omega_2 > \quad (18)$$

$$\dot{\boldsymbol{p}}(t) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\Omega_1^2 & -2\xi_1\Omega_1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \Omega_1^2 & 2\xi_1\Omega_1 & -\Omega_2^2 & -2\xi_2\Omega_2 \end{pmatrix} \boldsymbol{p}(t)$$

$$+ \begin{pmatrix} 0 \\ \omega(t)e(t) \\ 0 \\ 0 \end{pmatrix}$$

$$(19)$$

where $\Omega_1 = 15$ [rad/s], $\xi_1 = 0.8$, $\Omega_2 = 0.3$ [rad/s], and $\xi_2 = 0.995$ are the filter parameters; $\omega(t)$ denotes a white noise signal; and $e(t)$ is an envelope function:

$$e(t) = \begin{cases} t^2/16 & 0[s] \leq t < 4[s] \\ 1 & 4[s] \leq t < 10[s] \\ e^{-(t-10)^2} & 10[s] \leq t \leq 15[s] \end{cases} \quad (20)$$

A time discretization step equal to $\Delta t = 0.01$ [s] is used to model the ground acceleration. Thus, the discrete representation of the white noise signal is $\omega(t_k) = \sqrt{2\pi S/\Delta t}\, z_k$, $k = 1, \ldots, 1501$, where $S = 10^{-3}$ [m²/s³] is the spectral density of the white noise and z_k, $k = 1,\ldots, 1501$ are independent, identically distributed standard Gaussian variables.

The failure event is formulated as a first-excursion problem during the time of analysis; the structural responses to be controlled are the two interstory drift displacements and the roof

Sensitivity of First-Excursion Probabilities for Nonlinear Stochastic Dynamical Systems, Table 1 Estimates of first-excursion probability sensitivity (cov: coefficient of variation)

Proposed approach		Finite differences	
$\frac{\partial \tilde{P}_F(\boldsymbol{y})}{\partial y_1}$	$\frac{\partial \tilde{P}_F(\boldsymbol{y})}{\partial y_2}$	$\frac{\partial \tilde{P}_F(\boldsymbol{y})}{\partial y_1}$	$\frac{\partial \tilde{P}_F(\boldsymbol{y})}{\partial y_2}$
(cov)	(cov)		
−2.24	−1.11	−2.49	−1.19
(20.7)%	(24.8)%		

displacement. The threshold values are chosen equal to 0.2 % of the floor height for the interstory drift displacements and 0.1 % of the frame height for the roof displacement. The design variables refer to the cross sections of the columns of the RC frame, more specifically to the second moment of area of the cross section. These cross sections are grouped into two design variables linking the columns of the first and second floor, respectively.

Results

The sensitivity of the first-excursion probabilities is estimated using the approach described previously. In particular, SS is applied considering 2,000 samples of the uncertain ground acceleration at each simulation stage. The resulting first-excursion probability estimate is equal to $\tilde{P}_F = 10^{-3}$. Hence, a total of 6,000 samples of the uncertain variables are required in order to estimate the sought probability. Recall that these samples are also used to estimate the coefficients ψ_0 and ψ_1. In order to calibrate the approximate model of Eq. 7, 200 perturbed designs are analyzed. Hence, the estimation of the probability sensitivity demands only 200 extra structural analyses. In order to illustrate the variability associated with the approach for estimating first-excursion probability sensitivity, a total of 100 independent runs were generated. The results in terms of the mean of these runs and their corresponding coefficient of variation are shown in Table 1. In addition and in order to validate the results obtained in Table 1, the probability sensitivity is estimated using a central finite difference estimator. In order to ensure the probability

estimates used in the finite difference scheme are sufficiently accurate, the average of 100 independent runs is considered. The results associated with the finite differences are presented in Table 1 as well. It can be observed that the results presented are in good agreement, indicating the approach for sensitivity estimation reported herein provides appropriate results.

Summary

This contribution has presented an approach for sensitivity analysis of first-excursion probabilities associated with nonlinear dynamical systems. The basis of the approach is combining advanced simulation methods for probability estimation with a series of local approximations involving the normalized demand function.

A salient feature of the approach reported herein is that it is numerically efficient. This is achieved as the proposed local approximations take advantage of the results already available from probability estimation.

The results presented in this contribution indicate that the proposed approach for sensitivity is applicable for problems involving a large number of uncertain parameters; results presented elsewhere (see, e.g., Valdebenito and Schuëller 2011) indicate that the approach is also capable of including a considerable number of design variables.

Cross-References

References

Au S, Beck J (2001) Estimation of small failure probabilities in high dimensions by subset simulation. Probabilist Eng Mech 16(4):263–277

Bathe K (1996) Finite element procedures. Prentice Hall, New Jersey

Chopra A (1995) Dynamics of structures: theory and applications to earthquake engineering. Prentice Hall, Englewood Cliffs

Jensen H, Valdebenito M, Schuëller G, Kusanovic D (2009) Reliability-based optimization of stochastic systems using line search. Comput Method Appl Mech Eng 198(49–52):3915–3924

Kang B-S, Park G-J, Arora J (2006) A review of optimization of structures subjected to transient loads. Struct Multidiscip Optim 31(2):81–95

Metropolis N, Ulam S (1949) The Monte Carlo method. J Am Stat Assoc 44(247):335–341

Pradlwarter HJ, Schuëller GI (1993) Equivalent linearization – a suitable tool for analyzing MDOF systems. Probabilist Eng Mech 8(2):115–126

Schuëller G (1997) A state-of-the-art report on computational stochastic mechanics. Probabilist Eng Mech 12(4):197–321

Schuëller G, Pradlwarter H, Koutsourelakis P (2004) A critical appraisal of reliability estimation procedures for high dimensions. Probabilist Eng Mech 19(4):463–474

Soong T, Grigoriu M (1993) Random vibration of mechanical and structural systems. Prentice Hall, Englewood Cliffs

Valdebenito M, Schuëller G (2011) Efficient strategies for reliability-based optimization involving non linear, dynamical structures. Comput Struct 89(19–20): 1797–1811

Sensors, Calibration of

Jens Havskov[1] and Gerardo Alguacil[2]
[1]Department of Earth Science, University of Bergen, Bergen, Norway
[2]Instituto Andaluz de Geofísica, University of Granada, Granada, Spain

Synonyms

Accelerometer; Response function; Seismometer; Seismometer damping

S

Introduction

The output of a seismic sensor, a seismometer or an accelerometer, is a time-varying voltage, which is related to the ground motion by a differential equation in the time domain or by a transfer function in the frequency domain. This transfer function or response function is characterized by a number of parameters, which are assumed to be constant, at least in the short term.

A seismic sensor will have calibration information given by the manufacturer. These specifications will be used to correct the seismic signal from the sensor to produce the true ground motion. If, e.g., for a given harmonic ground velocity $X(\omega)$, the output from the sensor is $Y(\omega)$, the amplitude response or transfer function $A(\omega)$ is defined as the ratio $A(\omega) = Y(\omega)/X(\omega)$, where ω is the frequency in radian/s. So, if $A(\omega)$ is known, the input (ground motion) can then be calculated as

$$X(\omega) \,=\, Y(\omega)/A(\omega) \qquad (1)$$

The sensor will, in addition to changing the amplitude of the ground signal, also change the phase of the signal, so an additional function will be needed for the phase response. The phase and amplitude response functions can be combined in one complex function ($A(\omega)$ would then be complex); however, here, for simplicity, they will be treated separately.

The amplitude and phase response functions are usually calculated from the sensor parameters but can also be obtained by direct measurements.

With time, the sensor might degrade or it might develop a fault so there is a need to be able to calibrate the sensor. This means determining the instrument calibration parameters (parametric calibration) or determining $A(\omega)$ directly for each frequency of interest (empirical calibration).

Modern sensors can be very complicated in terms of electronics and mechanical construction so it might not be possible to determine all relevant parameters; however, some parameters can be obtained with simple tests. In this section,

basic parameters will be described, some tests used to obtain them as well as methods for obtaining $A(\omega)$ directly.

Seismic Sensors

Seismic sensors can be divided into two kinds: sensors measuring the ground velocity (velocity sensor or seismometer) and sensors measuring the ground acceleration (accelerometers). A seismometer can be passive, meaning there is no electronic parts or active with an electronic circuit integrated. All accelerometers for seismic recording are active sensors. Piezoelectric accelerometers are passive (although they may have a conditioning circuit built-in) and are widely used for structural vibrations monitoring, but seldom in seismology, due to their poor sensitivity at low frequency and low dynamic range.

Passive Seismometers: A passive seismometer consists of a swinging mass with a coil moving in magnetic field. It is also called an electromagnetic sensor. The swinging system has resonance frequency ω_0. An example is seen in Fig. 1.

When the mass is moving, the magnetic field will vary in the coil. An output voltage proportional to the velocity of the mass relative to the ground will then be produced. The proportionality constant is called the generator constant G and has the unit V/ms^{-1}. But the relative motion of the mass depends on the frequency of the ground motion. Thus the sensor will not be equally sensitive to ground motion for all frequencies. Qualitatively the response of the sensor can be understood as follows. If the ground moves with a very fast sinusoidal motion, the mass remains stationary in an inertial frame, and thus the ground sinusoidal velocity is measured directly. With the ground moving very slowly, the mass would have time to follow the ground motion, so there would be little relative motion and the gain would be low. At the resonance frequency, the mass could get a new push at the exact right time, so the mass would move with a larger and larger amplitude, only limited by the damping of the motion. The sensor amplitude response $A(\omega)$,
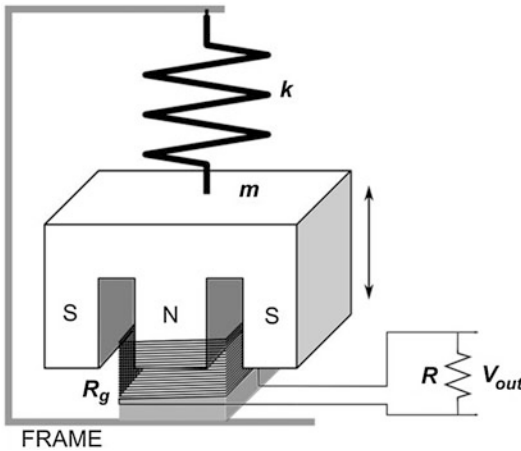
Sensors, Calibration of, Fig. 1 A model of an electromagnetic sensor. The mass m is also a magnet suspended by a spring k. The coil resistance is R_g, the damping resistor is R, and the voltage output is V_{out}. The mass motion is damped by the current through the coil (fixed to the frame) and the external damping resistor

which is the output voltage of the sensor as a function of the input ground velocity, can be obtained as (e.g., Lay and Wallace 1995)

$$A(\omega) = \frac{G\omega^2}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4h^2\omega^2\omega_0^2}} \qquad (2)$$

where h is the damping, ω is the frequency, and ω_0 is the natural frequency of the swinging system. Figure 2 shows examples of $A(\omega)$ with different damping.

It is seen that, as the damping decreases, the gain gets a larger and larger peak at the natural frequency. The flat curve represents a damping of $h = 0.707$ and is the desired value. The damping is achieved by shunting the signal coil with a damping resistor and thereby draining energy out of the swinging system.

Active Sensors: The heart of the active sensor is a device measuring the ground acceleration, the so-called force balanced accelerometer (FBA) (Fig. 3). The FBA has a feedback coil, which can exert a force equal and opposite to the inertia force due to the acceleration. The displacement transducer sends a current to this force coil through a resistor R in a negative feedback loop. The polarity of the current is such that it opposes any motion of the mass, and it will try to prevent the mass from moving at all with respect to the frame. A small permanent acceleration on the mass will therefore result in a small permanent current, and a large acceleration will need a large current. The current is in fact proportional to the ground acceleration, so the voltage over the resistor gives a direct measure of acceleration. This is how nearly all accelerometers work, and in practice the only constant of importance is the generator constant G in units of V/g, where g is the gravity acceleration.

The FBA principle is now the heart of nearly all modern strong motion and broadband sensors (BB, sensors recording in a large frequency band like 0.01–50 Hz). By connecting an integrating circuit after the output, the sensor can give out a voltage proportional to velocity like for passive sensors. However, due to the mechanical-electrical qualities of the sensor, there is in practice a low frequency limit for the flat velocity response. For lower frequencies, the amplitude response decreases proportional to frequency squared. This means that in practice, a BB sensor will have the equivalent of a free period and a damping, and its amplitude response can be approximated by Eq. 2, although a more exact model of its response should include additional parameters that influence the high-frequency behavior. In practice Eq. 2 must then be multiplied with a function which often represents a filter, like a Butterworth filter that can be represented by a simple function of filter frequency and number of poles of the filter; see Havskov and Alguacil (2010). An estimate for the filter parameters can be obtained by empirical calibration; see later.

The sensor parameters to measure can therefore be summarized as instrument natural frequency, damping, and generator constant. With these parameters, the amplitude response function Eq. 2 can be calculated.

Alternatively the amplitude response function can be determined directly by measuring the output of the sensor with a controlled input.

S
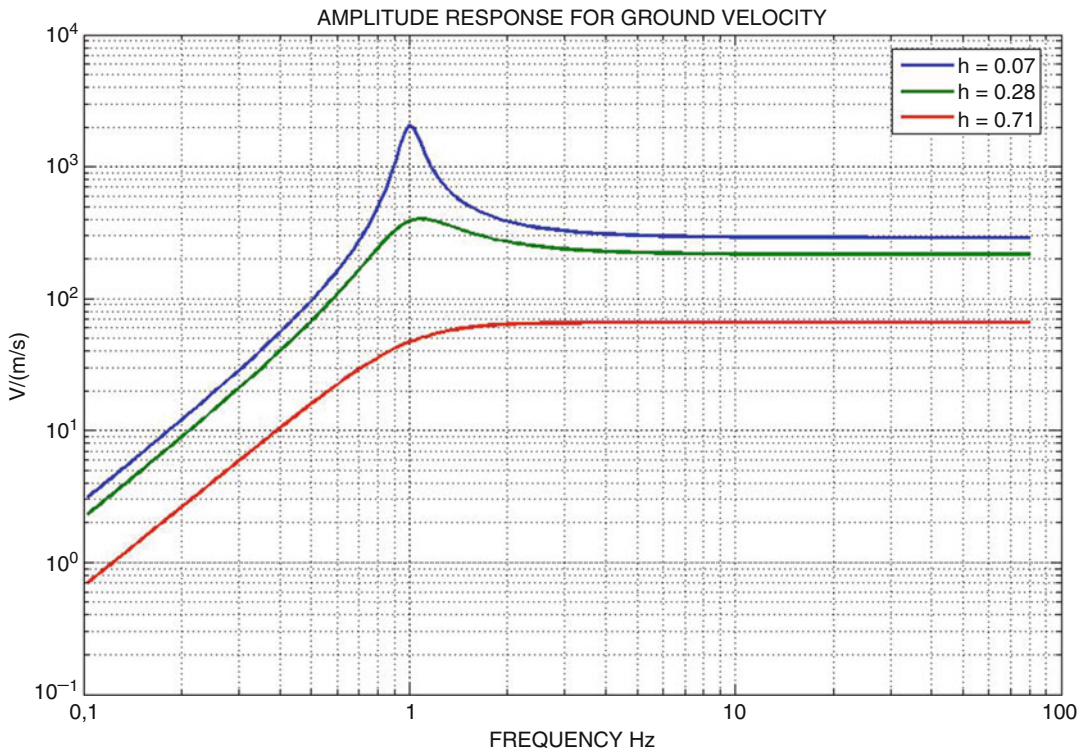
Sensors, Calibration of, Fig. 2 Amplitude response to ground velocity of a typical passive sensor with a natural frequency of 1 Hz for different damping values

Sensors, Calibration of, Fig. 3 Simplified principle behind the force balanced accelerometer. The displacement transducer normally uses a capacitor C, whose capacitance varies with the displacement of the mass. A current, proportional to the displacement transducer output, will force the mass to remain stationary relative to the frame (Figure from Havskov and Alguacil 2010)
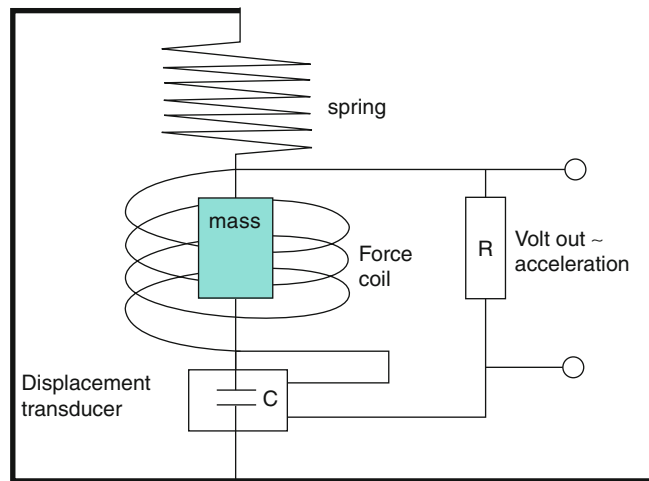
Determining Sensor Parameters

Natural Frequency

The frequency with which the seismometer mass is oscillating depends on the damping; more damping makes it swing at a lower frequency ω_1. The relation is (Havskov and Alguacil 2010)

$$\omega_1 = \sqrt{\omega_0^2(1 - h^2)} \qquad (3)$$
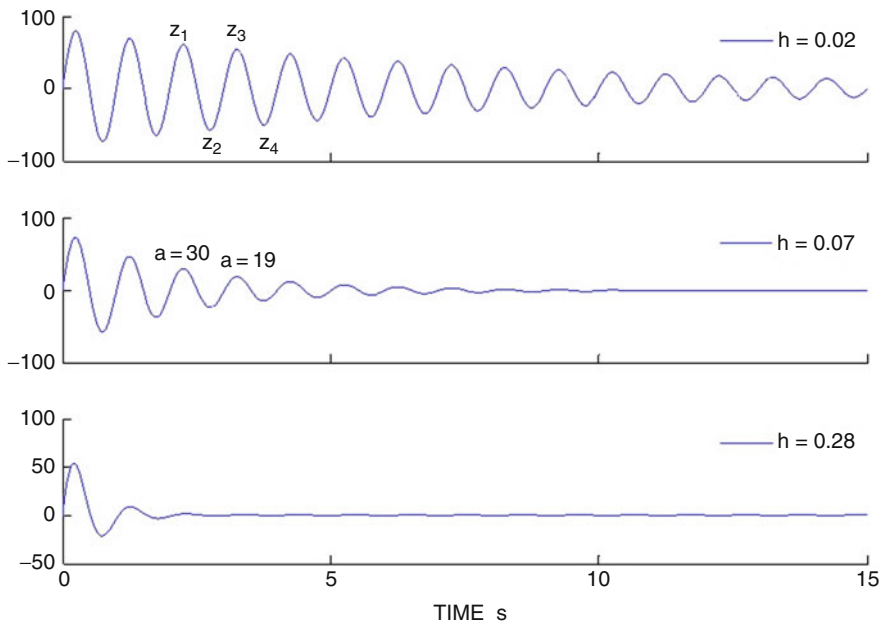
Sensors, Calibration of, Fig. 4 Free swinging of three typical 1 Hz seismometers with different open-circuit damping h. Traces have been generated synthetically. The decaying extremes (peak amplitudes) are labeled z_1, z_2, etc. (*top trace*). On the *middle trace*, the amplitudes of two maximums following each other are given in an arbitrary scale (Figure modified from Havskov and Alguacil 2010)

The natural frequency is therefore best measured without any external damping. The free period can be determined in several ways depending on the type of sensor.

Passive Sensor: Give the sensor an impulse or a step to make it swing, a small push or tilt will do. For a very few sensors, it is possible to see the mass swing, else the output signal has to be observed on an oscilloscope or recorded. Measure the time of several swings and calculate the average or measure the frequency from the recorded signal. Many short-period sensors (natural frequency higher than 1 Hz) have too high open-circuit damping (damping without a damping resistor) to observe more than a few swings of the signal (Fig. 4), and the signal must be recorded to measure the period.

BB Sensor: The response of a BB sensor is controlled mainly by its internal feedback loop. The user can only measure its "apparent free period"; however, this is difficult since the damping is fixed at 0.7 so there will be almost no swinging (see Fig. 9). Most BB sensors have a calibration input, which can be used to produce a calibration pulse from which the "apparent free period" and damping can be obtained; see section on calibration pulse later.

If the open-circuit damping is high, say larger than 0.15 (corresponding to about 1 % change in period), the measured free period should be corrected for damping.

Damping

The open-circuit damping h_m is caused by the friction in the mechanical system and can be determined from the signals from the free swinging system as shown in Fig. 4. Using the extreme amplitudes z_1 and z_n where z_1 is extreme 1 and z_n is extreme $n + 1$ (so $n = 1$ for two following extremes (Fig. 4)), the open-circuit damping can be determined as (Havskov and Alguacil 2010)

$$h_m = \frac{\ln(z_1/z_n)}{\sqrt{n^2\pi^2 + \ln^2(z_1/z_n)}} \qquad (4)$$

Measuring z_1 and z_2, h_m can now be determined directly. In Fig. 4, two following maxima ($n = 2$)

Sensors, Calibration of,
Fig. 5 Tilting an
accelerometer to determine
generator constant. It is
assumed that the sensor
horizontal direction is
toward the east
(Figure from Havskov and
Alguacil 2010)
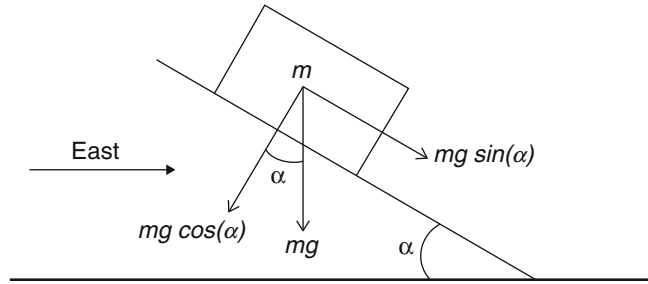
have approximate amplitudes 30 and 19, respectively. This gives a damping of

$$h_m = \frac{\ln(30/19)}{\sqrt{2^2\pi^2 + \ln^2(30/19)}} = 0.0725 \quad (5)$$

A sensor in operation should always have a damping resistor connected in order to get the desired total damping of 0.707 so as to get the most flat response. The additional damping caused by the electrical damping is called h_e, and the corresponding total resistance of the coil and damping resistor is R_T. The damping resistor cannot be determined experimentally as above, since with total damping of 0.707, there would be too little overswing. It can be shown (Havskov and Alguacil 2010) that the ratio between two electrical damping constants is inversely related to the ratio of the damping resistances:

$$R_{T2} = R_{T1}\frac{h_{e1}}{h_{e2}} \quad (6)$$

R_{T1} and R_{T2} are two different total damping resistances and h_{e1} and h_{e2} the corresponding damping coefficients caused by the resistances, respectively. In addition comes the open-circuit damping, so, the total damping with R_{T1} is h_1:

$$h_1 = h_{e1} + h_m \quad (7)$$

and writing Eq. 6 in terms of the total damping gives

$$R_{T2} = R_{T1}\frac{h_{e1}}{h_{e2}} = R_{T1}\frac{h_1 - h_m}{h_2 - h_m} \quad (8)$$

When $h_2 = 0.707$, the corresponding $R_{T2} = R_{0.707}$ is

$$R_{0.707} = R_{T1}\frac{h_1 - h_m}{0.707 - h_m} \quad (9)$$

so the procedure to determine $R_{0.707}$ is:

1. Determine h_m as described above.
2. Connect a damping resistor R_{T1} of "suitable size," meaning that the overswing is substantially less than when measuring h_m but large enough to get an accurate determination of h_1.
3. Use Eq. 9 to determine $R_{0.707}$.

For BB sensors or other active sensors, there is no damping resistor, and the damping is preset from the factory. These sensors have a buffered low-impedance voltage output, so an external load resistor would not influence the sensor at all.

Generator Constant

In order to determine the generator constant experimentally, a known input to the sensor must be used or the generator constant must be calculated from other known measurable parameters.

Accelerometer: For a static ground acceleration, an accelerometer has an output proportional to this acceleration. An accelerometer cannot distinguish any difference between an inertial force due to ground acceleration and an equivalent gravity force. By tilting the accelerometer, the effective force on the three components can be determined for different tilt angles and the sensitivity determined; see Fig. 5.

The accelerometer is tilted an angle α. The force in the horizontal direction is now $mg\sin(\alpha)$,
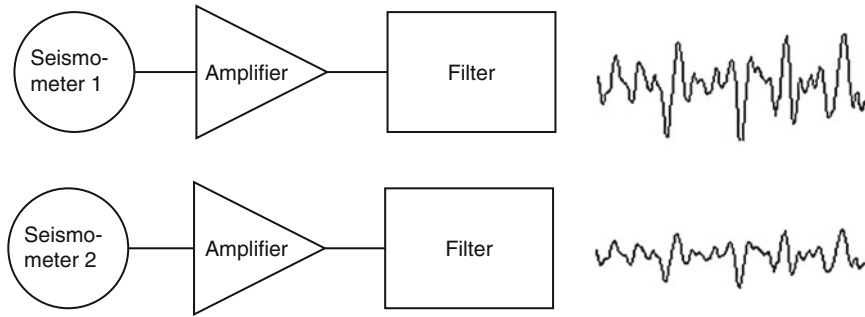
Sensors, Calibration of, Fig. 6 Comparison of signals recorded from two different seismometers with the same recording equipment. The filters and amplifiers are identical and filter out signals below the seismometers natural frequencies. The ratio of the output amplitudes (*right*) indicates the ratio of the sensor generator constants (Figure from Havskov and Alguacil 2010)

while in the vertical direction it is $mg\cos(\alpha)$. The voltage output for the vertical component is V_z, and for the horizontal component it is V_h. Considering that in the horizontal position, the output is supposed to be zero and that the vertical force has been *decreased* from mg to $mg\cos(\alpha)$, the generator constant can be calculated as

$$G_z = \frac{-V_z}{g(1 - \cos(\alpha))} \quad \text{and} \quad G_h = \frac{V_h}{g\sin(\alpha)} \quad (10)$$

V_z will be negative since the vertical force is decreased, and V_h will be positive since the force is in the east direction. This method is very simple and can determine the generator constant accurately, but will not give a dynamic calibration. The horizontal components can be tested for symmetry by inclining in the opposite direction. In the case that the instrument output is not zero in the horizontal position, this offset should be adjusted before any measurement is made; alternatively it may be subtracted from the output for each measurement.

Passive Velocity Seismometer: The generator constant can be calculated from

$$G = \sqrt{2m\omega_0 h_e R_T} \quad (11)$$

where m is the mass of the moving mass and h_e is the electrical damping corresponding to the total resistance R_T. In the previous section it has been described how R_T, h_e, and ω_0 can be determined.

All Sensors: A shaking table is a platform that moves with a controlled motion. Setting it up to move with a harmonic motion with maximum amplitude A (velocity or acceleration) at a frequency where the response curve is flat, the output from the sensor will also be a harmonic function with amplitude B. The generator constant is then

$$G = B/A \quad (12)$$

Unfortunately a shaking table is rarely available, so other methods must generally be used.

The ground is always moving so if the motion is known, it is equivalent to having a shaking table moving both horizontally and vertically. The problem is then to determine this motion in order to have an absolute measure. This is done by using a well-calibrated sensor. The experiment then consists of setting up the two sensors close to each other and assuming that they are subject to the same ground motion (Fig. 6). The data is filtered so as to use a frequency band where both sensors are measuring the signal in the flat part of their instrument response function, e.g., from 2 to 5 Hz using a BB sensor as a reference and a 1 Hz sensor to be calibrated.

The generator constant can then be determined as

$$G = G_R B/A \quad (13)$$

where G_R is the generator constant of the reference sensor, A is the amplitude of the reference sensor, and B is the amplitude of the unknown

sensor. The amplitude can be picked for any part of the two seismograms provided that they are identified as the same "swing"; see Fig. 6. It is here assumed that both sensors are of the same type like comparing a velocity sensor to a velocity senor or an accelerometer to an accelerometer. Note that many accelerometers are not sensitive enough to be able to use this method, so the tilt method should be used.

Calibration of Velocity Sensors by Stepwise Motion: The main idea behind this method is to (1) move the sensor to a known distance (like 1 mm) in its direction of sensitivity, (2) record the signal, and (3) correct the signal for the known displacement frequency response (response function relating sensor output to displacement) to get the displacement. Theoretically, the displacement is now calculated, which can be compared to the actual displacement, from which the generator constant can be calculated. The method is described in detail by Wielandt in chapter "▶ MEMS Sensors for Measurement of Structure Seismic Response and Their Application" in NMSOP (Bormann 2002), which claims that the method works well for broadband sensors (accuracy down to 1 %) and even 10 Hz seismometers. Although the method sounds simple, it is not trivial to correct for the instrument response down to DC, particularly for SP instruments.

Another problem is how to move the sensor, horizontally or vertically, at a controlled distance. A simple instrument that has been used for vertical motion is a mechanical balance (Fig. 7). Placing the sensor on the mechanical balance table, the seismometer can be moved at a controlled distance as measured on the balance arm vertical scale. In principle, it is enough to make one displacement, but usually several are made to take average measurements.

A simple portable calibration table which operates on this principle is commercially available from Lennartz Electronic (www.lennartz-electronic.de). it moves the vertical axis a known displacement. It also allows to tilt the sensor a known angle, which is sensed as an apparent step in acceleration along the horizontal axis, thus permitting the calibration of horizontal components as well.
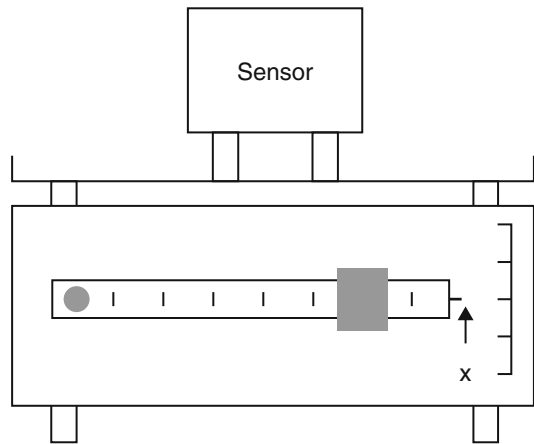
Sensors, Calibration of, Fig. 7 Moving a sensor at a controlled distance vertically with a mechanical balance. The ratio between the motion at x (where we can measure it) and on the balance table can be determined by placing a mass m_1 at x that will balance a mass m_2 on the balance table. The ratio of the two motions is then m_1/m_2 (Figure from Havskov and Alguacil 2010)

It should be noted that some of the triaxial BB seismometers have the three sensors arranged symmetrically, forming an orthogonal frame with its active axes inclined 54.7° with respect to horizontal, so each sensor is sensitive to vertical motions. Their outputs are then electronically combined to generate the conventional two horizontal and one vertical outputs. If the individual sensor outputs are available, they can be calibrated with only the vertical step motion.

Measuring the Complete Response Curve $A(\omega)$

The alternative to obtain the individual sensor parameters and then calculating $A(\omega)$ is to obtain $A(\omega)$ directly. This is an empirical calibration and can only be done by providing a known input to the sensor $X(\omega)$ and measuring the output $Y(\omega)$, from which the response function can be obtained as

$$A(\omega) = Y(\omega)/X(\omega) \qquad (14)$$

Shaking Table: Again a shaking table is the simplest way. The shaking table can move
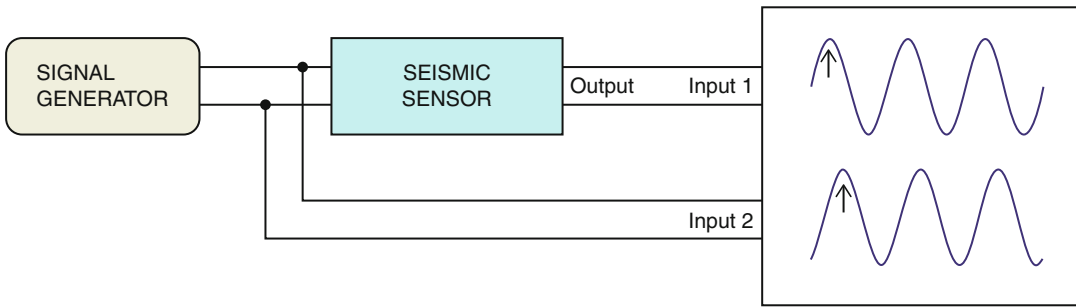
Sensors, Calibration of, Fig. 8 Setup for measuring the instrument response using the calibration input (Figure from Havskov and Alguacil 2010)

vertically and horizontally in a controlled fashion so the ground input is exactly known. All that has to be done is to measure the output, as a function of frequency, and divide by the known input to get the response function. It is not as simple as it sounds. Making a precise shaking table, particularly for horizontal motion, without introducing tilt, is complicated and expensive, so shaking tables are rarely available for the general user and only used in special laboratories.

Sensor Calibration Input: Many sensors have a calibration input. For passive seismometers, this is a calibration coil around the mass, and applying a current to the coil results in a force on the mass. Most active sensors do not have a calibration coil, but an equivalent test can be made when a voltage is applied to the feedback loop to produce a force on the mass. In both cases it is possible to accelerate the moving mass by applying a voltage to the calibration input. By using a sine wave, the sensors experience the equivalent of a ground acceleration, and, by varying the frequency, the complete frequency-dependent amplitude response can be determined, since the input $X(\omega)$ is known and $Y(\omega)$ is measured. This method assumes that the calibration circuit is accurate. An experimental setup is seen in Fig. 8.

A signal generator sends out a sine wave with a constant amplitude V_0. This voltage is applied to the sensor calibration input and exerts a force on the mass proportional to V_0. For the velocity sensor with a calibration coil, the amplitude of the force f_i will be

$$f_i = K_c V_0 / R_c \qquad (15)$$

where K_c is the calibration coil motor constant (N/A) and R_c is the resistance of the calibration coil. For the active sensor

$$f_i = K_a V_0 \qquad (16)$$

where K_a is the calibration input sensitivity (N/V). With a sensor mass of m, the equivalent ground acceleration amplitude is then f_i/m, the equivalent ground velocity is $f_i/m\omega$, and the equivalent ground displacement amplitude is $f_i/m\omega^2$. If the amplitude of the seismometer voltage output is V_s, then the velocity amplitude response (V/ms^{-1}) for this frequency is

$$A(\omega) = \frac{V_s m \omega}{f_i} \qquad (17)$$

and similarly for displacement response. By varying the frequency and measuring both input and output signals, a complete response curve can be obtained.

When the calibration coil is used to calibrate a passive sensor, an undesirable effect may be present: a spurious coupling between the calibration coil and the signal coil. This coupling affects mainly the higher frequencies and may be approximately corrected for by subtracting the signal output when the same current is injected with the sensor mass locked (Sauter and Dorman 1986; Steck and Prothero 1989).

From Fig. 8, it is seen that the output signal not only has been changed in amplitude but also has been delayed a little relative to the input signal; in

other words, there has been a phase shift. In this example, the phase shift is positive; see definition in Eq. 18. The complete frequency response of the sensor therefore consists of both the amplitude response function and the phase response function $\Phi(\omega)$. Considering a general input harmonic waveform $x(\omega,t) = X(\omega) \cdot cos\,(\omega t)$ at frequency ω, the output can be written as

$$y(\omega,t) = X(\omega) \cdot A(\omega) \cdot cos\,(\omega t + \Phi(\omega)) \quad (18)$$

The phase shift is here defined as a quantity being added to the phase as seen above. Thus comparing Fig. 8 and Eq. 18, it is seen that the phase shift is positive.

The phase response function can therefore also be measured using this harmonic drive method. It can similarly be obtained using a shaking table.

Knowing the complete response function $A(\omega)$, it is now possible to get an indication if the response function can be represented by Eq. 2 or there are additional filters to include.

Using the Ground Motion as a Shaking Table: The experiment consists of setting up two sensors close to each other and assuming that they are subject to the same ground motion. If the known and unknown sensors have recorded signals of the ground motion z_1 and z_2, respectively, and corresponding spectra are $Z_1(\omega)$ and $Z_2(\omega)$, the input ground motion (displacement, velocity or acceleration) is

$$X_1(\omega) = \frac{Z_1(\omega)}{A_1(\omega)} \quad (19)$$

where A_1 is the response function of the known sensor 1. The unknown response function A_2 is then

$$A_2(\omega) = A_1(\omega)\frac{Z_2(\omega)}{Z_1(\omega)} \quad (20)$$

Note that if A_1 is in displacement, then A_2 will also be in displacement and similarly if A_1 is in velocity or acceleration. It is, e.g., possible to use an accelerometer to calibrate a seismometer or

vice versa. Thus, by measuring the signals of the known and unknown sensor, it is in principle simple to determine the response function of the unknown sensor. In order for the method to work, the sensor self-noise must be well below the signal levels generated by the ground motion. That will clearly limit the usefulness of the method at lower frequencies, where, e.g., for a geophone, the output signal level will be low. So, contrary to normal sensor installation, these tests should be made in a high ground noise environment like the top floor of a building, although some care is required to avoid air-coupled acoustic noise, which may affect both sensors in different ways (Pavlis and Vernon 1994).

If the spectra $Z_1(\omega)$ and $Z_2(\omega)$ of the two sensors outputs are contaminated with instrumental noise (a noncausal output), a simple ratio Z_2/Z_1 may yield a biased or unreliable estimate of the response. By using only the correlated part of the signals, i.e., the part due to the common input of both sensors, the ground noise in this case, a more reliable estimate, will be obtained. This may be done by using a known relation between input and output of linear and causal systems: the input-output cross-spectrum is the product of the transfer function and the input power spectrum (see, e.g., Ljung 1999). Consider a linear system whose input is the output of seismometer 1 (Z_1) and its output is the output of seismometer 2 (Z_2). This system would have a transfer function $A_2/A_1 = P_{21}(\omega)/P_{11}(\omega)$, or Z_2/Z_1 if these two spectra were noise-free and generated by the same input. Therefore, a more robust estimation of the unknown response A_2 is obtained using

$$A_2(\omega) = A_1(\omega)\frac{P_{21}(\omega)}{P_{11}(\omega)} \quad (21)$$

where $P_{11}(\omega)$ represents the power spectrum of the output of sensor 1 and $P_{21}(\omega)$ is the cross-spectrum between the outputs of sensors 1 and 2. The advantage of this equation is that, ideally, it cancels out the contributions of uncorrelated components (i.e., self-noise) of the input and output signals. This method has been further
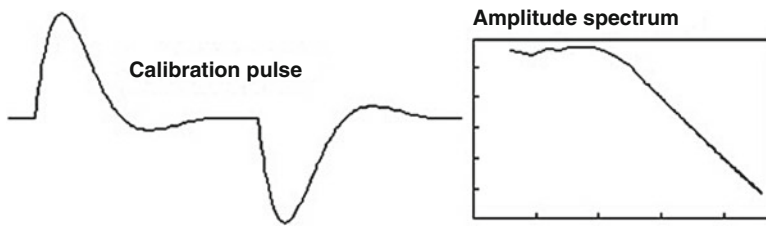
Sensors, Calibration of, Fig. 9 Calibration pulse from a velocity sensor. *Left*: The recorded calibration pulse due to an input step in acceleration. *Right*: The log-log spectrum of the calibration pulse. The sensor natural frequency ω_0 is indicated. The damping is 0.707 (Figure from Havskov and Alguacil 2010)

improved using three channels of data, resulting in more robust estimate of the relative response function (Sleeman et al. 2006). However, small misalignments between the three sensors may prevent the background noise to be fully canceled out, and a leak of such noise is then evaluated as sensor self-noise. Gerner and Bokelmann (2013) suggest a technique to numerically fix this effect by searching for the optimal alignment.

Calibration Pulse: It is common practice to generate a calibration pulse by applying a step current into the calibration coil of a passive sensor. For a BB sensor an equivalent test can be made when a voltage is applied to the feedback loop to produce a force on the mass. As the applied force is proportional to the coil current or the voltage, this input signal is equivalent to a ground step in acceleration (see also section above). The output pulse can be recorded (Fig. 9), and this pulse can be used in several ways to obtain sensor parameters.

From signal theory (e.g., Scherbaum 2007) it is known that the Fourier transform of the impulse response of a linear system is the frequency response function. However, what is generated here is a step. If the amplitude spectrum of the calibration pulse is $C(\omega)$, the response function will then be $A(\omega) = \omega C(\omega)$. This corresponds to the response function for acceleration since the input was an acceleration step so the velocity response for the velocity sensor is then $\omega^2 C(\omega)$. Thus, multiplying the spectrum in Fig. 9 with ω^2 will give the velocity response for $h = 0.707$ seen in Fig. 2. It is also possible to obtain the approximate natural frequency ω_0 as the frequency where the spectral amplitude (Fig. 9) has decreased from the flat level by a factor

$1/\sqrt{2} = 0.707$ or the frequency for which the low-frequency and high-frequency asymptotes intersect. It is then in principle simple to get the frequency response for the sensor by just doing spectral analysis of the calibration pulse; however, ground and electronic noise might make it inaccurate. In practice the signal-to-noise ratio is improved by summing several calibration pulses with their onsets well aligned. The background noise is decreased in this way, while the signal due to the calibration pulses is increased.

Summary

Calibration of seismic sensor can be done in many ways. The calibration can determine the frequency response of a sensor by determining the sensor parameters natural frequency, damping, and generator constant from which the frequency response can be calculated. The alternative is to experimentally determine the complete frequency response of the sensor using, e.g., a shaking table or the sensor calibration input. In most cases it will be possible for the general user to get a good idea about the response function using one of the methods described in this paper.

Cross-References

▶ Passive Seismometers
▶ Principles of Broadband Seismometry
▶ Recording Seismic Signals
▶ Seismic Accelerometers
▶ Seismometer Self-Noise and Measuring Methods

References

Bormann P (ed) (2002) IASPEI New manual of seismological observatory practice (NMSOP). GeoForchungsZentrum Potsdam, Potsdam

Gerner A, Bokelmann G (2013) Instrument self-noise and sensor misalignment. Adv Geosci 36:17–20. doi:10.5194/adgeo-36-17-2013

Havskov J, Alguacil G (2010) Instrumentation in earthquake seismology. Springer, Dordrecht, p 358

Lay T, Wallace TC (1995) Modern global seismology. Academic, San Diego, p 521. ISBN 0-12-732870-X

Ljung L (1999) System identification: theory for the user, 2nd edn. Prentice Hall PTR, Englewood Cliffs, p 609

Pavlis GL, Vernon FL (1994) Calibration of seismometers using ground noise. Bull Seismol Soc Am 84:1243–1255

Sauter AW, Dorman LM (1986) Instrument calibration of ocean bottom seismographs. Mar Geophys Res 8:265–275

Scherbaum F (2007) Of poles and zeros, fundamentals of digital seismology, revised 2nd edn. Springer, Dordrecht, p 271

Sleeman R, van Wettum A, Trampert J (2006) Three-channel correlation analysis: a new technique to measure instrumental noise of digitizers and seismic sensors. Bull Seismol Soc Am 96(1):258–271

Steck L, Prothero WA (1989) Seismic calibration using the simplex algorithm. Bull Seismol Soc Am 79:1618–1628

Site Response for Seismic Hazard Assessment

Gaetano Elia
School of Civil Engineering and Geosciences, Newcastle University, Newcastle Upon Tyne, UK

Introduction

Lessons learned worldwide from historical (e.g., Niigata, Japan, and Alaska in 1964) and recent strong earthquakes (e.g., L'Aquila in 2009, Chile in 2010, Tohoku, Japan, and Christchurch in 2011, among others) have distinguished site amplification and soil liquefaction as two of the main causes of damage to man-made and natural structures during seismic events. As illustrated in Fig. 1, the ground shaking observed at surface during an earthquake depends on the seismic source characteristics and focal mechanism, the deep wave propagation from the fault to the bedrock, and the local soil conditions. The first two phenomena are commonly studied by seismologists, geologists, and geophysicists, while the third one falls in the geotechnical earthquake engineering field, being strongly related to the mechanical behavior of soils subjected to dynamic loading. The term "site effects" refers to the overall set of modifications of the bedrock motion, in terms of amplitude, frequency content, and duration, during its (almost) vertical propagation through shallow geological deposits. Despite the traveling path of the propagating waves within the surface soils is often less than 100 m, local site conditions can produce significant ground motion amplification, slope instability, excessive building settlements (i.e., subsidence), and liquefaction in loose and saturated granular deposits, thus playing a crucial role in building and infrastructure damage (Fig. 1).

Ground response numerical analyses are, therefore, used by geotechnical engineers to predict the free-field motion, to determine permanent soil deformations (leading to subsidence and liquefaction), and to evaluate the risk of instability of earth structures (e.g., earthquake-induced landslides). They may also include the evaluation of basin and topographic effects on ground motion (e.g., King and Tucker 1984; Bard and Bouchon 1985; Geli et al. 1988; Bard 1994; Ashford et al. 1997; Bouckovalas and Papadimitriou 2005; Semblat et al. 2005; Pagliaroli et al. 2011). In addition, the results of these studies can be incorporated into microzonation and probabilistic seismic hazard analyses (e.g., Tsai 2000; Bazzurro and Cornell 2004; Papaspiliou et al. 2012).

Site response analyses have been traditionally performed using a one-dimensional (1D) frequency-domain numerical scheme based on the equivalent viscoelastic approach. This approach has been extensively adopted in the last 30 years, and it is widely accepted in the engineering practice, although its limitations are well known. Being based on a total stress formulation, it disregards the buildup of excess pore water pressures in the soil deposit. Additionally, the adopted equivalent viscoelastic material properties cannot properly represent the soil
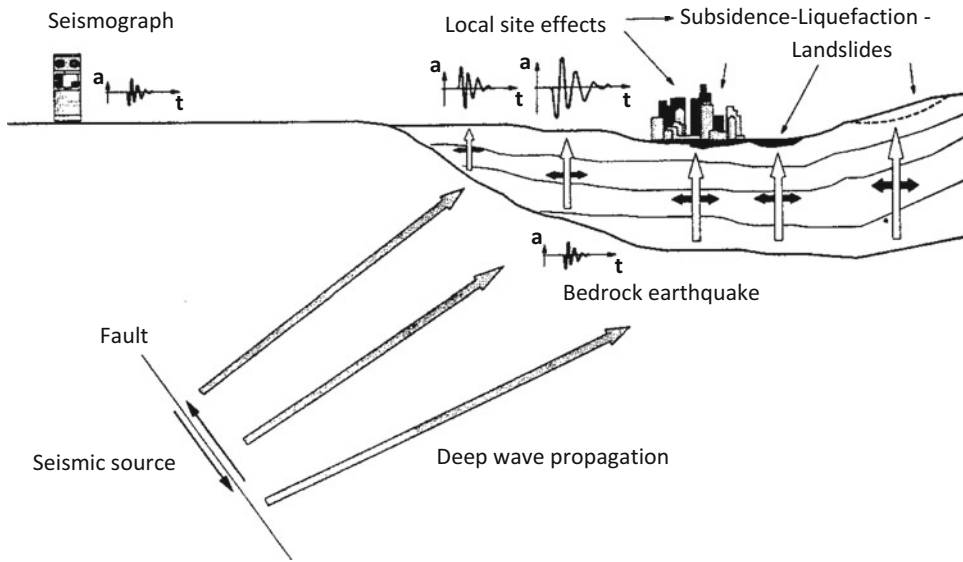
Site Response for Seismic Hazard Assessment, Fig. 1 Wave propagation from seismic source to ground surface and related geotechnical problems (Modified from Lanzo and Silvestri 1999)

behavior under cyclic loading for high seismic intensities at bedrock. Finally, the 1D scheme cannot take into account site effects related to surface and buried complex morphologies, i.e., topographic and valley effects. Time-domain schemes are nowadays available to solve the wave propagation problem in a more realistic way, accounting for the solid–fluid interaction by means of a coupled effective stress formulation. In those schemes, the behavior of the soil can be described using either simple or sophisticated nonlinear soil constitutive models of different level of complexity. In addition, time-domain analyses, usually performed with finite element codes, can also describe two- (2D) and three-dimensional (3D) complex geometries to model topographic and basin effects. Nevertheless, these nonlinear analyses are seldom adopted by nonexpert users because the calibration of advanced soil constitutive models can be challenging and the code usage protocols are often unclear or poorly documented in the literature.

This entry begins with a short overview of available methods for site characterization and evaluation of soil dynamic properties (section "Soil Dynamic Properties and Measurement Techniques"), factors that are essential for the assessment of site response effects. Ground response analyses are then discussed in the section "Ground Response Analysis," describing linear and nonlinear approaches for the study of one-dimensional wave propagation problems in free-field conditions. Soil–structure interaction and earthquake-induced ground failure problems, such as soil liquefaction, landslides, and retaining structure instability, are outside the scope of the entry and, therefore, are not discussed here. A review is given in Kramer (1996), Kramer and Stewart (2004), and Semblat and Pecker (2009). Finally, the section "Future Challenges" provides an overview of future challenges in the field of geotechnical earthquake engineering.

Soil Dynamic Properties and Measurement Techniques

Soil Dynamic Properties

As schematically indicated by Fig. 1, the propagation of waves from bedrock to ground surface is commonly considered as a vertical process. According to Snell's law (cf. Richart et al. 1970), in fact, waves traveling from higher-velocity materials into lower-velocity materials

Fig. 2 Mechanical
behavior of a soil element
subjected to an irregular
simple shear loading
history (Modified from
Lanzo and Silvestri 1999)
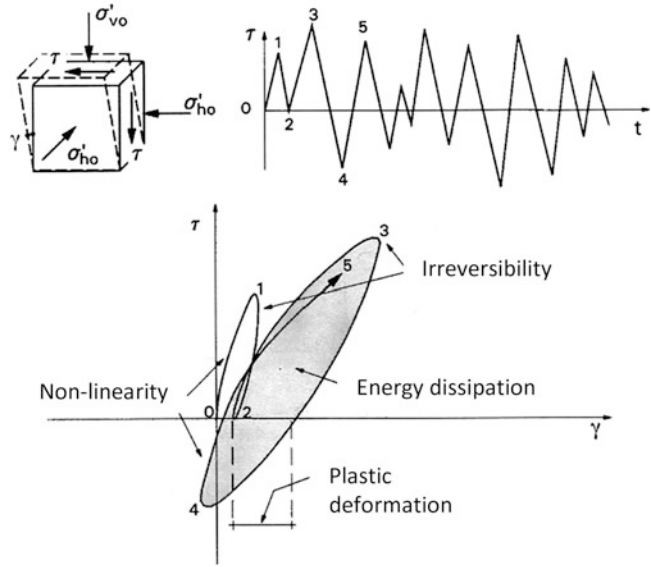
are refracted closer to the normal to the inter-
faces. Therefore, earthquake waves propagating
upward through horizontal layers characterized
by lower velocities and densities (such as in typ-
ical top soil deposits) are refracted closer to a
vertical path. Moreover, saturated soils subjected
to earthquake loading behave essentially in
undrained conditions, given the rapidity of the
seismic action. The volumetric deformations
induced by P-waves are, consequently, negligible
with respect to the distortional deformations
associated with S-waves. For the above reasons,
many of the methods of ground response analysis
presented in the section "Ground Response Anal-
ysis" simulate the seismic event as a loading
process induced by SH-waves only, characterized
by a vertical traveling path associated with soil
particle motion in the horizontal plane. This phe-
nomenon requires the analysis of the mechanical
behavior of soils under simple cyclic shear load-
ing conditions. Considering a generic soil ele-
ment within the deposit in geostatic conditions
(when only the vertical, σ'_{v0}, and the horizontal,
σ'_{h0}, effective stresses are applied), the earthquake
action induces an additional simple shear stress
$\tau(t)$ changing irregularly with time (Fig. 2).

The corresponding stress–strain ($\tau - \gamma$) curve
under this cyclic loading history is typically

characterized by state dependency, early irrevers-
ibility, nonlinearity, buildup of excess pore pres-
sures, decrease of nominal stiffness, and related
hysteretic dissipation (e.g., Sangrey et al. 1969;
Hardin and Drnevich 1972; Castro and Christian
1976; Vucetic and Dobry 1991). Under a sym-
metric cyclic loading condition, the hysteresis
loop in the $\tau - \gamma$ plane can be effectively
described by two parameters: the shear modulus
and the damping ratio. As shown in Fig. 3, the
tangent shear modulus G_{tan} represents the soil
stiffness for a specific point of the loop, and it
changes continuously throughout the cycle.

An overall indication of the average soil stiff-
ness is, instead, represented by the secant shear
modulus G_{sec}, defined as:

$$G_{sec} = \frac{\tau_c}{\gamma_c} \qquad (1)$$

where τ_c and γ_c are the maximum shear stress and
shear strain amplitudes, respectively. The area
included in the hysteresis loop is a measure of
the energy dissipated by the soil during the cycle
and is described by the damping ratio D:

$$D = \frac{W_D}{4\pi W_S} = \frac{A_{loop}}{2\pi G_{sec}\gamma_c^2} \qquad (2)$$
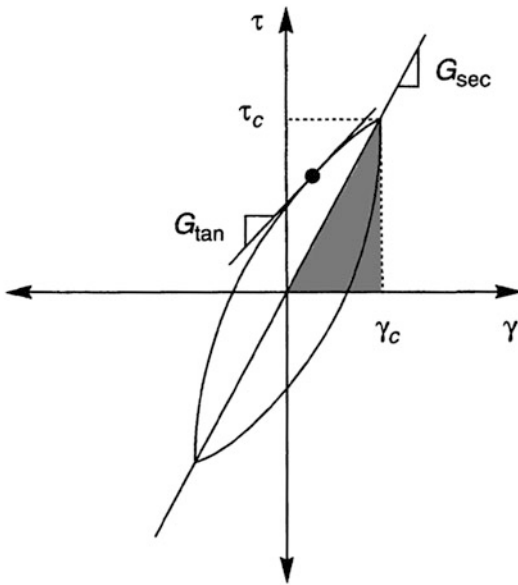
Site Response for Seismic Hazard Assessment, Fig. 3 Evaluation of shear moduli and damping ratio from hysteresis loop (Modified from Kramer 1996)

where W_D is the dissipated energy (equal to the area of the hysteresis loop A_{loop}) and W_S is the maximum strain energy represented by the shaded area in Fig. 3.

Staring from a maximum value, i.e., G_0 or G_{max}, the soil stiffness tends to decrease with increasing shear strains. Its variation with cyclic shear strain amplitude (γ_c) is described graphically by a normalized modulus reduction curve $G/G_0 - \gamma_c$. Conversely, the energy dissipation provided by the soil, increasing with loop amplitude due to plasticity effects, is typically depicted by a $D - \gamma_c$ curve. Therefore, the mechanical characterization of the soil stiffness and damping requires both the evaluation of G and D at very low strains and the way in which the two properties change with cyclic shear strain amplitude. Laboratory tests have shown that soil stiffness and damping are also influenced by other factors, such as mean effective confining pressure, plasticity index (PI), overconsolidation ratio (OCR), and number of imposed cycles (N). In particular, extensive laboratory investigations on the cyclic response of normally consolidated and slightly

overconsolidated reconstituted clays presented by Vucetic and Dobry (1991) indicated that (Fig. 4):

(a) The values of OCR and effective consolidation stress have almost no effect on the position and shape of G/G_0 and D curves.
(b) The plasticity index PI is the key factor controlling the dimensionless parameters G/G_0 and D.
(c) The number of cycles N does affect both the value of G/G_0 (due to the degradation of the shear modulus with N) and the $D - \gamma_c$ curve at high cyclic strains (i.e., after the volumetric threshold).

It should be noted that the modulus reduction and damping curves reported in Fig. 4 for $PI = 0$ are nearly identical to the average curves commonly used for sands (Seed and Idriss 1970). This suggests that the curves shown in Fig. 4 can be used for both fine- and coarse-grained soils. Modulus reduction and damping ratio curves are influenced by the mean effective stress for cohesionless and low-plasticity soils, but this influence decreases with increasing plasticity index, being generally not significant for $PI \geq 30$ (Ishibashi 1992).

Measurement Techniques

The measurement of soil dynamic properties is a crucial task in the solution of any geotechnical earthquake engineering problem. Particularly for the evaluation of local site effects, the characterization of the soil deposit above bedrock in terms of variation of G_0 and D with depth and $G/G_0 - \gamma_c$ and $D - \gamma_c$ curves is essential. A variety of field and laboratory techniques are available, each oriented toward the measurement of low-strain properties and characterized by different advantages and limitations. A complete review of the existing techniques is outside the scope of this entry, and only the most significant ones are discussed in the following.

Low-strain field tests induce seismic waves in the soil and measure the velocities at which these waves propagate. The maximum shear modulus can be computed using the measured shear wave

Site Response for Seismic Hazard Assessment, Fig. 4 Nonlinear behavior as function of plasticity index in terms of: (**a**) normalized shear modulus and (**b**) damping ratio (Vucetic and Dobry 1991)
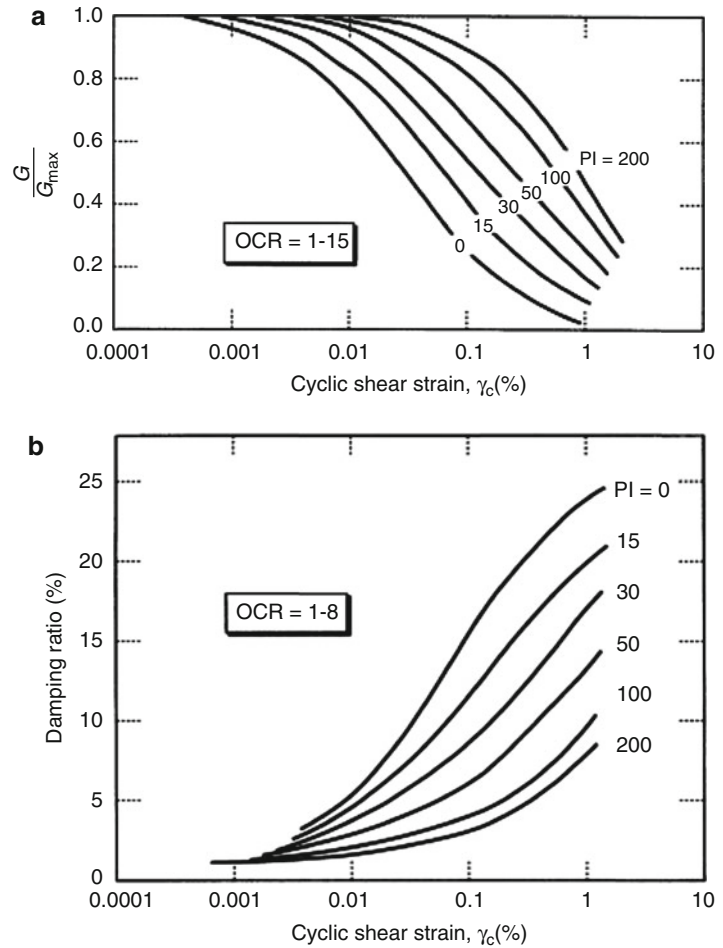
velocity (V_S) and the soil density (ρ) through the following equation based on the elasticity theory:

$$G_{max} = \rho V_S^2 \tag{3}$$

Shear wave velocities can be measured in situ by several seismic tests, including cross-hole and down-hole, seismic cone penetrometer, suspension logger, SASW (spectral analysis of surface waves), and MASW (multichannel analysis of surface waves). A review of these test methods is given in Woods (1994) and Kramer (1996). Their accuracy can be sensitive to procedural details, soil conditions, and interpretation techniques. Fig. 5 shows the layouts and principles of three established geophysical tests: the continuous surface wave (CSW) method, the down-hole test, and the cross-hole technique. In the CSW test (Fig. 5a), a mechanical, servo-hydraulic, or electromagnetic vibrator applies a single-frequency sinusoidal force to the ground surface. Rayleigh waves traveling through the soil are detected by a series of geophones (usually two) displaced at a range of distances from the source. By changing the input frequency, a profile of phase velocity against wavelength is obtained, and, consequently, a stiffness profile with depth can be computed. Although less economical than an SASW test, the CSW method has been proved to provide better data, as background noise can be easily recognized and filtered. Likewise the CSW test, the MASW approach uses a multiple of equally spaced receivers (usually 12 to 60) that are deployed on the surface along a survey line.

Each receiver is connected to a common multichannel recording instrument (i.e., a seismograph). This is the most significant difference between the CSW and the MASW techniques, as CSW is usually based on a two-receiver approach. Also the MASW method generally uses an impulsive source, such as a sledgehammer, to produce surface waves, whereas the CSW technique makes use of a frequency-controlled vibrator.

Down-hole and cross-hole tests are alternative low-strain techniques which require one or more borings. In a down-hole test, a vibration source is placed on the ground surface adjacent to a borehole. The arrival of seismic energy is detected at depth either by geophones secured against the borehole sides or by geophones within a seismic CPT (Fig. 5b). The test is repeated changing the depth of the geophones (typically at 1 m intervals) to plot the shear wave travel time as function of depth. The average shear velocity can be computed by knowing the distance between the source and the receiver. The cross-hole test makes use of more than one boring (usually two, less commonly three as this latter option is more expensive): a source is placed in one boring and a receiver is placed at the same depth in each of the other boreholes (Fig. 5c). An impulsive disturbance is applied at the source and the travel time to each receiver is measured. A borehole verticality survey is required in order to calculate the actual distance between the boreholes at each test depth (usually 1 m intervals). The wave propagation velocity is, in fact, computed by knowing the distances between receivers. The use of two sets of receivers (three boreholes) avoids the issue of trigger accuracy, but increases the cost of the test. While CSW and down-hole tests allow to determine the shear modulus for distortion in vertical plane (G_v) as the source produces a vertically polarized horizontally traveling shear wave, the cross-hole technique allows to calculate also the shear modulus in the horizontal plane (G_h) by using horizontally polarized shear wave sources.

The standard penetration test (SPT) and the cone penetration test (CPT), although originally developed for the measurement of soil properties mobilized at large strains, can be indirectly used to determine shear modulus profiles in situ by using empirical correlations between penetration resistance (N_{SPT}) or tip resistance (q_c) values and G_{max} (a review is given in Kramer 1996). These estimates are affected by high uncertainties and should be used very cautiously, given the scatter in the data on which they are based and the variability in the results obtained by different correlations. Such correlations should be adopted only for preliminary estimates of G_{max} in the framework of simplified approaches.

Finally, standard spectral ratio (SSR) and horizontal-to-vertical spectral ratio (HVSR) methods for the determination of the deposit fundamental frequencies are becoming increasingly popular not only in the research field but also in professional practice (i.e., for microzonation studies). Site amplification factors can, in fact, be inferred, at least in the linear strain range, using the SSR technique described by Field and Jacob (1993). HVSR amplifications obtained experimentally can, instead, be used to validate numerical model results (e.g., SESAME 2004; Lanzo et al. 2011).

In general, field tests have the advantage to describe the dynamic properties of the soil as it is in situ. Laboratory testing methods are, on the contrary, usually performed on relatively small specimens that can be disturbed by the sampling technique and may not be representative of the larger body of soil from which they are retrieved. Nevertheless, cyclic and dynamic laboratory tests are complementary to field methods as they can provide the description of soil nonlinearity over a wide range of shear strains in terms of $G/G_0 - \gamma_c$ and $D - \gamma_c$ curves. Typical low-strain element techniques are the resonant column (RC) and the bender element (BE) tests. A schematic view of the resonant column apparatus is presented in Fig. 6: a solid or hollow cylindrical specimen is subjected to harmonic torsional or axial loading by an electromagnetic loading system. The system usually applies a harmonic load for which the frequency and amplitude can be controlled. The fundamental frequency of the specimen can be identified by gradually increasing the loading frequency. Given the mass polar moment of

Site Response for Seismic Hazard Assessment, Fig. 5 Field seismic tests: (**a**) continuous surface wave (*CSW*), (**b**) down-hole, and (**c**) cross-hole (Clayton 2011)
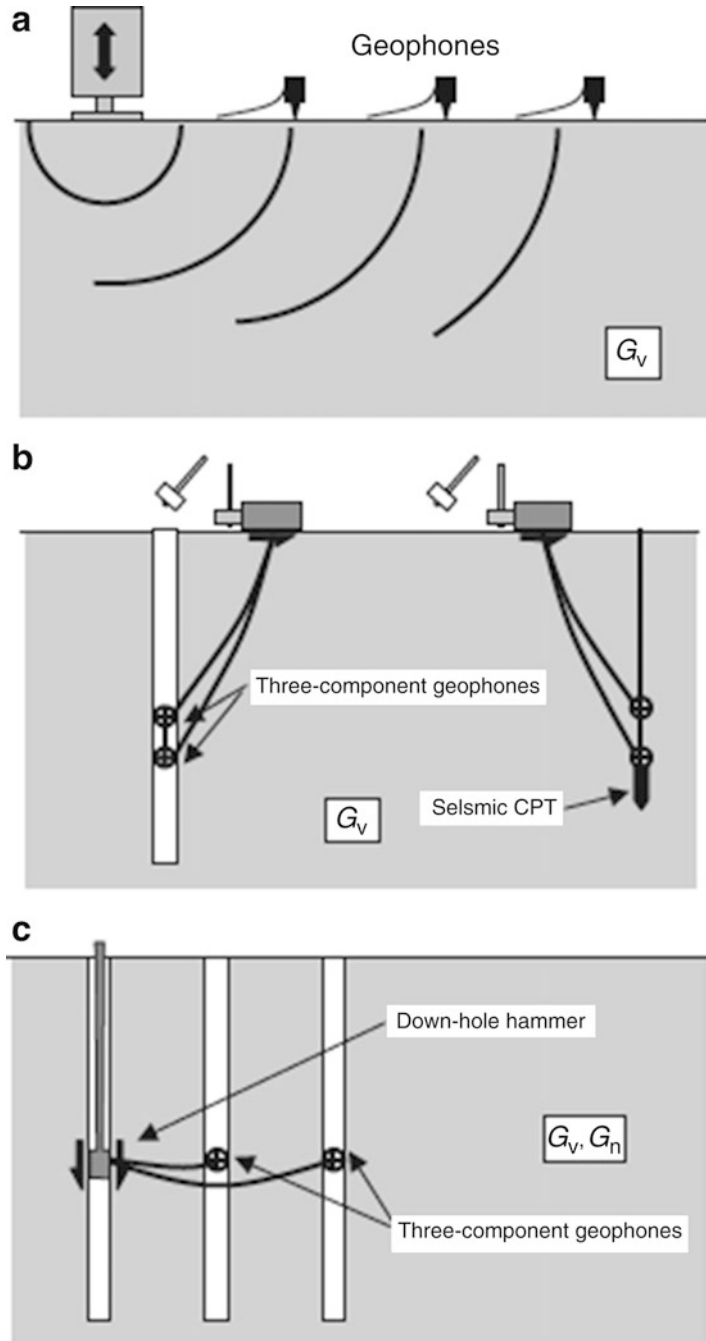
inertia of the loading head and the specimen mass and dimensions, the shear modulus of the soil can be calculated, assuming linear elasticity (i.e., using Eq. 3). Repeating the test with increasing loading amplitude, the variation of secant shear modulus with shear strain can also be measured.

Damping can be determined using the half-power bandwidth method or from the logarithmic decrement by placing the sample in free vibration (ASTM 2007). This technique has been modified to allow cyclic torsional shear testing to strain levels above those typically achieved during a
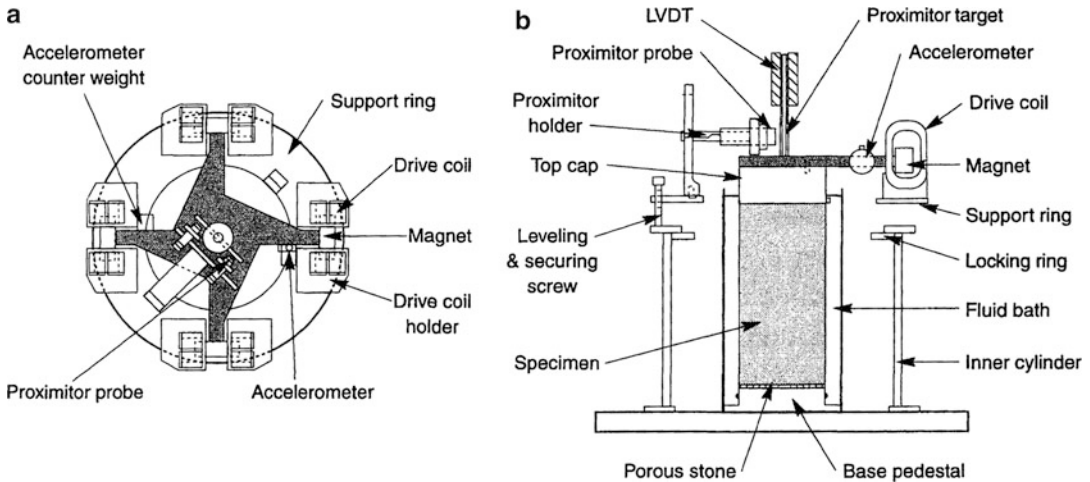
Site Response for Seismic Hazard Assessment, Fig. 6 Schematic drawing of a typical resonant column apparatus: (**a**) top view and (**b**) profile view (Kramer 1996)

conventional resonant column test. It is well established that the large strain rates related to the high frequencies applied in an RC test can affect the measured small-strain shear modulus and damping ratio. Dry cohesionless soils usually do not exhibit rate effects, while high strain rates can produce an increase of shear stiffness at small strains in fine-grained plastic soils leading to a G_0 overprediction with respect to cyclic simple shear tests (e.g., Lo Presti et al. 1997; Cavallaro et al. 2003). Moreover, the laboratory results indicate that some energy is always dissipated by the soil, even at very low strains, so that the damping ratio is never equal to zero. This mechanism cannot be justified by plasticity theory, as no hysteretic dissipation of energy associated to the development of plastic deformations takes place at strain levels within the elastic domain. Therefore, the initial damping ratio D_0 observed in resonant column experiments can be probably attributed to material viscous effects and/or the inertia of the resonant column apparatus (e.g., Meng and Rix 2003; Lo Presti et al. 2007).

Another laboratory technique that allows the measurement of the shear wave velocity of a soil sample is the bender element testing method. Bender elements consist of two piezo-ceramic plates bonded together in such a way that application of a voltage causes one plate to expand and the other to contract. This generates seismic waves in the soil sample in which the device is embedded. At the same time, a lateral disturbance of a bender element produces a voltage, thus allowing the detection of incoming waves. Therefore, BE can be used as both transmitters and receivers of S-waves within a soil sample (typically triaxial). By measuring the time required for the wave to travel from the source to the receiver and knowing the distance between each, the shear wave velocity of the specimen can be measured.

Finally, laboratory tests able to measure the dynamic soil properties at high strain levels have been derived from conventional tests by adding cyclic loading capabilities to the testing apparatus. Examples are represented by cyclic triaxial tests with local strain measurements and cyclic simple shear tests.

Table 1 summarizes the range and applicability of the most common field and laboratory cyclic and dynamic techniques, with particular reference to the induced shear strain level, characteristic frequency, and dynamic soil properties which can be obtained from the tests.

Ground Response Analysis

Analytical solutions have been developed to solve one-dimensional wave propagation problems, assuming that the input motion is harmonic and

Site Response for Seismic Hazard Assessment, Table 1 Range and applicability of field and laboratory dynamic tests (Modified from Vinale et al. 1996)

Test type			Shear strain γ (%)	Frequency f (Hz)	Shear modulus	Damping ratio
Field	Standard	SPT	–	–	$N_{SPT} \to V_S \to G_0$	–
		CPT			$q_c \to V_S \to G_0$	
	Geophysics	Down-hole	$<10^{-3}$	$10 \div 100$	$V_S \to G_0$	Possible
		Cross-hole			$V_S \to G_0$	
		SASW			$V_R \to V_S \to G_0$	
Lab	Cyclic	Triaxial	$>10^{-2}$	$0.01 \div 1$	$\sigma - \varepsilon \to E \to G$	Hysteresis loop $\to D$
		Simple shear	$>10^{-2}$	$0.01 \div 1$	$\tau - \gamma \to G$	
		Torsional shear	$10^{-4} \div 1$	$0.01 \div 1$	$\tau - \gamma \to G, G_0$	
	Dynamic	Bender elements	$<10^{-3}$	>100	$V_S \to G_0$	Possible
		Resonant column	$10^{-5} \div 10^{-1}$	>10	$f_r \to G, G_0$	H-p or log. decrement $\to D$
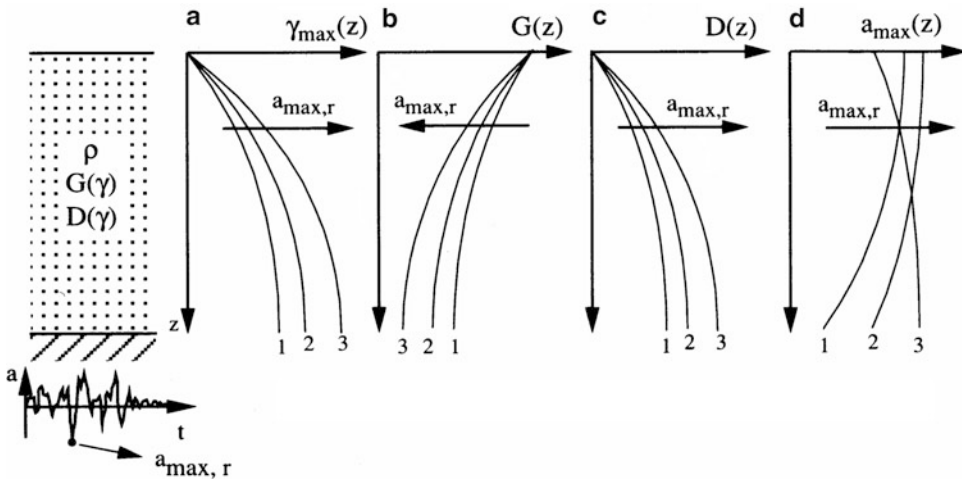
Site Response for Seismic Hazard Assessment, Fig. 7 Profiles with increasing bedrock acceleration of (**a**) maximum shear strain, (**b**) shear modulus, (**c**) damping ratio, and (**d**) maximum acceleration mobilized within the soil deposit (Lanzo and Silvestri 1999)

the soil is a homogeneous elastic or viscoelastic material (e.g., Roesset 1977). They make use of amplification functions to describe the dynamic response of a soil deposit (considered as a single-degree-of-freedom system) in the frequency domain. This approach is limited to the analysis of linear systems, as it relies on the principle of superposition. When the soil deposit is layered, heterogeneous (i.e., G_0 changes with depth), and nonlinear (i.e., G and D are functions of the shear strain induced by the earthquake), the amplification function is no more constant (as in the analytical solutions) but site specific. The effects of soil nonlinearity on wave propagation are qualitatively illustrated in Fig. 7, where a homogeneous deposit overlying a horizontal bedrock is subjected to an input motion characterized by increasing values of the maximum acceleration ($a_{max,r}$).

Site Response for Seismic Hazard Assessment, Fig. 8 Peak acceleration on soft soil sites as function of bedrock maximum acceleration (Idriss 1990)
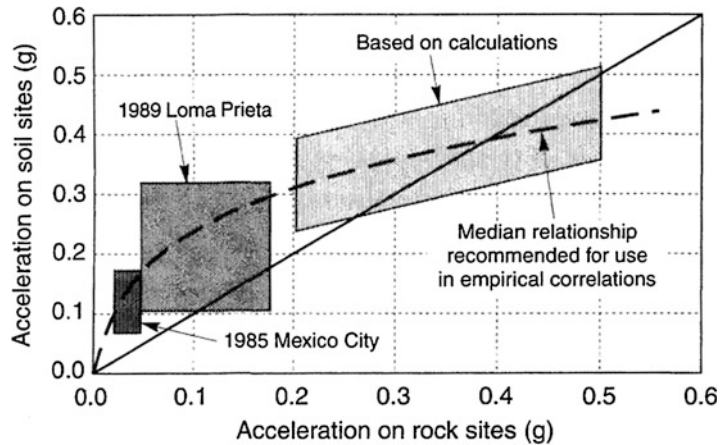
Increasing the peak ground acceleration at bedrock:

- The mobilized shear strain increases (Fig. 7a).
- The shear modulus reduces and the damping ratio increases correspondingly (Fig. 7b, c).
- For low input energy, the peak acceleration along the soil profile increases moving from the bedrock to the ground surface (profiles 1 and 2 in Fig. 7d).
- For high input energy, the peak acceleration along the profile can reduce because the nonlinear behavior reduces the soil shear stiffness, thus preventing from the transmission of high frequencies, while the corresponding increase of damping ratio reduces the displacement and acceleration amplitude peaks (profile 3 in Fig. 7d).

In terms of amplification factor (given as the ratio between the max acceleration at surface to the max acceleration at bedrock), nonlinearity effects can lead to a reduction of the amplification at surface with increasing acceleration levels at bedrock. On the basis of data recorded during Mexico City (1985) and San Francisco Bay area (1989) earthquakes, and of additional ground response analyses, Idriss (1990) related peak accelerations on soft soil deposits to those on rock sites, as reported in Fig. 8. For acceleration levels lower than about 0.4 g, the peak acceleration at soft soil surface is likely to be higher than on rock. At higher acceleration levels, on the contrary, nonlinearity can prevent from the development of peak accelerations at surface as large as those observed on rock sites.

Local site conditions not only influence the peak acceleration amplitudes but can also strongly affect the frequency content of surface motions. Fig. 9 shows the effects of local soil conditions on the shape of the normalized response spectra computed from ground motions recorded on different sites: for periods above 0.5 s, spectral accelerations for soil sites are higher than those for rock sites. The figure clearly indicates that deep and soft soil deposits enhance the transmission of low frequencies (high periods). The results also show that the use of a single response spectrum shape for all site conditions is not appropriate. This evidence has been incorporated in a large number of seismic codes worldwide which propose the use of different spectral shapes for different subsoil conditions.

Equivalent Linear Approach

Nonlinear soil behavior can be approximated by an equivalent linear characterization of soil dynamic properties. The method makes use of the exact continuum solution of wave propagation in horizontally layered viscoelastic materials subjected to vertically propagating transient motions (e.g., Roesset 1977). It models the nonlinear variation of soil shear modulus and

Fig. 9 Normalized
response spectra (5 %
damping) for different local
site conditions (Seed
et al. 1976)
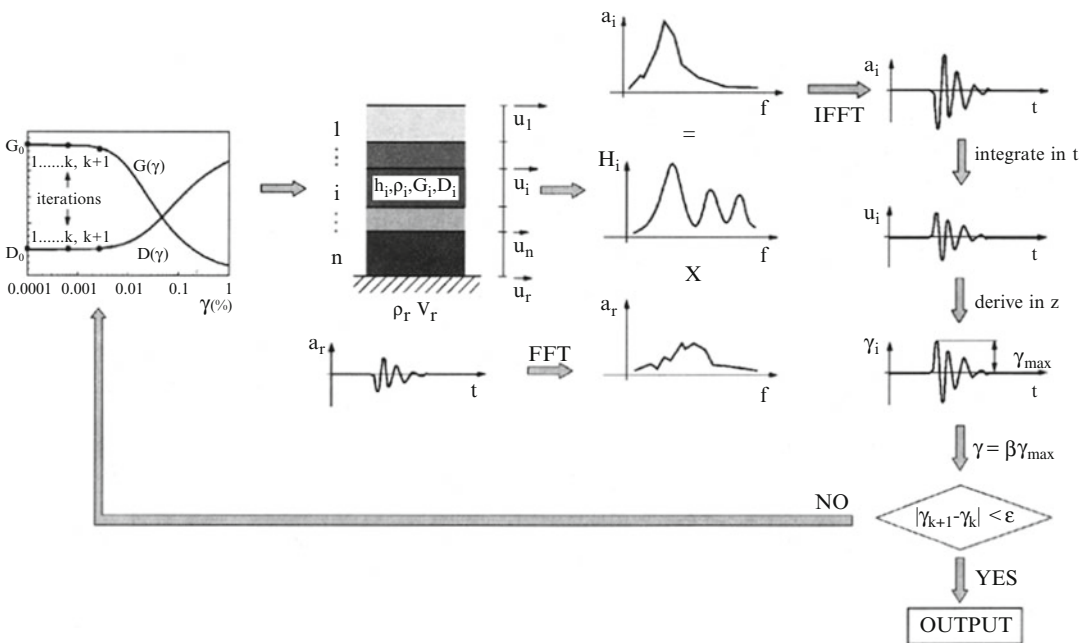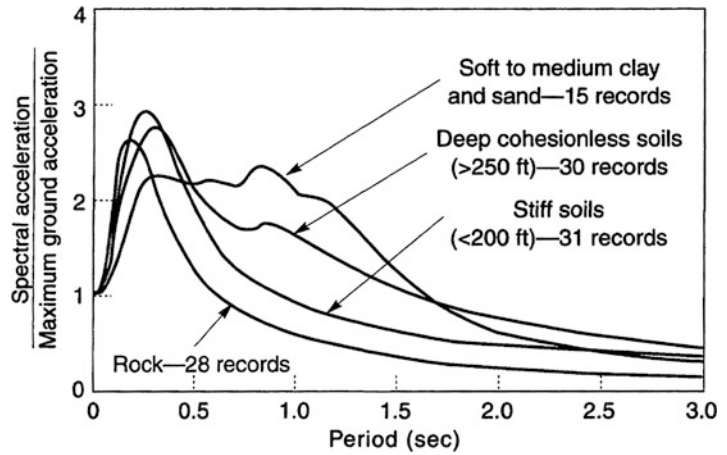
Site Response for Seismic Hazard Assessment, Fig. 10 Equivalent linear algorithm (Modified from Lanzo and
Silvestri 1999)

damping with shear strain through a sequence of
linear analyses with iterative update of stiffness
and damping parameters. For a given soil layer,
G and D are assumed to be constant with time
during the shaking. Therefore, an iterative proce-
dure is needed to ensure that the properties used
in the linear dynamic analyses are consistent with
the level of strain induced in each layer by the
input motion.

The iterative procedure, illustrated in Fig. 10,
operates as follows:

1. Initial estimates of G and D are made for each
   layer at low-strain values.
2. These initial values are used in a linear
   frequency-domain dynamic analysis where
   the acceleration time history at bedrock is
   transformed through a fast Fourier transform

(FFT) algorithm and the transfer functions between the bedrock and each layer (H_i in Fig. 10) are computed.

3. The inverse fast Fourier transform (IFFT) algorithm is employed to obtain acceleration, displacement, and shear strain time histories at each layer.

4. The effective shear strain (γ) in each layer is determined from the maximum shear strain (γ_{max}) as follows:

$$\gamma = \beta\gamma_{max} \qquad (4)$$

where $\beta = (M - 1)/10$ is usually related to the expected earthquake magnitude M.

5. From this effective shear strain, new equivalent linear values of G and D are chosen for the next iteration.

6. Steps 2–5 are repeated until a convergence criterion is satisfied.

The most widely used computer software for one-dimensional ground response analysis, based on the equivalent linear approach described above, is SHAKE (Schnabel et al. 1972) and its modified version SHAKE91 (Idriss and Sun 1992). For two-dimensional geometries, the equivalent linear approach has been implemented in a number of codes such as FLUSH (Lysmer et al. 1975) and QUAD4M (Hudson et al. 1994), among others. Equivalent linear methods are extensively adopted in engineering practice for their simplicity, flexibility, and low computational requirements. Nevertheless, their limitations are well known. The analysis is performed adopting a total stress approach. This means that excess pore water pressures induced by the earthquake cannot be predicted and the displacements due to consolidation processes cannot be calculated. The model employed to describe the mechanical behavior of soils is viscoelastic. Therefore, the method cannot predict permanent soil displacements or cumulated strains at the end of the analysis. Even though the iteration process allows to approximate nonlinear soil behavior, the approach is still a linear method of analysis. The strain-compatible soil properties are constant throughout the duration of the earthquake. The method is thus not capable of representing changes in soil stiffness and hysteretic damping during the seismic action.

For problems where strain levels remain low (stiff soil profile and/or relatively weak input motions), the equivalent linear method can produce reasonable estimates of ground response. For high seismic intensities at bedrock, nonlinear time-domain analyses should be preferred as they are likely to provide better results (section "Verification and Parametric Studies").

Nonlinear Approach

In recent years, the use of fully coupled effective stress formulations, based on Biot's theory for solid–fluid interaction (Biot 1941; Zienkiewicz et al. 1999), is becoming increasingly popular as the theoretical basis for the dynamic analysis of geotechnical structures (e.g., Dewoolkar et al. 2001; Elgamal et al. 2002a; Aydingun and Adalier 2003; Dakoulas and Gazetas 2005; Liu and Song 2005; Madabhushi and Zeng 2007; Sica et al. 2008; Alyami et al. 2009; Shahrour et al. 2010; Elia et al. 2011; Kontoe et al. 2011; Elia and Rouainia 2013, 2014). Specifically for ground response analysis, nonlinear approaches are able to solve the wave propagation problem by direct numerical integration in the time-domain, accounting for the development of plastic deformations and buildup of excess pore water pressures within the soil deposit induced by the earthquake. In those schemes, the mechanical behavior of the soil can be described using either simple or sophisticated nonlinear constitutive models of different level of complexity. In terms of spatial discretization, most numerical approaches are modelling the problem in 1D or 2D plane strain conditions. Nonlinear three-dimensional (3D) ground response analyses are rarely found in the literature, being only performed for critical structures, such as tunnels, bridges, and earth dams (e.g., Elgamal 1992; Stamos and Beskos 1995; Elgamal et al. 2008; Ou 2009; Hatzigeorgiou and Beskos 2010).

A summary of the available computer software, employing equivalent linear and nonlinear approaches to solve the wave propagation problem in one- and two-dimensional conditions, is reported in Table 2.

Geometry	Software	Type of analysis	
1D	SHAKE (Schnabel et al. 1972)	TS	EL
	SHAKE91 (Idriss and Sun 1992)		
	PROSHAKE (EduPro Civil Systems 1998)		
	SHAKE2000 (www.shake2000.com)		
	EERA (Bardet et al. 2000)		
	TESS (Pyke 1992)	TS	NL
	NERA (Bardet and Tobita 2001)		
	DEEPSOIL (Hashash 2009)		
	DESRA_2 (Lee and Finn 1978)	ES	
	DESRAMOD (Vucetic 1986)		
	D-MOD_2 (Matasovic 1995)		
	D-MOD2000 (GeoMotions 2007)		
	SUMDES (Li et al. 1992)		
	CYCLIC 1D (www.soilquake.net)		
2D	QUAD4 (Idriss et al. 1973)	TS	EL
	QUAD4M (Hudson et al. 1994)		
	FLUSH (Lysmer et al. 1975)		
	QUAKE/W (GeoSlope 2002)		
	SPECTRA (Borja and Wu 1994)	TS	NL
	DYNAFLOW (Prevost 2002)	ES	NL
	GEFDYN (Aubry and Modaressi 1996)		
	TARA-3 (Finn et al. 1986)		
	FLAC 2D (Itasca 2002)		
	PLAXIS 2D (www.plaxis.nl)		
	SWANDYNE II (Chan 1995)		
	OpenSees (McKenna and Fenves 2001)		

TS total stress, *ES* effective stress, *EL* equivalent linear, *NL* nonlinear

Verification and Parametric Studies

The effectiveness of the different numerical approaches, briefly described in the previous sections, to simulate the complex wave propagation process has been tested over the last decades using real vertical array data and/or centrifuge test results. In these cases, in fact, the input motion is reasonably well defined, as it is directly recorded on rock or it is imposed during the test. In addition, validation studies have been conducted in several international benchmark projects, such as ESG-IASPEI/IAEE 1992 (Turkey Flat and Ashigara Valley arrays), ESG-IASPEI/IAEE 2006 (Grenoble basin), Turkey Flat 2008, and E2VP 2010 (EuroSeisTest, Greece).

As an example, Borja et al. (1999) presented a verification study using the recordings from accelerometer arrays installed at the Lotung Large-Scale

Seismic Test (LSST) site in Taiwan. The site was established in 1985 to study the dynamic behavior of two scaled-down nuclear plant containment structures constructed by the Electric Power Research Institute (EPRI) and the Taiwan Power Company.

Borja and coworkers studied the ground response at Lotung employing both a standard linear equivalent method (SHAKE) and an advanced nonlinear approach, where a bounding surface soil constitutive model (formulated in terms of total stresses) was implemented into the three-dimensional finite element (FE) code SPECTRA. Fig. 11a, b compares the recorded ground surface accelerations along the EW and NS directions (i.e., FA1-5/EW and FA1-5/NS) with the corresponding predictions obtained through SHAKE. The peak values of both EW and NS accelerations recorded at ground surface

Site Response for Seismic Hazard Assessment, Fig. 11 Comparison of ground surface accelerations recorded at Lotung (EW and NS components) and predictions by SHAKE and SPECTRA (Borja et al. 1999)
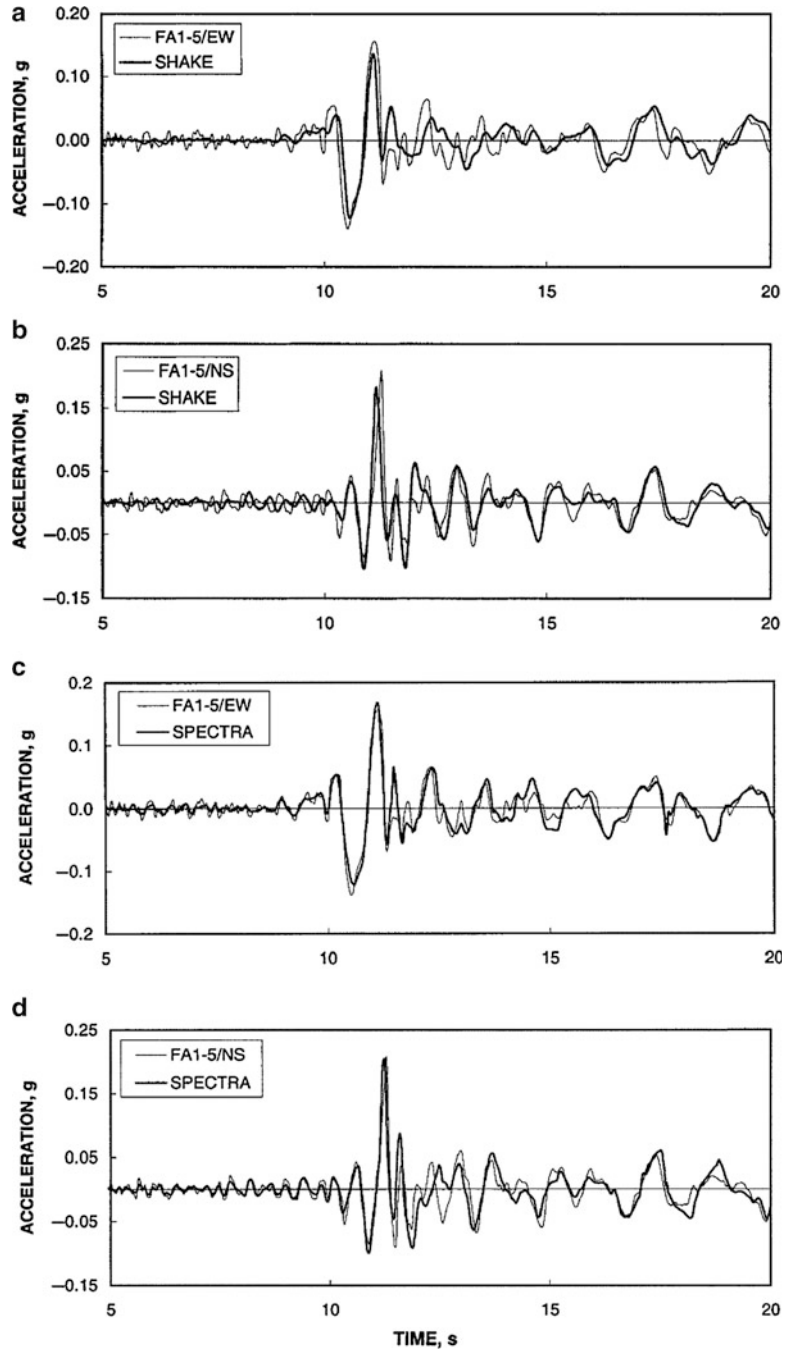
were slightly underpredicted. Moreover, a time shift in the acceleration peak between the recorded and computed NS motions can be observed. On the contrary, Fig. 11c, d shows that peak accelerations were predicted by the nonlinear model within 10–20 % error in both directions. An excellent match also in terms of zero crossings was obtained using the advanced nonlinear time-domain approach. The advantage of the described approach lies in its capacity to extend the analysis to three dimensions, thus allowing the simultaneous application of the

Site Response for Seismic Hazard Assessment,
Fig. 12 VELACS model 1 recorded and computed acceleration time histories (Elgamal et al. 2002b)
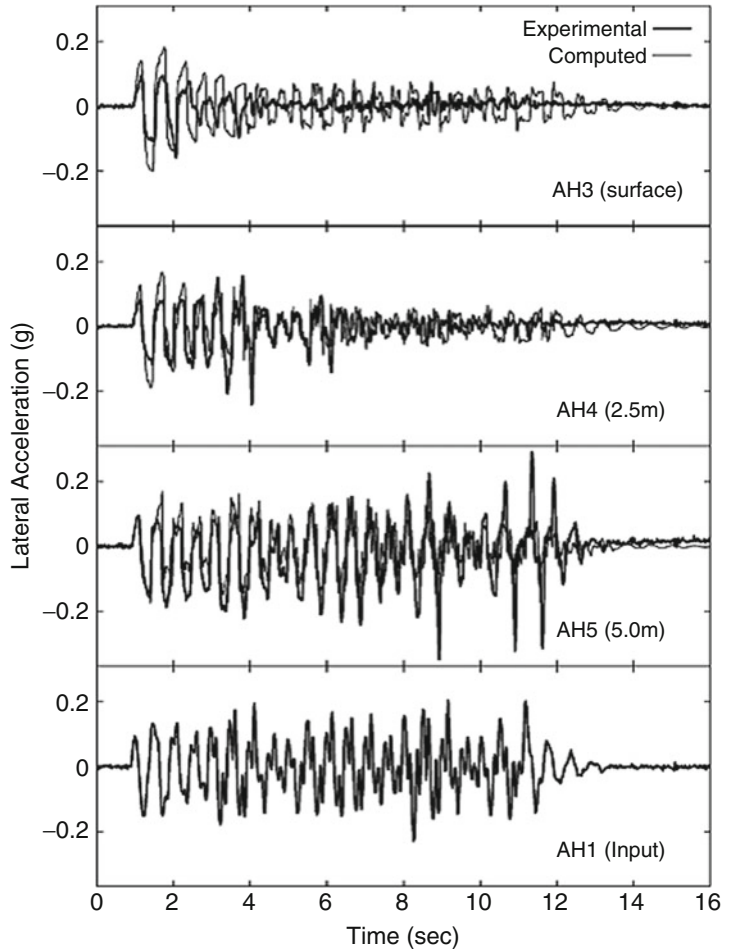
three components of the earthquake (one vertical and two horizontal) at the base of the FE column.

Another example of verification study, this time using centrifuge testing data, is provided by Elgamal et al. (2002b). The authors adopted an advanced plasticity-based constitutive model to simulate the cyclic mobility response mechanism typically observed in medium-dense sands subjected to cyclic loading. The model was implemented into an unnamed fully coupled finite element code to predict the dynamic response observed during different centrifuge tests (VELACS project, Arulanandan and Scott 1993). In particular, Fig. 12 shows the comparison between the computed acceleration time histories and the ones recorded for VELACS model 1 (representing a level site composed of Nevada sand and subjected to a 2Hz harmonic base excitation). Good agreement was achieved between the computed and recorded responses at different depths within the soil deposit in terms of accelerations, although the numerically predicted settlement due to liquefaction was generally smaller than observations. In addition, the work by Elgamal and coworkers demonstrates how, under level ground conditions, shear strains can be relatively small with minor cyclic mobility effects, while mildly sloping ground may result in large cyclic shear strain accumulation leading to unbounded flow failure. The importance of the dominant excitation frequency on post-liquefaction soil response is also highlighted.

The two examples described above are representative of the state of the art in the field of

nonlinear ground response analysis, but other verification studies are reported by Kramer and Stewart (2004). Although a direct comparison with recorded ground response during seismic actions is invaluable, the performance of advanced numerical approaches can also be usefully evaluated through parametric analyses. These numerical studies can, in fact, provide helpful guidelines to nonexpert users of time-domain nonlinear schemes by clarifying the importance of the different ingredients required to perform sophisticated simulations. One of the most controversial factors affecting FE dynamic simulations is the calibration of the viscous damping, usually needed in nonlinear analyses when the adopted soil constitutive model is unable to provide enough hysteretic energy dissipation through plasticity. Viscous damping is typically introduced in FE codes by means of the Rayleigh formulation, whose damping matrix is defined as follows:

$$[C] = \alpha_R[M] + \beta_R[K] \qquad (5)$$

where $[M]$ and $[K]$ are the mass and the stiffness matrix of the system, respectively. The coefficients α_R and β_R are obtained considering the following relationship with the damping ratio D:

$$\begin{Bmatrix} \alpha_R \\ \beta_R \end{Bmatrix} = \frac{2D}{\omega_m + \omega_n} \begin{Bmatrix} \omega_m \omega_n \\ 1 \end{Bmatrix} \qquad (6)$$

where ω_m and ω_n are the angular frequencies related to the frequency interval $f_m \div f_n$ over which the viscous damping is equal to or lower than D. The amount of Rayleigh damping to be introduced in an FE dynamic analysis is difficult to quantify a priori, but, at the same time, its magnitude can play a crucial role on the results of the simulations (Woodward and Griffiths 1996). Different possible calibration procedures have been proposed in the literature to identify the interval $f_m \div f_n$. A well-established one (e.g., Hudson et al. 1994) suggests to select f_m as the first natural frequency of the deposit f_1, while f_n is assumed equal to n times f_m, where n is the closest odd integer

larger than the ratio f_p/f_1 between the predominant frequency of the input earthquake motion (f_p) and the fundamental frequency of the soil deposit (f_1). This latter assumption was based on the evidence that the higher modes of a shear beam are odd multiples of the fundamental mode of the beam. Recently, Kwok et al. (2007) used linear frequency-domain solutions to provide guidelines on the specification of viscous damping adopted in four 1D nonlinear codes (namely, D-MOD_2, DEEPSOIL, OpenSees, and SUMDES). The idea was to assume the results of 1D linear equivalent analyses as target solutions for the nonlinear approaches, thus guiding the user in the calibration of the viscous damping parameters. The main conclusions of this parametric investigation can be summarized as follows:

- The target damping ratio should be set to the small-strain material damping.
- The two target frequencies should be set, as a first approximation, to the first mode site frequency and five times the site frequency (i.e., $n = 5$).
- More generally, the target frequencies should be established through an iterative process by which linear time-domain and frequency-domain solutions are matched.

Similarly to Kwok et al. (2007), Amorosi et al. (2010) performed a parametric investigation using two different finite element codes (i.e., PLAXIS 2D and SWANDYNE II) and compared the results of 2D FE viscoelastic simulations with 1D linear equivalent solutions (i.e., EERA). In order to provide a useful framework for finite element users, some of the factors potentially influencing the numerical results were critically discussed. In particular, the amount of viscous damping adopted in FE viscoelastic analyses, the spatial and time discretization, and the nature of boundary conditions were examined. The investigation showed that:

- The traditionally adopted procedures for the calibration of the Rayleigh coefficients (e.g., Hudson et al. 1994) can lead to large

overestimation of the peak ground acceleration with respect to the EERA solutions. Therefore, a novel calibration procedure was proposed based on the linear equivalent amplification function and the frequency content of the input motion. It was suggested that the first target frequency (f_m) should be set to the first mode site frequency significantly excited by the seismic motion, while the second (f_n) should be selected equal to the frequency where the amplification function gets lower than one.

- As regards the effect of boundary conditions adopted in the FE simulations, tied nodes at the lateral boundaries of the mesh should be preferred, as they are properly representative of a 1D condition and allow to obtain a perfect match with EERA. When using Lysmer and Kuhlemeyer (1969) boundaries, the 2D mesh should be characterized by a width to height ratio between five and eight to avoid spurious wave reflections at the vertical boundaries.

- In terms of spatial discretization, the FE mesh should always satisfy the condition that the spacing between the nodes, Δl_{node}, must be smaller than approximately one-tenth to one-eighth of the wavelength associated with the maximum frequency component f_{max} of the input wave.

- The time-step algorithm adopted in the FE code can introduce some numerical damping in the integration of the governing equations of motion, and, therefore, the time-stepping coefficients need to be selected carefully in order not to lose accuracy.

Two alternative soil damping formulations for small and large strains have been recently implemented by Phillips and Hashash (2009) in the nonlinear one-dimensional site response analysis code DEEPSOIL. The first one introduces an approach to construct a frequency-independent viscous damping matrix which reduces the overdamping at high frequencies and, therefore, the filtering at those frequencies. A good match with frequency-domain analysis results can be obtained with this new approach, which, nonetheless, requires the calculation of the eigenvalues

and eigenvectors of the matrix $[M]^{-1}[K]$. The second formulation introduces a reduction factor that modifies the extended Masing loading/unloading stress–strain relationship (Masing 1926) to match measured modulus reduction and damping curves simultaneously over a wide range of shear strains. This modified hysteretic model allows to reduce the typically observed overestimation of damping predicted by advanced soil constitutive models for large strains (see next section). The proposed models have been implemented in a total stress nonlinear code and, therefore, cannot predict accumulation of excess pore water pressures during the shaking.

Future Challenges

While during recent years important advances have been made in the field of geotechnical earthquake engineering, some challenges still remain to be addressed.

In situ and laboratory characterization of soil dynamic properties represents a crucial aspect of any seismic design of geotechnical structures. Techniques for more accurate characterization of in situ soil shear wave velocity and damping ratio are required (e.g., Hall and Bodare 2000). Although expensive and time consuming, advanced laboratory cyclic and/or dynamic tests are essential for the correct understanding of soil behavior during seismic excitations and for the proper calibration of sophisticated constitutive assumptions introduced in nonlinear ground response analyses. At the same time, these tests should be interpreted within the framework of plasticity theory, as irreversible deformations in soil samples can develop also for small imposed strains.

With respect to site response analysis, the use of nonlinear time-domain approaches should be encouraged as a viable alternative to the standard linear equivalent method. It can, in fact, better predict soil deformation, degradation of stiffness, and accumulation of excess pore water pressures throughout the shaking. The issue of the accurate prediction of hysteretic damping with advanced soil constitutive models still remains controversial, as these models can significantly

underpredict the damping ratio for small strains and overestimate it in the large strain range. Some adjustments in the mathematical formulation of advanced soil models have been proposed recently (e.g., Phillips and Hashash 2009; Seidalinov and Taiebat 2014) to address this point, but more research work is needed. In general, clear guidance and protocols should be provided to nonexpert users of time-domain codes in order to improve the reliability of nonlinear ground response analysis results.

Another open challenge in the area of seismic risk assessment is related to the selection of appropriate input motions for nonlinear ground response simulations. The correct definition of the design seismic actions, based on seismic hazard and site response analyses, is, indeed, essential to fully implement the performance-based design approach proposed by code prescriptions (e.g., Eurocode 8). The options available to engineers in terms of input acceleration time series are represented by artificial, synthetic, and real accelerograms. Spectrum-compatible artificial records are characterized by excessive number of cycles of strong motion, and consequently they possess unreasonably high-energy content. It is now widely accepted that the use of these artificial records is problematic and not suitable particularly for nonlinear analyses (e.g., Bommer and Acevedo 2004). Synthetic accelerograms generated from seismological source models are highly sensitive to the definition of earthquake source parameters, which involve a significant degree of expert judgment. On the contrary, real earthquake accelerograms represent a more viable option for providing input to dynamic analyses, being more realistic than artificial records and easier to obtain than synthetic accelerograms. In any case, current practice for selecting and scaling ground motions for linear and nonlinear response-history analyses is based largely on engineering judgment. Very few systematic studies provide impartial guidance to geotechnical engineers regarding appropriate methods to use in site-specific applications. The majority of the research has been focused in the last years on structural response more than on site response (e.g., NIST 2011). Practitioners often select

ground motion records based only on distance, site soil conditions, and magnitude of the characteristic event expected to dominate the seismic hazard. However, many other factors, such as directivity of the rupture and presence of basin, affect the ground motion intensity and frequency content, which ultimately govern the nonlinear response and damage in geotechnical and structural systems (e.g., Frankel et al. 2011). Despite this, current practice is not adequately equipped to fully incorporate near-fault directivity, basin, and duration effects in the design process. Methods to implicitly consider inelastic demands by amplifying the design spectra do not provide a reliable basis for representing realistic ground input motions (Kalkan and Kunnath 2006).

More research is still needed to address all these issues, requiring a better integration between the work of geotechnical engineers, seismologists, geologists, and structural engineers. Finally, an ongoing transfer of research knowledge into practice represents a significant requisite to support performance-based earthquake engineering design and quantitative seismic risk assessment.

Summary

Experience from past strong earthquakes worldwide has demonstrated the relevance of local soil conditions on seismic ground response. The changes in amplitude, frequency content, and duration of the seismic motion during its propagation in soil deposits, commonly referred to as site effects, have a crucial impact on buildings and infrastructure response during earthquakes. Numerical methods allow geotechnical engineers to quantify the effects of soil deposits on the wave propagation process from bedrock to ground surface. These methods can be divided into: (i) frequency-domain analyses (using the linear equivalent method) and (ii) time-domain schemes (usually performed with finite element codes). The benefits and limitations of the two approaches are extensively discussed in this entry, together with a short description of the available methods for site

S

characterization and evaluation of soil dynamic properties. An overview of future challenges in the field of geotechnical earthquake engineering is also presented at the end of the entry.

List of symbols

a_{max}	Earthquake maximum acceleration
A_{loop}	Area of the hysteresis loop
D	Damping ratio
f	Frequency
f_1	Fundamental frequency of the soil deposit
f_p	Predominant frequency of the input earthquake motion
f_r	Resonance frequency
G_{sec}	Secant shear modulus
G_{tan}	Tangent shear modulus
G_0 or G_{max}	Initial shear modulus
M	Earthquake magnitude
N	Number of imposed cycles
N_{SPT}	SPT number
OCR	Overconsolidation ratio
PI	Plasticity index
q_c	CPT test tip resistance
V_R	Rayleigh wave velocity
V_S	Shear wave velocity
W_D	Energy dissipated in one hysteresis loop
W_S	Maximum strain energy in one hysteresis loop
α_R	Rayleigh damping coefficient
β	Ratio of effective shear strain to maximum shear strain
β_R	Rayleigh damping coefficient
γ	Shear strain
γ_c	Cyclic shear strain
ρ	Soil density
ω	Angular frequency

References

Alyami M, Rouainia M, Wilkinson SM (2009) Numerical analysis of deformation behaviour of quay walls under earthquake loading. Soil Dyn Earthq Eng 29(3):525–536

Amorosi A, Boldini D, Elia G (2010) Parametric study on seismic ground response by finite element modelling. Comput Geotech 37(4):515–528

Arulanandan K, Scott RF (eds) (1993) Proceedings of VELACS symposium. A.A., Balkema, Rotterdam

Ashford SA, Sitar N, Lysmer J, Deng N (1997) Topographic effects on the seismic response of steep slopes. Bull Seismol Soc Am 87(3):701–709

Associazione Geotecnica Italiana (2005) Aspetti Geotecnici della Progettazione in Zona Sismica. Patron Editore, Bologna

ASTM (2007) Standard test methods for modulus and damping of soils by resonant-column method. ASTM D4015-07. ASTM International, West Conshohocken

Aubry D, Modaressi A (1996) GEFDYN – Manuel Scientifique. Ecole Centrale Paris, Châtenay-Malabry

Aydingun O, Adalier K (2003) Numerical analysis of seismically induced liquefaction in earth embankment foundations. Part I. Benchmark model. Can Geotech J 40(4):753–765

Bard PY (1994) Effects of surface geology on ground motion: recent results and remaining issues. In: Proceedings of the X European conference on earthquake engineering, vol 1, Vienna

Bard PY, Bouchon M (1985) The two-dimensional resonance of sediment-filled valleys. Bull Seismol Soc Am 75(2):519–541

Bardet JP, Tobita T (2001) NERA – a computer program for Nonlinear Earthquake site Response Analyses of layered soil deposits. Department of Civil Engineering, University of Southern California, Los Angeles

Bardet JP, Ichii K, Lin CH (2000) EERA – a computer program for Equivalent-linear Earthquake site Response Analyses of layered soils deposits. University of Southern California, Department of Civil Engineering, Los Angeles. User manual

Bazzurro P, Cornell CA (2004) Nonlinear soil-site effects in probabilistic seismic hazard analysis. Bull Seismol Soc Am 94:2110–2123

Biot MA (1941) General theory of three-dimensional consolidation. J Appl Phys 12:155–164

Bommer JJ, Acevedo AB (2004) The use of real earthquake accelerograms as input to dynamic analysis. J Earthq Eng 8(4):1–50

Borja RI, Wu WH (1994) Vibration of foundations on incompressible soils with no elastic region. J Geotech Eng ASCE 120(9):1570–1592

Borja RI, Chao HY, Montáns FJ, Lin CH (1999) Nonlinear ground response at Lotung LSST site. J Geotech Geoenviron Eng ASCE 125(3):187–197

Bouckovalas GD, Papadimitriou AG (2005) Numerical evaluation of slope topography effects on seismic ground motion. Soil Dyn Earthq Eng 25:547–558

Castro G, Christian JT (1976) Shear strength of soils and cyclic loading. J Geotech Eng Div ASCE 102(GT9):887–894

Cavallaro A, Lanzo G, Pagliaroli A, Maugeri M, Lo Presti DCF (2003) A comparative study on shear modulus and damping ratio of cohesive soil from laboratory tests. In: Di Benedetto H et al (eds) Deformation characteristics of geomaterials. Swets & Zeitlinger, Lisse, pp 257–265

Chan AHC (1995) User manual for DIANA-SWANDYNE II. School of Engineering, University of Birmingham, Birmingham

Clayton CRI (2011) Stiffness at small strain: research and practice. Géotechnique 61(1):5–37

Dakoulas P, Gazetas G (2005) Seismic effective-stress analysis of caisson quay walls: application to Kobe. Soils Found 45(4):133–147

Dewoolkar MM, Ko H-Y, Pak RYS (2001) Seismic behaviour of cantilever retaining walls with liquefiable backfills. J Geotech Geoenviron Eng ASCE 127(5):424–435

EduPro Civil Systems Inc. (1998) ProShake – ground response analysis program. Redmond, Washington. User's manual

Elgamal A (1992) Three-dimensional seismic analysis of La Villita dam. J Geotech Eng ASCE 118(12):1932–1958

Elgamal A, Parra E, Yang Z, Adalier K (2002a) Numerical analysis of embankment foundation liquefaction countermeasures. J Earthq Eng 6(4):447–471

Elgamal A, Yang Z, Parra E (2002b) Computational modelling of cyclic mobility and post-liquefaction site response. Soil Dyn Earthq Eng 22:259–271

Elgamal A, Yan L, Yang Z, Conte J (2008) Three-dimensional seismic response of humboldt bay bridge-foundation-ground system. J Struct Eng ASCE 134(7):1165–1176

Elia G, Rouainia M (2013) Seismic performance of earth embankment using simple and advanced numerical approaches. J Geotech Geoenviron Eng ASCE 139(7):1115–1129

Elia G, Rouainia M (2014) Performance evaluation of a shallow foundation built on structured clays under seismic loading. Bull Earthq Eng 12(4):1537–1561

Elia G, Amorosi A, Chan AHC, Kavvadas M (2011) Fully coupled dynamic analysis of an earth dam. Géotechnique 61(7):549–563

Field EH, Jacob KH (1993) The theoretical response of sedimentary layers to ambient seismic noise. Geophys Res Lett 20(24):2925–2928

Finn WDL, Yogendrakumar M, Yoshida N, Yoshida H (1986) TARA-3: a program for nonlinear static and dynamic effective stress analysis. Soil Dynamics Group, University of British Columbia, Vancouver

Frankel A, Stephenson W, Carver D, Odum J, Williams R, Rhea S (2011) Probabilistic seismic hazard maps for seattle: 3D sedimentary basin effects, nonlinear site response, and uncertainties from random velocity variation. In: Proceedings of the 4th IASPEI/IAEE international symposium: effects of surface geology on seismic motion, University of California, Santa Barbara

Geli L, Bard PY, Jullien B (1988) The effect of topography on earthquake ground motion: a review and new results. Bull Seismol Soc Am 78(1):42–63

GeoMotions (2007) D-MOD2000 – a computer program package for seismic response analysis of horizontally layered soil deposits, earthfill dams, and solid waste landfills. GeoMotions LLC, Washington. User's manual

GeoSlope Int. Ltd. (2002) QUAKE/W for finite element dynamic earthquake analysis. User's Guide, Calgary

Hall L, Bodare A (2000) Analyses of the cross-hole method for determining shear wave velocities and damping ratios. Soil Dyn Earthq Eng 20:167–175

Hardin B, Drnevich VP (1972) Shear modulus and damping in soils: measurements and parameter effects. J Soil Mech Div ASCE 98:603–624

Hashash YMA (2009) DEEPSOIL V 3.7, Tutorial and user manual. 2002–2009. University of Illinois at Urbana-Champaign, Urbana

Hatzigeorgiou GD, Beskos DE (2010) Soil-structure interaction effects on seismic inelastic analysis of 3-D tunnels. Soil Dyn Earthq Eng 30:851–861

Hudson M, Idriss IM, Beikae M (1994) QUAD4M: a computer program to evaluate the seismic response of soil structures using finite element procedures and incorporating a compliant base. Center for Geotechnical Modeling, University of California, Davis

Idriss IM (1990) Influence of local site conditions on earthquake ground motions. In: Proceedings of IV U.-S. national conference on earthquake engineering, vol 1, Palm Springs

Idriss IM, Sun JI (1992) SHAKE91: a computer program for conducting equivalent linear seismic response analyses of horizontally layered soils deposits. Center for Geotechnical Modeling, University of California, Davis

Idriss IM, Lysmer J, Hwang R, Seed HB (1973) QUAD-4: a computer program for evaluating the seismic response of soil structures by variable damping finite element procedures. Report No EERC 73-16, Earthquake Engineering Research Center, University of California, Berkeley

Ishibashi I (1992) Discussion to "effect of soil plasticity on cyclic response", by M. Vucetic and R. Dobry. J Geotech Eng ASCE 118(5):830–832

Itasca Consulting Group Inc. (2002) FLAC – Fast Lagrangian Analysis of Continua. Minneapolis. User's manual

Kalkan E, Kunnath SK (2006) Effects of fling step and forward directivity on seismic response of buildings. Earthq Spectra 22(2):367–390

King JL, Tucker BE (1984) Observed variations of earthquake motion across a sediment-filled valley. Bull Seismol Soc Am 74(1):137–151

Kontoe S, Zdravkovic L, Potts DM, Menkiti CO (2011) On the relative merits of simple and advanced constitutive models in dynamic analysis of tunnels. Géotechnique 61:815–829

Kramer SL (1996) Geotechnical earthquake engineering. Prentice Hall, Upper Saddle River, New Jersey

Kramer SL, Stewart JP (2004) Geotechnical aspects of seismic hazards. In: Bozorgnia Y, Bertero VV (eds) Earthquake engineering – from engineering seismology to performance-based engineering. CRC Press LLC, Boca Raton

S

Kwok AOL, Stewart JP, Hashash YMA, Matasovic N, Pyke R, Wang Z, Yang Z (2007) Use of exact solutions of wave propagation problems to guide implementation of nonlinear seismic ground response analysis procedures. J Geotech Geoenviron Eng ASCE 133(11):1385–1398

Lanzo G, Silvestri F (1999) Risposta Sismica Locale (Teoria ed esperienze). Hevelius Edizioni srl, Benevento

Lanzo G, Silvestri F, Costanzo A, d'Onofrio A, Martelli L, Pagliaroli A, Sica S, Simonelli A (2011) Site response studies and seismic microzoning in the Middle Aterno valley (L'Aquila, Central Italy). Bull Earthq Eng 9(5):1417–1442

Lee MKW, Finn WDL (1978) DESRA-2, dynamic effective stress response analysis of soil deposits with energy transmitting boundary including assessment of liquefaction potential, Soil mechanics series no 38. Department of Civil Engineering, University of British Columbia, Vancouver

Li XS, Wang ZL, Shen CK (1992) SUMDES: a nonlinear procedure for response analysis of horizontally layered sites subjected to multi-directional earthquake loading. Department of Civil Engineering, University of California, Davis

Liu H, Song E (2005) Seismic response of large underground structures in liquefiable soils subjected to horizontal and vertical earthquake excitations. Comput Geotech 32(4):223–244

Lo Presti DCF, Jamiolkowski M, Pallara O, Cavallaro A, Pedroni S (1997) Shear modulus and damping of soils. Géotechnique 47(3):603–617

Lo Presti DFC, Pallara O, Mensi E (2007) Characterization of soil deposits for seismic response analysis. In: Ling HI et al (eds) Soil stress–strain behavior: measurement, modeling and analysis. Springer, Dordrecht, The Netherlands, pp 109–157

Lysmer J, Kuhlemeyer RL (1969) Finite dynamic model for infinite media. J Eng Mech Div ASCE 95(EM4):859–877

Lysmer J, Udaka T, Tsai C-F, Seed HB (1975) FLUSH: a computer program for approximate 3-D analysis of soil-structure interaction problems. Report EERC 75-30, University of California, Berkeley

Madabhushi SPG, Zeng X (2007) Simulating seismic response of cantilever retaining walls. J Geotech Geoenviron Eng ASCE 133(5):539–549

Masing G (1926) Eigenspannungen und verfestigung beim messing. In: Second international congress on applied mechanics, Zurich, pp 332–335

Matasovic N (1995) D-MOD_2 A computer program for seismic response analyses of horizontally layered soil deposits, earthfill dams and solid waste landfills. GeoSyntec Consultants, Huntington Beach

McKenna F, Fenves GL (2001) The OpenSees command language manual, version 1.2. Pacific Earthquake Engineering. Research Center, University of California, Berkeley

Meng J, Rix GJ (2003) Reduction of equipment-generated damping in resonant column measurements. Géotechnique 53(5):503–512

NIST (2011) Selecting and scaling earthquake ground motions for performing response-history analyses. NIST GCR 11-917-15, National Institute of Standards and Technology, Gaithersburg

Ou JH (2009) Three-dimensional numerical modelling of interaction between soil and pore fluid. PhD thesis, School of Engineering, University of Birmingham

Pagliaroli A, Lanzo G, D'Elia B (2011) Numerical evaluation of topographic effects at the Nicastro ridge in Southern Italy. J Earthq Eng 15(3):404–432

Papaspiliou M, Kontoe S, Bommer JJ (2012) An exploration of incorporating site response into PSHA-part II: sensitivity of hazard estimates to site response approaches. Soil Dyn Earthq Eng 42:316–330

Phillips C, Hashash YMA (2009) Damping formulation for nonlinear 1D site response analyses. Soil Dyn Earthq Eng 29:1143–1158

Prevost JH (2002) DYNAFLOW – a nonlinear transient finite element analysis program. Version 2002. Release 01.A, Department of Civil Engineering & Operation Research, Princeton University, Princeton

Pyke RM (1992) TESS: a computer program for nonlinear ground response analyses. TAGA Engineering Systems and Software, Lafayette

Richart FE, Hall JR, Woods RD (1970) Vibrations of soils and foundations. Prentice-Hall, Englewood Cliffs

Roesset JM (1977) Soil amplification of earthquakes. In: Desai CS, Christian JT (eds) Numerical methods in geotechnical engineering. McGraw-Hill, New York, pp 639–682

Sangrey DA, Henkel DJ, Esrig MI (1969) The effective stress response of a saturated clay soil to repeated loading. Can Geotech J 6(3):241–252

Schnabel PB, Lysmer J, Seed HB (1972) SHAKE: a computer program for earthquake response analysis of horizontally layered sites. Report no EERC72-12, Earthquake Engineering Research Center, University of California, Berkeley

Seed HB, Idriss IM (1970) Soil moduli and damping factors for dynamic response analyses. Report EERC 70-10, Earthquake Engineering Research Center, University California, Berkeley

Seed HB, Ugas C, Lysmer J (1976) Site dependent spectra for earthquake resistant design. Bull Seismol Soc Am 66(1):221–243

Seidalinov G, Taiebat M (2014) Bounding surface SANICLAY plasticity model for cyclic clay behavior. Int J Numer Anal Methods Geomech 38(7):702–724

Semblat JF, Pecker A (2009) Waves and vibrations in soils: earthquakes, traffic, shocks, construction works. IUSS Press, Pavia

Semblat JF, Kham M, Parara E, Bard PY, Pitilakis K, Makra K, Raptakis D (2005) Seismic wave amplification: basin geometry vs soil layering. Soil Dyn Earthq Eng 25:529–538

SESAME (2004) Guidelines for the implementation of the H/V spectral ratio technique on ambient vibrations measurements, processing and interpretation. SESAME European Research Project. WP12 – Deliverable D23.12

Shahrour I, Khoshnoudian F, Sadek M, Mroueh H (2010) Elastoplastic analysis of the seismic response of tunnels in soft soils. Tunn Undergr Space Technol 25: 478–482

Sica S, Pagano L, Modaressi A (2008) Influence of past loading history on the seismic response of earth dams. Comput Geotech 35(1):61–85

Stamos AA, Beskos DE (1995) Dynamic analysis of large 3-D underground structures by the BEM. Earthq Eng Struct Dyn 24(6):917–934

Tsai C-CP (2000) Probabilistic seismic hazard analysis considering nonlinear site effect. Bull Seismol Soc Am 90:66–72

Vinale F, Mancuso C, Silvestri F (1996) Dinamica dei terreni. Manuale di Ingegneria Civile Cremonese. Zanichelli/ESAC, Roma

Vucetic M (1986) Pore pressure buildup and liquefaction at level sand sites during earthquakes. Rensselaer Polytechnic Institute, Troy

Vucetic M, Dobry R (1991) Effects of the soil plasticity on cyclic response. J Geotech Eng Div ASCE 117(1): 89–107

Woods RD (1994) Geophysical characterization of sites. Balkema, Rotterdam

Woodward PK, Griffiths DV (1996) Influence of viscous damping in the dynamic analysis of an earth dam using simple constitutive models. Comput Geotech 19(3): 245–263

Zienkiewicz OC, Chan AHC, Pastor M, Schrefler BA, Shiomi T (1999) Computational geomechanics (with special reference to earthquake engineering). Wiley, Chichester

Site Response: 1-D Time Domain Analyses

Neven Matasovic
Geosyntec Consultants, Huntington Beach, CA, USA

Synonyms

1-D nonlinear seismic site response analysis; 1-D nonlinear site response analysis; One-dimensional (1-D) time domain analysis

Introduction

One-dimensional (1-D) time domain analysis is one of several currently available methods to evaluate the influence of local site conditions on input (i.e., bedrock) ground motions. However, this method is favored by Geotechnical Earthquake Engineers as it takes into account the unique geotechnical characteristic of a given site, including soil nonlinearity and hysteretic behavior and porewater pressure generation and dissipation. It can also be applied for a wide range of shaking intensities, and it is relatively easy to implement in practice as it is coded in commercially available software.

Through back analysis of numerous case histories, the 1-D time domain analysis has been shown to work well when analyzed soil deposits are horizontally layered, there is a significant impedance contrast within the profile, and when material (model) parameters are established through a reasonable site investigation and characterization efforts.

There are many circumstances under which a 1-D time domain site response analysis is performed. Such an analysis is commonly performed in the following situations: (i) when soil conditions cannot be reasonably categorized into one of the standard site conditions, (ii) when empirical site factors for the site are not available (e.g., the International Building Code Site Class F), (iii) when special ground conditions govern the design (e.g., soil liquefaction, seismic settlement and/or lateral spreading, or slope stability), (iv) for any case where the objective is to obtain ground motions considered to be more representative of the local geologic and seismic site conditions than motions obtained from other approaches, and/or (v) when calibration of a more advanced, multidimensional models with or without structural elements is required.

In general, site response analysis methods can be classified based upon the domain in which calculations are performed (frequency domain or time domain), the sophistication of the constitutive model employed (linear, equivalent linear, and/or nonlinear), whether the effects of

S

porewater pressure generation are neglected or not (total-stress and effective-stress analyses, respectively), and the dimensionality of the space in which the analysis is performed (one-dimensional, 1-D; quasi 2-D, 2-D, and 3-D). Other considerations in classifying site response analysis methods include modeling of cyclic reduction and cyclic degradation in a total-stress mode. This chapter focuses upon time domain nonlinear total- and effective-stress 1-D site response analysis which has significant advantages over its frequency domain linear and equivalent-linear counterparts in areas of very high seismicity (e.g., PGA $\geq$ 0.4 g) and/or when the effects of soft and/or potentially liquefiable soils in the profile cannot be ignored.

Theoretical Background

Dynamic Response Model

In 1-D time domain analysis, seismic response of a horizontally layered soil deposit is computed by solving the dynamic equation of motion. The dynamic equation of motion is commonly written as

$$[M]\ \{\ddot{u}\} \ + \ [C]\ \{\mathring{u}\} \ + \ [K]\ \{u\}$$
$$= \ -[M]\ \{\ddot{u}_g\} \tag{1}$$

where [M], [C], and [K] are the mass matrix, viscous damping matrix, and nonlinear stiffness matrix, respectively; $\{u\}$, $\{\mathring{u}\}$, and $\{\ddot{u}\}$ are the displacements, velocities, and accelerations of the mass [M] relative to the base; and $\{\ddot{u}_g\}$ is the acceleration of the base.

The stiffness matrix [K] is derived from the nonlinear (and effective-stress) soil constitutive model selected to represent a cyclic response of soil. In principle, all damping in the soil can be captured through the hysteretic loops of the constitutive model used to model cyclic soil behavior. However, most available soil constitutive models cannot properly represent soil damping at low strains. Therefore, it is necessary to add damping through the use of velocity proportional viscous damping, as discussed in detail below.

The dynamic equation of motion can be solved by numerical integration. The numerical integration calls for temporal discretization (i.e., system of coupled equations is discretized temporally), hence the term "solution in time domain." The dynamic equation is solved by a time-stepping scheme. Examples of time-stepping schemes include Wilson's theta (θ) algorithm (Clough and Penzien 1993) and several variations of the Newmark's β algorithms (Newmark 1959).

In order to solve the dynamic equation of motion, it is necessary to properly discretize the domain of interest, which in this case is the 1-D soil column. Two different approaches for discretization of the soil column are available: (i) lumped mass discretization and (ii) finite element discretization.

Figure 1 shows a lumped mass discretization (i.e., lumped mass model) that of a horizontally layered soil deposit. Soil mass (m) is lumped at the layer interfaces (layers 1 through n). In each layer, soil stiffness is represented by a nonlinear spring. In addition to hysteretic damping inherent to a nonlinear spring, a viscous damping coefficient (c) is also part of a model shown in Fig. 1, as schematically represented by dashpots.

The finite element discretization requires solving the dynamic equation of motion (Eq. 1) by means of an explicit time marching integration algorithm. In this type of discretization, mass can be distributed over layer thickness.

The thickness of the layers in both lumped mass and finite element discretization have an impact on calculated site response. The layer thickness determines the maximum and minimum frequency that can be propagated through a soil column. If the layer is too thick, important components of the input ground motion may be filtered, and thus the calculated ground response may be underestimated. If layer thickness is too small, representation of the actual site conditions might be inadequate (e.g., soil lens is modeled as a soil layer).

Nonlinear Constitutive Models with Hysteretic Damping

One-dimensional nonlinear total-stress site response analysis is generally conducted using
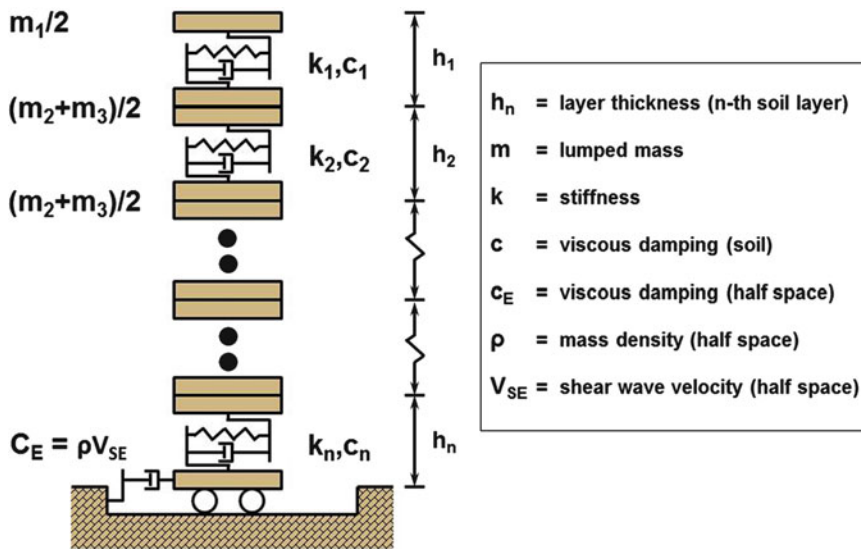
Site Response: 1-D Time Domain Analyses, Fig. 1 Lumped mass discretization for 1-D dynamic response model (Matasovic 1993)

relatively simplified constitutive models of soil layers. These models evolved from the early stress–strain relationships of Ramberg and Osgood (1943) and Kondner and Zelasko (1963). The hyperbolic model, introduced by Duncan and Chang (1970) for axial soil behavior, was based upon the basic formulation of the above-cited shear stress and strain behavior models. It was accompanied by sets of generic material properties and hence allowed for an elegant and simple way to capture soil nonlinearity at small axial strains. All three models provided the basis for constitutive models that are presently in use. These include models by Pyke (1979), Matasovic (1993), Matasovic and Vucetic (1993), and Darendeli (2001) which, in addition to better simulation of nonlinear stress–strain behavior, allow for simulation of cyclic loading and reloading in accordance with certain rules. The stress–strain relationship in these models is generally established by: an initial loading curve; a series of rules that describe the backbone curve (see Fig. 2 for definition of backbone curve); and unloading–reloading behavior rules required to establish cyclic loops. The most widely used rules are the Masing rules (Masing 1926) and extended Masing rules

(Pyke 1979; Vucetic 1990). Recently, Phillips and Hashash (2009) developed a set of rules that compensate for one of the most common deficiency of nonlinear models with hysteretic damping – overestimation of damping at large shear strains.

Bounding surface plasticity models may be used to simulate nonlinear stress–strain behavior in one dimension. Example of such models and their application in site response analysis can be found in Borja et al. (2002).

Viscous Damping Models

The viscous damping models are an essential part of nonlinear site response programs. This is because: (i) viscous damping is a necessary input into a linear-elastic solution that is typically an integral part of a time domain site response analysis program; (ii) viscous damping readily increases numerical stability of several numerical schemes used to solve Eq. 1; and (iii) damping calculated by a hyperbolic model at very low strains (strains less than 0.1 %) is very small compared to the values obtained by laboratory measurements. Viscous damping compensates for that difference.

Given the above, most engineers introduce a small amount of viscous damping into site
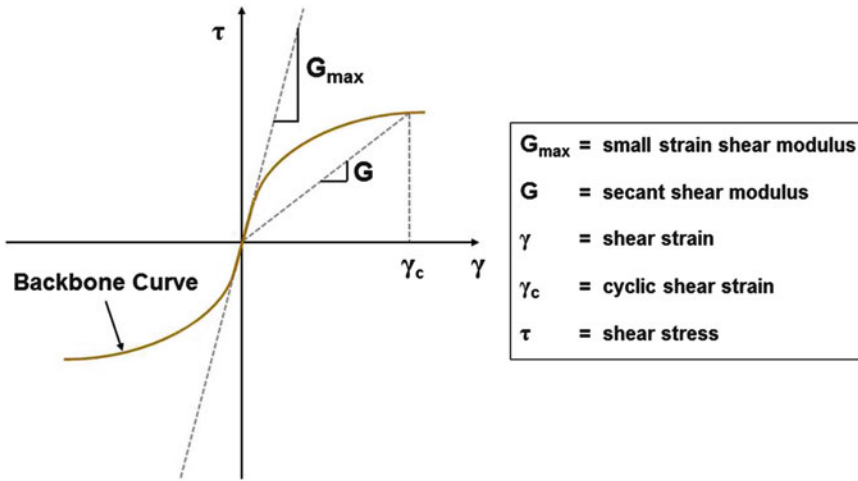
Site Response: 1-D Time Domain Analyses, Fig. 2 Backbone curve as stress–shear strain relationship for monotonic loading

Site Response: 1-D Time Domain Analyses, Fig. 3 Schematic illustration of viscous damping change with frequency
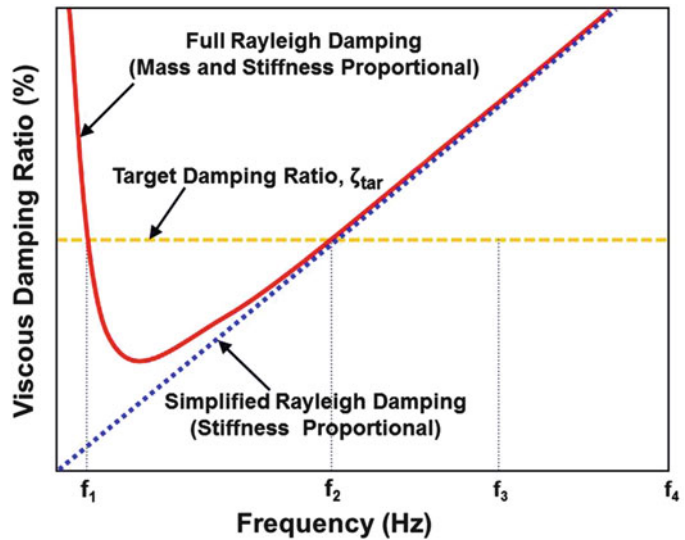
response analysis. The value of viscous damping introduced is typically small (0.5–5 %).

The most commonly used formulation for evaluation of the viscous damping coefficient (c) is the Rayleigh damping (model). The viscous damping coefficient can be evaluated as

$$c = \alpha_R \mathbf{m} + \beta_R \mathbf{k} \qquad (2)$$

where α_R and β_R are the Rayleigh damping coefficients (Rayleigh and Lindsay 1945)

and $\mathbf{m}$ and $\mathbf{k}$ are elements of the mass and stiffness matrices, respectively.

Figure 3 illustrates how Rayleigh damping, expressed through c, changes with frequency. The viscous damping ratio can be brought closer to a constant value of the target damping ratio (ξ_{tar}) by specifying c at only one frequency (e.g., at f_2 in Fig. 3), which is termed the *simplified* Rayleigh damping formulation, and/or at two frequencies (at f_1 and f_2), which is termed the *full* Rayleigh damping formulation.

Kwok et al. (2007) recommended, for most practical applications, use of the *full* Rayleigh damping formulation in nonlinear (total-stress) site response analysis whereby the first frequency is equal to the fundamental frequency of the soil column, while the second frequency is equal to five times the fundamental frequency.

Time Domain Analysis with Porewater Pressure Change

General

The cyclic loading of saturated soils is accompanied by porewater pressure change. This change includes simultaneous porewater pressure generation and dissipation. If the generated porewater pressures are sufficiently large, the soil stiffness and strength are significantly reduced, and ultimately, in some soils, liquefaction can occur.

In nonlinear site response analysis with porewater pressure generation (i.e., in effective-stress analysis), the response of the soil to cyclic loading accounts for the generation of excess porewater pressure (PWP) during cyclic shearing of the soil as well as dissipation of these excess porewater pressures during and after the cyclic loading. The representation of dissipation/redistribution of porewater pressure influences soil stiffness (modulus) and strength (shear stress) during shaking, thus resulting in a more realistic simulation of site response. The PWP dissipation/redistribution is discussed later.

The influence of PWP changes during cyclic loading is incorporated in soil constitutive models in two ways: (i) semiempirical PWP generation models used in combination with total-stress soil models and (ii) advanced effective-stress models whereby the model formulation is in terms of effective stress, and the PWP change is computed as the change between total stresses (or loads) and effective stresses via the soil constitutive model.

Semiempirical Porewater Pressure Generation Models

In this class of porewater pressure generation models, PWP generation in a soil profile is computed in a semiempirical way. At the beginning of shaking (i.e., at time t = 0), stress–strain relationships of the soil are identical to that of the total-stress models since PWP is zero. As shaking progresses, PWP is generated and cyclic degradation (of clay microstructure) starts. Subsequently, the effects of porewater pressure generation and, in some models, of cyclic degradation are included by degradation of soil strength and stiffness. Different models for degradation of soil strength and stiffness are used in some of the models. For example, Matasovic (1993) and Matasovic and Vucetic (1995b) proposed degradation index functions that degrade strength and stiffness of sands at different rates, while the concept of the degradation index (Idriss et al. 1978) is generally used to degrade strength and stiffness of soft clays.

A number of PWP generation models have been developed starting with Martin and Seed (1978) and Martin et al. (1975). A more recent example of the semiempirical porewater pressure model for saturated sand is Dobry et al. (1985) model. This model was developed based upon strain-controlled cyclic direct simple shear and cyclic triaxial testing. The model was later modified by Vucetic (1986) to allow for quasi-two-dimensional shaking and further by Matasovic (1993) to more accurately model porewater pressure-induced degradation of shear modulus and shear stress. The porewater pressure generation models described thus require the use of an equivalent number of cycles to represent earthquake shaking. Polito et al. (2008) introduced an energy-based model (GMP model) for the generation of porewater pressure based on a large number of laboratory tests, which does not require the development of an equivalent number of cycles.

With exception of the modified Dobry et al. (1985) model, there is limited information available to guide the user in selecting the appropriate PWP model parameters.

The effect of cyclic degradation on soil stiffness and strength is illustrated in Fig. 4 for the MKZ constitutive model (Matasovic 1993; Matasovic and Vucetic 1993). The initial hysteresis loop shown in the figure refers to the first cycle of cyclic loading (i.e., at time t = 0).
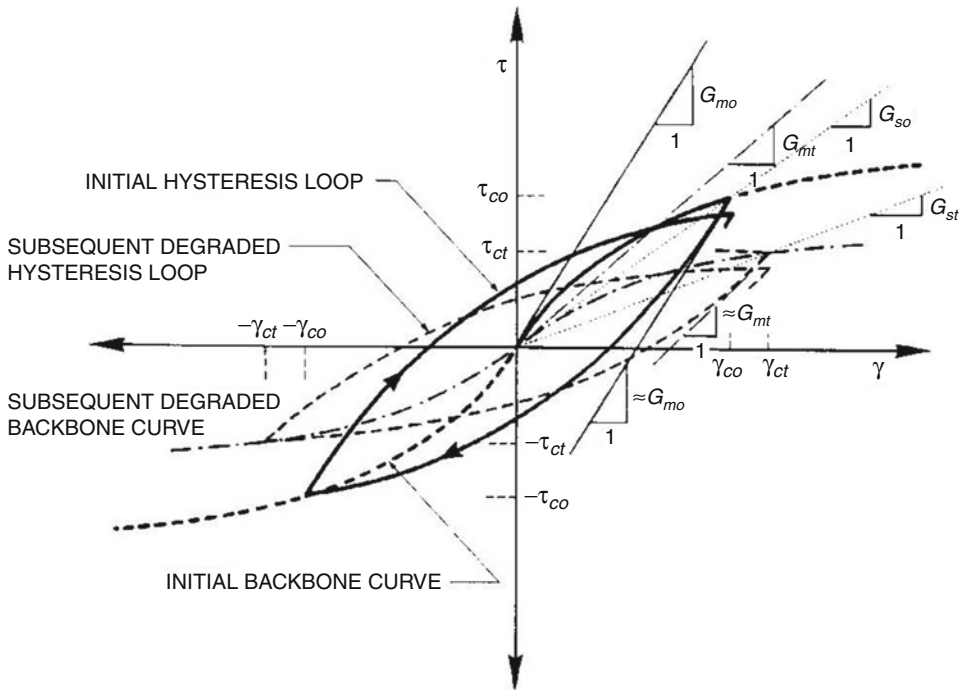
Site Response: 1-D Time Domain Analyses, Fig. 4 Stress–strain behavior modeling illustrating stiffness degradation due to porewater pressure buildup (Matasovic 1993)

The subsequent degraded hysteretic loop refers to any subsequent cycle (i.e., at time t) for which enough porewater pressure has built up to degrade both initial shear modulus G_{mo} and initial shear stress τ_{co} at corresponding shear strain γ_{co}.

An example of a porewater pressure model for clay is the Matasovic and Vucetic (1995a) model. This model was developed based upon the results of cyclic simple shear testing. It should be noted that porewater pressure in clay is of much lower intensity than in sand and that in overconsolidated clay, both positive and negative (suction) porewater pressures may develop (e.g., Matasovic and Vucetic 1992).

Advanced Effective-Stress-Based Models
Another class of soil constitutive models used in site response analysis is effective-stress models. In these models, the formulation of the constitutive law is developed in effective-stress space, and porewater pressures are computed as the difference between effective stresses and total stresses in the domain of interest. Examples of plasticity-based constitutive models include Roscoe and Schofield (1963), Pestana (1994), and Elgamal et al. (2001). These advanced constitutive models are capable of simulating complex soil behavior under a variety of loading conditions. Key elements of these models include yield surfaces, flow rules, and hardening (or softening) laws. A review of advanced constitutive models with application in site response analysis is provided in Potts and Zdravković (1999).

Generic material parameters for advanced constitutive models are often not available. Evaluation of material parameters for these models requires significant expertise and detailed site-specific soil properties.

Porewater Pressure Dissipation and Redistribution Models

The layers in a soil column have finite, saturated hydraulic conductivities. During ground shaking, even though loading is relatively rapid, it is possible that porewater dissipation and redistribution can occur due to differences in hydraulic gradients and hydraulic conductivities between layers.

An early model for porewater pressure dissipation and redistribution was introduced by Martin and Seed (1978). Input parameters include (effective-stress-dependent) constrained rebound modulus and (saturated) hydraulic conductivity. A porewater pressure dissipation model for composite soil deposits (alternating sand and clay layers) was developed by Matasovic and Vucetic (1995a).

Porewater pressure redistribution occurs and is often modeled in accordance with the principles of Terzaghi's theory of consolidation.

Time Domain Analysis Computer Programs

There are about 25 time domain analysis computer programs that are available globally, either commercially or through direct contact with the author(s). However, only a handful of these programs have a provision to perform an effective-stress analysis, including 2-D and 3-D programs which can be used for analysis of 1-D geometry.

Kwok et al. (2007) performed a comprehensive survey of available time domain analysis computer programs in the United States and found that only five 1-D nonlinear programs were used at the time, including OpenSees (Ragheb 1994) which is primarily used for analysis of 2-D and 3-D problems. Other programs identified by Kwok et al. (2007) include DEEPSOIL (Hashash and Park 2001), D-MOD_2 (Matasovic 2006; subsequently updated with graphical user interface as D-MOD2000), SUMDES (Wang 1990), and

TESS (Pyke 2000), with last three programs also having a provision for effective-stress analysis. The study compared the programs in total-stress mode and found that, when input is properly controlled, most of the software provided similar results.

Matasovic and Hashash (2012) performed a survey of the United States Department of Transportation (DOT) practices. Part of the survey was an inquiry on what site response analysis programs are used, both by DOT-s and their consultants. The survey revealed that the most used program was D-MOD2000 (Matasovic and Ordonez 2007) followed by DEEPSOIL (Hashash et al. 2011). However, D-MOD2000 was the only program cited to have been used in effective-stress mode. Coincidently, these two programs are the only time domain analysis programs available in United States that have a graphical user interface and the effective-stress analysis provision.

Calibration and Benchmarking Studies

A number of individuals and groups have conducted back-analysis exercises to evaluate site response analysis procedures (and programs) versus measured response to strong ground shaking. Most of these studies include use of recordings from downhole vertical arrays.

Researchers have used a range of inverse analysis techniques that include ad hoc system identification (e.g., Zeghal and Elgamal 1993; Assimaki and Steidl 2007) and evolutionary soil models (e.g., Tsai and Hashash 2007) to calibrate soil constitutive models in site response analysis actual (in situ) soil behavior.

A number of benchmarking exercises have also been conducted to evaluate site response analysis tools. A recent key exercise is the PEER benchmarking exercise for total-stress site response analyses (Kwok et al. 2007). Another interesting benchmarking exercise was the evaluation of site response at Turkey Flat (e.g., Kramer 2009). The interesting outcome of

this exercise is in highlighting the challenges encountered in computing the response at a well-constrained relatively stiff soil site.

Summary

One-dimensional time domain analysis is a powerful tool for evaluation of site response of horizontally layered soil deposits. Compared to other types of 1-D analysis, it has a provision to simulate cyclic soil behavior in a more realistic manner than its frequency domain counterparts, and it can accommodate models for seismically induced porewater pressure buildup. Compared to its 2-D and 3-D counterparts, the advantage of this type of analysis lies in simple constitutive models for which material parameters can be readily evaluated, or generic sets of material parameters are available to practicing engineers. Furthermore, compared to its 2-D and 3-D counterparts, 1-D analysis is far better calibrated and validated. Consequently, consistent with current trends, it is anticipated that the role of 1-D analysis will evolve into calibration of 2-D and 3-D models.

One-dimensional time domain analysis is not perfect. Further calibration and validation of 1-D models are required both in effective-stress mode and at high levels of shaking (>0.4 g). There are also several issues related to the application of this type of analysis in engineering practice and research, as summarized by Hashash et al. (2010). Future trends will certainly address these issues, yet also include improvement of graphical user interfaces that are essential not only for effective use of the programs, but for evaluation and validation of the results.

Cross-References

▶ Selection of Ground Motions for Response History Analysis
▶ Site Response: Comparison Between Theory and Observation

References

Assimaki D, Steidl J (2007) Inverse analysis of weak and strong motion downhole array data from the Mw 7.0 Sanriku-Minami earthquake. Soil Dyn Earthq Eng 27:73–92

Borja RI, Duvernay BG, Lin CH (2002) Ground response in Lotung: total stress analyses and parametric studies. J Geotech GeoEnviron Eng 128(1):54–63

Clough RW, Penzien J (1993) Dynamics of structures, 2nd edn. McGraw Hill, London

Darendeli MB (2001) Development of a new family of normalized modulus reduction and material damping curves. PhD thesis, University of Texas at Austin, Austin

Dobry R, Pierce WG, Dyvik R, Thomas GE, Ladd RS (1985) Pore pressure model for cyclic straining of sand. Research Report, Civil Engineering Department, Rensselaer Polytechnic Institute, Troy, 56 p

Duncan JM, Chang C-Y (1970) Nonlinear analysis of stress and strain in soils. J Soil Mech Found Div 96(SM5):1629–1653

Elgamal A, Lai T, Yang Z, He L (2001) Dynamic soil properties, seismic downhole arrays and applications in practice. In: Proceedings of the 4th international conference on recent advances in geotechnical earthquake engineering and soil dynamics, San Diego

Hashash YMA, Park D (2001) Nonlinear one-dimensional seismic ground motion propagation in the Mississippi embayment. Eng Geol (Amst) 62(1–3):185–206

Hashash YMA, Phillips C, Groholski D (2010) Recent advances in non-linear site response analysis. In: Proceedings of the 5th international conference on recent advances in geotechnical earthquake engineering and soil dynamics, San Diego

Hashash YMA, Groholski DR, Phillips CA, Park D, Musgrove M (2011) DEEPSOIL 4.0, User Manual and Tutorial. 98 p

Idriss IM, Dobry R, Singh RD (1978) Nonlinear behavior of soft clays during cyclic loading. J Geotech Eng Div ASCE 104(12):1427–1447

Kondner RL, Zelasko JS (1963) A hyperbolic stress–strain formulation of sands. In: Proceedings of the 2nd Pan American conference on soil mechanics and foundation engineering, Sao Paulo, pp 289–324

Kramer SL (2009) Analysis of Turkey flat ground motion prediction experiment – lessons learned and implications for practice. In: SMIP09 seminar on utilization of strong-motion data, San Francisco, pp 1–22

Kwok O-LA, Stewart JP, Hashash YMA, Matasovic N, Pyke R, Wang Z, Yang Z (2007) Use of exact solutions of wave propagation problems to guide implementation of nonlinear ground response analysis procedures. ASCE J Geotechn Geoenviron Eng 133(11): 1385–1398

Martin PP, Seed HB (1978) APOLLO, a computer program for the analysis of pore pressure generation and dissipation in horizontal sand layers during cyclic

earthquake loading. Report No. EERC 78–21, Earthquake Engineering Research Center, University of California, Berkeley

Martin GR, Finn WDL, Seed HB (1975) Fundamentals of liquefaction under cyclic loading. J Geotech Eng Div ASCE 101(GT5):423–438

Masing G (1926) Eigenspannungen und Verfestigung beim Messing. In: Proceedings of the 2nd international congress on applied mechanics, Zürich, pp 332-335

Matasovic N (1993) Seismic response of composite horizontally layered soil deposits. PhD dissertation, Civil and Environmental Engineering Department, University of California, Los Angeles, 452 p

Matasovic N (2006) D-MOD_2 – a computer program for seismic response analysis of horizontally layered soil deposits, earthfill dams, and solid waste landfills. User's Manual, GeoMotions, LLC, Lacey, 20 p (plus Appendices)

Matasovic N, Hashash YMA (2012) Site response analysis in transportation engineering practice – a TRB Survey. In: Proceedings of the GeoCongress 2012, Oakland, CD-ROM paper, pp 1789–1798

Matasovic N, Ordonez GA (2007) D-MOD2000 – a computer program package for seismic response analysis of horizontally layered soil deposits, earthfill dams, and solid waste landfills. User's Manual, GeoMotions, LLC, Lacey, Washington, 182 p. http://www.geomotions.com

Matasovic N, Vucetic M (1992) A pore pressure model for cyclic straining of clay. Soils Found J JSSMFE 32(3):156–173

Matasovic N, Vucetic M (1993) Cyclic characterization of liquefiable sands. ASCE J Geotech Eng 119(11):1805–1822

Matasovic N, Vucetic M (1995a) Generalized cyclic degradation-pore pressure generation model for clays. ASCE J Geotech Eng 121(1):33–42

Matasovic N, Vucetic M (1995b) Seismic response of soil deposits composed of fully saturated clay and sand layers. In: Kenji Ishihara AA (ed) Proceedings of the IS-Tokyo'95, the first international conference on earthquake geotechnical engineering, Balkema, Rotterdam

Newmark NM (1959) A method of computation for structural dynamics. J Eng Mech Div 85:67–94

Pestana JM (1994) A unified constitutive model for clays and sands. ScD thesis, MIT, Cambridge, MA

Phillips C, Hashash YMA (2009) Damping formulation for nonlinear 1D site response analyses. Soil Dyn Earthq Eng 29(7):1143–1158

Polito CP, Green RA, Lee JH (2008) Pore pressure generation models for sands and silty soils subjected to cyclic loading. J Geotech Geoenviron 134(10):1490–1500

Potts DM, Zdravković L (1999) Finite element analysis in geotechnical engineering: theory. Thomas Telford, London

Pyke RM (1979) Nonlinear soil models for irregular cyclic loadings. J Geotech Eng Div, ASCE 105(GT6):715–726

Pyke RM (2000) TESS: a computer program for nonlinear ground response analyses. TAGA Engineering Systems & Software, Lafayette, http://www.tagasoft.com

Ragheb AM (1994) Numerical analysis of seismically induced deformations in saturated granular soil strata. PhD dissertation, Department of Civil Engineering, Rensselaer Polytechnic Institute, Troy

Ramberg W, Osgood WR (1943) Description of stress–strain curves by three parameters. Technical Note 902, National Advisory Committee for Aeronautics, Washington, DC

Rayleigh JWS, Lindsay RB (1945) The theory of sound. Dover, New York

Roscoe KH, Schofield AN (1963) Mechanical behavior of an idealized 'wet' clay. In: Proceedings of the 2nd European conference on soil mechanics, vol 1, Wiesbaden, pp 47–54

Tsai C-C, Hashash YMA (2007) A novel framework integrating downhole array data and site response analysis to extract dynamic soil behavior. Soil Dyn Earthq Eng 28(3):181–197

Vucetic M (1986) Pore pressure buildup and liquefaction of level sandy sites during earthquakes. PhD dissertation, Rensselaer Polytechnic Institute, Troy, 616 p

Vucetic M (1990) Normalized behavior of clay under irregular cyclic loading. Can Geotech J 27:29–46

Wang ZL (1990) Bounding surface hypoplasticity model for granular soils and its applications. PhD dissertation, University of California at Davis

Zeghal M, Elgamal AW (1993) Lotung sites: downhole seismic data analysis. Electric Power Research Institute, Palo Alto

Site Response: Comparison Between Theory and Observation

Takaji Kokusho
Department of Civil & Environmental
Engineering, Chuo University, Tokyo, Japan

Synonyms

1/4 wavelength formula; Two-layer system; Downhole array; Multilayer system; Multireflection theory; Nonlinear soil properties; Seismic site amplification; Soil damping model; Surface array; S-wave velocity

Introduction

Kanai et al. (1959) discovered from earthquake observations that a horizontal component in site response can be explained by multi-reflection of SH-waves propagating vertically in soil layers. Horizontal displacement of the vertically propagating SH-wave is expressed as

$$u = Af(z - V_s t) + Bf(z + V_s t) \qquad (1)$$

where z = vertical coordinate upward positive; t = time; V_s = S-wave velocity; $f()$, $g()$ = an arbitrary functions; and A, B = amplitudes of upward and downward waves, respectively. Site response governed by this equation is largely dependent on the composition of soil layers, S-wave velocities, and soil damping ratios of those layers and their strain-dependent variations. In the following, site response observations in array systems are compared with the SH-wave multi-reflection theory to clarify its applicability in terms of dominant frequencies and spectrum amplifications with special emphasis on soil damping mechanism and strain-dependent soil nonlinearity.

Basic Mechanism on Site Amplification

For a harmonic wave with angular frequency ω and amplitudes A, B, Eq. 1 is replaced by

$$u = Ae^{i(\omega t - kz)} + Be^{-i(\omega t + kz)} \qquad (2)$$

Here, k is the wave number expressed as

$$k = \omega/V_s = \omega/(G/\rho)^{1/2} \qquad (3)$$

Basic mechanisms of site amplification may be simplified by a two-layer model as illustrated in Fig. 1, where a surface layer with the thickness H overlies a base layer of infinite thickness. Pertinent properties are: ρ_1, ρ_2 = soil density; V_{s1}, V_{s2} = S-wave velocity; G_1, G_2 = shear modulus; and D_1, D_2 = damping ratio of the surface
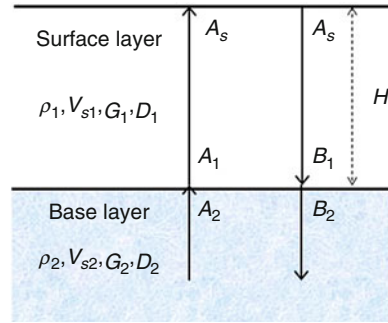
Site Response: Comparison Between Theory and Observation, Fig. 1 A simplified two-layer model to calculate site amplification

and base layer, respectively. The wave equations in the surface and base layers are expressed as

$$\begin{aligned} u_1 &= A_1 e^{i(\omega t - k_1 z)} + B_1 e^{i(\omega t + k_1 z)} \\ u_2 &= A_2 e^{i(\omega t - k_2 z)} + B_2 e^{i(\omega t + k_2 z)} \end{aligned} \qquad (4)$$

where A_1, B_1 = amplitudes of upward and downward waves in the surface layer and A_2, B_2 = those in the base layer.

Utilizing the boundary conditions that at $z = 0$, $u_1 = u_2$ and $G_1 \frac{\partial u_1}{\partial z} = G_2 \frac{\partial u_2}{\partial z}$ and at $z = H$, $G_1 \partial u_1/\partial z = 0$, and also introducing an impedance ratio α as

$$\alpha = \frac{k_1 G_1}{k_2 G_2} = \frac{\rho_1 V_{s1}}{\rho_2 V_{s2}}, \qquad (5)$$

the transfer function $(2A_s)/(2A_2)$ is obtained as

$$\frac{2A_s}{2A_2} = \frac{2}{(1 + \alpha)e^{ik_1 H} + (1 - \alpha)e^{-ik_1 H}} \qquad (6)$$

Here, $2A_s$ is the wave amplitude to be observed at the ground surface, and $2A_2$ is that at the base layer if it were outcropped and free from the overburden. The site amplification to be used in micro-zonation mapping is determined by Eq. 6.

There are two types of site response monitoring: (a) surface array and (b) downhole array as illustrated in Fig. 2. In the surface array (a), ground motion is monitored at two different
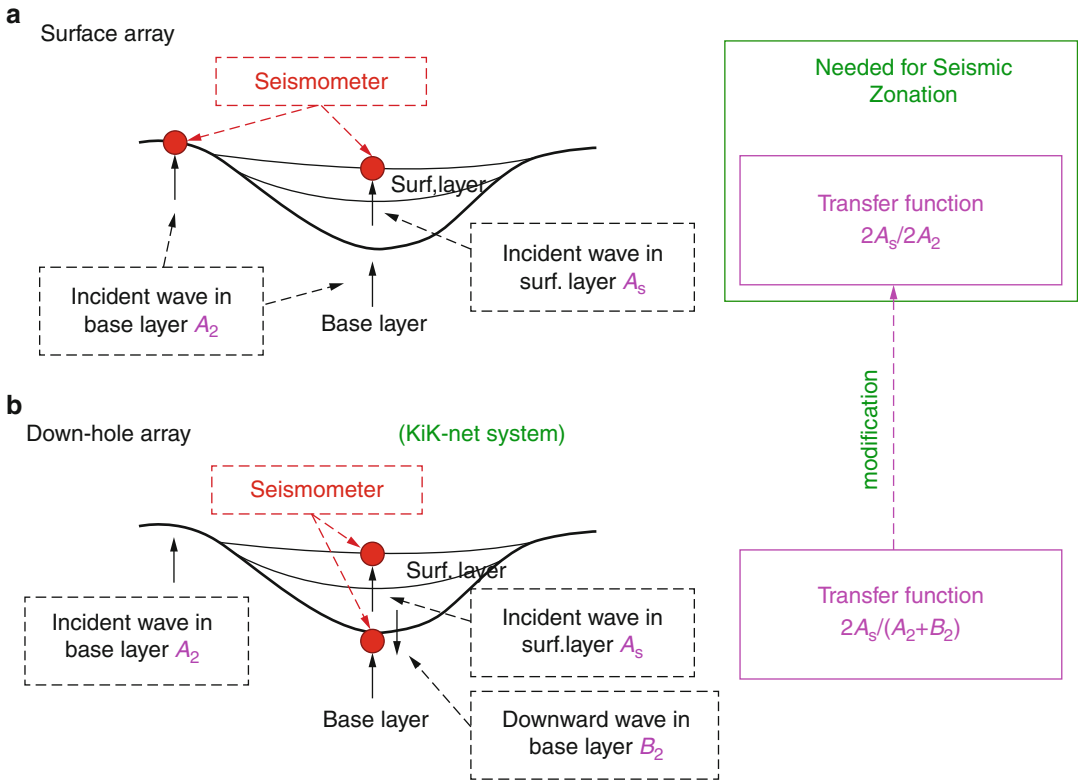
Site Response: Comparison Between Theory and Observation, Fig. 2 Two types of earthquake observation array systems to measure site amplification between ground surface and base layer

surface locations with different geologies, overlying soft layer and outcropping stiff base layer. If the soft layer is underlain by the same base layer and the upward wave in the base layer A_2 is assumed the same at the two places, the site amplification in the soft soil site with respect to outcropping base layer is given by Eq. 6.

In the downhole array (b), surface and downhole seismometers can evaluate the site amplification exactly at the same location, and its transfer function can be formulated as

$$\frac{2A_s}{A_2 + B_2} = \frac{2}{e^{ik_1 H} + e^{-ik_1 H}} \qquad (7)$$

Here, B_2 is the amplitude of the downward wave in the base layer, which is influenced by the dynamic response of the surface layer. In order

to derive the transfer function for micro-zonation $(2A_s)/(2A_2)$, Eq. 7 cannot directly be used, but some modification is necessary to extract the outcropping base motion $2A_2$ from observed base motion $(A_2 + B_2)$.

In a two-layer system, the resonant frequency can be computed by the next equation

$$f = \omega/2\pi = (2n - 1)\frac{V_{s1}}{4H} \qquad (8)$$

where n = the order of resonance. The most important is the first-order resonant frequency $n = 1$, and the equation $f_1 = V_{s1}/(4H)$ is named as the 1/4 wavelength formula. In many cases, the site amplification in real site conditions with multilayer systems can be simplified by a
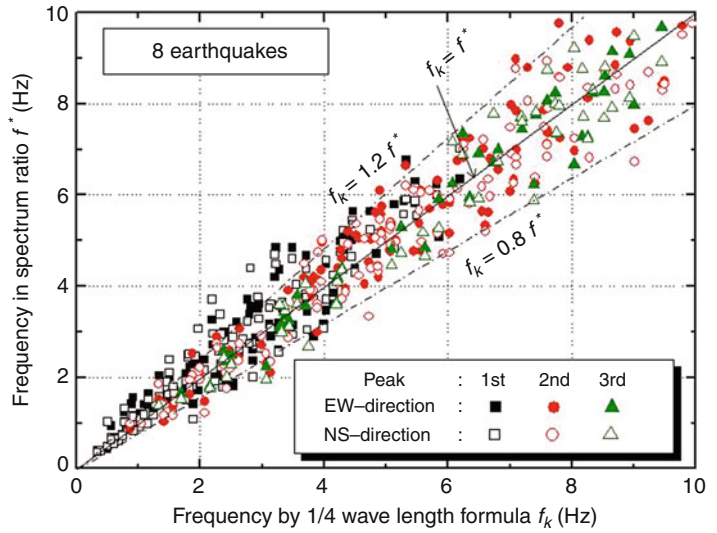
two-layer model consisting of a soft surface layer underlain by a stiff base layer. In this case, too, the 1/4 wavelength formula can be expressed in an extended form as

$$f_j = 1 \left/ \left[4 \sum_{i=1}^{j} (H_i/V_{si}) \right] \right. \qquad (9)$$

Here, j is the number of surface layers involved to create a specific resonant frequency. In Fig. 3, frequencies f_j obtained from the above equation are compared with peak frequencies f^* in the spectrum ratio of observed motions at a number of downhole array sites in Japan during recent strong earthquakes. The figure shows that the value f^* are mostly within 0.8–1.2 times f_k, indicating that the extended 1/4 wavelength formula Eq. 9 may be useful to roughly estimate the resonant frequency by simplifying actual multilayer soil systems by the two-layer system. Also note in the diagram that Eq. 9 holds for not only the first- but also second- and third-order peak frequencies, which are generated by combinations of base layers of different depths and corresponding overlying layers.

In Fig. 4, two types of transfer functions of the two-layer system in Fig. 1 calculated by Eqs. 6 and 7 are compared in the same diagram. It is apparent from the two equations that the properties of the base layer are included in $(2A_s)/(2A_2)$ in terms of impedance ratio α but not in $2A_s/(A_2 + B_2)$. When $\alpha = \rho_1 V_{s1}/\rho_2 V_{s2} = 0$ in Eq. 6, indicating $\rho_2 V_{s2} \to \infty$ (rigid base layer), Eqs. 6 and 7 are identical. Hence, in the downhole array transfer function, the base layer is equivalent to a rigid base with the prescribed motion $A_2 + B_2$ when no radiation damping occurs, resulting in infinite amplification in resonant frequencies if there is no soil damping in the surface layer, $D_1=0$. In contrast, $(2A_s)/(2A_2)$ produces a certain finite peak values corresponding to α even for $D_1=0$, because energy radiation into the base layer is possible. It is noteworthy that in a multilayer system, the base layer where the downhole seismometer is installed also serves as the rigid base, whose properties have nothing to do with the downhole array amplification (Schnabel et al. 1972).

Soil Damping Models and Site Amplification

In order to incorporate soil damping in site amplifications, the variables included in Eqs. 6 and 7, k,

Site Response: Comparison Between Theory and Observation, Fig. 4 Comparison to two types of transfer function $2A_s/2A_2$ and $2A_s/(A_2 + B_2)$ in a two-layer system
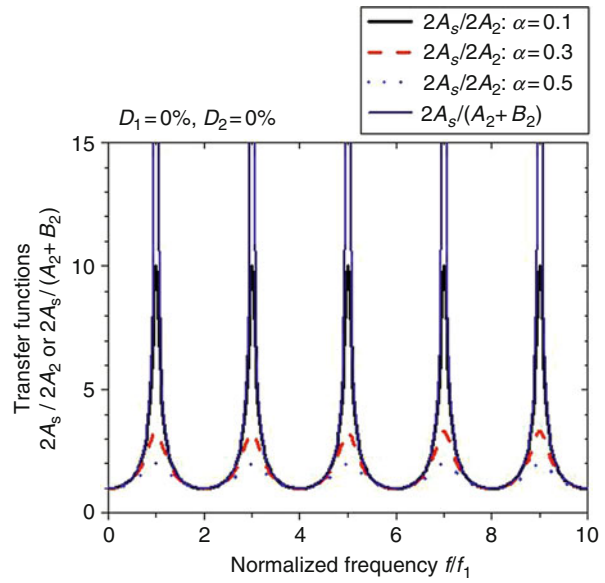
Site Response: Comparison Between Theory and Observation, Fig. 5 Three mechanical models for soil damping
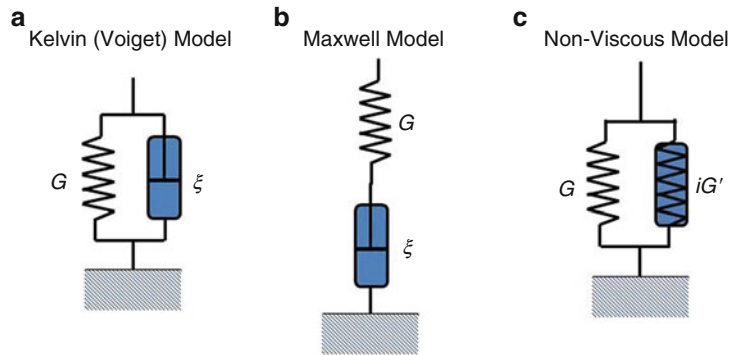
V_s, and G, should be replaced by complex wave number k^*, complex S-wave velocity V_s^*, and complex shear modulus G^*, respectively, so that Eq. 3 is modified as

$$k^* = \omega/V_s^* = \omega/(G^*/\rho)^{1/2} \qquad (10)$$

Three mechanical models for soil damping can be incorporated in site amplification: (a) Kelvin (Voigt) model, (b) Maxwell model, and (c) nonviscous model as illustrated in Fig. 5, consisting of a spring with shear modulus G and a dashpot with viscosity ξ. For model (c), the dashpot is replaced by a nonviscous

(time independent) dashpot with an imaginary constant iG'. For each model, shear stress τ versus shear strain γ relationship is expressed as

$$\tau = G^*\gamma \qquad (11)$$

where G^* is complex shear modulus considering soil damping. The complex shear modulus G^* normalized by G is formulated for each damping models a, b and c using $i = \sqrt{-1}$ as follows (Ishihara 1996):

$$G^*/G = 1 + i(\xi/G)\omega \qquad (12a)$$
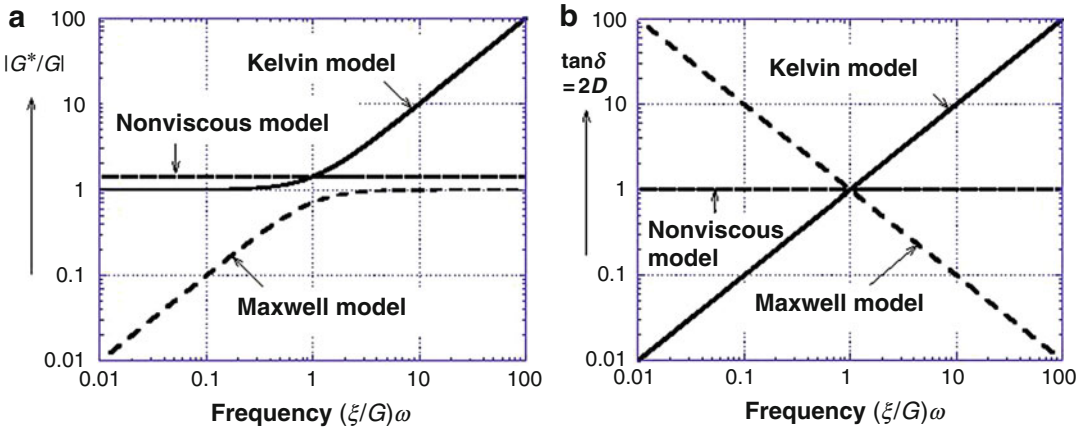
Site Response: Comparison Between Theory and Observation, Fig. 6 Variations of normalized shear modulus |G^*/G| and tan δ along with normalized angular frequency $(\xi/G)\omega$

$$G^*/G = \frac{i\omega}{1/(\xi/G) + i\omega} \tag{12b}$$

$$G^*/G = 1 + iG'/G \tag{12c}$$

In these equations, the normalized complex shear modulus can be expressed as

$$G^*/G = |G^*/G|e^{i\delta} \tag{13}$$

where δ is the phase lag angle between stress–strain relationships of the models, and tan δ is called a loss coefficient. In Fig. 6, the variations of normalized complex shear modulus |G^*/G| and tan δ are shown versus dimensionless angular frequency $(\xi/G)\omega$ for models (a) and (b), and constant values are shown for model (c). The damping ratio is correlated with the loss coefficient as follows (Ishihara 1996):

$$D = (\tan\delta)/2 \tag{14}$$

Hence, from Eqs. 12a, b, and c:

$$D = \tan\delta/2 = \omega\xi/(2G) \tag{15a}$$

$$D = \tan\delta/2 = G/(2\omega\xi) \tag{15b}$$

$$D = \tan\delta/2 = G'/(2G) \tag{15c}$$

From Eq. 2, the equation for a wave including soil damping propagating unidirectionally toward $+z$ is written as

$$u = Ae^{i(\omega t - k^* z)} \tag{16}$$

If k^* in this equation is substituted by Eq. 10, and G^* in Eq. 10 is further substituted by Eqs. 12a, b, and c, Eq. 16 can be transformed as follows:

$$u = Ae^{-\beta z}e^{i\omega(t - z/V_s)} \tag{17}$$

Here, β is a positive real number called a wave attenuation coefficient by internal damping, because it determines how the wave attenuates as it propagates by distance z. If *delta* is small, *beta* is correlated to the damping ratio D as

$$\beta = \omega D/V_s \tag{18}$$

Substituting D in Eq. 15a, b and c into Eq. 18 gives β for the corresponding damping models:

$$\beta = \omega^2\xi/(2\rho V_s^3) \tag{19a}$$

$$\beta = \rho V_s/(2\xi) \tag{19b}$$

$$\beta = \omega G'/(2\rho V_s^3) \tag{19c}$$
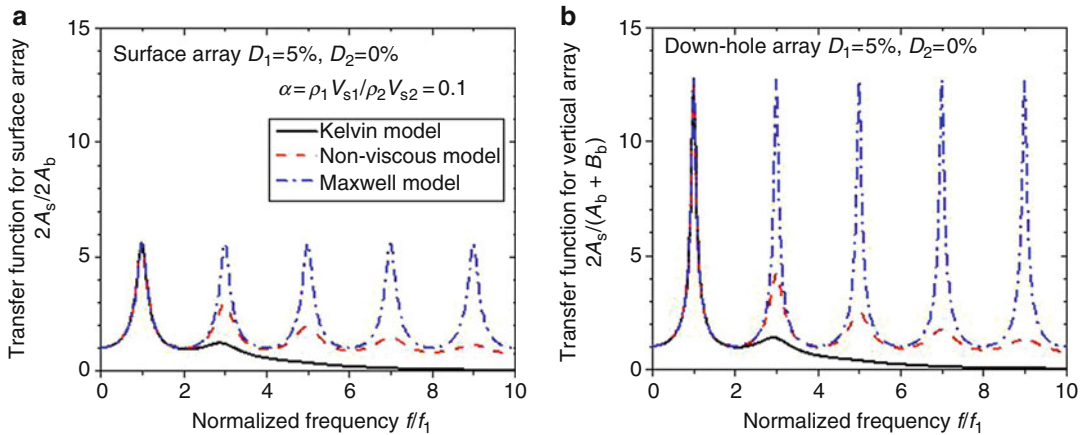
Site Response: Comparison Between Theory and Observation, Fig. 7 Transfer functions of a two-layer model using three types of soil damping models: (**a**) $2A_s/$ $2A_b$ ($\alpha = 0.1$) for surface array and (**b**) $2A_s/(A_b + B_b)$ for downhole array

In Fig. 7, transfer functions of the two-layer model in Fig. 1 are shown for the three damping models assuming the damping ratios in the surface and base layers $D_1 = 5\%$ and $D_2 = 0\%$, respectively. Figure 7a is for the surface array $(2A_s)/(2A_b)$ with the impedance ratio $\alpha = (\rho_1 V_{s1})/(\rho_2 V_{s2}) = 0.1$, and Fig. 7b is for the downhole array $2A_s/(A_b + B_b)$. The transfer functions are identical in the first peak but obviously differ in higher-order peaks corresponding to the different damping models. In the Kelvin model, the amplification in higher-order peaks tend to reduce more drastically than in the nonviscous model, and in the Maxwell model, it is unchanged at all peaks. These differences arise from the different formulations of β in Eqs. 19a, b, and c with respect to ω. In cyclic loading tests on soil elements, it is widely accepted that soil damping is almost frequency independent (Hardin 1965; Hardin and Drnevich 1972). Hence, in most engineering practice, the soil damping is assumed nonviscous as in Eq. 15c. In this case, the peak amplification of the transfer function becomes lower for higher-order peaks because the wave attenuation coefficient β is proportional to ω as shown in Eq. 19c.

In Figs. 8a–e, some typical Fourier spectrum ratios calculated between surface and base

records at six downhole array sites where PGA exceeds 0.2 g during strong earthquakes are depicted with thin lines in EW and NS directions. For each site, the soil profile with density ρ and S-wave velocity V_s is tabulated in Table 1 together with installation levels of seismometers indicated with arrows. Using the properties in the table, transfer functions for each site, $2A_s/(A_b + B_b)$ for the downhole array and $(2A_s)/(2A_b)$ for the surface array, are calculated and superposed in the diagram with thick curves, where A_b and B_b = amplitudes of upward and downward waves in the base layer. In the calculation, the nonviscous damping model is used as normal engineering practice, and a uniform damping ratio $D = 2.5\%$ is assumed tentatively for all layers. If the observed spectrum ratios are compared with corresponding transfer functions for downhole arrays, $2A_s/(A_b + B_b)$, a fairly well correspondence in peak frequencies can be recognized between the two in most sites despite some minor effect of soil nonlinearity. This indicates the applicability of one-dimensional soil models in these sites to a certain extent.

If $2A_s/(A_b + B_b)$ is compared with $(2A_s)/(2A_b)$ at each site, the difference in the peak values between the two transfer functions is obviously large.

a EQ.3 (TKCH08)

Layer No.	Thickness H[m]	Depth [m]	Density [t/m²]	Vs [m/s]
1	4	4	2.1	130
2	32	36	2.1	480
3	42	78	2.1	590
4	22	100	2.6	2800

b EQ.3 (TKCH11)

Layer No.	Thickness H[m]	Depth [m]	Density [t/m²]	Vs [m/s]
1	1	1	1.8	150
2	11	12	21	470
3	10	22	2.1	470
4	23	45	2.2	570
5	6	51	2.3	730
6	13	64	2.4	970
7	14	78	2.5	1500
8	22	100	2.6	2680

c EQ.3 (NMRH02)

Layer No.	Thickness H[m]	Depth [m]	Density [t/m²]	Vs [m/s]
1	4	4	2	110
2	6	10	2.1	320
3	30	40	2.1	500
4	8	48	1.6	430
5	34	82	2.1	510
6	12	94	2.3	650
7	9	103	2.4	870

d EQ.3 (TKCH04)

Layer No.	Thickness H[m]	Depth [m]	Density [t/m²]	Vs [m/s]
1	4	4	2.1	380
2	16	20	2.1	380
3	18	38	2.1	680
4	46	84	2.1	890
5	88	172	2.1	1020

e EQ.5 (FKOH03)

Layer No.	Thickness H[m]	Depth [m]	Density [t/m²]	Vs [m/s]
1	6	6	2.1	250
2	8	14	2.2	440
3	11	25	2.3	700
4	75	100	2.6	2030

f EQ.5 (SAGH03)

Layer No.	Thickness H[m]	Depth [m]	Density [t/m²]	Vs [m/s]
1	4	4	2	160
2	10	14	2.1	380
3	20	34	2.4	900
4	69	103	2.6	1980

Site Response: Comparison Between Theory and Observation, Fig. 8 Transfer functions of a multilayer system $2A_s/(A_b + B_b)$ compared with $2A_s/2A_b$ and observed spectrum ratios at six downhole array sites during a strong earthquake

However, the coincidence in peak frequencies is almost perfect in (a) and good in (b) but gets poorer in (c) and (d) and very poor in (e) and (f). The reason may be gleaned by examining the soil profiles. In (a) and (b), the V_s-value at the downhole seismometer is much larger than the upper layers, and the seismometer depth is not so deep from the boundary of clear V_s-contrast. In (c) and (d), the V_s-value at the base layer is not so different from the upper layers, and the seismometer depth is not so deep from a boundary of major V_s-contrast. In (e) and (f), though the V_s-value at the base layer is much larger than the upper layers, the depth of seismometer is too deep from the upper boundary of clear V_s-contrast to properly detect the response of the upper layers. This observation tells us a significance of choosing appropriate seismometer depth in deploying a downhole array system considering site specific soil profiles.

Another observation may be made on amplification values at individual peaks of observed spectrum ratios compared with the downhole array transfer functions $2A_s/(A_b + B_b)$.

Site Response: Comparison Between Theory and Observation, Table 1 Soil profiles and properties at two vertical array sites corresponding to spectrum ratios in Figs. 9 and 10

For Fig. 9				
IBUH03	Soil profiles and properties			
Layer no.	Thickness h (m)	Depth H (m)	Density (t/m^3)	V_s (m/s)
1	2	2	21	60
2	16	18	21	90
3	10	28	20	190
4	12	40	21	320
5	12	52	21	210
6	24	76	21	310
7	56	132	21	430
8	21	153	22	520
For Fig. 10				
NMRH02	Soil profiles and properties			
Layer no.	Thickness h (m)	Depth H (m)	Density (t/m^3)	V_s (m/s)
1	4	4	2.0	110
2	6	10	2.1	320
3	30	40	2.1	500
4	8	48	1.6	430
5	34	82	2.1	510
6	12	94	2.3	650
7	9	103	2.4	870

As already mentioned, the damping ratio is tentatively set as 2.5 % in calculating $2A_s/(A_b + B_b)$. At a glance, the peak amplification values of the transfer functions are almost monotonically decreasing with increasing peak frequencies, presumably reflecting the damping characteristics of the non-viscous model used here, as previously explained. In contrast, the peak values of observed spectrum ratios seem to be quite site specific and show no consistent increasing or decreasing trend. This indicates that the soil damping is not perfectly nonviscous (frequency independent) but to some extent frequency dependent (the damping ratio D decreases with increasing frequency f) in a similar manner as the Maxwell model. It is clear that the Kelvin model, for which the peak amplification tends to decrease more drastically with increasing f, has even less applicability than the nonviscous model. Thus, the frequency-independent damping found in laboratory soil tests may not perfectly fit the actual performance of site amplification based on earthquake observations.

One major reason for this deviation may be wave scattering in heterogeneous soils. The wave scattering effect on frequency dependency of damping in elastic waves propagating in heterogeneous earth crust is an important topic in seismology (e.g., Frankel and Clayton 1986; Wu 1982). The effect tends to be more conspicuous with decreasing strain in a stiff rock site. The damping ratio D by wave scattering may be expressed as

$$D = D_r(f/f_r)^{-m} \qquad (20)$$

where f_r = a reference frequency, D_r = reference damping ratio, and $D = D_r$ at $f = f_r$. The power constant m is a positive number, and $m = 1$ corresponds to the Maxwell model, as obviously seen in Eq. 15b. In general, the damping ratio may be expressed as the sum of the frequency-independent hysteretic part D_0 and the frequency-dependent part as

$$D = D_0 + D_r(f/f_r)^{-m} \qquad (21)$$

S
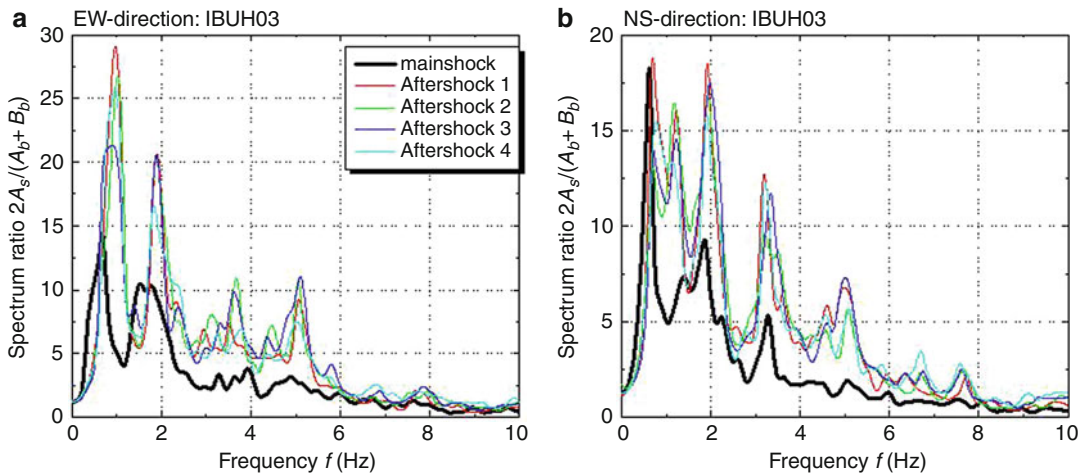
Site Response: Comparison Between Theory and Observation, Fig. 9 Fourier spectrum ratios $2A_s/(A_b + B_b)$ observed at a very soft soil site during main shock and aftershocks in directions EW (**a**) and NS (**b**)

However, the frequency-dependency of damping is normally ignored in geotechnical engineering practice, because hysteretic nature of soil damping becomes more dominant during strong earthquakes as soil gets softer and strain-dependent nonlinearity becomes more dominant.

Soil Nonlinearity

Seismically induced shear strain γ is calculated from particle velocity $\dot{u}$:

$$\gamma = -\dot{u}/V_s \qquad (22)$$

if upward propagating wave is chosen in Eq. 1, for instance. This indicates that the shear strain tends to be greater for strong earthquakes with larger particle velocity amplitudes and in soft soils with smaller V_s-values. Because stress–strain behavior of soil is nonlinear and hence V_s and D are highly strain dependent, the site response during strong earthquakes may greatly differ from that during small earthquakes, particularly in soft soil sites. The soil nonlinearity effect is normally evaluated using the equivalent linear approximation method in engineering practice (Schnabel et al. 1972).

In each of Figs. 9 and 10, Fourier spectrum ratios for a strong main shock are compared with

those for weaker aftershocks in directions EW (a) and NS (b) using earthquake records obtained in the same downhole arrays. As indicated in Table 1, the soil condition for the former site is very soft ($V_s < 100$ m/s at top 18 m), while that of the latter site are relatively stiff ($V_s > 320$ m/s deeper than 4 m). Thick curves for the main shock are obviously different with thin curves for the aftershocks with respect to peak values and peak frequencies. The difference of peak values is greater in Fig. 9 than in Fig. 10 because of the softer soil conditions, and the difference in the same site tends to be greater in higher-order peaks than in the first peak because softer layers near the ground surface tends to contribute more.

Basic effects of soil nonlinearity on site amplification can be examined using the two-layer system shown in Fig. 1. Three levels of induced equivalent strain amplitude are assumed, and corresponding shear modulus degradation G/G_0 and equivalent damping ratio D (nonviscous) are determined based on empirical curves (Fig. 11) often used in engineering practice (Seed and Idriss 1970). Figures 12a and b depict the transfer functions $2A_s/(A_b + B_b)$ for the downhole array and $(2A_s)/(2A_b)$ for the surface array calculated for the three levels of nonlinear soil properties.

Obviously, soil nonlinear properties have great effects on the peak frequencies and the peak amplifications not only in $2A_s/(A_b + B_b)$ as
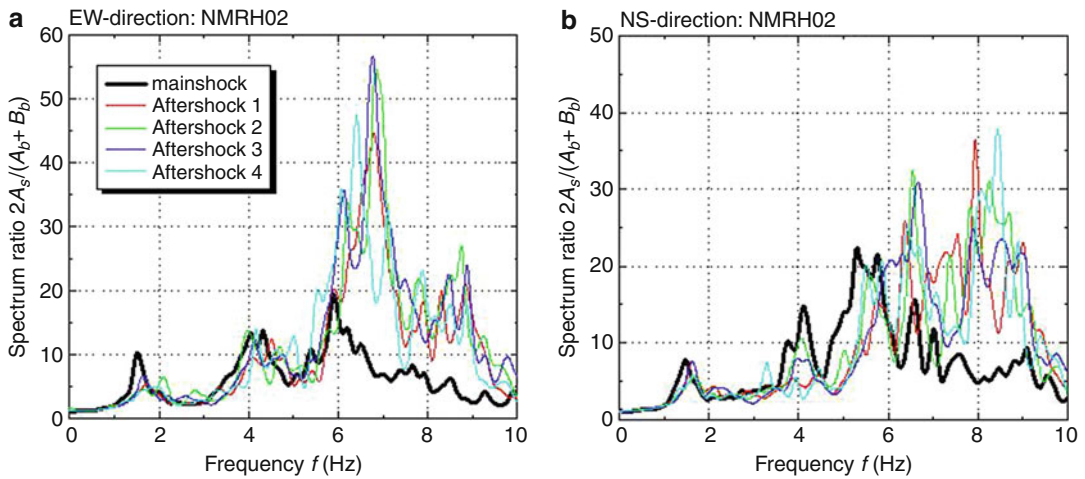
Site Response: Comparison Between Theory and Observation, Fig. 10 Fourier spectrum ratios $2A_s/(A_b + B_b)$ observed at a stiff soil site during main shock and aftershocks in directions EW (**a**) and NS (**b**)
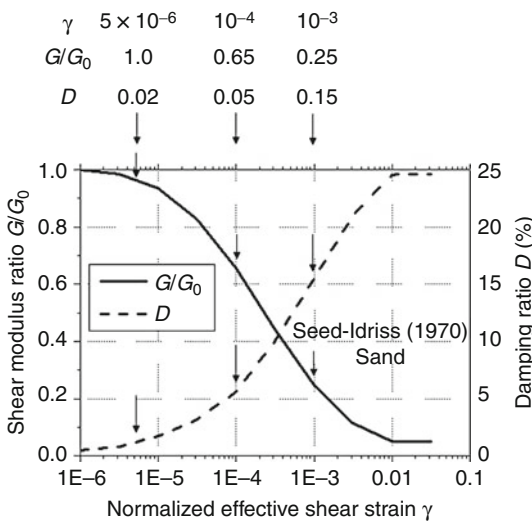
Site Response: Comparison Between Theory and Observation, Fig. 11 Soil nonlinearity curves (G/G_0-γ, D-γ) used in calculating amplifications in a two-layer system

recognized in actual site amplifications in Figs. 9 and 10 but also in $(2A_s)/(2A_b)$. However, the difference in the peak amplifications due to strain level is less pronounced in $(2A_s)/(2A_b)$ than in $2A_s/(A_b + B_b)$ for the first peak in particular. It is because the radiation damping effect presented by the impedance ratio α affects $(2A_s)/(2A_b)$ in Eq. 6, whereas no effect of α is involved in

$2A_s/(A_b + B_b)$ as indicated in Eq. 7. Under the paramount effect of radiation damping associated with α, the difference in the amplification $(2A_s)/(2A_b)$ due to strain-dependent nonlinear properties becomes less conspicuous. Furthermore, the impedance ratio, $\alpha = \rho_1 Vs_1/\rho_2 Vs_2$, which becomes smaller with degraded modulus or degraded S-wave velocity V_{s1} in the surface layer, tends to give larger amplification compensating the effect of increased damping ratio in the surface layer during strong earthquakes. Thus, the difference in soil nonlinearity between the main shock and aftershocks seems to have smaller influence on the amplification in $(2A_s)/(2A_b)$ than in $2A_s/(A_b + B_b)$ as indicated from the comparison of Fig. 12a, b.

The Fourier spectrum ratio corresponding to the transfer function $(2A_s)/(2A_b)$ cannot be obtained directly from downhole array records, but the peak value can be calculated from $2A_s/(A_b + B_b)$. Figure 13 shows an example how this calculation is carried out. First, a transfer function $2A_s/(A_b + B_b)$ is calculated at a downhole array site based on the multi-reflection theory. Among soil properties needed, the S-wave velocities of individual layers are given by in situ logging tests if strain-dependent soil nonlinearity is not so significant, and soil densities are judged from previous experiences.

S

Site Response: Comparison Between Theory and Observation, Fig. 12 Transfer functions of a two-layer system (**a**) $2A_s/(A_b + B_b)$ for downhole array and (**b**) $2A_s/2A_b$ for surface array in different induced strain levels
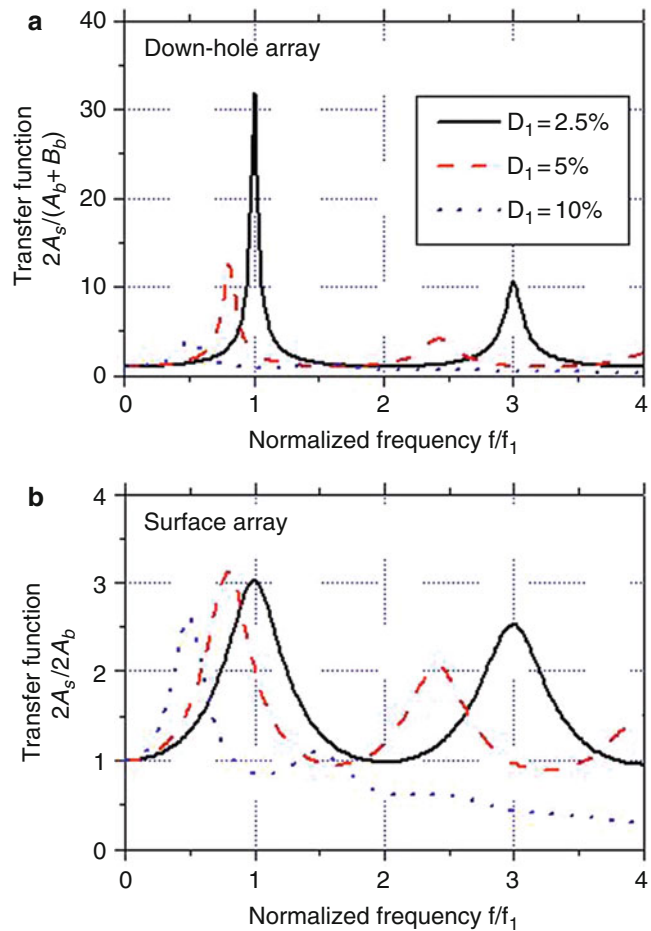
The damping ratio D is tentatively assumed 2.5 % in all layers and also postulated to be nonviscous or frequency independent.

Then, the calculated transfer function $2A_s/(A_b + B_b)$ is compared with the corresponding spectrum ratio obtained from downhole array records. If a peak in the transfer function can be found at about the same frequency in the spectrum ratio of recorded motions, it is identified as the corresponding peak, and the damping ratio assumed as $D = 2.5$ % previously is modified to have the same peak value, by using the equation $D = Q_1/Q_2 \times 2.5(\%)$, where Q_1 is the peak value of the calculated transfer function and Q_2 is that of spectrum ratio based on the actual records as indicated in Fig. 13. Not only the first peak but also the higher-order peaks are

compared in this manner if possible, and the values of D in the two directions, EW and NS, are averaged for individual peaks. Then, the transfer function $(2A_s)/(2A_b)$ is computed using the modified damping ratio D based on the same multilayer system. In this way, the peak amplifications for $(2A_s)/(2A_b)$ between surface and base, to be used in seismic zonation study for surface soil resting on the common base layer, can be obtained using the downhole array records.

In Fig. 14, the peak amplification values in the first peak of the spectrum ratios are compared between small aftershocks (in the horizontal axis) where shaking is relatively smaller (PGA ≤ 0.1 g) and main shock (in the vertical axis) with PGA $\approx 0.1 \sim 2.4$ g, based on vertical array records during strong earthquakes and

Site Response: Comparison Between Theory and Observation, Fig. 13 How to obtain spectrum peak $2A_s/2A_b$ from $2A_s/(A_b + B_b)$ of vertical array observed motions at the same site
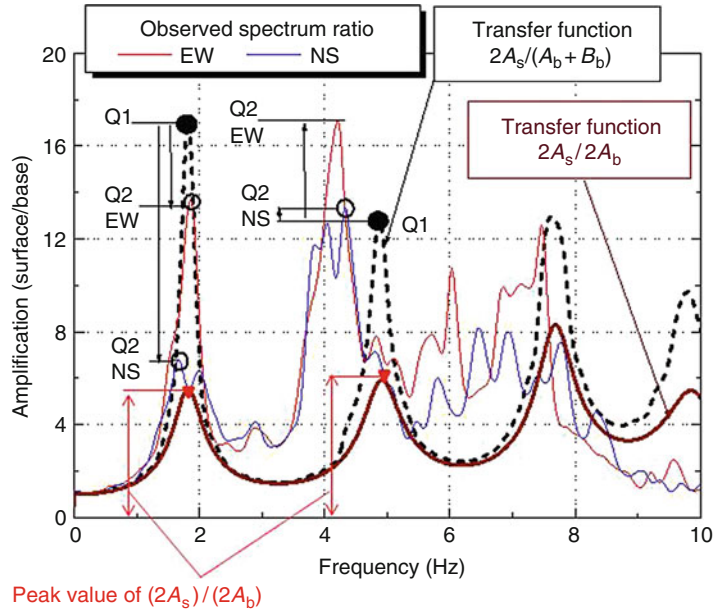
Site Response: Comparison Between Theory and Observation, Fig. 14 Comparison of spectrum peak amplifications between aftershocks and main shocks of downhole arrays $2A_s/(A_b + B_b)$ and surface arrays in $2A_s/2A_b$
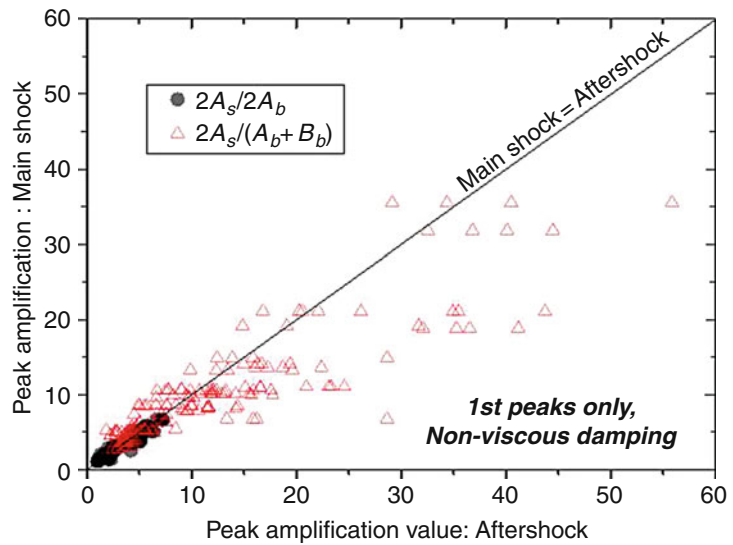
associated aftershocks recorded at many vertical array sites in Japan (Kokusho and Sato 2008; Kokusho 2013). The triangular symbols for downhole arrays $2A_s/(A_b + B_b)$ are dispersed in a wide range, and the majority is plotted around or below the diagonal line (main shock = aftershock), indicating that during strong shaking the site amplification in terms of $2A_s/(A_b + B_b)$

possibly reduces due to soil nonlinearity. In contrast, the solid circular symbols corresponding to $(2A_s)/(2A_b)$ in surface arrays, though the amplification values are small, concentrate near the diagonal line, indicating that the effect of soil nonlinearity is less dominant in the first peak site amplification, as demonstrated in Fig. 12 by using a simplified two-layer model.

Thus, the soil nonlinearity affects the site amplification considerably with respect to peak frequency and peak amplification. However, for the amplification of the first peaks in the surface array, the nonlinear effect in obviously minor. This may simplify the seismic zonation procedure considering strong shaking.

Summary

1. Two types of site amplification monitoring can be defined and implemented in earthquake observations, surface array and downhole array. For micro-zonation mapping, the amplification for the surface array is needed.
2. For reproduction of site amplification observed in the field, it is possible to simplify a complicated multilayer system to a two-layer system composed of a surface layer of a certain thickness underlain by a base layer of infinite thickness. Resonant frequencies of the multilayer system can roughly be evaluated using the 1/4 wavelength formula.
3. In the downhole array, the installation level of the downhole seismometer serves as a rigid boundary, and soil properties below have nothing to do with the amplification of the ground surface with respect to the base motion.
4. Among three soil damping models, Kelvin, Maxwell, and nonviscous, the nonviscous model is normally used for site amplification evaluation in engineering practice. In this case, the peak amplification of the transfer function becomes lower for higher-frequency peaks because the wave attenuation coefficient β is proportional to the frequency.
5. There are evidences from earthquake observations to suspect that some degree of frequency-dependent damping similar to Maxwell damping due to a wave-scattering mechanism exists in the field. However, it is normally ignored in geotechnical engineering, because hysteretic (nonviscous) nature of soil damping becomes more dominant during strong earthquakes as soil gets softer and more strain-dependent nonlinear.
6. Site amplification observations show that soil nonlinearity affects the site amplification considerably with respect to peak frequency and peak amplification. However, for the amplification of the first peak in the surface array, the nonlinear effect is obviously minor, which may simplify the seismic zonation procedure even for strong earthquakes.

Cross-References

▶ Downhole Seismometers
▶ Dynamic Soil Properties: In Situ Characterization Using Penetration Tests
▶ Site Response: 1-D Time Domain Analyses

References

Frankel A, Clayton RW (1986) Finite difference simulations of seismic scattering: implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity. J Geophys Res 91(B6):6465–6489

Hardin BO (1965) The nature of damping in sands. Proc Soil Mech Found Div, ASCE 91(SM1):63–97

Hardin BO, Drnevich VP (1972) Shear modulus and damping in soils: measurement and parameter effects. J Soil Mech Found Div, ASCE 98(SM6):603–624

Ishihara K (1996) The representation of stress–strain relations in cyclic loading, Chapter 3. In: Ishihara K - (ed) Soil behaviour in earthquake geotechnics, Oxford engineering science series, 46. Clarendon, Oxford, pp 16–39

Kanai K, Tanaka T, Yoshizawa S (1959) Comparative studies of earthquake motions on the ground and underground. Bulletin of the Earthquake Research Institute, Tokyo University, Vol. 37, 53–87.

Kokusho T (2013) Site amplification formula using average vs in equivalent surface layer based on vertical array strong motion records. ICEGE, Istanbul

Kokusho T, Sato K (2008) Surface-to-base amplification evaluated from KiK-net vertical array strong motion records. Soil Dyn Earthq Eng 28:707–716

Schnabel PB, Lysmer J, Seed HB (1972) SHAKE – a computer program for earthquake response analysis of horizontally layered sites, Report EERC 72-12. University of California, Berkeley

Seed HB, Idriss IM (1970) Soil Moduli and damping factors for dynamic response analysis, Report EERC 70-10. University of California, Berkeley

Wu R-S (1982) Attenuation of short period seismic waves due to scattering. Geophys Res Lett 9(1):9–12

Social Media Benefits and Risks in Earthquake Events

Hamish McLean
School of Humanities, Griffith University,
Nathan, QLD, Australia

Synonyms

Citizen seismologists; Disaster communication analysis; Earthquake education; Enhanced situational awareness; Post-disaster recovery; Psychological first aid; Real-time warnings on Twitter

Introduction

The rapidly evolving social media platforms, with an estimated 1.9 billion users worldwide, offer a myriad of communication benefits and risks in the context of a disaster. Social media generally refers to internet-based technologies that enable people to interact and share resources and information using either text or multimedia applications (Lindsay 2011; Dabner 2011). Advances in mobile devices allow access to anyone who has the ability to connect online (Abbasi et al. 2012). For example, the microblogging platform Twitter allows followers to track what an account holder is doing and thinking in real time within the confines of 140 characters (Kaigo 2012). Tweets can be sent from a variety of platforms ranging from cell phones to computers. Other examples of social media platforms today are Facebook, YouTube, Qzone, Pinterest, Instagram, and Flickr. This entry describes how social media, in particular Twitter, can be used for a variety of applications before, during, and after an earthquake, both inside and outside the area of impact. These include detection, warnings, connecting to survivors, situational awareness, notifying responders of where help is needed, and galvanizing humanitarian aid. Importantly, the increasing participation of "citizen seismologists" via social media is filling the information gap with field observations immediately after an earthquake (Young et al. 2013) for both first responders and survivors. In fact, social media, in particular Twitter, may be the only immediate source of data from locations with limited sensors or other scientific instruments. Obviously, not all earthquakes are reported on social media as many events are in remote areas or undersea and in countries with limited social media access or the magnitude is too small to be felt. A unique benefit of social media is that it is user generated – disaster agencies, seismologists, and other parties do not have to motivate citizens to tweet – they will do it anyway, potentially by the thousands in a significant earthquake. The challenge is how to transform the rapidly spreading flood of real-time information, some of it inaccurate, into reliable, useful, and valuable data. Part of the solution is to train Twitter users to tweet messages that can be more easily analyzed both manually and automatically using a crisis-specific syntax (Starbird and Stamberger 2010). This approach, among others discussed in this entry, will help guide disaster response, galvanize ongoing humanitarian efforts, and add value to the expanding body of earthquake sciences gathered since the development of the modern seismograph in the late 1800s.

Sourcing Information in Disasters

People facing a disaster seldom act on one source of information. Hunting for firsthand local information, they will "channel swap" across the traditional media, go online to news websites and social media, and contact family and friends. Without proof from a variety of trusted sources, people will believe there is no immediate threat or that the situation does not apply to them. The delay in taking action – known as milling – can increase the risk of death or injury from a hazard before steps are taken for self-protection. The milling effect is increased if warnings are vague and conflicting across various channels or the credibility of the source is in doubt. Under these circumstances, social media becomes a double-edged sword. The speed of social media can reduce the decision-making lag time while, at

the same time, increasing confusion and uncertainty if the information is wrong or conflicting. It is generally accepted that trust in social media information remains well below that of traditional media. The 2014 Edelman Trust Barometer found that 47 % of people trusted social media sources, while 65 % trusted the traditional media. In practice this was reflected in the 2011 Japan earthquake. Although most people turned to social media as their most reliable source of information, one third described the "lack of trust in formation as the greatest problem associated with social media use in the disaster" (Perry et al. 2012, p. 15). For disaster agencies, the challenge of earning trust and building credibility and relationships can be overcome by actively engaging online before the disaster. As noted by the US-based Centers for Disease Control and Prevention (CDC): "Organisations need to be regular users of social media before the crisis. If not, social media users will go to other sources and groups with whom they already have relationships for information" (2012, p. 268). Engaging with social media before a disaster can also build preparedness and education for survival. For example, more than 1.3 million people in New Zealand took part in the "ShakeOut" earthquake drill in 2012, driven by a social media campaign. Many organizations took the drill realistically. For example, a hospital in the country's North Island shut off its power and water for 24 h to simulate the impact of an earthquake. Portable cooking, lighting, and other equipment maintained hospital services (Civil Defence Emergency Management 2012). Similar events are held in other countries to prepare communities for an earthquake. Although disaster agencies are using social media to educate communities and disseminate information, they have been historically more reluctant to rely on social media data in their incident command systems. The reasons include fears about misinformation (Williams et al. 2012), the speed and spread of the information which makes validation difficult (Gowing 2009), and a lack of understanding about how they can make use of social media (Duffy 2012; Palen 2008). For example, Tapia et al. (2013) argue that a major issue for

disaster agencies is how social media data can be effectively incorporated into time-critical decision-making processes. As Tapia et al. (2013, p. 770) note, "while data quality continues to be a barrier, what is far more important to organizational use is the serving of this data at the appropriate time, in the appropriate form to the appropriate person and the appropriate level of confidence."

Online Support on the Ground

Driven by developing mobile technology, social media is taking on an increasing role in connecting people in disasters. Part of the reason is that people reach out to both their "offline" and "online" communities during a crisis or disaster (Dutta-Bergman 2006). In other words, they parallel their physical world with their virtual world to garner "social support and gather information, and vice versa." In turn, on a much wider scale, this online convergence builds and strengthens community resilience through "people power" (Duffy 2012). A bank of social capital is developed by the exchange of information "during difficult times" to build relationships between people (Kaigo 2012). Therefore, social media in disaster impact areas is frequently driven by the community to share knowledge and as a form of empowerment toward recovery. The CDC (2012) notes that the public uses social media on a greater scale in the hours after a disaster than official agencies. Williams et al. (2012), who have produced a practical guide to community-based social media in disasters based on lessons from a series of tornadoes and floods in the United States, found that post-disaster social media was generated and driven by citizens rather than emergency agencies. In most cases, community-managed social media was the primary source of information for those impacted by the disaster. For example, a University of Missouri Extension Facebook site – Branson Tornado Info – attracted 14,000 followers within 12 h of a tornado in February 2012. One victim posted: "For the first few days after the storm, this Facebook page was our main source of

Social Media Benefits and Risks in Earthquake Events, Table 1 The expanding role of social media in earthquake events

2008 Sichuan (China) earthquake	The first alert of the severity of the earthquake was reported on Twitter (Moore 2008; Cellan-Jones 2008)
2010 Haiti earthquake	First major earthquake in the age of social media. The purpose-built Ushahidi crisis map became an emergency reporting system to locate people texting for help. Social media-generated significant humanitarian aid (Meier 2012; MacLeod 2010)
2011 Christchurch (NZ) earthquake	Social media became "the Church or meeting hall" for people to support each other and share information (Mathewson 2012)
2011 Japan earthquake and tsunami	Social media platforms, particularly Facebook, Twitter, and Japan's own social media site Mixi, provided information and connected families when telephone communication was damaged or became congested (Wallop 2011)
2011 Turkey earthquake	Facebook was used to coordinate donations and aid requests. People trapped in rubble tweeted for help (Turgut 2011)
2013 Ya'an (China) earthquake	Social media platforms become a place for mourning when people turned their profile photographs to gray in remembrance of the victims (Hui 2013)

information. Volunteers here answered our questions about where to go to get help, what resources were available and what we needed to do next" (Williams et al. 2012, p. 18). In the context of earthquakes, social media has played an important and increasingly sophisticated role, as the following Table 1 demonstrates:

Twitter Faster than Earthquakes

It is well documented that the real-time speed and user-input microblogging capability of Twitter make it one of the most useful social media tools for disaster management agencies to gain a rapid snapshot of the earthquake aftermath and

the level of response and resources that may be required. Earthquake alerts are now within the scope of Twitter. Crooks et al. (2013) found that the velocity of Twitter can be used as a warning system in large-scale events. Tweets and re-tweets spread from the epicenter faster than the physical effects to distant locations. For example, tweets about the 2011 Virginia earthquake were read in New York 30 s before it was felt there, "showing that information moves faster through networks than the earthquakes themselves" (Perry et al. 2012, p. 6). Automatic earthquake warnings sourced from Twitter data are being developed using algorithms to analyze keywords, the number of words, and their context (Sakaki et al. 2010). Importantly, Tweets can include locations, which is fundamental in sensing earthquake events (Sakaki et al. 2010). Another real-time online method of detecting earthquakes in a general location is to trace the IP addresses of visitors to earthquake information websites. For example, surges in visitor traffic to the European-Mediterranean Seismological Centre from people wanting information about what they had felt can provide a snapshot within 2 min into the location and potential damage from an earthquake (Bossu et al. 2011).

Greater Situational Awareness

Tweets within the first minutes of an earthquake are mostly generated from around the epicenter and provide potentially useful situational awareness for both emergency responders, seismologists, and, importantly, survivors. In the 2011 Christchurch, New Zealand, earthquake, the first tweet was within 30 s, the first photo was within 4 min, and the first video was uploaded to YouTube in 40 min. It is widely accepted that information such as this supplements, rather than replaces, data sourced from scientific instruments. As Earle et al. (2011, p 709) note: "The qualitative descriptions contained in the tweets are available at the same time as the seismically-derived earthquake parameters and sometimes provide a responding seismologist with a quick indication of the severity of the earthquake

effects." Although Twitter's 140 characters may be limiting, the narratives often produce a consensus on the intensity of the earthquake because "citizens tend to report very similar experiences" (Young et al. 2013, p. 19). Hashtags are quickly formulated to spread information. For example, within minutes of the 2011 Christchurch, New Zealand, earthquake, the hashtags #eqnz and #chch helped to share images and videos of the damage (Edmond 2013). Within 2 min of a tremor in Victoria, Australia, more than 100 tweets were posted, giving an indication of strength and reports of minor damage (Anderson 2012). In the 2011 Japan earthquake, Twitter was more effective in providing information in devastated areas than traditional media and websites (Kaigo 2012). It has led the Japanese authorities to consider making social media networks part of the country's emergency call system (Dugan 2012). Scanning other social media platforms can also strengthen situational awareness. There are a myriad of social media platforms that can convey video, photographs, audio, and written accounts of the earthquake event. For example, the following chart outlines key benefits of Twitter, YouTube, Facebook, and blogs (Table 2).

Social Media Clutter, Misinformation, and Rumors

Social media is flooded with information in a disaster event. For example, an estimated 2.3 million tweets mentioned Haiti or the Red Cross in the 24 h following the 2010 Haiti earthquake. In Japan, more than 2000 tweets were posted every second of the day following the 2011 earthquake and tsunami (Meier 2013). Significant numbers of tweets are re-tweeted, adding to the social media "clutter" and potential spread of misinformation in the aftermath of a disaster. To make sense of the situation and to gather reliable data, a number of social mapping and analysis projects are under development. They include machine-learning approaches for classifying and extracting "informative" Twitter messages to augment situational awareness (Imran et al. 2013)

Social Media Benefits and Risks in Earthquake Events, Table 2 The benefits of various social media platforms for communicating during a disaster

Twitter (140 characters)	Instant messaging One-to-many receivers (followers) Monitoring first impressions of the shaking in real time Issuing warnings and alerts rapidly Integrating mass or interpersonal communication Initiating situational awareness from the field Channels to dispel rumors and correct information Interactive mapping Rapid updating of traditional media Alternative method of seeking aid for survivors Linking to more detailed information on other platforms Ability to "snowball" information by re-tweeting
YouTube (and others)	Broadcasting live vision from a location Providing a channel to group videos Upload vision/audio for traditional and online media consumption Facilitating updates
Facebook	Offers two-way communication Connecting people inside and outside the disaster area Providing more information than Twitter Immediate updates Delivering vision/audio/images Linking to other agencies and sources of information Accessible to anyone with an email address
Blogs	Rapid updates Allows discussion Space for more information Linking to other agencies and sources of information Providing opportunities to share stories and experiences

and volunteers, such as micromappers.com, to quickly filter social media data during a disaster using apps to tag tweets and photographs. Advances in technology are providing dividends. Another project, the Artificial Intelligence for Disaster Response, reports that 40–80 % of tweets containing disaster information can be detected automatically, with an 80–90 %

accuracy rate on whether the tweet was from an eyewitness. To obtain more structured firsthand accounts from social media users, crowdsourcing approaches are utilized by dedicated not-for-profit and government-based earthquake reporting and information sites. For example, the US Geological Survey actively seeks contributions through its "Did You Feel It" project by "asking people where they were, what they observed and what they experienced during the earthquake" (Young et al. 2013, p. 2). Earthquake-Report.com, which describes itself as the "best independent earthquake reporting site in the world," utilizes multiple social media platforms to share firsthand accounts of earthquakes in real time along with merging data from scientific sources. Crowdsourcing questions from the impact area include location, scale of intensity, and a brief description of the experience, including damage.

Squashing Viral Rumors and False Information

People power – termed the Wikipedia effect – is a self-correcting social media. Many of the rumors and misinformation in a disaster are identified and corrected by social media users. Traditional journalists, unable to compete with the speed of social media, have assumed the mantle of "fact-checkers" to validate information. For example, fake information and images distributed on Twitter during Hurricane Sandy in 2012 were quickly ousted by other social media users and traditional media outlets. Although traditional media outlets have published inaccurate information sourced from social media, it is often quickly corrected. Twitter and Facebook were used extensively by the Queensland Police Service during the South East Queensland, Australia, floods in 2011 to identify rumors and respond with factual information. Those platforms became an important resource for traditional media. As Bruns et al. (2012, p. 8) observed: "Additionally, @QPSMedia also played a crucial role in enabling affected locals and more distant onlookers to begin the difficult process of making

sense and coming to terms with these events, even while they were still unfolding."

Psychological Support for Survivors

Social media plays a key role in the earthquake recovery phase. Unlike the limitations of the one-way traditional media model, social media's two-way interaction helps rebuild communities and bring together families, friends, and neighbors. Importantly, social media provides "psychological first aid" where people "reported feeling a sense of connectedness and usefulness, felt supported by others and felt encouraged by the help and support being given to people" (Taylor et al. 2012, p. 25). In the weeks and months following the 2011 Christchurch earthquake, Dabner (2011) found that online discussion provided support and information, with one participant describing it as a lifeline "that helped her (and therefore her children) cope with aftershocks by realizing normally (sic) would eventually return" (2011, p. 10). Due to the level of destruction, the role of churches was resumed through social media, with one researcher observing: "Social media was really a way for people to feel like they weren't being forgotten or like they were part of a larger community. As far as someone sitting at home alone at 10 pm, they were not able to go out for a cuppa. That's where social media really kicked in" (Chapman-Smith 2012).

Summary

Social media continues to rapidly evolve as a useful tool in earthquake communication. Developing technology will increase the accuracy of information from "citizen seismologists" to enhance situational awareness, improve warnings, coordinate aid and recovery, and galvanize humanitarian relief efforts. Impacted communities will increasingly turn to social media as a way to communicate lifesaving information, gather support, and empower each other in the recovery process.

Cross-References

References

Abbasi M, Kumar S, Filho A, Liu H (2012) Lesson learned in using social media for disaster relief – ASU Crisis Response Game. http://www.public.asu.edu/~huanliu/papers/SBP12Game.pdf

Anderson M (2012) Earthquake social media response. http://www.cfaconnect.net.au/news/earthquake-social-media-response.html

Bossu R, Gilles S, Mazet-Roux G, Roussel F, Kamb L (2011) Flash sourcing, or rapid detection and characterization of earthquake effects through website traffic analysis. Ann Geophys 54(6):716–727

Bruns A, Burgess J, Crawford K, Shaw F (2012) #qldfloods and @QPSMedia: Crisis Communication on Twitter in the 2011 South East Queensland Floods. Media Ecologies Project. ARC Centre of Excellence for Creative Industries & Innovation, Brisbane, Australia

Cellan-Jones R (2008) Twitter and the China earthquake, BBC blog, 12 May 2008. http://www.bbc.co.uk/blogs/legacy/technology/2008/05/twitter_and_the_china_earthqua.html

Centres for Disease Control (2012) Crisis and emergency communication. http://emergency.cdc.gov/cerc/pdf/CERC_2012edition.pdf

Chapman-Smith B (2012) Social media took 'church' role in quakes. http://www.nzherald.co.nz/technology/news/article.cfm?buffer_share=a6e0e&c_id=5&objectid=10828401

Civil Defence Emergency Management (2012) New Zealand ShakeOut, more than 1.3 million involved. http://www.scoop.co.nz/stories/AK1209/S00584/new-zealand-shakeout-more-than-13-million-involved.htm

Crooks A, Croitoru A, Stefanidis A, Radzikowski K (2013) #Earthquake: Twitter as a distributed sensor system. Trans GIS 17(1):124–147

Dabner N (2011) Breaking ground in the use of social media: a case study of University earthquake response to inform education design with Facebook. www.educ.utas.edu.au/users/afluck/.../wg332011_submission_12.doc

Duffy N (2012) Using social media to build community resilience. Aust J Emerg Manag 27(1):40–45

Dugan L (2012) Would you place a 911 call on Twitter? http://www.mediabistro.com/alltwitter/911-call-on-twitter_b27752

Dutta-Bergman MJ (2006) Community participation and Internet use after September 11: complementarity in channel consumption. J Comput Mediat Commun 11(2):659–673

Earle PS, Bowden D, Guy M (2011) Twitter earthquake detection: earthquake monitoring in a social world. Ann Geophys 54(6):708–715

Edmond S (2013) Social media pre, during and post-earthquake. http://blog.potentia.co.nz/social-media-quake/

Gowing N (2009) Skyful of lies and black swans. The new tyranny of shifting information power in crises. The Reuters Institute of Technology, Oxford

Hui N (2013) Social media's role in Ya'an earthquake aftermath is revealing. http://www.chinafile.com/social-medias-role-yaan-earthquake-aftermath-revealing

Imran M, Elbassuoni S, Castillo C, Diaz F, Meier P (2013) Extracting information nuggets from disaster-related messages in social media. In: Proceedings of the 10th international ISCRAM conference, Baden

Kaigo M (2012) Social media usage during disasters and social capital: Twitter and the Great East Japan Earthquake. Keio Commun Rev 34:19–35

Lindsay BR (2011) Social media and disasters: current uses, future options, and policy considerations. Congressional Research Services Report 7-5700. http://fpc.state.gov/documents/organization/174191.pdf

MacLeod L (2010) New media vital in breaking Haiti earthquake story. http://www.bbc.co.uk/worldservice/worldagenda/2010/01/100122_worldagenda_haiti_monitoring.shtml

Mathewson N (2012) Social media excel after quakes. http://www.stuff.co.nz/the-press/news/christchurch-earthquake-2011/7517088/Social-media-excel-after-quakes

Meier P (2012) How crisis mapping saved lives in Haiti. http://voices.nationalgeographic.com/2012/07/02/crisis-mapping-haiti/

Meier P (2013) Automatically extracting disaster-relevant information. http://irevolution.net/2013/04/01/auto-extracting-disaster-info/

Moore M (2008) China earthquake brings out citizen journalists. http://www.telegraph.co.uk/news/worldnews/asia/china/1950212/China-earthquake-brings-out-citizen-journalists.html

Palen L (2008) Online social media in crisis events. EDUCAUSE Q 31(3):76–78

Perry B, Shaw R, Takeuchi Y (2012) Utilisation of social media in the East Japan earthquake and tsunami and its effectiveness. J Nat Disaster Sci 34(1):3–18

Sakaki T, Okazaki M, Matsuo Y (2010) Earthquake shakes Twitter users: real-time event detection by social sensors. In: Proceedings of the 19th international WWW conference (WWW2010), Raleigh

Starbird K, Stamberger J (2010) Tweak the Tweet: leveraging microblogging proliferation with a prescriptive syntax to support citizen reporting. In: Proceedings of the 7th international ISCRAM conference, Seattle

Tapia A, Moore C, Johnson N (2013) Beyond the trustworthy Tweet: a deeper understanding of microblogged

data use by disaster response and humanitarian relief organisations. In: Proceedings of the 10th international ISCRAM conference, Baden

Taylor M, Well G, Howell G, Raphael B (2012) The role of social media as psychological first aid as a support to community resilience building. A Facebook study for Cyclone Yasi Update. Aust J Emerg Manag 27(1):20–26

Turgut P (2011) Turkey's earthquake: social media to the rescue. http://world.time.com/2011/10/24/turkeys-earthquake-social-media-to-the-rescue/

Wallop H (2011) Japan earthquake: how Twitter and Facebook helped. http://www.telegraph.co.uk/technology/twitter/8379101/Japan-earthquake-how-Twitter-and-Facebook-helped.html

Williams R, Williams G, Burton D (2012) The use of social media in disaster recovery. http://extension.missouri.edu/greene/documents/PlansReports/using%20social%20media%20in%20disasters.pdf

Young JC, Wald D, Earle P, Shanley L (2013) Transforming earthquake detection and science through citizen seismology. Woodrow Wilson International Centre for Scholars, Washington, DC

Soil-Structure Interaction

Christos Vrettos
Division of Soil Mechanics and Foundation Engineering, Technical University of Kaiserslautern, Kaiserslautern, Germany

Synonyms

Dynamic soil-structure interaction; Seismic soil-structure interaction; SSI

Introduction

Seismic waves propagating through the soil impinge upon structures founded on the soil surface or embedded into it. Displacements are then produced both in the structure and in the soil. The mutual dependency of the displacements is called soil-structure interaction, abbreviated as SSI. Consequently, the motion occurring at the base of the structure is different compared to the free-field motion (motion in the absence of the structure). Soil-structure interaction characteristics depend on several factors:

– Intensity, wavelength, and angle of incidence of the seismic waves
– Soil stratigraphy
– Stiffness and hysteretic damping of the particular soil layers
– Geometry and rigidity of the foundation
– Embedment depth of the structure
– Inertia characteristics, slenderness, and natural vibration period (eigenperiod) of the superstructure
– Presence of nearby structures

Various effects are associated to this phenomenon:

– A building founded on compliant ground has different vibrational characteristics, for example, higher natural period compared to the same building on rigid base (solid rock). The softer the soil, the larger the difference.
– A part of the vibrational energy emanating from the compliant structure foundation is transmitted into the surrounding soil through wave radiation in the unbounded soil medium and hysteretic energy dissipation. Such effects do not occur in a rigidly supported structure.
– Due to the compliance of the foundation, the motion at the foundation base contains rocking and torsional components in addition to the translational components.

The mechanisms governing soil-structure interaction can be divided into two distinct interactions: inertial and kinematic interaction.

Kinematic interaction is the deviation of the soil response from the free-field motion due to the resistance of the stiffer foundation to conform to the distortions of the soil imposed by the traveling seismic waves. It is commonly expressed in terms of frequency-dependent transfer functions relating the disturbed motion at the interface foundation/soil to the free-field motion.

Inertial interaction arises as the structure responds to the soil motion induced by kinematic interaction at the foundation level. Inertial forces

S

are developed in the structure being transmitted to the compliant soil. Frequency-dependent impedance functions are used to represent the stiffness of the foundation/soil system and the associated radiation damping.

The relative impact of each contribution is a function of the characteristics of the incoming waves, the foundation geometry and rigidity, and the soil conditions.

The analysis is particularly challenging due to the semi-infinite extent of the soil medium, the nonlinearity of the soil behavior, the inherent variability of the soil stratigraphy, and the dependency of the response on frequency. Several procedures of different degrees of complexity have been proposed during the past five decades. A historical overview is given by Kausel (2010). The book by Wolf (1985) provides a rigorous and comprehensive treatment of the topic including applications to seismic problems.

Even with the computational facilities available today, such analyses are associated with a major effort, both for modeling the soil-structure system and carrying out the calculation. In particular during the early design stage, parametric studies are necessary in order to assess the influence of the various parameters and optimize the system for purposes of cost estimation. This necessitates the application of simplified methods that capture the essential features of the system response. The next sections provide a brief overview with emphasis on such simplified methods.

Soil-Foundation-Structure Analysis Models

Two general approaches are commonly used for the analysis of soil-structure interaction problems.

Direct Approach
The soil and the structure are treated together in a combined analysis by modeling them using finite elements or finite differences in two or three dimensions. This offers the advantage that inelastic behavior, particularly for the soil, can be taken

into account by the step-by-step numerical integration of the equations of motion within a time-domain algorithm. A drawback is the necessity to specify the input motion at the base of the model, where it is not known a priori. Since the design seismic motion is usually given at the free surface or at outcropping rock, a deconvolution is necessary to obtain the compatible bedrock motion. Often the bedrock is located at large depths thus prohibiting the modeling of the entire soil layer, and some artificial boundary is defined at a shallower depth. The deconvolution then involves an iterative procedure. For convenience, the deconvolution is often carried out using algorithms that are based on 1D vertical shear wave propagation, thus requiring an adjustment of the model parameters in order to achieve compatible solutions between 1D and 2D analyses. Attention is further required in the selection of appropriate boundary conditions at the side boundaries of the discretized domain to avoid spurious reflections that would contaminate the results. The composite soil-structure model is finally subjected to the previously determined base rock motion, and the evolution in time of displacement and stresses is computed.

Substructure Approach
The underlying calculation method comprises three steps. At first the seismic motion acting at the foundation level is determined assuming a rigid but massless foundation. This is referred to as foundation input motion (FIM), and for an embedded structure, it will include both translational and rotational components. In the second step, the complex-valued frequency-dependent impedances for the foundation/soil system are determined. The real part of the impedance function represents a linear spring and the imaginary part, a viscous dashpot accounting for the energy radiation into the soil medium. Finally, the structure supported by the frequency-dependent springs and dashpots is subjected to the foundation input motion computed in the first analysis step.

While impedance functions are sufficient for rigid foundations, distributed springs and dashpots placed around the foundation are used for

nonrigid embedded foundations when the distribution of sectional forces is sought. In this case, due to the vertical variation of ground motion, the imposed differential ground displacements vary over the height of the basement walls.

The validity of this approach – often called superposition theorem – is shown by Kausel and Roesset (1974). The main advantage of the method is that each step can be handled independently and with different algorithms. Further, it allows an insight into the contributions from each analysis step and is particularly suitable for parametric studies.

The application of the principle of superposition requires linear behavior. Inelastic behavior is implemented using equivalent linearization by selecting soil modulus and radiation damping to correspond to the likely effective strain level the soil will experience under the specific loading. This is achieved by means of an iterative procedure. Superposition is shown to be a reasonable approximation even when inertia forces produce large strains in the vicinity of the foundation, since shear strains due to kinematic interaction effects are usually significant in deeper soil regions.

Inertial Interaction

Shallow Foundations

The illustration of the concepts is made on the basis of a simple structure-soil system being composed by a linear structure of height h, mass m, lateral stiffness k, and damping ratio β_{str} that is connected to a rigid foundation of radius r resting on the surface of a homogeneous elastic half-space. The half-space is used to represent the unbounded soil medium and is characterized by its shear modulus G, Poisson's ratio v, and mass density ρ. Mass and moment of inertia of the foundation are neglected for simplification. The compliance of the soil is modeled by two frequency-dependent springs placed underneath the rigid foundation: a horizontal translation spring of stiffness K_x and a rotational spring of stiffness K_θ. Energy dissipation in the soil due to friction within the material (hysteretic damping)

and wave radiation in the unbounded medium is modeled by a pair of frequency-dependent dashpots with coefficients C_x and C_θ attached parallel to the respective springs. This model may be viewed as a single- or multistory building after an appropriate reduction of the degrees of freedom.

Springs and dashpots for each degree of freedom j can be condensed to complex-valued impedances that are expressed in two equivalent forms:

$$\tilde{K} = K_j + i\omega C_j = K_j\big(1 + i2\beta_j\big) \qquad (1)$$

where ω is the circular frequency of the excitation, i is the imaginary unit, and β_j is a damping coefficient that is related to the viscous dashpot coefficient of a simple oscillator by

$$\beta_j(\omega) = \frac{\mathrm{Im}\big(\tilde{K}_j\big)}{2\mathrm{Re}\big(\tilde{K}_j\big)} = \frac{\omega C_j}{2K_j} \qquad (2)$$

The use of β_j has the advantage that at resonance of the compliant system β_j corresponds to the percentage of critical damping.

The undamped natural vibration period of the structure in its fixed-base condition is

$$T = 2\pi\sqrt{\frac{k}{m}} \qquad (3)$$

For the case of a compliant base, it can be shown that the respective natural period is (Veletsos and Meek 1974)

$$\tilde{T} = T\sqrt{1 + \frac{k}{K_x} + \frac{kh^2}{K_\theta}} \qquad (4)$$

Hence, the period of the flexibly supported structure is higher than that on rigid base. Since the spring stiffnesses are in general frequency dependent, an iterative procedure is necessary to evaluate the period $\tilde{T}$. A reasonable approximation consists in using the spring values corresponding to the fixed-base natural period, and even simpler is to use the static values of the springs.
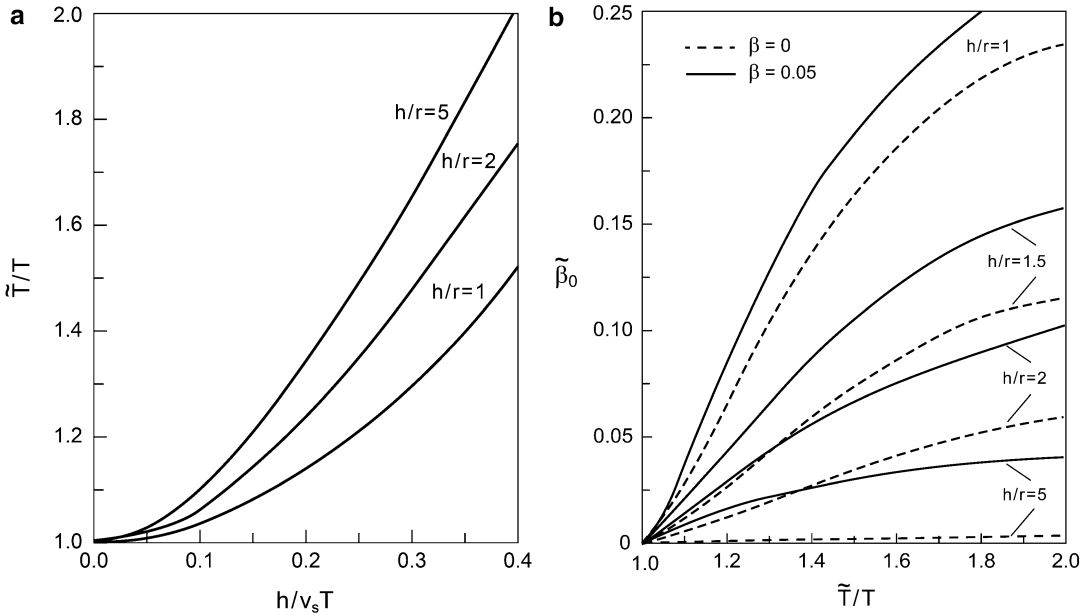
Soil-Structure Interaction, Fig. 1 Effects of soil-structure interaction: (**a**) elongation of natural period in dependency on the ratio of structure-to-soil stiffness $\bar{s}$; (**b**) increase in effective damping in dependency on the natural period ratio $\tilde{T}/T$. Curves are for mass ratio $\bar{m} = 0.15$ and different values of the slenderness ratio $\bar{h}$. Poisson's ratio $\nu = 0.45$ (Adapted from Veletsos (1977). In: Hall WJ (ed) Structural and geotechnical mechanics, 1st edn, © 1977. Reprinted by permission of Pearson, Inc., Upper Saddle River, NJ)

The dimensionless parameters controlling the period lengthening are

$$\text{Stiffness ratio structure-to-soil} \quad \bar{s} = \frac{h}{T v_s} \quad (5)$$

$$\text{Slenderness ratio} \quad \bar{h} = \frac{h}{r} \quad (6)$$

$$\text{Mass ratio} \quad \bar{m} = \frac{m}{\rho \pi r^2 h} \quad (7)$$

with

$$v_s = \sqrt{G/\rho} \quad (8)$$

denoting the shear wave velocity in the soil.

The stiffness ratio will be larger for stiff structural systems such as shear walls and smaller for flexible systems such as moment frames. For soil and weathered rock sites, this term is typically smaller than 0.1 for flexible systems such as moment frames and between approximately 0.1 and 0.5 for stiff systems such as shear wall and braced frame structures. The period lengthening variation with the stiffness ratio is shown in Fig. 1a for typical values of the parameters involved.

For the overall effective damping ratio of the system, several approaches have been proposed in the literature differing in the degree of approximation involved. Usually products of damping ratios are neglected as higher-order terms. The most widespread among these solutions – that also entered design codes – is that derived by Veletsos and Meek (1974). Assuming structural damping of viscous nature, the overall effective damping becomes

$$\tilde{\beta} = \beta_0 + \frac{\beta_{\text{str}}}{\left(\tilde{T}/T\right)^3} \quad (9)$$

where β_0 represents the contribution from the soil-structure interaction – being referred to as foundation damping – that includes both material

and radiation damping (Veletsos 1977). The respective expression is written here in the more general form

$$\beta_0 = \left(\frac{T}{\tilde{T}}\right)^3 \left| k \frac{\tilde{T}}{T} \left(\frac{\beta_x}{\tilde{K}_x} + \frac{\beta_\theta}{\tilde{K}_\theta} h^2\right) \right| \quad (10)$$

From Eq. 9 it is evident that the effectiveness of the structural damping is reduced by soil-structure interaction as the period ratio $\tilde{T}/T$ increases. This may lead to a decrease in overall damping unless this reduction is compensated by the increase in the foundation damping. In practice, effective damping is taken higher than the structural damping, the value 5 % used in the development of design provisions being considered as a lower bound. Figure 1b shows the significant increase of the foundation damping with decreasing slenderness ratio h/r: rocking motion that is characterized by small radiation damping dominates the response of slender structures, whereas for squat structures the prevailing motion is horizontal translation that radiates energy into the soil more efficiently.

Observations based on data from instrumented buildings confirmed the analytical findings. For the majority of structures, the stiffness ratio h/Tv_s will be less than 0.5 and the mass ratio will range between 0.1 and 0.2 with a typical average of 0.15 (Stewart et al. 2003). The case studies analyzed revealed that the governing parameter for inertial interaction is the stiffness ratio and that these effects can be neglected for values less than 0.1.

Impedances for Shallow Foundations

Frequency-dependent springs and dashpots for shallow foundations have been determined in the last decades by several authors for different geometries and soil stratigraphies. In most cases radiation damping is expressed in terms of the dashpot coefficient C_j, as given in Eq. 1. The stiffness K_j at zero frequency is referred to as the static foundation stiffness and is denoted by K_j^0. The effects of frequency on the spring values for the particular vibrational mode j are then given by stiffness modifiers such that

$$K_j = K_j^0 \; k_j \quad (11)$$

Exact closed-form solutions are available only for perfectly rigid circular foundations and relaxed boundary conditions at the soil-foundation interface, i.e., normal stresses are neglected for swaying and shear stresses for rocking. These solutions are

Horizontal translation

$$K_x^0 = \frac{8}{2-v} Gr \quad (12)$$

Rocking

$$K_\theta^0 = \frac{8}{3(1-v)} Gr^3 \quad (13)$$

These expressions may be used for square foundations – and also for rectangular foundations with aspect ratio less than 3 – by replacing the radius by an equivalent value that yields the same footprint area for swaying and equal moments of inertia for rocking.

A review of available solutions for foundation impedances is presented by Pais and Kausel (1988) and Gazetas (1991) and the update by Mylonakis et al. (2006). Approximate expressions and graphs are compiled for various configurations and for all six modes of vibration. They include static values for rectangular foundations, stiffness modifiers, and expressions for the radiation damping.

We restrict here the presentation of results to swaying and rocking motion for rectangular foundations with footprint area $2a \times 2b$ with $a \geq b$ with the x-axis running parallel to the longer foundation side. The subscripts θ_x and θ_y in the impedances indicate rotation around the x- and y-axis, respectively. The weak coupling between translational and rocking mode is neglected. The frequency dependency is captured by the dimensionless parameter

$$a_0 = \frac{\omega b}{v_s} \quad (14)$$

and the foundation aspect ratio is denoted by

$$\ell = \frac{a}{b} \geq 1 \qquad (15)$$

The approximate expressions obtained by Pais and Kausel (1988) are displayed in the following.

Surface Foundations
The static solutions are:

$$\text{Swaying } K_x^0 = \frac{Gb}{2-v}\left[6.8\ell^{0.65} + 2.4\right] \qquad (16)$$

$$\text{Swaying } K_y^0 = \frac{Gb}{2-v}\left[6.8\ell^{0.65} + 0.8\ell + 1.6\right] \qquad (17)$$

$$\text{Rocking } K_{\theta_x}^0 = \frac{Gb^3}{1-v}\left[3.2\ell + 0.8\right] \qquad (18)$$

$$\text{Rocking } K_{\theta_y}^0 = \frac{Gb^3}{1-v}\left[3.73\ell^{2.4} + 0.27\right] \qquad (19)$$

The frequency-dependent stiffness modifiers are:

$$\text{Swaying } k_x = 1 \qquad (20)$$

$$\text{Swaying } k_y = 1 \qquad (21)$$

$$\text{Rocking } k_{\theta_x} = 1 - \left[\frac{0.55a_0^2}{0.6 + \dfrac{1.4}{\ell^3} + a_0^2}\right] \qquad (22)$$

$$\text{Rocking } k_{\theta_y} = 1 - \left[\frac{(0.55 + 0.01\sqrt{\ell})a_0^2}{2.4 - \dfrac{0.4}{\ell^3} + a_0^2}\right] \qquad (23)$$

The viscous damping coefficients accounting for radiation damping as determined from the dashpot coefficients using Eq. 2 are:

$$\text{Swaying } \beta_x = \left[\frac{4\ell}{K_x^0/Gb}\right]\left[\frac{a_0}{2k_x}\right] \qquad (24)$$

$$\text{Swaying } \beta_y = \left[\frac{4\ell}{K_y^0/Gb}\right]\left[\frac{a_0}{2k_y}\right] \qquad (25)$$

$$\text{Rocking } \beta_{\theta_x} = \left[\frac{(4/3)\bar{v}\ell a_0^2}{\left(K_{\theta_x}^0/Gb^3\right)\left[\left(2.2 - \dfrac{0.4}{\ell^3}\right) + a_0^2\right]}\right]\left[\frac{a_0}{2k_{\theta_x}}\right] \qquad (26)$$

$$\text{Rocking } \beta_{\theta_y} = \left[\frac{(4/3)\bar{v}\ell^3 a_0^2}{\left(K_{\theta_y}^0/Gb^3\right)\left[\left(\dfrac{1.8}{1 + 1.75(\ell-1)}\right) + a_0^2\right]}\right]\left[\frac{a_0}{2k_{\theta_y}}\right] \qquad (27)$$

where

$$\bar{v} = \sqrt{2(1-v)/(1-2v)} \qquad \bar{v} \leq 2.5 \qquad (28)$$

is the ratio of compressional wave velocity to shear wave velocity in the soil.

It should be mentioned that the exact curves for the stiffness modifiers and the damping factors have in general a smooth wavy form, and the expressions given above consist approximations to these curves.

Key features of the system behavior are:

Dynamic modifiers for translational stiffness are almost unity independent of the foundation aspect ratio, whereas rocking modifiers are significantly reduced with frequency in a very weak dependency on the foundation aspect ratio.

Radiation damping for the horizontal translational mode is only modestly influenced by the direction of vibration or the foundation aspect ratio. For rocking on the other hand, the damping is strongly affected by the aspect ratio and the direction of vibration, increasing with the foundation aspect ratio. At low frequencies damping in rocking motion is smaller compared to that in horizontal translation due to interference phenomena; it only overweighs translational damping at higher frequencies and for elongated foundations when excited in the direction of the longer foundation side. Hence, translational foundation movement may often be predominant with respect to radiation damping.

Embedded Foundations

The references cited above contain also information for embedded foundations. Embedment increases static foundation stiffness. According to the review by Pais and Kausel (1988), dynamic stiffness modifiers remain largely unaffected. The dynamic analyses for obtaining such impedances usually assume a perfect contact between the soil and the basement walls, a situation that seldom occurs in reality. This yields higher damping values as observed from the actual response of buildings. A practical, conservative approach consists in considering the embedment effects only for the static stiffness and applying the dynamic modifiers of surface foundations. Alternatively, one may use the formulae given by Gazetas (1991) and by Mylonakis et al. (2006) that consider an effective height of the contact zone along the perimeter of the embedded foundation.

Soil Layering

Impedance functions for multilayered soils can only be determined with specialized software that is not easily accessible to practicing engineers. Available algorithms are mostly based on finite element procedures incorporating efficient consistent boundaries for the proper energy radiation at the domain boundaries.

A particular case constitutes a soil layer of finite depth on rock where a cutoff frequency exists, below which there is no radiation damping. The respective formulae given in the above references may be used for a two-layer system when the shear wave velocity in the top layer is less than half of that of the underlying stratum. Impedances for square foundations on uniform or nonuniform soil layer overlying a half-space are tabulated by Wong and Luco (1985).

Parameters for Soil Behavior

The expressions given above assume linear elastic or viscoelastic soil behavior. However, for moderate or strong seismic excitations, the nonlinearity of the soil must be taken into account. Hence, the values of the shear modulus entering the equations for the SSI effects must be adjusted to reflect the strain level in the ground associated with the stipulated design ground motion. In critical projects seismic site response analyses are carried out with the soil properties being determined from special dynamic laboratory tests on undisturbed samples. First-order estimates for the strain-compatible values are given in some code provisions. Typical values as recommended by Eurocode 8, Part 5 (CEN 2004), are tabulated below in terms of their small-strain amplitude values G_0 and v_{s0} in dependency on the effective ground acceleration defined as the spectral acceleration at the plateau of the response spectrum divided by 2.5. Guide values for the hysteretic soil damping are also given.

The small-strain values of the soil shear modulus or the shear wave velocity may be determined by a variety of methods, the choice depending on the variability of the soil conditions, available knowledge on the material behavior, and the importance of the structure. These methods include: (i) empirical relationships in terms of the SPT blow count or of the tip resistance of the CPT tests, (ii) geophysical field methods based on wave propagation, and (iii) dynamic laboratory tests. An overview of the testing procedures and available design equations is summarized by Kramer (1996).

	Effective spectral ground acceleration [g]		
	0.10	0.20	0.30
G/G_0	0.80 (±0.10)	0.50 (±0.20)	0.36 (±0.20)
v_s/v_{s0}	0.90 (±0.07)	0.70 (±0.15)	0.60 (±0.15)
Damping ratio	0.03	0.06	0.10

Adaption in Design Codes and Implication for the Design

The implications of inertial SSI for design are illustrated in Fig. 2 with reference to the acceleration response spectrum used for evaluating seismic base shear forces in buildings. Idealized envelope spectra in modern codes initially increase with period, attaining a plateau value, and start decreasing monotonically after a certain period that is in the order of 0.4 to 1.0 s.

Soil-Structure Interaction, Fig. 2 Effect of natural period elongation and foundation damping on a typical acceleration design spectrum (Adapted from Stewart et al. (2003) by permission of the Earthquake Engineering Research Institute)
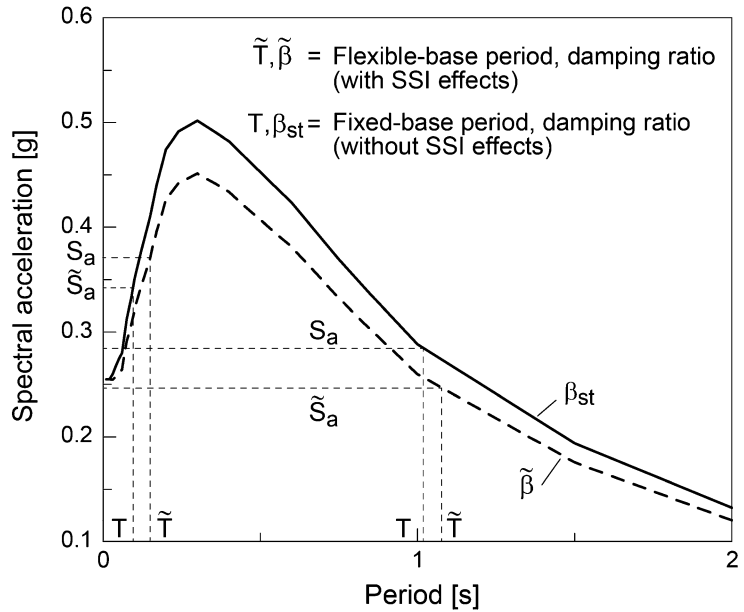
Soil-Structure Interaction, Fig. 3 Ratio of spectral acceleration to peak ground acceleration for 5 % structural damping for some severe earthquakes with long-period components compared to that of a typical code for soft soil (Adapted from Gazetas and Mylonakis (1998) by permission of the American Society of Civil Engineers)
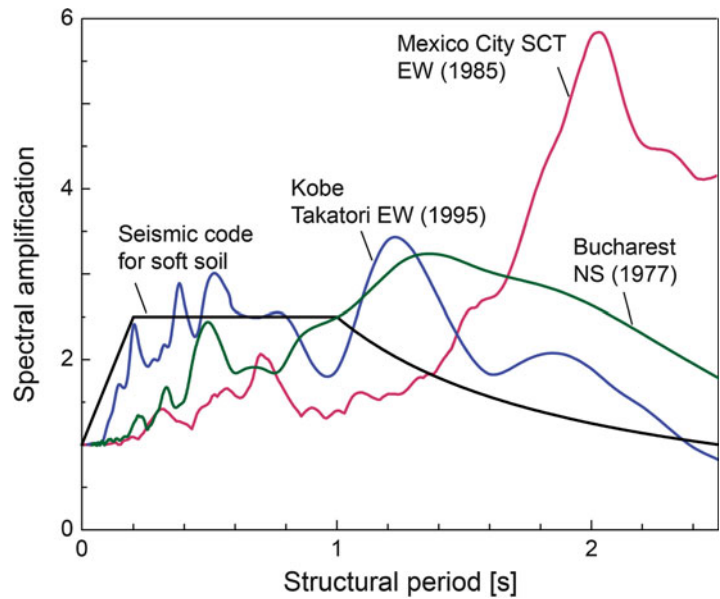
For buildings with periods larger than about 0.5 s, consideration of period elongation and flexible base damping will lead to a reduction of the base shear demand. Hence, in most cases, SSI effects are neglected in the frame of conservative design.

However, there are various seismic environments with recorded response spectra exhibiting their peak at periods greater than 1.0 s. Spectra from some prominent records are contrasted to a typical design spectrum for soil in Fig. 3. SSI phenomena in these earthquakes had detrimental effects as revealed by analyses linking site conditions and building natural periods to observed damage. In the 1985 Mexico City earthquake, for example, due to SSI effects, the natural period of

10–12 story buildings founded on soft clay was altered from about 1.0–1.5 s to nearly 2.0 s, thus coinciding with the peak of the response spectrum at the particular site. The associated phenomena are elucidated among others by Gazetas and Mylonakis (1998).

Hence, proper assessment of both the anticipated seismic input and the prevailing soil conditions is an indispensable prerequisite in any SSI analysis. In modern seismic codes the site characterization for deep soil deposits is based almost exclusively on the near-surface region of the soil (often the top 30 m), disregarding the depth of the underlying rock. The representative average shear wave velocity to this depth in this deposit is used as parameter for the classification; cf. Dobry et al. (2000).

Pile Foundations

Single Pile

Consider a pile horizontally loaded at its head at the ground surface. The deformed shape of the pile extends down to a so-called active (or effective) length below which it becomes negligible. This length depends on the pile diameter, the elastic modulus of the soil, the ratio of pile modulus to soil modulus, and the fixity conditions. Expressions for static and dynamic loading are given by Gazetas (1991). For static loads this length is of the order of 10–20 pile diameters, while for dynamic loading this length will be greater due to the wave propagation. With respect to flexural response, the pile can be modeled without significant error as an infinite-long beam when its length is greater than the active length. Two models are commonly used for the analysis: elastic continuum theory or Winkler spring models (Pender 1993).

Following the same principles as for shallow foundations, the horizontally loaded soil-pile system may be represented by three impedances corresponding to swaying, rocking, and cross-swaying-rocking. Consideration of the latter is necessary since the reference level is located at the pile head and the resultant of the reactions acts at a specific depth thus inducing a bending moment at the pile head. Expressions synthesized

from results by various authors are summarized by Gazetas (1991). The static stiffnesses are expressed in terms of the pile diameter d and Young's moduli of the soil and the pile E and E_p, respectively:

$$\text{Swaying } K_x^0 = dE \left(\frac{E_p}{E} \right)^{0.21} \qquad (29)$$

$$\text{Rocking } K_\theta^0 = 0.15 d^3 E \left(\frac{E_p}{E} \right)^{0.75} \qquad (30)$$

$$\text{Cross-swaying-rocking } K_{x\theta}^0 = -0.22 d^2 E \left(\frac{E_p}{E} \right)^{0.50} \qquad (31)$$

The dynamic modifiers are approximately equal to unity:

$$k_x = k_\theta = k_{x\theta} \approx 1 \qquad (32)$$

The expressions for the radiation damping β_j as defined by Eq. 2 are given in dependency on the dimensionless frequency

$$a_0 = \frac{\omega \, d/2}{v_s} \qquad (33)$$

as follows:

$$\text{Swaying } \beta_x = 0.35 \, a_0 \left(\frac{E_p}{E} \right)^{0.17} \qquad (34)$$

$$\text{Rocking } \beta_\theta = 0.11 \, a_0 \left(\frac{E_p}{E} \right)^{0.20} \qquad (35)$$

$$\text{Cross-swaying-rocking } \beta_{x\theta} = 0.27 \, a_0 \left(\frac{E_p}{E} \right)^{0.18} \qquad (36)$$

Pile Groups

Building foundations are always constructed as groups of piles. In evaluating the dynamic stiffness of a pile group, the interactions between the piles must be taken into consideration, just like in the case of static loading. However, the cross-interaction of individual piles is strongly

dependent on frequency, thus precluding description by simple explicit formulae. The rigorous solution methods available are based on the thin-layer method (Kaynia and Kausel 1982; Waas and Hartmann 1984). Fortunately, a remarkable simple solution procedure was discovered by Dobry and Gazetas (1988) that is straightforward to implement, thus facilitating the assessment of the associated SSI effects with a very good accuracy. The respective interaction coefficients between the piles are given in terms of pile spacing, excitation frequency, and the wave velocity through the soil between the piles. The values for the overall stiffness and damping of the pile group are then assembled using the respective values of the single piles and these interaction factors.

Kinematic Interaction

Shallow Foundations
Kinematic interaction is induced by the presence of a stiff foundation that forces the foundation motions to deviate from the free-field motions. The associated phenomena are due to (i) base-slab averaging of inclined or incoherent seismic waves, and (ii) embedment of the foundation.

Base-Slab Averaging
Seismic waves impinging at directions other than vertical arrive at different points along the foundation at different times giving rise to the so-called wave passage effects. The apparent propagation velocity of the waves is in the order of 1.5–3.5 km/s and is controlled by the wave propagation in the underlying rock. In addition to this, ground motion is in most cases inherently incoherent resulting from inhomogeneities along the travel path from the source to the site.

Studies conducted hitherto mainly address the wave passage problem that is amenable to analytical treatment. They show that the slab due to its stiffness and flexural rigidity averages the free-field displacement pattern by reducing the translational motions and at the same time introducing rotational motions. The latter include rocking in the presence of inclined SV-, P-, or

Rayleigh waves, and torsion in the presence of SH- or Love waves. The torsion of symmetrical buildings observed in earthquakes is a consequence of obliquely incident seismic waves. Further, the modification of the seismic motion depends on the frequency content of the seismic motion with high-frequency components being filtered out by the slab when the respective apparent wavelength is shorter than an effective length of the foundation slab (the diameter for circular foundations).

Kinematic interaction effects are expressed in terms of transfer functions relating the amplitude of the foundation input motion to that of the free-field motion. The system considered consists of a rectangular foundation with area $2a \times 2b$, $a \geq b$ that is subjected to harmonic SH waves of circular frequency ω with particle motion in the direction of the x-axis impinging on the foundation at an angle α_v with the vertical and propagating along the positive y-axis. The transfer functions derived by Veletsos et al. (1997) include both coherent and incoherent seismic motions. They are given in dependency on the dimensionless parameter

$$\tilde{a}_0 = \frac{\omega b_e}{v_s} \sqrt{\kappa^2 + \sin^2\alpha_v \left(\frac{b}{b_e}\right)^2} \qquad (37)$$

where $b_e = \sqrt{ab}$ is the half-side length of an equivalent square foundation, v_s is the shear wave velocity, and κ is a ground motion incoherence parameter. The curves shown in Fig. 4 represent the two limiting cases with $\kappa = 0$ and $\alpha_v = 0$, respectively. The transfer functions for torsional motions are referred to the foundation edge being the product of foundation half-width b and rotational angular distortion.

Lateral transfer functions are for both types of wave motion only very weak dependent on the aspect ratio a/b suggesting that the governing parameter is the foundation area. The induced torsional component, however, is very sensitive both to the aspect ratio and the type of wave motion.

Recent observations on buildings indicate that the apparent value of κ (denoted by κ_a) is nearly
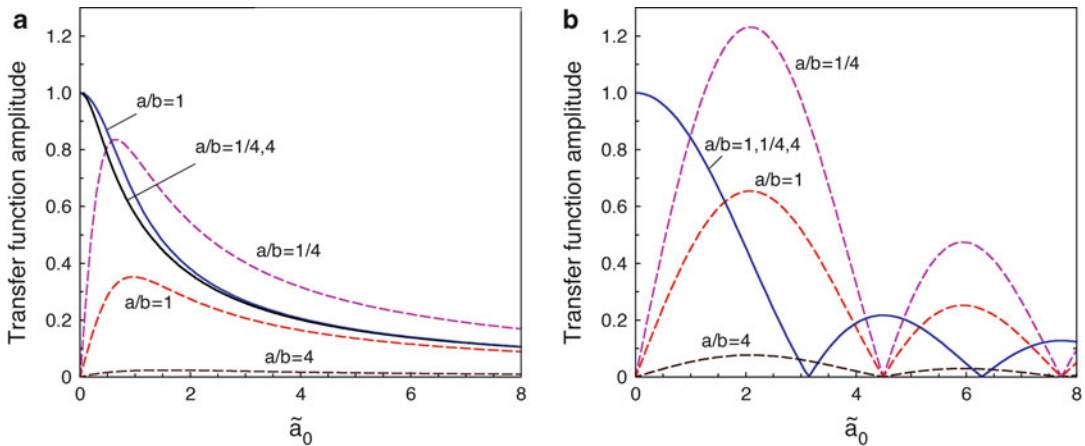
Soil-Structure Interaction, Fig. 4 Amplitude of transfer functions between free-field and foundation input motion for rectangular foundations subjected to obliquely incident shear waves: (**a**) vertically incident, incoherent waves; (**b**) non-vertically incident, coherent waves. The *solid lines* are for the horizontal motion and the *dashed lines* for the induced torsional component. Curves computed from expressions in Veletsos et al. (1997)

proportional to the small-strain shear wave velocity v_s yielding roughly $\kappa_a = 0.2$ at a typical value $v_s = 250$ m/s, Kim and Stewart (2003).

Foundation Embedment

Embedment effects result from the scattering of incoming waves. Rocking motions develop due to nonuniformly distributed tractions against the side walls. Assessment is made by means of transfer functions relating the base-slab translational and rocking motions to the free-field motions. An accurate numerical solution for cylindrical foundations subjected to coherent shear wave motions is provided by Day (1977). The embedment depth e is normalized with respect to the radius of the foundation r, and the frequency dependency is captured by the dimensionless parameter $a_0 = \omega r / v_s$. Figure 5 shows typical patterns for vertically propagating waves.

Piles

Piles embedded in a soil stratum respond to incident vertical shear waves in dependence on their flexural rigidity in relation to the stiffness of the surrounding soil. The incoming wave field is modified, the displacement at the pile head differs from that of the free field, and pile-head rotation is induced. The displacement reduction depends on the ratio of pile modulus to soil modulus, the slenderness of the pile, and the frequency of excitation with high-frequency components being filtered out especially by relatively short, rigid piles.

Analyses of kinematic interaction that include variable stratigraphy are carried out mostly by using the Winkler spring model. The ends of the springs and dashpots that capture SSI effects are connected to the free field where the soil response is imposed. The latter is computed independently. The accuracy of this simplified approach depends on the selection of the springs and dashpots that are obtained adopting physically justified approximations (Pender 1993).

Attention deserves the bending moment induced in the pile during the passage of seismic waves. The maximum value occurs, as expected, at soil layer interfaces, strongly increasing with the contrast in shear wave velocity between the bottom and top layer (Nikolaou et al. 2001).

The numerical study by Fan et al. (1991) using the continuum model by Kaynia and Kausel (1982) for pile groups excited by vertically propagating shear waves provides graphs showing the effects of pile rigidity relative to that of the soil, pile slenderness, pile spacing, number of piles, and pile-head fixity conditions. An idealized general shape of the frequency dependence of the kinematic response is defined in terms of a

S

Soil-Structure Interaction, **Fig. 5** Amplitude of transfer functions between free-field and foundation input motion for cylindrical, embedded foundation subjected to vertically incident coherent shear waves for different normalized embedment depths. (**a**) horizontal translation; (**b**) rocking component (Redrawn from Day (1977))
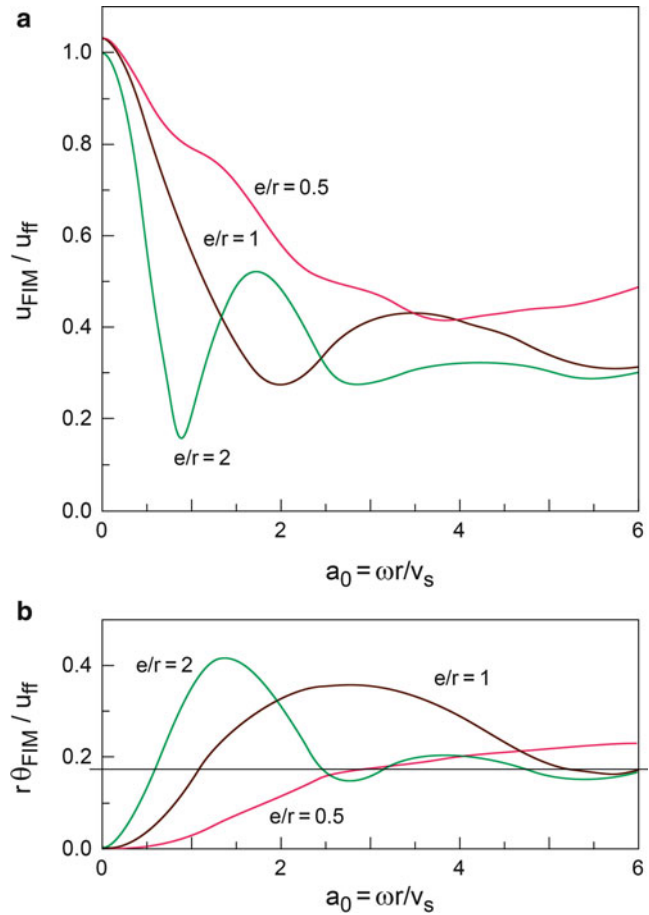
displacement factor relating the pile-head displacement to that of the free field. This factor is approximately unity at low frequencies with the pile closely following the ground movement; in the medium frequency range, it decreases with frequency and beyond a distinct frequency fluctuates around a constant value of 0.2–0.4. The difficulty consists in defining the transition frequencies for the particular system layout.

It must be realized though that there is no simple means for evaluating kinematic interaction for pile groups. In noncritical situations, however, this difficulty may be circumvented by neglecting kinematic interaction. This is justified by findings that kinematic effects for pile groups are similar to those for individual piles, in particular for horizontal translation and to a lesser extent for torsional and rocking vibration modes.

Concluding Remarks

Despite its inherent complexity, the theory of linear soil-structure interaction and the implications in structural performance are now well understood. Refinements, optimization, and validation studies are subjects of ongoing research. The variability in the stratigraphy of soil deposits, the nonlinearity of the soil behavior, the frequency dependency of the response, and the limited availability of specialized software for the analysis make the proper assessment of the SSI effects still a difficult task, requiring physical insight when applying such concepts. It should be self-evident that the effective implementation in an integrated structural design asks for a close collaboration between structural and geotechnical engineers.

Summary

The main effects of soil-structure interaction on the seismic response of structures founded on compliant ground are presented. The modeling concepts to capture the associated modification of the building natural period and the energy dissipation due to radiation damping are highlighted by reference to relatively simple structures. Both kinematic and inertial actions are treated. Available expressions for the dynamic impedance functions are summarized both for shallow foundations and piles. A brief account is given of the implications in seismic design provisions for buildings.

Cross-References

▶ Building Codes and Standards

▶ Earthquake Response Spectra and Design Spectra

▶ Engineering Characterization of Earthquake Ground Motions

▶ Response Spectrum Analysis of Structures Subjected to Seismic Actions

▶ Seismic Analysis of Masonry Buildings: Numerical Modeling

▶ Selection of Ground Motions for Response History Analysis

▶ Site Response: 1-D Time Domain Analyses

▶ Substructuring Methods for Finite Element Analysis

References

CEN-European Committee for Standardization (2004) Eurocode 8: design of structures for earthquake resistance part 5: foundations, retaining structures and geotechnical aspects

Day SM (1977) Finite element analysis of seismic scattering problems. PhD thesis, University of California, San Diego

Dobry R, Gazetas G (1988) Simple method for dynamic stiffness and damping of floating pile groups. Géotechnique 38:557–574

Dobry R, Borcherdt RD, Crouse CB, Idriss IM, Joyner WB, Martin GR, Power MS, Rinne EE, Seed RB (2000) New site coefficients and site classification

system used in recent building seismic code provisions. Earthq Spectra 16(1):41–67

Fan K, Gazetas G, Kaynia A, Kausel E (1991) Kinematic seismic response of single piles and pile groups. J Geotech Eng, ASCE 117(12):1860–1879

Gazetas G (1991) Foundation vibrations. In: Fang HY (ed) Foundation engineering handbook, 2nd edn. Van Nostrand Reinhold, New York, pp 553–593

Gazetas G, Mylonakis G (1998) Seismic soil structure interaction: new evidence and emerging issues. In: Geotechnical earthquake engineering and soil dynamics III, ASCE, Reston, pp 1119–1174

Kausel E (2010) Early history of soil-structure interaction. Soil Dyn Earthq Eng 30(9):822–832

Kausel E, Roesset JM (1974) Soil-structure interaction problems for nuclear containment structures. In: ASCE power division specialty conference, electric power and the civil engineer, Boulder, pp 469–498

Kaynia AM, Kausel E (1982) Dynamic stiffness and seismic response of pile groups. In: Research report R82-03, Massachusetts Institute of Technology

Kim S, Stewart JP (2003) Kinematic soil-structure interaction from strong motion recordings. J Geotech Geoenviron Eng, ASCE 129(4):323–335

Kramer SL (1996) Geotechnical earthquake engineering. Prentice Hall, Upper Saddle River

Mylonakis G, Nikolaou A, Gazetas G (2006) Footings under seismic loading: analysis and design issues with emphasis on bridge foundations. Soil Dyn Earthq Eng 26:824–853

Nikolaou S, Mylonakis G, Gazetas G, Tazoh T (2001) Kinematic pile bending during earthquakes: analysis and field measurements. Géotechnique 51(5):425–440

Pais A, Kausel E (1988) Approximate formulas for dynamic stiffnesses of rigid foundations. Soil Dyn Earthq Eng 7(4):213–227

Pender M (1993) Aseismic pile foundation design analysis. Bull NZ Nat Soc Earthq Eng 26(1):49–160

Stewart JP, Kim S, Bielak J, Dobry R, Power MS (2003) Revisions to soil-structure interaction procedures in NEHRP design provisions. Earthq Spectra 19(3):677–696

Veletsos AS (1977) Dynamics of structure-foundation systems. In: Hall WJ (ed) Structural and geotechnical mechanics. Prentice-Hall, Englewood Cliffs, pp 333–361

Veletsos AS, Meek JW (1974) Dynamic behavior of building-foundation systems. Earthq Eng Struct Dyn 3:121–138

Veletsos AS, Prasad AM, Wu WH (1997) Transfer functions for rigid rectangular foundations. Earthq Eng Struct Dyn 26(1):5–17

Waas G, Hartmann HG (1984) Seismic analysis of pile foundations including pile-soil-pile interaction. In: Proceedings, 8th World conference on earthquake engineering, San Francisco, vol V, pp 55–62

Wolf JP (1985) Dynamic soil-structure interaction. Prentice-Hall, Englewood Cliffs

Wong HL, Luco JE (1985) Table of impedance functions for square foundations on layered media. Soil Dyn Earthq Eng 4(2):64–81

S

Source Characterization for Earthquake Early Warning

Aldo Zollo, Gaetano Festa, Antonio Emolo and
Simona Colombelli
Department of Physics, University of Naples
Federico II, Federico II – AMRA S.c.ar.l,
Napoli, Italy

Synonyms

Earthquake early warning; Earthquake ground
motion; Earthquake source observation; Real-
time location; Real-time magnitude

Introduction

Earthquake Early Warning Systems (EEWS) are
real-time, seismic monitoring infrastructures that
are able to provide a rapid notification of the
potential damaging effects of an impending
earthquake. This objective is achieved through
the fast telemetry and processing of data from
dense instrument arrays deployed in the source
region of the event of concern (regional EEWS)
or surrounding/at the target infrastructure (front-
detection or site-specific EEWS).

A regional EEWS is based on a dense sensor
network covering a portion or the entire area that
is threatened by earthquakes. The relevant source
parameters (event location and magnitude) are
estimated from the early portion of recorded sig-
nals (initial P-waves) and are used to predict, with
a quantified confidence, a ground-motion inten-
sity measure at a distant site where a target struc-
ture of interest is located. Site-specific
(or on-site) EEWS consist of a single sensor or
an array of sensors deployed in the proximity of
the target structure that is to be alerted and whose
measurements of amplitude and predominant
period on the initial P-wave motion are used to
predict the ensuing peak ground motion (mainly
related to the arrival of S- and surface waves) at
the same site. Front-detection EEWS is essen-
tially a variant of the on-site approach, where a
barrier-shaped, accelerometric network is
deployed between the source region and the tar-
get site to be protected. The alert is issued when
two or more nodes of the array record a ground
acceleration amplitude larger than a default
threshold value. For typical regional distances,
the peak acceleration at the barrier nodes is
expected to be associated with the S-wave train,
so that the distance between the network and the
target is set to maximize the lead time (i.e., the
time available for warning before the arrival of
strong ground shaking at the target sites), which
is, in this case, the travel time of S-waves from
the barrier to the target site.

EEWS have experienced a very rapid
improvement and a wide diffusion in many active
seismic regions of the world in the last three
decades (Fig. 1). They are operating in Japan,
Taiwan, Mexico, and California. Many other sys-
tems are under development and testing in other
regions of the world such as in Italy, Turkey,
Romania, and China. Most of these existing
EEWS essentially operate in the two different
configurations described above, i.e., regional
and on-site, depending on the source-to-site dis-
tance and on the geometry of the considered
network with respect to the source area. The
"front-detection" EEWS such as the barrier-
type, Seismic Alert System (Espinosa-Aranda
et al. 2011) for Mexico City can be particularly
advantageous when the only potential seismic
sources are at some distance from the strategic
target to be protected.

The regional EEWS approach is based on the
detection of the initial P-wave signal at a number
of near-source stations, typically 4 to 6. Several
methodologies have been proposed for the real-
time estimation of the earthquake location and
magnitude and are now implemented in the EW
algorithms, such as ElarmS (Allen et al. 2009),
Virtual Seismologist (Cua et al. 2009), and
PRESTo (Satriano et al. 2010) presently running
in California, Switzerland, and Southern Italy,
respectively. In the framework of EU REAKT
(Strategies and Tools for Real-time Earthquake
Risk Reduction, FP7:ENV2011.1.3.1-1) and
international collaboration projects, testing of
PRESTo early warning platform is performed in
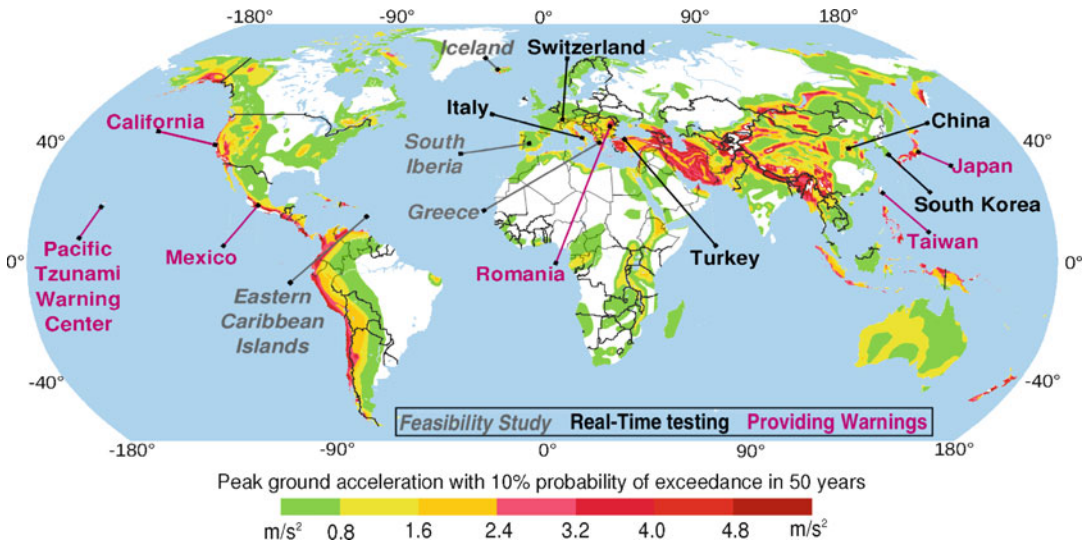
Source Characterization for Earthquake Early Warning, Fig. 1 The map shows the distribution of Earthquake Early Warning Systems around the world, with a color indicating the status of the system. In *purple*, the operative systems, which are providing warnings to public users. In *black*, the systems which are currently under real-time testing. *Gray* color is finally used for those countries where feasibility studies are currently being doing

Romania, Greece, Turkey, Spain, and South Korea. The real-time magnitude estimation is generally inferred from the measurement of peak displacement amplitude and/or the predominant period measured in the first few seconds of the recorded P-signal, typically 3–4 s. Although the saturation of the P-wave parameters has been observed for M > 6.5–7 earthquakes, several methodologies making use of longer time windows of the P-wave and/or the S-wave to update magnitude estimates have been shown to be efficient in minimizing the problem of magnitude underestimation (Colombelli et al. 2012b). The source location and magnitude estimations, which are continuously updated by adding new station data, as the P-wave front propagates through the regional EW network, are then used to predict the severity of ground shaking at sites far away from the source, by using regional-specific, ground-motion prediction equations.

The on-site early warning approaches are generally aimed at estimating the expected peak ground shaking, associated with S- and surface waves, directly from the recorded early P-wave signal. This is achieved through the use of empirical regressions between measurements performed on the initial P-wave signal and the final peak ground motion. Wu and Kanamori (2005) first showed that the maximum amplitude of a high-pass filtered vertical displacement, measured on the initial 3 s of the P-wave (named P_d), can be used to estimate the peak ground velocity (PGV) at the same site, through a power-law relationship. The main advantage is that this relationship does not require an independent estimate of the magnitude as for regional EEWS. Although initially observed for near-source records (distances < 30 km), further analyses on independent datasets have confirmed that log PGV vs. log P_d scaling is still valid at relatively large distances (distances < 300 km) (Zollo et al. 2010; Colombelli et al. 2012a). Most of the currently operating on-site EEWS are threshold-based, alert methodologies: the alert is issued as the measured initial P-wave peak amplitude overcomes a given threshold which is arbitrarily set according to the predicted S-wave peak ground-motion amplitude. Since small magnitude earthquakes may have very large amplitudes driven by high-frequency spikes, such a basic threshold system can produce frequent false alarms. A more robust approach is to combine

the P-wave peak (which scales with distance and magnitude) and P-wave predominant period (which scales with the magnitude), into a single proxy to be used for on-site warning (Wu and Kanamori 2005). Following this idea, Zollo et al. (2010) and Colombelli et al. (2012a) have proposed a threshold-based EW method based on the real-time measurement of the period (τ_c) and peak displacement (P_d) parameters at stations located at increasing distances from the earthquake epicenter. The measured values of early warning parameters are compared to threshold values, which are set for a given minimum magnitude and instrumental intensity. At each recording site an alert level is assigned based on a decisional table with four levels defined upon threshold values of the parameters P_d and τ_c. Given a real-time, evolutionary estimation of earthquake location from first P arrivals, the method also furnishes an estimation of the extent of potential damage zone as inferred from continuously updated averages of the period parameter and from mapping of the alert levels determined at the near-source accelerometer stations.

P-wave-based, regional, and on-site EW methods can be integrated in a unique alert system (as actually done, e.g., in the new version of PRESTo, e.g., PRESTo Plus, Zollo et al. 2014), which can be used in the very first seconds after a moderate-to-large earthquake to determine the earthquake location and magnitude and to map the most probable damaged zone, using data from receivers located at increasing distances from the source.

Methodologies for regional earthquake early warning assume a point-source model of the earthquake source and isotropic wave amplitude attenuation. These assumptions may be inadequate to describe the earthquake source of large earthquakes and wave amplitude attenuation effects, and they can introduce significant biases in the real-time estimation of earthquake location and magnitude. This issue is critically related to the EEWS performances in terms of expected lead time and of uncertainties in predicting the peak ground motion at the site of interest. Within this context, new developments have been

proposed, such as the strategy of expanding the P-wave time window for the real-time signal processing, the 2D mapping of the potential damage zone, and the use of continuous GPS measurements and methodologies to estimate fault rupture extent in real time by classifying stations into near source and far source. These innovative aspects of early warning will be discussed in the present review, with a specific focus on methods for rapid and reliable source characterization for early warning applications.

Methodology

Point-Source Characterization for an Earthquake Early Warning System
In EEWS the strong motion is generally synthesized by a single parameter (in most cases the peak ground velocity, PGV, the peak ground acceleration, PGA, or the peak ground displacement, PGD, Fig. 2), which can be directly related to the damage that a building or an infrastructure may undergo because of the earthquake. Two possible approaches can be explored for the prediction/estimation of ground-motion parameters at a given site.

A first possibility is to relate the ground shaking to simplified macroscopic description of the source, yielding ground-motion prediction equations. In such a case, indicating with PGX the selected ground-motion parameter, the simplest attenuation relationship relates the logarithm of PGX with the earthquake-to-site distance R and the magnitude M:

$$\log PGX = f(M, R) \qquad (1)$$

By definition of magnitude, a linear function of M is in most cases adequate to describe the influence of the earthquake size on the ground motion. The distance effect is instead accounted for by two terms describing the decay of the amplitude owing to geometrical spreading and inelastic processes within the upper crust. Nevertheless, more complex ground-motion prediction equations may be used containing high-order terms, focal depth dependence, and site effects. It is worth to
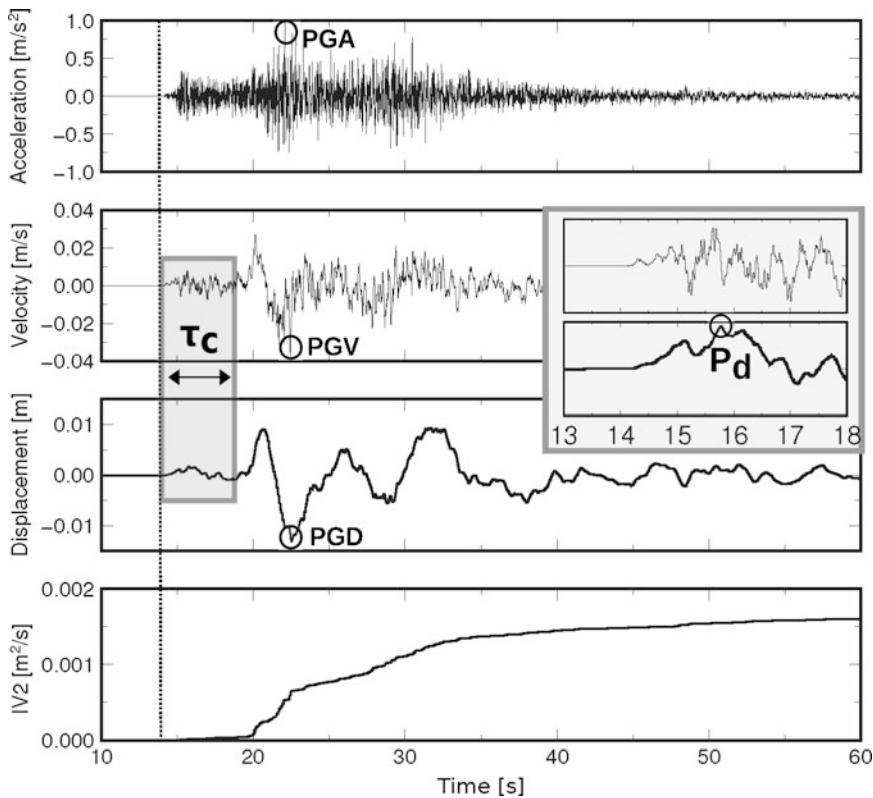
Source Characterization for Earthquake Early Warning, Fig. 2 Schematic illustration of early warning parameters. From *top* to *bottom*, an example of the vertical component of acceleration, velocity, displacement, and integral of squared velocity (IV2) signals. The ground-motion parameters PGA, PGV, and PGD are measured as the absolute maximum along the signal, using acceleration, velocity, and displacement records, respectively.

The *gray small box* shows a zoom on the first few seconds of P-wave on velocity (*top*) and displacement (*bottom*) records. The initial peak displacement (P_d) is measured as the absolute maximum of displacement waveform on the early portion of P-wave (typically 2–4 s) while the period parameter τ_c is measured from the ratio between initial displacement and velocity waveforms in the same time window

note that while magnitude is an ensemble measure for earthquake size, the definition of the distance requires a specific metric, which is sensitive to the ratio between the source-to-site distance and the earthquake size. At distances significantly larger than the source size, a point-source approximation for the earthquake can be generally assumed, and R refers to the epicentral or the hypocentral distance. In the near-source range, instead, finite-fault effects may be relevant and a different metric for the distance could be required. Anyhow, whatever choice of attenuation relationships and distance metrics are done, the ground-motion prediction requires the knowledge of earthquake location and size. This model

is used by regional early warning systems, for which the characterization of the source is performed by a network installed in the source vicinity.

A complementary approach is based on empirical relationships between a ground-motion parameter Py measured in the early portion of the P-wave train and the final PGX. This is physically grounded on the first-order approximation that log Py has the same magnitude and distance dependence of log PGX, being differences concentrated only on static and possibly frequency-dependent effects. In such a case, estimation of source parameters is hidden in the common dependence and the uncertainty may be

S

significantly reduced by avoiding the estimation of magnitude and distance. Additionally, this approach does not require a seismic network to constrain the source parameters and can be efficiently used for single stations. Such a model is the one implemented in on-site early warning systems.

Finally both models can be combined together, with source parameters estimated by a regional network and ground motion locally verified at a specific site.

Real-Time Location

A main concern for any early warning system is the reliable estimation of earthquake hypocenter in real time. Recently, Satriano et al. (2008) have developed an evolutionary approach aimed at constraining the earthquake location, which starts when the first station is triggered by the seismic event and is updated as time passes. This technique is based on the equal differential time (EDT) formulation and provides a probabilistic density function for the earthquake location in a three-dimensional space accounting, at each time step, for information from both triggered and not-yet-triggered stations. In particular, with a single, initial, recorded arrival time, the hypocenter has to belong to the Voronoi cell containing the triggered station, which is created using the travel times to the not-yet-triggered stations. As more and more stations identify the seismic event, the location is constrained in the volume defined by the intersection of the Voronoi cells for the remaining not-yet-triggered stations and the EDT surfaces evaluated for all pairs of already-triggered stations.

Let us consider a seismic network (Fig. 3a) with N operational recording stations $S_1, \ldots, S_N$ and a gridded volume V containing both the network and the earthquake source. Preliminarily, we computed the travel times from each grid point (i, j, k) in the volume V to each station of the network. If the earthquake hypocenter is at the node (i^*, j^*, k^*) of the searching grid, the classical EDT formulation prescribes that the difference between the theoretical travel times ttt_m and ttt_n from the event source to two stations S_m and S_n of the network is exactly equal to the difference between the observed arrival times ot_m and ot_n at the same stations, since they share the same earthquake origin time:

$$(ttt_m - ttt_n)_{i^*, j^*, k^*} = ot_m - ot_n, \qquad \text{with} \quad m \neq n \tag{2}$$

If a homogeneous velocity model is assumed, the previous equation defines a 3D hyperbolic surface whose symmetry axis passes through the two stations. With N_T triggered stations, we obtain $N_T(N_T - 1)/2$ surfaces and the hypocenter is then located at the region crossed by the maximum number of EDT surfaces.

This description is modified in the evolutionary approach introduced by Satriano et al. (2008) in which, at each time step, EDTs are evaluated not only for pairs of triggered stations but also for those pairs where only one station has already triggered. This means that when the first station, let us say S_n, is triggered by the earthquake at the time t_n, it is immediately possible to limit the hypocentral position (Fig. 3b) from the EDT surfaces defined considering that any operational but not-yet-triggered receiver S_l will identify the seismic arrival at a later time t_l (with $t_l \geq t_n$) such that

$$(ttt_l - ttt_n)_{i^*, j^*, k^*} \geq 0, \qquad \text{with} \quad l \neq n \tag{3}$$

The equality in the previous equation corresponds to the set of points in the volume of interest for which the travel time at the first station is equal to the travel time at any remaining, not-yet-triggered, receivers (conditional EDT). As a consequence, the inequality delimitates the region in the volume V, bounded by the conditional EDT, that must contain the earthquake hypocenter. In the case of a homogeneous propagation medium, the region bounded at the time t_n by the conditional EDTs evaluated for all the stations corresponds to the Voronoi cell for the receiver S_n. As the time goes on, the source volume is bounded by the system of equations

$$(ttt_l - ttt_n)_{i^*, j^*, k^*} \geq t_{\text{cur}} - t_n, \qquad \text{with } l \neq n \tag{4}$$
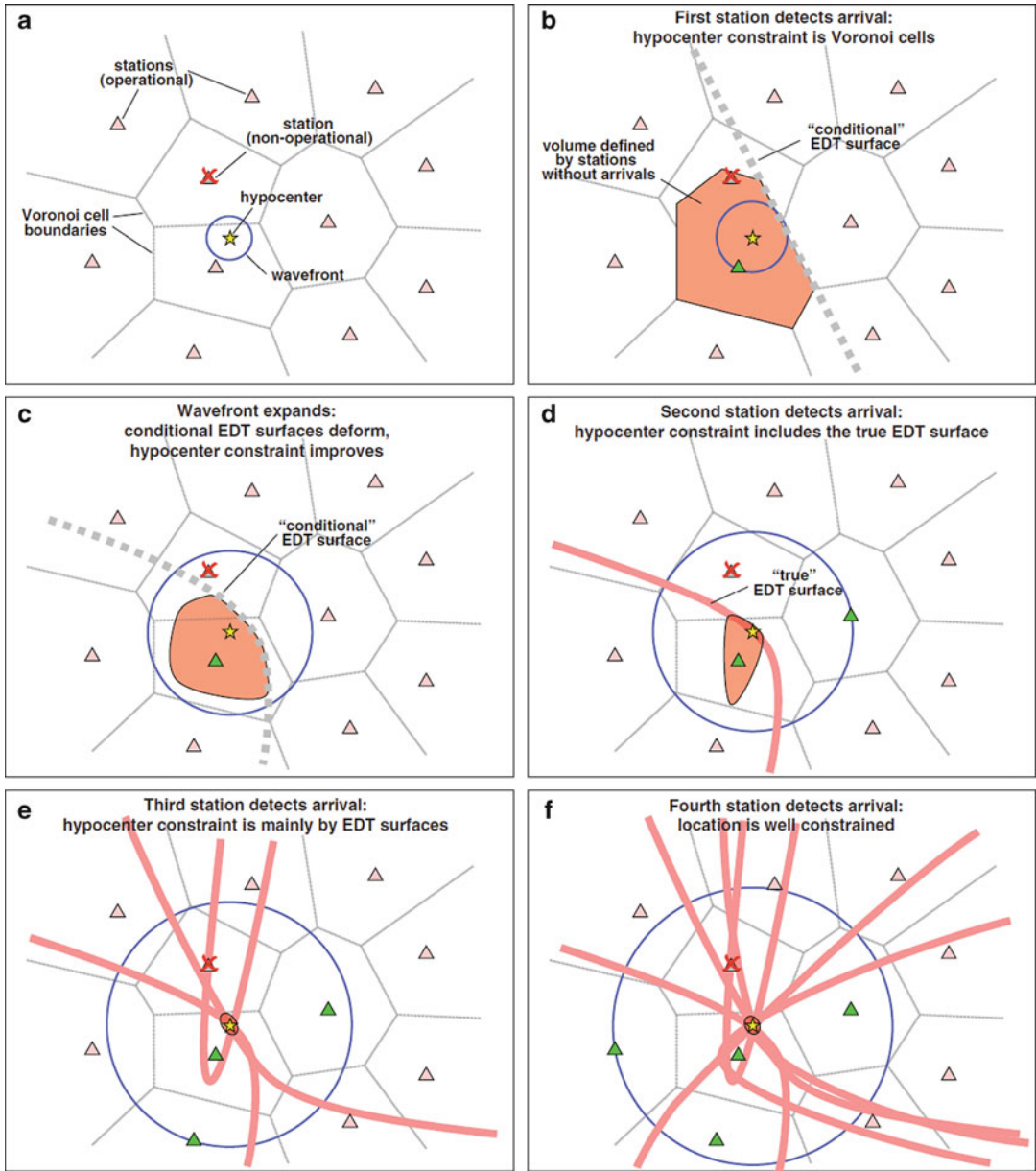
Source Characterization for Earthquake Early Warning, Fig. 3 Sketch of the evolutionary earthquake location approach. For simplicity of the presentation, it is shown only a plane view of the epicentral location. (**a**) The Voronoi cells associated with each station (operational and not) of the network are a priori known. (**b**) As the first station triggers, it is possible to identify a volume that is bounded by the conditional EDT surface on which the travel time for the first receiver is equal to the travel times for each not-yet-triggered station. This volume is likely to contain the true earthquake location. (**c**) As the time passes, further information from station that has not yet triggered is gained and the EDT surface bends around the first triggered station. As a consequence, the source volume decreases. (**d**) As the second station triggers, a true EDT surface is defined. The intersection between this surface and the conditional EDT surfaces limits the likely source volume whose dimension is continuing to decrease. (**e**) Two more true EDT surfaces become available when the third station triggers, thus better constraining the earthquake location. (**f**) As more and more stations trigger, the earthquake location converges to the standard EDT solution which is based on true EDT only (After Satriano et al. (2008))

provided that, at the current time t_{cur}, only the station S_n has triggered at the previous time t_n. As t_{cur} increases, the hypocentral volume becomes smaller since the conditional EDT surfaces fold around the station S_n (Fig. 3c).

It is possible to prescribe a probability density function (pdf) for the hypocentral volume associated with each inequality in the previous equations and for each grid point. We assign a value $p_{n,l}(i, j, k) = 1$ if the inequality is satisfied and a value $p_{n,l}(i, j, k) = 0$ if it is not. Summing up over all the stations, we obtain, for each grid point, a non-normalized PDF $P(i, j, k)$ whose maximum value is $(N-1)$ for those grid points for which all inequalities in the equation are satisfied.

When the second station (and, as the time progresses, further receivers) triggers, the equation is evaluated for all the possible pairs of triggered and not-yet-triggered stations. Then, true EDT surfaces are defined for each pair (n, m) of triggered stations computing, for each grid point, the quantity

$$q_{n,m}(i,j,k) = \exp\left\{\frac{\left[(ttt_n - ttt_m)_{i,j,k} - (ot_n - ot_m)\right]^2}{2\sigma^2}\right\},$$

$$\text{with } n \neq m$$

$$(5)$$

where the expression in square brackets corresponds to the standard EDT and σ represents the uncertainty associated with the arrival time picking and travel-time computation. The quantity $q_{n,m}(i, j, k)$ varies between 0 and 1. We then sum the $q_{n,m}$'s with the updated $p_{n,l}$'s to obtain a new $P(i, j, k)$ which now has a maximum value equal to

$$P_{max} = (N - N_T)N_T - \frac{N_T(N_T - 1)}{2} \quad (6)$$

being N_T the number of triggered stations. Once we have evaluated P, we can define the function

$$Q(i,j,k) = \left(\frac{P(i,j,k)}{P_{max}}\right)^N \quad (7)$$

which varies in the range [0, 1] and represents the PDF for hypocentral location at the grid cell (i, j, k).

As the time increases and more and more stations trigger, the number of not-yet-triggered stations becomes smaller and smaller, and the earthquake location converges toward the hypocenter which would be obtained through the standard EDT approach using data from all operational stations of the network (Figs. 3d–f). Tests performed on both synthetic and real data have shown that when a dense seismic network (i.e., with a mean station spacing of about 10 km) is deployed around the fault zone, a location accuracy is achieved within 1–3 s after the first arrival detection.

Real-Time Magnitude Estimation and Earthquake Rupture

Magnitude estimation for early warning applications is based on empirical relationships relating the earthquake size with parameters measured in the early portion of the P- and S-wave trains. These parameters are generally associated to the low-frequency content of the data, which is sensitive to the seismic moment, and can be related to the maximum amplitude, the dominant frequency, or the energy released by the event. Associated proxies are peak values, predominant period and integrated measurements, respectively. Several authors showed that the initial portion of recorded P-waves carries information about the event magnitude, both through its frequency content and amplitude (Allen and Kanamori 2003; Kanamori 2005; Zollo et al. 2006; Wu and Zhao 2006; Böse et al. 2007; Wu and Kanamori 2008). A review of the common used parameters is detailed in the next section; here we want to note that peak and energy estimates are dependent on the source-to-site distance, while predominant period is pretty insensitive to the epicentral location (Allen and Kanamori 2003). In all cases, when dealing with dense regional networks surrounding the fault that generated the earthquake, location is generally available before or at the same time the first estimates of magnitude are performed. Therefore, it is not a

disadvantage to use distance-dependent parameters as compared to distance-independent ones. All of the parameters are estimated from measures in the early portion of the signal: with the goal of issuing an early warning for earthquakes with magnitude larger than 5.5, the time scale in which performing the measurements is few seconds (2–4 s generally).

The definition of a time window t_0 for the measurements corresponds to image a specific area on the fault plane, delimited by the corresponding isochrone. It is defined as the set of points whose radiation arrives at a given station at the same time t_0. Hence, the portion of the fault enlightened by the first few seconds of signal depends both on the relative location of the station as compared to the fault and on the phase, either P or S. Specifically, for the same time window, S-waves image a larger area on the fault, because their speed is closer to the rupture speed as compared to the P-waves, and the regions explored by the different stations do not overlap as much as P-waves. We then expect that inclusion of early portion of S-waves may significantly constrain the estimation of the earthquake magnitude, as compared to the only use of P-waves. This approach is possible for close stations (with epicentral distances smaller than 30 km) for which the analysis of S-waves does not significantly affect the lead time for early warning. Additionally, isochrone mapping enhances that few seconds of P/S-waves correspond to an earthquake size of magnitude 6–6.5. This indicates that the early portion of the signal captures almost the whole rupture process up to this magnitude, while effective prediction is possible beyond that limit. Using a kinematic description of the rupture, we can argue that early warning parameters up to magnitude 6–6.5 change because both the ruptured area and the total average slip increase. Beyond that threshold, these parameters image almost the same portion of fault and any change with magnitude is only ascribed to any increase of the average slip in that region.

From this consideration we can retain that scaling of early warning parameters may be different for different magnitude ranges, and a two-slope behavior is expected around the deterministic threshold. Hence, the use of a single relationship for a broad magnitude range should be statistically checked to avoid under/over estimations of the magnitude at the limits of the investigated range. Additionally, if any scaling occurs for events with magnitude larger than 6, this indicates a different initiation process for earthquakes that reach different sizes. Specifically, since the scaling of parameters is observed up to magnitude 7.5, this implies that when we look at the earthquake rupture on a given space scale characteristic of a magnitude 6.0 event, we can probabilistically forecast if this rupture will soon stop or it will grow up to a larger space scale. Beyond that limit, standard regression laws saturate and a different approach is required to capture the earthquake size for very big-size events. As a final comment, since the parameters have a different sensitivity to the slip increase, the uncertainty associated to the magnitude estimation may change. Generally, peak parameters are more sensitive to slip changes, the slope of the scaling is larger, and the uncertainty in the magnitude estimation is smaller as compared to predominant period estimations.

We now discuss a real-time approach for magnitude estimation based on a probabilistic evolutionary approach (Lancieri and Zollo 2008). This approach generalizes standard averages coming from single-station estimates of magnitude and can include also a priori probability density functions for magnitude characterization. As discussed before, we assume a linear scaling between the logarithm of the early warning parameter Py and the final magnitude

$$\log Py^{\text{theo}} = A + BM + C(R/R_0) \qquad (8)$$

where C is a function of the distance R and R_0 a reference distance. Coefficients are generally different for P- and S-waves. For low-frequency estimators that depend on the distance, the predominant effect comes from geometrical spreading and we can assume $C(R) = K \log R$. After defining the duration of the window for P- and S-waves, the computation of Py is performed by isolating the corresponding time windows after

the P- and S-wave arrivals. P-wave arrival is automatically picked, while S-wave arrival is computed using a theoretical travel-time database derived by a 1D or 3D velocity model suitable for the area of investigation. Since S-wave estimates are significantly different from P-wave estimates, we need to be sure that the selected P-wave window is not contaminated by the following S-wave, to avoid any bias in the magnitude estimation. If the S-wave arrival is expected in the P-wave window, that P-wave measurement is discarded. For a single Py measurement at a given station and for a given phase, the magnitude can be estimated using the Bayes theorem:

$$P(M|Py) = \frac{P(Py|M)P_a(M)}{\displaystyle\int_M P(Py|M)P_a(M)} \qquad (9)$$

In the above formula the probability $P_a(M)$ represents the a priori probability on the magnitude, the probability $P(Py|M)$ comes from the regression relationship, and, assuming a Gaussian distribution for $\log(Py)$, this writes

$$P(Py|M) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[\frac{1}{2\sigma}\left(\log Py - \log Py^{\text{theo}}\right)^2\right] \qquad (10)$$

(σ is the standard deviation of the fit). Finally the denominator is a normalization factor. This approach can be directly plugged into a real-time evolutionary magnitude estimation where the probability is updated as new information comes from additional stations or phases. Without any information, we assume an a priori magnitude distribution coming from the long-term earthquake catalog inspection (generally a Gutenberg-Richter distribution). To build up an evolutionary scheme, assuming to have computed the probability density function for the magnitude after $n-1$ measurements of Py, this function is then used as the a priori distribution for the n th measurement. It is worth to note that this scheme allows accounting for magnitude saturation. In this case the probability

distribution $P(Py|M)$ is assumed uniform beyond the magnitude saturation threshold.

Parameters for Magnitude Estimation

In the context of real-time applications, different amplitude and period parameters have been proposed to get independent estimates of the earthquake size. We present here an overview of the parameters used by the different systems for the magnitude estimation.

Following an original idea of Nakamura (1988), Allen and Kanamori (2003) first proposed the use of a period parameter, measured on the first few seconds of P-wave signal, to infer the size of the ongoing earthquake. The predominant period is measured on the vertical component of velocity record and is defined as

$$T_i^p = 2\pi\sqrt{X_i/D_i} \qquad (11)$$

where

$$X_i = \alpha X_{i-1} + x_i^2 \qquad (12)$$

$$D_i = \alpha D_{i-1} + \left(\frac{dx}{dt}\right)_i^2 \qquad (13)$$

where T_i^p is the predominant period at the i th sample, x_i is the recorded ground velocity, X_i is the smoothed squared ground velocity, D_i is the smoothed squared velocity derivative, and α is a smoothing constant. They showed that an empirical log-linear correlation exists between the event magnitude and the maximum of T_i^p (named $T_{i\ \max}^p$) within 2–4 s after the P-wave arrival. Using a dataset of Californian earthquakes, they derived two linear relations between $T_{i\ \max}^p$ and magnitude, for small earthquakes (in the magnitude range 3–5) and for larger earthquakes (magnitude range 5–7.3). In particular, for the upper magnitude class they found a log-linear relationship of the form:

$$M = 7.0\log\left(T_{i\ \max}^p\right) + 5.9 \qquad (14)$$
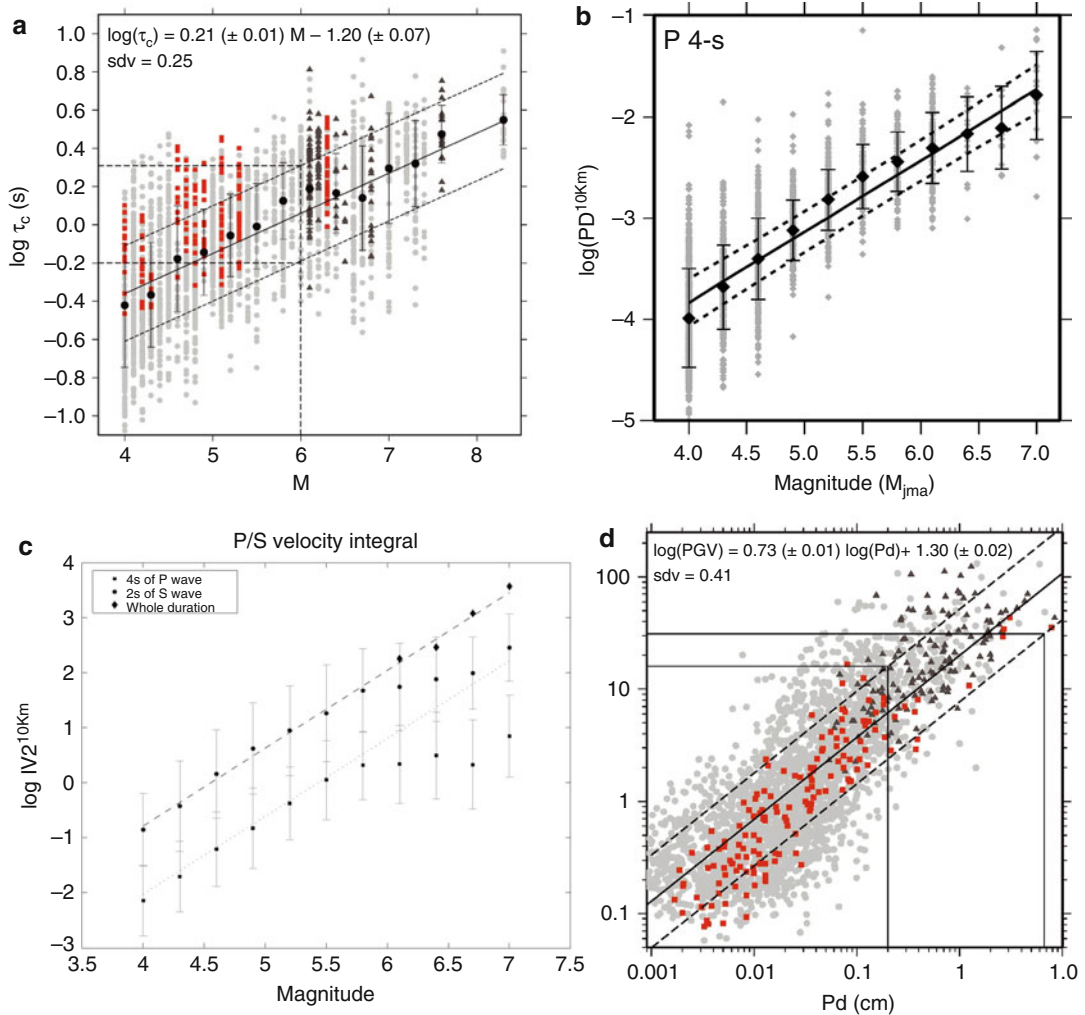
Source Characterization for Earthquake Early Warning, Fig. 4 Empirical scaling relationships between early warning parameters and magnitude. (**a**) The average period as a function of magnitude (After Zollo et al. (2010)); (**b**) the initial peak displacement as a function of magnitude (After Lancieri and Zollo (2008)); (**c**) the correlation between P_d and PGV (After Zollo et al. (2010)); and (**d**) the scaling of the integral of squared velocity (IV2) with magnitude (After Festa et al. (2008))

A similar period parameter has been proposed by Kanamori (2005), who defined the average period (τ_c) of the first seconds of P-wave signal as

$$\tau_c = 2\pi \sqrt{\frac{\int_0^{\tau_0} u^2(t)dt}{\int_0^{\tau_0} v^2(t)dt}} \quad (15)$$

where u and v are displacement and velocity, respectively; the integrals are computed over a time window $(0, \tau_0)$ starting from the P-wave arrival and τ_0 is generally set at 3 s.

Using a database of Taiwan, Japan, and Italy earthquakes (magnitude range $4 < M < 8.3$), Zollo et al. (2010) determined the relationship between average period and magnitude (Fig. 4a). By measuring τ_c in seconds, they found

$$\log(\tau_c) = 0.21(\pm 0.01)M - 1.19(\pm 0.08) \quad (16)$$

Both period parameters ($T^p_{i\,\text{max}}$ and τ_c) are empirically related to the event magnitude but are pretty independent of the distance.

An alternative estimate of the earthquake size can be obtained using amplitude parameters. Zollo et al. (2006) showed that the low-pass filtered, peak displacement amplitude of initial P- and S-wave seismic signals correlates with the earthquake magnitude (Fig. 4b). The P- and the S-peak amplitudes are measured in a short time window (2–4 s) after the arrival times of P- and S-waves, respectively. The initial peak amplitude can be measured on the single vertical component or on the modulus of the displacement vector, as proposed by Lancieri and Zollo 2008. In both cases the functional dependence of the peak amplitude follows the more general relationship, including also the dependence on the distance (Eq. 8), with the reference distance $R_0 = 10$ km. Hereinafter, we will adopt the initial P-peak amplitude of the vertical component and will refer to it as P_d.

Finally, Festa et al. (2008) investigated the scaling of the early-radiated energy, with the final size of the event. The radiated energy can be inferred from the squared velocity integral ($IV2$), which is measured on the initial portion of P- and S-wave signals and is defined as follows:

$$IV2_c = \int_{t_c}^{t_c + \Delta t_c} v_c^2(t)dt \quad (17)$$

where the subscript c refers to the P- or S-phase, t_c is the corresponding first arrival, and v_c is the particle velocity measured on the seismograms. Moreover Δt_c is the length of the signal along which the analysis is performed. They found an evident log-linear scaling of $IV2$ with magnitude for both P- and S-wave data, up to M = 5.8 (Fig. 4c). Beyond this value, the early energy increases less with respect to the final magnitude. Thus, they suggested that early-radiated energy can be used to discriminate whether the event has a magnitude larger or smaller than 5.8, and only in the latter case it allows for real-time magnitude

estimation. For larger magnitudes a saturation effect of $IV2$ prevents from a correct evaluation of the event size. The saturation effect disappears when the velocity integral is evaluated along the whole signal duration, showing a robust log-linear correlation up to larger magnitudes (~7).

In addition to the event magnitude, another relevant and complementary piece of information to be provided in real time is the estimate of the expected ground shaking at target sites. This latter represents an important aspect of the practical implementation of an EEWS and is crucial for the decision-making processes and the prompt activation of security actions and emergency procedures. With this aim, an empirical correlation between the initial peak displacement (P_d) and the final observed peak ground velocity (PGV) can be used. Analyzing a database of earthquake records from Japan, Taiwan, and Italy (2009 Mw 6.3 L'Aquila mainshock and aftershocks) and considering a maximum recorded distance of 60 km, Zollo et al. (2010) obtained the following regression relation (Fig. 4d):

$$\log(PGV) = 0.73(\pm 0.01)\log(P_d) + 1.30(\pm 0.02) \quad (18)$$

where PGV is in cm/s and P_d is in centimeters.

Amplitude and period parameters thus provide complementary information about the ongoing earthquake, being them related to the expected ground shaking at recording sites and to the earthquake magnitude, respectively. With this in mind, Wu and Kanamori (2008) proposed an original, on-site alert-level scheme, based on the combination of measured P_d and τ_c parameters. The idea is to rapidly distinguish the case of a small/large and close/faraway event from the observed values of P_d and τ_c at each recording site. Following this idea, Zollo et al. (2010) proposed a threshold-based approach to EEWS which is aimed at the setup of local alert levels based on a decision table. The key element of the method is the real-time, simultaneous measurement of initial peak displacement (P_d) and period parameter (τ_c) in a 3 s window after the first P-arrival time and on the use of the initial peak displacement as a proxy for the PGV.
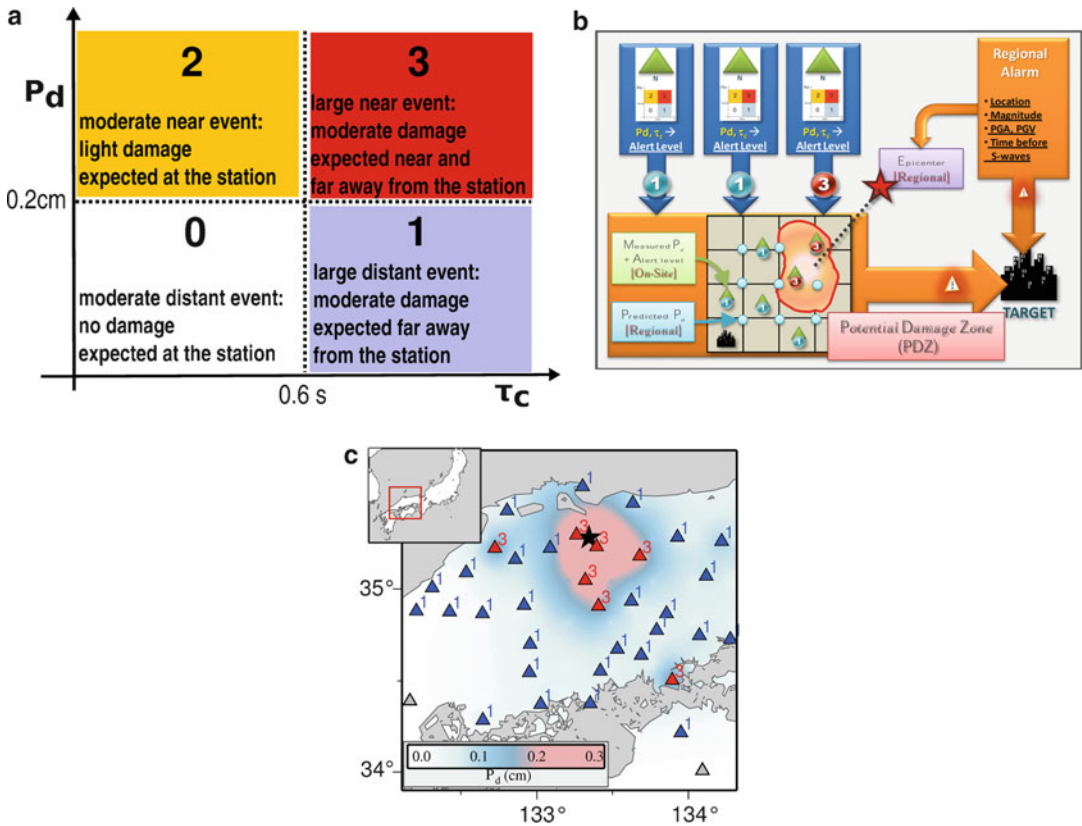
Source Characterization for Earthquake Early Warning, Fig. 5 Alert levels and threshold values for observed early warning parameters (After Zollo et al. (2010)). (**a**) P_d versus τ_c diagram showing the chosen threshold values and the regions delimiting the different alert levels. Level 3 = damage expected nearby and far away from the station; level 2 = damage expected only nearby the station; level 1 = damage expected only far away from the station; level 0 = no expected damage (After Zollo et al. (2010)). (**b**) Conceptual scheme for the potential damage zone definition. Measured and predicted Pd values over the area are interpolated and the PDZ is obtained by delimiting the $Pd = 0.2$ isoline. (**c**) Example of potential damage zone resulting from the interpolation of measured and predicted P_d values. *Gray triangles* represent the stations triggered by the earthquake, while *red* and *blue triangles* show the alert level recorded at each station, as soon as 3 s of signal after the P-picking are available. The color transition from *light blue* to *red* delimits the potential damage, PDZ (After Colombelli et al. (2012a))

From the analysis of strong-motion data of Japan, Taiwan, and Italy, Zollo et al. (2010) calibrated the threshold values, for the definition of four alert levels (0, 1, 2, 3) (Fig. 5a). The threshold values have been established according to the P_d vs. PGV and to the τ_c vs. M empirical relationships (Fig. 4a, d). The threshold values correspond to a minimum magnitude $M = 6$ and to an instrumental intensity $I_{MM} = 7$, assuming that the peak ground velocity provides the instrumental intensity through the relationship of Wald et al. (1999). The alert-level scheme can be interpreted in terms of potential damaging effects nearby the recording station and far away from it. For example, following the scheme of Fig. 5a, the maximum alert level (level 3, i.e., $\tau_c \geq 0.6$ s and $P_d \geq 0.2$ cm) corresponds to an earthquake with predicted magnitude $M > 6$ and with an expected instrumental intensity (at the site) $I_{MM} \geq 7$. This means that the earthquake is likely to have a large size and to be located close to the recording site and a high level of damage is therefore expected either nearby or far away from the recording station. On the contrary, in

case of a recorded alert level equal to $0\,(\tau_c < 0.6\,\mathrm{s}$ and $P_d < 0.2\,\mathrm{cm})$, the event is likely to be small and far from the site, thus no damage is expected either close or far away from the station. The application of this method to a series of large Japanese events $(M > 6)$ (Colombelli et al. 2012a) confirmed that the threshold-based approach is a robust strategy to rapidly predict the expected damage at recording sites, showing a very good matching between the real-time assigned local alert levels and the final, observed peak ground velocity, carried by the later arrival of S-waves.

Extended Source Approach

Expansion of the P-Wave Time Window for Large Magnitude Events

There are no concerns, in fact, on the effectiveness of EEWS for the real-time characterization of small and moderate events, for which the fracture process is concluded within a few seconds. Anyhow, regression relationships between peak displacement, predominant period, and other integrated parameters based on 2–4 s of P-wave saturate beyond a magnitude threshold that can be estimated into the range 7–7.5. This saturation is owed to the limited window range that images a too small portion of the fault plane. When an earthquake extends for hundreds of kilometers on the fault, its dynamics is expected to be controlled by large asperities that the rupture can meet during its propagation also far away from the hypocenter. Another aspect to be carefully checked is the filter used to get a reliable low-frequency signal. Generally the high-pass cutoff frequency is 0.075 Hz that is comparable with the corner frequency of an $M\,7$ event. This means that up to magnitude 7, the cutoff filter still includes the ensemble radiation emitted by the fault while for larger magnitudes the selected frequency range extracts only the constructive interference coming from specific regions of the fault. Investigation of large magnitude events hence requires a progressive extension of the P-wave time window, with update of early warning parameters at individual stations. Since at fixed time close stations experienced a large time window when compared to farther stations, an efficient procedure should manage in real time both the increasing of the time window at single stations and different time windows, when combining information coming from different stations.

The analysis of strong-motion data of the recent Mw 9.0, 2011 Tohoku-Oki mega-thrust earthquake confirmed the necessity of considering larger time windows to overcome the problem of parameter saturation. Colombelli et al. (2012b) proposed an evolutionary approach to early warning, in which amplitude and period parameters (P_d and τ_c, respectively) are measured in progressively increasing P-wave time windows. The real-time measurements are no more done within a fixed portion of signal but starting from the P-wave arrival and expanding the time window up to the arrival of the S-waves at each recording site. With such an approach, the standard methodologies and the empirical regression relationships can be extended to very large earthquakes, provided that appropriate time windows are selected for the measurements. With this approach data from different stations require a specific weight depending on the available P-wave time window (PTW). Specifically, a weight proportional to the square of PTW was shown to provide a good compromise between the window length and the number of stations. As shown in Fig. 6a, that peak displacement increases with PTW, indicating a final level of 8.5. Final underestimation may be related in this case to the regression relationship, validated for smaller magnitude events and shorter distances, to the cutoff frequency, or to the specific dynamics of the event, which exhibited a frequency-dependent radiation (Fig. 6b). On the other hand, predominant period does not monotonically increase with larger time window, indicating that its use in real time could be sensitive to the specific event to be analyzed.

Real-Time Estimation of Fault Rupture Extent Using Strong-Motion Data

Even if we improve the magnitude estimation for very large earthquakes, distance metrics might be inadequate, mostly for targets that are close to the causative fault but far from the epicenter. In such a
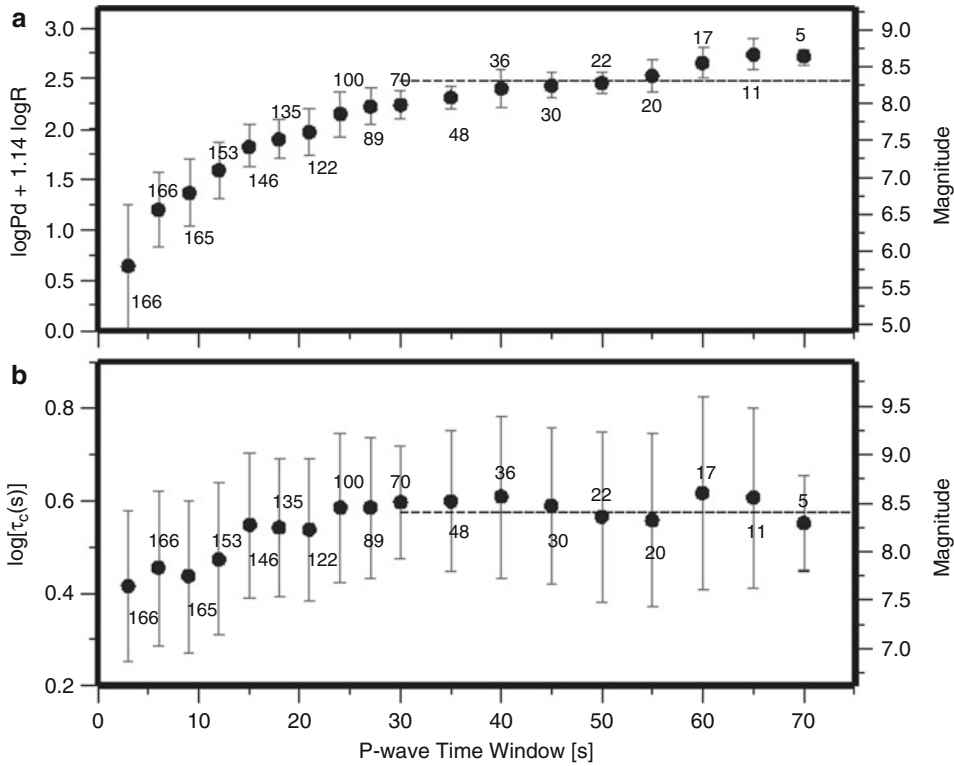
Source Characterization for Earthquake Early Warning, Fig. 6 Real-time evolution of average values of peak displacement (P_d) (**a**) and predominant period (τ_c) (**b**) as a function of the P-wave time window used. Error bars are computed as the standard deviation associated to each value; the gray number close to each point represents the number of stations used for each considered time window. Both parameters exhibit saturation when a 25–30 s PTW is used: the saturation level is shown by the *gray-dashed lines*. For each plot, the corresponding magnitude scale is also represented; this has been derived based on the coefficients of Eq. 1 and on the τ_c vs. M relationship determined by Zollo et al. (2010) (After Colombelli et al. (2012b))

case ground-motion prediction could be significantly underestimated. To this end, Yamada et al. (2007) proposed a method to classify recording stations in near- and far-source distance. If a dense seismic network is available, this subdivision can be used to infer information about the fault geometry. Analyzing strong-motion data from past events, the authors found that the combination of vertical acceleration and horizontal velocity produces the best performance for stations classification. Their discriminant is given by

$$F_i = 6.046 \log Za_i + 7.885 \log Hv_i - 27.091 \tag{19}$$

where Za_i and Hv_i represent the peak acceleration on the vertical component (in cm/s^2) and the

square root of the sum of the squares of the peak velocity on the horizontal components, recorded at the ith station, respectively. Thus, the quantity

$$P_i = \frac{1}{1 + e^{-F_i}} \tag{20}$$

provides the probability that the i th station is in the near- or far-source region. Specifically, the recording site is classified as a near-source station if the probability is greater than 1/2.

For real-time application, peak values used for recording site classification are computed from incoming data every 10 s for each station and then used in the discriminant function. As example, in Fig. 7 the procedure is applied to the case of
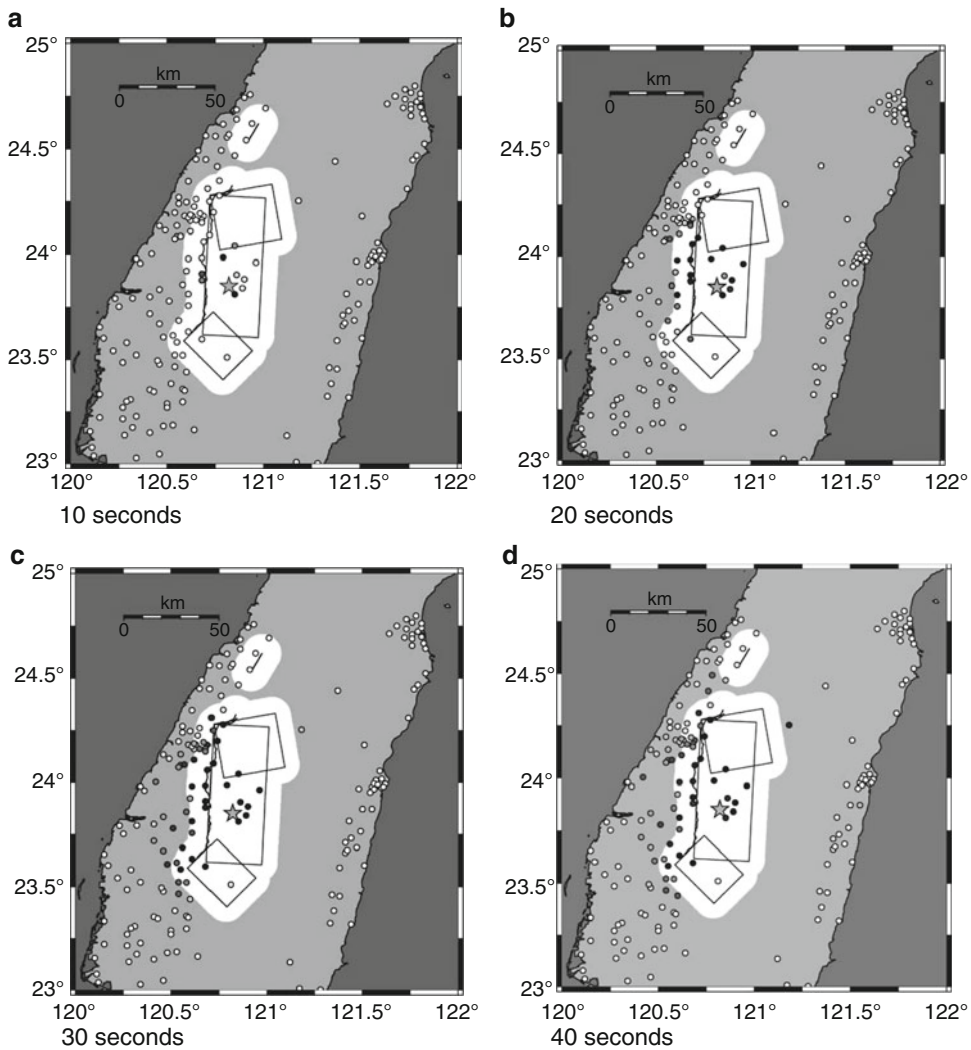
Source Characterization for Earthquake Early Warning, Fig. 7 Snapshots showing the recording stations identified as near-source sites in the case of the 1999 M7.0 Chi-Chi (Taiwan) earthquake. The *star* marks the earthquake epicenter. *Circles* represent the recording stations. *White* zones correspond to the source area and *rectangle* s identify the surface projection of the causative faults. The darker is the marker, the higher is the probability that the corresponding site is located in near source (After Yamada et al. (2007))

the 1999, M7.0, Chi-Chi (Taiwan) earthquake. As time goes on, the near-fault stations identify quite well the fault extent.

In a recent paper, Yamada (2014) proposed an improved version of this discriminant. This approach identifies the fault rupture geometry by classifying stations into near-source and faraway source and provides reasonably good estimates of the extent of the near-source area.

Potential Damaged Zone

As for the point-source analysis, on-site approaches can be used to predict finite source effects, such as directivity and azimuthal changes on the ground motion directly analyzing the early motion at the single sites, without explicitly estimating any source parameters. The idea is to use the threshold-based approach to compute threshold levels at single stations and to interpolate the

levels to define the potential damage zone associated with the earthquake (Kanamori 2005; Zollo et al. 2010).

At a single site, the real-time measure of the P-wave parameters (P_d and τ_c) thus provide a rough but rapid alert notification, although no information is given about the accurate earthquake size and location. Considering that the same measure can be performed at different nodes of a dense array of stations, deployed nearby and far away from the earthquake epicenter, the mapping of the recorded maximum alert levels (3 or 2) provides a preliminary estimation of the extent of the potential damage zone (PDZ), i.e., the area where the damages are expected to be equal or greater to those predicted by the level VII according to the instrumental intensity scale.

The same authors proposed that a rough but very rapid estimation of the PDZ extent can also be made from the updated averages of the period parameter τ_c which is recorded at strong-motion sites in the near-source region. Fixing the parameter P_d at its threshold value and using the progressively updated estimations of the period parameter, the empirical equation

$$\log P_d = A + B \log \tau_c + C \log R \qquad (21)$$

(R is the hypocentral distance; A,B,C are determined by the regression analysis) can be used to determine the radius of the area within which the strong ground motion is expected to produce an instrumental intensity $I_{MM} > 7$. The off-line application of the method showed a very consistent match between the rapidly predicted (within a few sec from the first recorded P-wave) and observed damage zone, the latter being mapped from detailed macroseismic surveys a few days after the event. A more refined technique for the real-time mapping of the PDZ has been proposed by Colombelli et al. (2012a), who tested the method on data from ten M > 6 Japanese earthquakes occurred in the period 2000–2009 and recorded at the K-Net/Kik-Net Japan accelerometric arrays. The method is based on the recursive use of the ground-motion prediction equation for P_d (Eq. 21), with real-time, updated estimate of the earthquake location and

characteristic period τ_c. The area covered by stations is divided into cells, using a prefixed spatial grid, which is needed to fill the gaps where stations are not available. At those stations where the first 3 s of signal after the P-picking are available, an alert level is locally assigned, based on P_d and τ_c measurements. At the same time, the event location is obtained by using the available P-picks and a real-time location method. Furthermore, the expected P_d value can be predicted at each node of the grid, through Eq. 21 and the τ_c averaged over the available measurements at the considered time step. Measured and predicted P_d values are thus interpolated; the area within which the highest level of damage is expected can be delimited by the isoline corresponding to $P_d = P_d^{th} = 0.2$ cm.

The predicted PGV values at stations and at grid nodes can be computed using the equation relating $\log(PGV)$ and $\log(P_d)$ (Eq. 1 in Zollo et al. 2010) and finally converted into an instrumental intensity measure using the relationship of Wald et al. (1999). An interpolation is performed between all intensity values to produce a real-time, continuously updated, intensity map (Fig. 5b,c). This procedure is repeated every time step (typically 1 s); as the waves propagate within the area and trigger other stations, the event location is refined by using P-pickings, average τ_c value is updated, and more data are used for the interpolation procedure. The studied cases for Japan earthquakes displayed a very good matching between the rapidly predicted earthquake potential damage zone inferred from initial P-peak displacement amplitudes and the instrumental intensity map, the latter being mapped after the event, using peak ground velocity and/or acceleration, or from field macroseismic surveys. The performance of the method has been evaluated by defining *successful*, *missed*, and *false alarms* in terms of observed versus predicted instrumental intensities for all the analyzed events and by counting their relative percentage. A very high percentage (88 %) of alert levels has been correctly assigned, and most of maximum alert levels correspond to the area within which the highest level of damage has been observed.
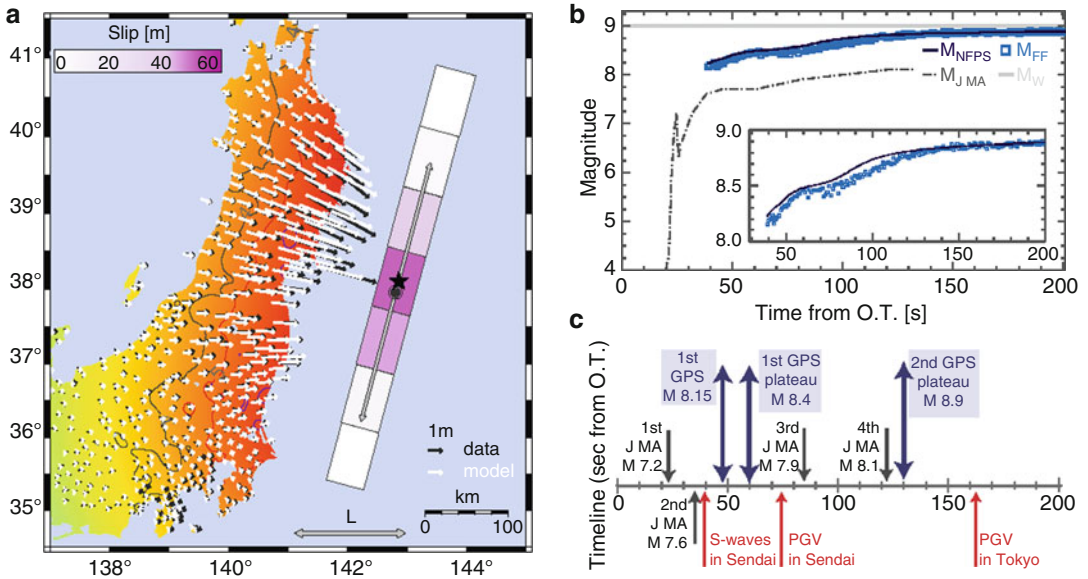
Source Characterization for Earthquake Early Warning, Fig. 8 (**a**) Example of snapshot of the GPS-based strategy for the Mw 9.0, 2011 Tohoku-Oki earthquake. The background color represents the predicted intensity distribution using the final magnitude value and the distance from the finite fault. The purple color scale shows the slip distribution on the seven-patch slip model. The total length estimate (L) is also plotted as vectors on the fault plane. Black vectors represent the observed horizontal offset while white vectors show the static displacement resulting from the inversion algorithm. (**b**) Magnitude estimate as a function of time from the origin time. The magnitude is obtained by using the static displacement

provided by GPS data using the near-field, point source approximation (dark blue solid line) and resulting from the slip inversion (small blue squares). For comparison, the evolution of magnitude estimate provided by the JMA early warning system is also shown as a dotted gray line, and the continuous gray line represents the real moment magnitude value. (**c**) Warning timeline for the Tohoku-Oki earthquake showing when the GPS information is available with respect to the time at which the strongest shaking occurs in the Sendai and Tokyo regions and with respect to the JMA warnings. After Colombelli et al (2013)

Summary

An approach that may be more robust is the inversion for the final slip on the fault plane, which allows consideration of the contributions from the entire fault plane and may provide a more realistic estimate of both size and potential damage of the ongoing event.

One way to recover the slip distribution on the fault plane is to use the permanent ground deformation, which is directly related to the earthquake magnitude. The static component of ground motion could, in principle, be obtained using dynamic range, accelerometric sensors, which are able to record unsaturated signals in a broad range of frequencies $(0 \div 100$ Hz). Accelerometric records are integrated twice to obtain displacement time series. Unfortunately, for near-field records, this operation may introduce artificial effects and long-period drifts (Boore et al. 2002), which are usually removed by applying a high-pass filter. The application of the filter, while removing the artificial distortions, reduces the low-frequency content of the recorded waveforms, resulting in the complete loss of the low-frequency energy radiated by the source and of the static offset, which is the most relevant piece of information for a large earthquake.

GPS stations are able to provide a direct and evolutionary measurement of the permanent ground deformation, i.e., of the resulting co-seismic displacement after the dynamic vibration has finished. With the increasing diffusion of

high-rate 1Hz GPS stations, the seismological community has begun looking at GPS data as a valid complement to the high-frequency information provided by seismic data.

Once the permanent static offset is extracted from GPS displacement time series, a real-time static slip inversion scheme can be used to infer the slip distribution on the fault plane for both the rapid determination of the event size and for the near-real-time estimation of the rupture area (Fig. 8). Many authors have recently started applying GPS data to EEW (Allen and Ziv 2011; Crowell et al. 2012; Wright et al. 2012; Ohta et al. 2012; Colombelli et al. 2013); they showed that 1 Hz GPS data provide a rapid and remarkably robust magnitude estimate and can be used for the real-time estimation of the rupture area, which, in turns, would allow for an improved prediction of the earthquake damaging potential.

Another possibility would be a fast kinematic inversion of the rupture process, searching for the fault geometry, the focal mechanism, and the slip distribution on the fault plane. A first approach was proposed by Dreger et al. (2005), who combined preliminary information on location and magnitude with scaling laws for source size to determine both the focal mechanism and a kinematic rupture model from inversion of strong-motion integrated displacement waveforms. The approach revealed useful for estimating the ground shaking in the near-fault domain for the 2003, M 6.5, San Simeon, California, earthquake. Nevertheless, computational times were estimated however of the order of minutes, making the results available in near real time but not in real time.

Cross-References

References

Allen RM, Kanamori H (2003) The potential for earthquake early warning in southern California. Science 3:685–848

Allen RM, Ziv A (2011) Application of real-time GPS to earthquake early warning. Geophys Res Lett 38, L16310. doi:10.1029/2011GL047947

Allen RM, Brown H, Hellweg M, Khainovski O, Lombard P, Neuhauser D (2009) Real-time earthquake detection and hazard assessment by ElarmS across California. Geophys Res Lett 36, L00B08. doi:10.1029/2008GL036766

Boore DM, Stephens CD, Joyner WB (2002) Comments on baseline correction of digital strong-motion data: examples from the 1999 hector mine, California, earthquake. Bull Seism Soc Am 92:1543–1560. doi:10.1785/0120000926

Böse M, Ionescu C, Wenzel F (2007) Earthquake early warning for Bucharest, Romania: novel and revised scaling relations. Geophys Res Lett 34, L07302. doi:10.1029/2007GL029396

Colombelli S, Amoroso O, Zollo A, Kanamori H (2012a) Test of a threshold-based earthquake early warning using Japanese data. Bull Seismol Soc Am 102:1266–1275. doi:10.1785/0120110149

Colombelli S, Zollo A, Festa G, Kanamori H (2012b) Early magnitude and potential damage zone estimates for the great Mw 9 Tohoku-Oki earthquake. Geophys Res Lett 39, L22306. doi:10.1029/2012GL053923

Colombelli S, Allen RM, Zollo A (2013) Application of real-time GPS to earthquake early warning in subduction and strike-slip environments. J Geophys Res 118:3448–3461. doi:10.1002/jgrb.50242

Crowell BW, Bock Y, Melgar D (2012) Real-time inversion of GPS data for finite fault modeling and rapid hazard assessment. Geophys Res Lett 39, L09305. doi:10.1029/2012GL051318

Cua G, Fisher M, Heaton T, Wiemer S (2009) Real-time performance of the virtual seismologist earthquake early warning algorithm in southern California *seism*. Res Lett 80:740–747. doi:10.1785/gssrl.80.5.740

Dreger DS, Gee L, Lombard P, Murray MH, Romanowicz B (2005) Rapid finite-source analysis and near-fault strong ground motions: application to the 2003 M_w 6.5 San Simeon and 2004 M_w 6.0 Parkfield Earthquakes. Seism Res Lett 76:40–48. doi:10.1785/gssrl.76.1.40

Espinosa-Aranda JM, Cuéllar A, Rodríguez FH, Frontana B, Ibarrola G, Islas R, García A (2011) The seismic alert system of Mexico (SASMEX): progress and its current applications. Soil Dyn Earth Eng 31:2154–2162. doi:10.1016/j.soildyn.2010.09.011

Festa G, Lancieri M, Zollo A (2008) Magnitude estimation from early radiated energy. Geophys Res Lett 35, L22307. doi:10.1029/2008GL035576

Kanamori H (2005) Real-time seismology and earthquake damage mitigation. Ann Rev Earth Planet Sci 33: 195–214

S

Lancieri M, Zollo A (2008) A Bayesian approach to the real time estimation of magnitude from the early P- and S-wave displacement peaks. J Geophys Res 113, B12302. doi:10.1029/2007JB005386

Nakamura Y (1988) On the urgent earthquake detection and alarm system (UrEDAS) In: Proceedings 9th world conference on earthquake engineering, Tokyo-Kyoto, Japan vol 7, pp 673–678

Ohta Y et al (2012) Quasi real-time fault model estimation for near-field tsunami forecasting based on RTK-GPS analysis: application to the 2011 Tohoku-Oki earthquake (Mw 9.0). J Geophys Res 117:B02311. doi:10.1029/2011JB008750

Satriano C, Lomax A, Zollo A (2008) Real-time evolutionary earthquake location for seismic early warning. Bull Seimc Soc Am 98(3):1482–1494

Satriano C, Elia L, Martino C, Lancieri M, Zollo A, Iannaccone G (2010) PRESTo, the earthquake early warning system for southern Italy: concepts, capabilities and future perspectives. Soil Dyn Earthq Eng 31(2):137–153. doi:10.1016/j.soildyn.2010.06.008

Wald DJ, Quitoriano V, Heaton T, Kanamori H (1999) Relationships between peak ground acceleration, peak ground velocity and modified mercalli intensity in California. Earthq Spectra 15:557–564

Wright TJ, Houlié N, Hildyard M, Iwabuchi T (2012) Real-time, reliable magnitudes for large earthquakes from 1 Hz GPS precise point positioning: the 2011 Tohoku-Oki (Japan) earthquake. Geophys Res Lett 39, L12302

Wu YM, Kanamori H (2005) Rapid assessment of damage potential of earthquake in Taiwan from beginning of P-waves. Bull Seism Soc Am 95:1181–1185. doi:10.1785/0120040193

Wu YM, Kanamori H (2008) Development of an earthquake early warning system using real-time strong motion signals. Sensors 8:1–9

Wu YM, Zhao L (2006) Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning. Geophys Res Lett 33, L16312. doi:10.1029/2006GL026871

Yamada M (2014) Estimation of fault rupture extent using near-source records for earthquake early warning. In Wenzel F, Zschau J (eds) Early warning for geological disasters. Advanced technologies in earth sciences. Springer, Berlin/Heidelberg, p 29ff. doi:10.1007/978-3-642-12233-0_2

Yamada M, Heaton T, Beck J (2007) Real-time estimation of fault rupture extent using neat-source versus far-source classification. Bull Seism Soc Am 97:1890–1910. doi:10.1785/0120060243

Zollo A, Lancieri M, Nielsen S (2006) Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion. Geophys Res Lett 33, L23312. doi:10.1029/2006GL027795

Zollo A, Amoroso O, Lancieri M, Wu YM, Kanamori H (2010) A threshold-based earthquake early warning using dense accelerometer networks. Geophys J Int 183:963–974

Zollo A, Colombelli S, Elia L, Emolo A, Festa G, Iannaccone G, Martino C, Gasparini P (2014) An integrated regional and on-site earthquake early warning system for southern Italy: concepts, methodologies and performances. In: Wenzel F, Zschau J (eds) Early warning for geological disasters - scientific methods and current practices. Springer-Verlag Berlin Heidelberg, Dordrecht

Spatial Filtering for Structural Health Monitoring

G. Tondreau[1] and A. Deraemaeker[2]
[1]Postdoctoral Researcher, Building Architecture and Town Planning (BATir), Brussels, Belgium
[2]FNRS Research Associate, Building Architecture and Town Planning (BATir), Brussels, Belgium

Synonyms

Control charts; Damage localization; Damage detection; Sensor networks; Spatial filtering; Structural health monitoring (SHM)

Introduction

Assessing the integrity of the structures in real time is a very important topic for which many methods have been developed in the last decades. Today, structural health monitoring (SHM) is gaining more and more attention: in the case of bridges, the maximum loads tend to increase (increase of the vehicle weights), while most of the structures are coming to the end of their theoretical lifetime. In addition, exceptional events such as collisions or earthquakes can cause more severe and fast deteriorations. Optimal maintenance calls for an early detection of small damage in structures, as it is well known that limited and frequent repairs are much less costly than major repairs or total rebuilding after

collapse. Current monitoring practice consists in scheduled maintenances including visual inspections, ultrasounds, eddy current, magnetic field, or radiography techniques (Hellier 2003). All these experimental methods require however that the vicinity of the flaw is known and that the proximity to be inspected is accessible. Moreover, these local inspections are labor intensive and therefore very expensive. A major problem is that traditional monitoring is noncontinuous which means that if a critical damage occurs between two inspections, it might lead to catastrophic structural failure. One of the most relevant examples is the I-35W Mississippi River Bridge case (Rofidal 2007): this bridge collapsed in August 2007 killing 13 people and injuring 145, despite annual inspection.

A general trend for new structures and bridges is a lighter and more slender design, which tends to increase the levels of vibrations under ambient excitation. While these levels of vibrations need to be controlled as they could be detrimental to the lifetime of the structure, they can also be used for the continuous monitoring of the structure without disruption or decrease of functionality. The basic idea is that the occurrence of damage alters the structural parameters which in turn affect the vibration characteristics. Vibration SHM of civil engineering structures relies on ambient vibrations, as artificial excitation of such large structures is often unpractical.

Based on this basic concept, many vibration-based SHM techniques have been developed in the last decades using mainly eigenfrequencies, damping ratios, or mode shapes (Doebling et al. 1998). The reason of this popularity is the ease of measuring modal parameters or frequency responses on real structures thanks to recent advances in sensors and sensing systems and in the development of efficient operational modal analysis techniques (Reynders and De Roeck 2008; Reynders et al. 2012). Such advances are so important that more and more very large bridges are instrumented with larger and larger sensor networks. China has been the driving force in this direction with the massive instrumentation of bridges in the Hong Kong area, the largest one

being the Stonecutters Bridge with more than 1,500 sensors among which are 58 accelerometers and 853 dynamic and static strain sensors (Ni et al. 2012). In Europe, the Messina bridge project, designed to be the largest cable-stayed bridge in the world would include a very large monitoring system with more than 3,000 sensors (De Neumann et al. 2011).

These technological advances have opened the way to real-time automated SHM of bridges. A major problem related to the use of such very large sensor networks is to find adequate techniques to post-process the data: intelligent methods are needed in order to take advantage of the enormous amount of information provided by these large networks. In fact, operational modal analysis is not yet fit for automated modal analysis using very large sensor networks.

With that perspective, rather than identifying online the full set of modal properties of the structure, an alternative is to condense the measured data while keeping the information about the potential damage occurring in the structure. The technology presented in this entry is spatial filtering, which consists in using a linear combiner to condense the information from very large networks of sensors into one or just a few "virtual sensors." Such virtual sensors can be designed in order to react strongly to damage while being insensitive to environmental changes or even to react to a damage occurring in a specific location along the bridge.

Rytter (Rytter 1993) has proposed a hierarchical decomposition of the SHM process in four levels, which has been widely accepted in the SHM literature: detection (level 1), localization (level 2), quantification (level 3) of the damage, and prediction of the remaining service life of the damaged structure (level 4). As the level of SHM increases, the knowledge about the damage increases and, usually, the complexity of the method increases as well. The method based on spatial filtering can deal with levels 1 and 2 by comparing data measured in the current unknown state with data measured from the structure assumed to be undamaged.

The general scheme of the method is shown in Fig. 1. It is divided in three parts: (i) the

SHM system
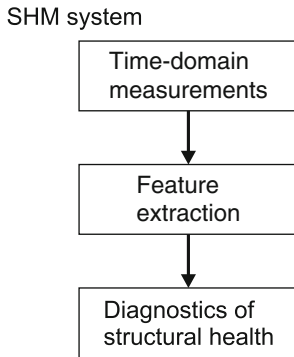
Spatial Filtering for Structural Health Monitoring, Fig. 1 General scheme of vibration-based SHM

measurement of raw time-domain data, including data reduction, (ii) the transformation of the data into information using feature extraction, and (iii) the diagnostics of the structural health based on the monitored features.

The first section of this entry deals with measurement of raw time-domain data and details the spatial filtering technique for data reduction. The second section is devoted to feature extraction, and the third section deals with the diagnostics, both for damage detection and localization. The last section presents an experimental illustration of the SHM technique on a 3.78 m long steel I-beam tested in the laboratory.

Measurement of Raw Time-Domain Data

The first building block of the SHM system is the sensor network. As stated earlier, a current trend is to implement very large sensor networks on large civil engineering structures. In such cases, it is often necessary to perform data reduction in order to decrease the power consumption and the bandwidth needed to transmit the data and to facilitate the data storage and post-processing. For SHM applications, an optimal reduction is one that significantly reduces the amount of data while keeping most of the information about the damage. The technique presented in this entry for data reduction is spatial filtering.

Spatial and Modal Filtering

Consider a structure excited with an ambient force $f(t)$ equipped with a network of n sensors whose time-domain output is denoted by $y_k(t)$ as shown in Fig. 2. The dynamic time-domain response at each sensor can be decomposed into a sum of contributions of the N mode shapes excited by the ambient force:

$$y_k(t) = \sum_{i=1}^{N} a_i(t)\phi_{ki} \tag{1}$$

where $a_i(t)$ is the modal amplitude of mode i and ϕ_{ki} is the projection of mode shape i on sensor k. The application of spatial filtering with coefficients α_k leads to a single sensor output $g(t)$:

$$g(t) = \sum_{k=1}^{n} a_k y_k(t) = \sum_{i=1}^{N}\sum_{k=1}^{n} a_k \phi_{ki} a_i(t) \tag{2}$$

The general scheme of spatial filtering can be used for different purposes. A first idea is to condense the information into the modal coordinates of the undamaged structure, which allows to reduce the data $y_k(t)$ from a very large network of n sensors to a limited set of $a_i(t)$ time series from N modal sensors. This idea is motivated by the fact that the vibration of structures typically involves only a few mode shapes which are excited in a given frequency band of interest.

To design a modal filter, the vector of the linear combiner α_k must be orthogonal to all the modes of the structure in a frequency band of interest, except mode l:

$$\sum_{k=1}^{n} \alpha_k \phi_{ki} = \delta_{li} \tag{3}$$

Equation 2 then reduces to

$$g(t) = a_l(t) \tag{4}$$

Equation 3 can be written in a matrix form:

$$[C]^T \{\alpha\} = e_l, \tag{5}$$

Spatial Filtering for Structural Health Monitoring, Fig. 2 Principle of spatial filtering on a network of n sensors
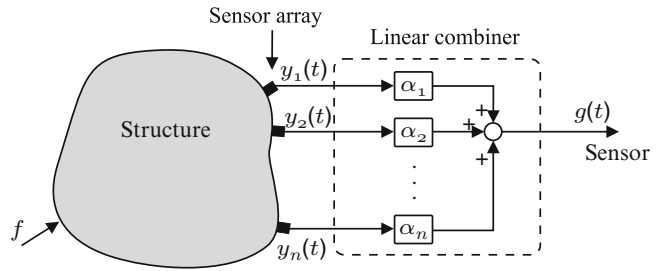
where $\{\alpha\} = \{\alpha_1 \cdots \alpha_n\}^T$ and $e_l = \{0\ 0\ \cdots\ 1 \cdots\ 0\ 0\}^T$ (all components set to 0 except the lth component) and $[C]$ is a matrix whose columns correspond to the N mode shapes projected on the n sensors of the array. Equation 5 can be solved only if there are at least as many sensors as there are mode shapes ($n \geq N$). In this case, matrix $[C]$ is rectangular and the system of equations is overdetermined (several solutions exist which satisfy Eq. 5). The minimum norm solution is usually adopted by computing the pseudo-inverse (regularized with singular value decomposition) of $[C]^T$ (Deraemaeker et al. 2008).

The modal filter can be tuned to any of the N mode shapes in the frequency band of interest. In the frequency domain, the FRF of the modal filter tuned on mode l is given by

$$G(\omega) = \frac{b_l}{\left(\omega_l^2 - \omega^2 + 2j\xi_l\omega_l\omega\right)}, \qquad (6)$$

where b_l depends on the excitation level and position. It corresponds to the FRF of a single degree of freedom system which presents a single peak at frequency ω_l.

Using modal filtering, the amount of data from a large network of n sensors can be significantly reduced to just a few modal filters. Such modal filters are virtual sensors which measure the amplitude of vibration of each mode separately. Typically, for a bridge excited by the ambiance, only a few (up to ten) mode shapes are relevant, and the information can be drastically reduced. The computation of the linear combiner coefficients α_k is based on Eq. 5 which requires the knowledge of matrix $[C]$. This matrix should be built using experimentally identified mode shapes in order to avoid the need for
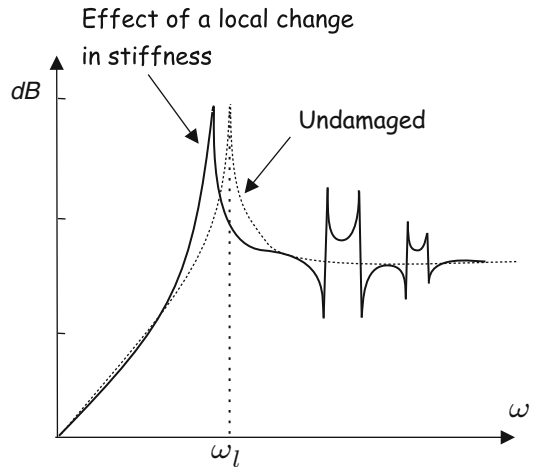
Spatial Filtering for Structural Health Monitoring, Fig. 3 Effect of a structural change on the modal filter tuned on mode l (Deraemaeker et al. 2008)

a numerical model of the structure to be monitored. This can be achieved thanks to operational modal analysis, using, for instance, stochastic subspace-based methods (Reynders and De Roeck 2008).

Effect of Damage on Modal Filters

Suppose now that damage initiates in the structure. This damage will alter the stiffness matrix, affecting the eigenfrequencies and mode shapes of the structure. The change of mode shapes will be reflected in matrix $[C]^T$ so that Eq. 5 will now be violated. In other terms, the damage will alter the mode shapes, and the coefficients of the linear combiner will not be tuned anymore. This will result in the reappearance of the filtered peaks, as illustrated in Fig. 3. This is the central idea of vibration-based SHM based on modal filters, as detailed in Deraemaeker et al. (2008).

Spatial Filtering for Structural Health Monitoring, Fig. 4 Example of I_{Peak} values for an increasing peak
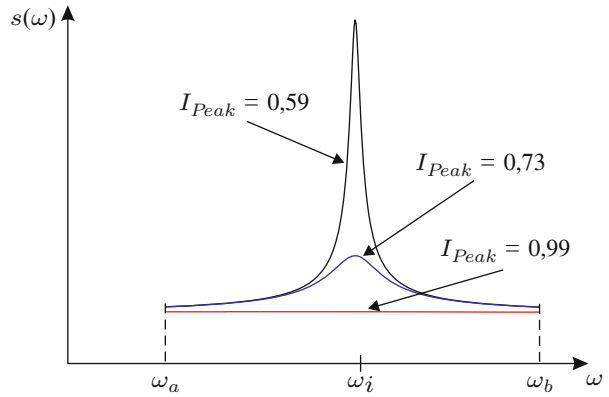
From Data to Information

The raw time-domain output of modal filters is not exploitable as such, as the changes in the time-domain response due to damage will be very small. The transformation in the frequency domain is an important step which allows to enhance these small changes by focusing on the frequency bands away from the main peak where the filtered peaks will reappear. In the case of ambient vibrations, the input force is not known, and the power spectral density (PSD) $S_{gg}(\omega)$ of $g(t)$ should be computed. This quantity is directly related to the amplitude of the FRF $G(\omega)$ as follows (Ewins 1984):

$$S_{gg}(\omega) = |G(\omega)|^2 S_{ff}(\omega) \qquad (7)$$

This equation shows that if peaks are filtered in the FRF $G(\omega)$, they will be filtered in $S_{gg}(\omega)$ as well. When damage occurs, the reappearance of spurious peaks should be monitored based on the power spectral density of the output of the modal filter $S_{gg}(\omega)$.

Feature Extraction

Because the spurious peaks are expected to appear around the initial eigenfrequencies of the structure, the strategy consists in extracting one feature in each frequency band around them. Let $s(\omega)$ be the frequency dependent amplitude in the frequency range (ω_a, ω_b) (Fig. 4). For ambient vibrations, $s(\omega)$ is the PSD $S_{gg}(\omega)$.

The frequency band is typically defined by $\omega_a = 0.95\omega_i$ and $\omega_b = 1.05\omega_i$, where ω_i is the angular eigenfrequency. A peak indicator is then computed in this frequency interval:

$$I_{\text{Peak}} = \frac{2\sqrt{3RVF}}{\omega_b - \omega_a}, \qquad (8)$$

where RVF is the root variance frequency defined by

$$RVF = \sqrt{\frac{\int_{\omega^a}^{\omega_b} (\omega - FC)^2 s(\omega) d\omega}{\int_{\omega^a}^{\omega_b} s(\omega) d\omega}}, \qquad (9)$$

and FC is the frequency center defined by

$$FC = \frac{\int_{\omega^a}^{\omega_b} \omega s(\omega) d\omega}{\int_{\omega^a}^{\omega_b} s(\omega) d\omega} \qquad (10)$$

Theoretically, I_{Peak} is equal to 1 if $s(\omega)$ is constant and decreases when the peak grows. Figure 4 gives an example of I_{Peak} values computed between ω_a and ω_b when a spurious peak grows around ω_i. The advantage of that feature is that it is very sensitive to the peak growth but not to the level of the excitation force, which is particularly needed for ambient vibrations. More details on the peak indicator computation can be found in Deraemaeker and Worden (2010).
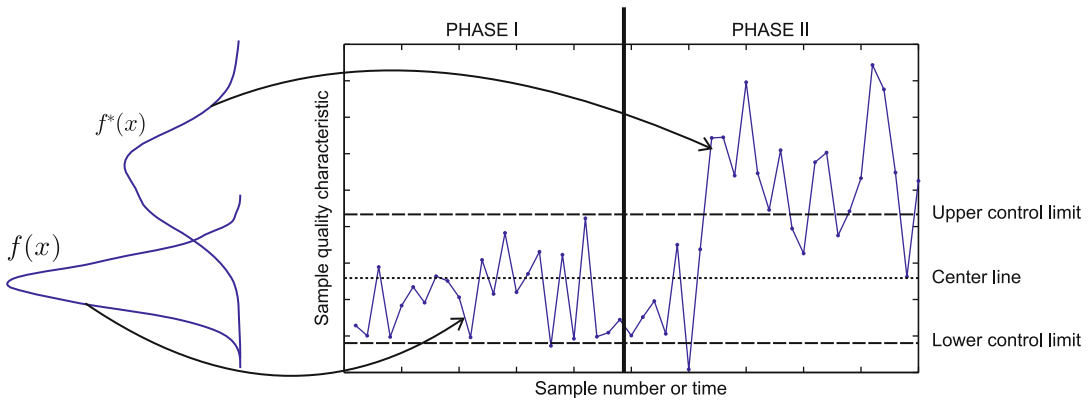
Spatial Filtering for Structural Health Monitoring, Fig. 5 A typical control chart

Diagnostics

The last building block of the SHM system is the diagnostics. It consists in assessing, based on the monitored features (here the peak indicators) whether the structure is healthy or damaged, and possibly in giving indication on the location of damage.

Statistical Analysis of the Features

When the excitations are random, the peak indicators behave like random variables. They will therefore follow a statistical distribution which can be inferred from several undamaged samples. Many tools have been developed to detect a change in that statistical distribution such as outlier analysis or hypothesis testing. In this contribution, control charts (Montgomery 2009; Ryan 2000) are presented. This tool of statistical quality control plots the features or quantities representative of their statistical distribution as a function of the samples. Different univariate or multivariate control charts exist but all these control charts are based on the same principle which is summarized in Fig. 5.

In phase I, a set of samples are collected and analyzed to infer statistical characteristics of the process when it is assumed to be in control (i.e., when the structure is undamaged). The aim of this step is to compute the control limits (upper control limit *UCL* and/or lower control limits *LCL*) between which the feature should be included if

the process stays in control. Those limits are governed by the statistical distribution $f(x)$ of the quality characteristic and the probability $1 - \gamma$ that any in-control sample will fall inside the control limits. There are control limits that can be computed to detect a shift of the mean value of the process or a shift of the variance of the process.

Once a set of reliable control charts has been established (phase I), the process is under monitoring (phase II). The process state is unknown (it might be in or out of control), and if a sample falls outside the control limits previously computed, it is considered as an abnormal value, and a warning is triggered. Phase I fixes the probability of type I (false alarms) and type II (missing alarms) errors. Because the control limit values are based on the number of samples in the in-control set of data, the statistical distribution $f(x)$, and the γ value, the statistical analysis must be done very carefully.

Typically, one can find two families of control charts in the literature: the univariate control charts and the multivariate control charts. The first family will be used if there is only one feature to be monitored, while the second one is used when several features are monitored at the same time. In the present application, this means that if one checks the appearance of only one spurious peak around one given natural frequency, the univariate control chart will be applied on that feature, while a multivariate

control chart will be used if one checks the spurious peaks around several eigenfrequencies in each modal filter. Finally, there are two categories of control charts in each family: the Shewhart control chart and the time-weighted control charts. The first category monitors each sample independently while the second category considers the previous samples to monitor the current sample, which allows to detect smaller shifts. It has been found that the best results are obtained when the Shewart control charts are applied, because the time weighted control charts increased too much the number of type I errors. For this reason, only the univariate and the Hotelling T^2 control charts are presented.

Individual control chart Consider that only one feature x following a normal distribution is monitored (e.g., one I_{Peak} value in each modal filter). The individual control chart will monitor that individual feature x. The control limits are

$$\text{Upper Control Limit: } UCL = \bar{x} + 3\frac{\overline{MR}}{d_2}$$
$$\text{Lower Control Limit: } LCL = \bar{x} - 3\frac{\overline{MR}}{d_2} \quad (11)$$

If the number of samples in phase I is n, then

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad (12)$$

$$\overline{MR} = \frac{1}{n}\sum_{i=2}^{n} |x_i - x_{i-1}|, \quad (13)$$

and $d_2 = 1.128$. In fact, $\frac{\overline{MR}}{d_2}$ is an estimate of the standard deviation σ of x assumed to follow a normal distribution in phase I. Equation 11 is therefore based on a choice of $\gamma = 0.027$. The individual control chart is designed to detect a shift of $\bar{x}$.

Hotelling T^2 control chart If two or more features are monitored at the same time, monitoring these two quantities independently by applying two or more univariate control charts can be very misleading, especially if those features are correlated. On the opposite, the Hotelling T^2 control chart is designed for the monitoring of several features simultaneously. Consider p features following a p-normal distribution. The Hotelling T^2 control chart monitors the Mahalanobis distance T^2:

$$T^2 = (x - \bar{x})^T \sum{}^{-1}(x - \bar{x}), \quad (14)$$

where Σ is the $p \times p$ estimated covariance matrix of features, x is the current $p \times 1$ feature vector, and $\bar{x}$ is the $p \times 1$ vector of estimated mean values of x vectors (only the undamaged samples are considered to obtain Σ and $\bar{x}$). Since the Mahalanobis distance is always positive, only the upper control limit UCL based on an F distribution is considered:

$$UCL = \frac{p(m+1)(m-1)}{(m^2 - mp)}F_{\gamma, p, m-p}, \quad (15)$$

where p is the number of variables, m is the number of samples in the set of data in phase I, and γ is such that there is a probability of $1 - \gamma$ that any in-control sample will fall between the control limits. Like the individual control chart, the Hotelling T^2 control chart detects a change of $\bar{x}$.

Damage Detection and Localization

When condensing all the sensors into a single virtual modal sensor as shown in Fig. 2, the statistical analysis can only give an indication on the deviation from the normal condition on the structure as a whole, leading to damage detection. The methodology can be extended to damage localization: consider now that the n sensors installed on the structure are grouped in several smaller sensor networks, each consisting of m sensors. Modal filters can be built for each of these local sensor networks resulting in independent *local modal filters* (Fig. 6).

If the local network I contains sensors $y_{1,I}, \ldots, y_{m,I}$, the output of its modal filter tuned to mode l is given by

$$g_I(t) = \sum_{k=1}^{m} \alpha_{k,I} y_{k,I}(t), \quad (16)$$
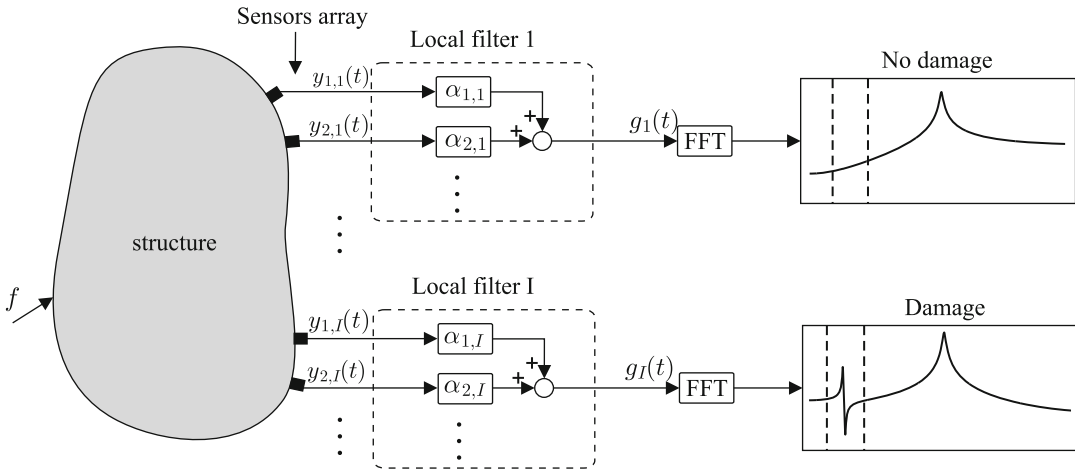
Spatial Filtering for Structural Health Monitoring, Fig. 6 Principle of damage localization using local modal filters

where the $\alpha_{k,I}$ coefficients are computed in order to satisfy the following condition:

$$\sum_{k=1}^{m} \alpha_{k,I}\phi_{(k,I)i} = \delta_{li}, \qquad (17)$$

where $\phi_{(k,I)i}$ is the kth ($k = 1, \ldots, m$) component of the ith mode shape projected on the Ith local sensor network. If a damage occurs under spatial filter I and if the sensor responses are locally sensitive to damage, the mode shape will only be altered in that spatial filter. As a result, only the spatial filter I will have spurious peaks, indicating the location of the damage.

The efficiency of the approach relies therefore on a very strong assumption: damage in a local filter will cause a local change of the mode shape which is limited to the very close vicinity of the damage location. The fulfillment of this requirement depends on the type of measured quantity which is considered. Two different approaches coexist in the literature: in Mendrok and Uhl (2010), accelerometers are used, while in Tondreau and Deraemaeker (2013), dynamic strain sensors are used. These two approaches have been compared in Tondreau and Deraemaeker (2011), showing that the method based on strain sensors has a higher sensitivity to damage and better

localization capabilities. This highlights the importance of the choice of the type of sensor which is part of the first building block of the SHM system. Applications of spatial filtering techniques have been limited so far to accelerometers or dynamic strain sensors.

Illustrative Example

Description of the Case Study
The experimental application consists in a 3.78 m long steel I-beam which is bolted on two big concrete cubes. The structure is excited with a Modal 110 electrodynamic shaker from MB Dynamics, and a network of 20 13 mm × 170 mm × 50 μm low-cost PVDF sensors have been fixed with double-coated tape, providing a continuous measurement of the dynamic strains along the beam between sensors 1 and 20. A National Instrument PXIe-1082 data acquisition system is used to measure the sensor responses with a sampling frequency of 6,400 Hz, as well as to generate a band-limited white noise between 0 and 500 Hz (not measured) which drives the shaker. Figure 7 shows the experimental setup as well as the definition of the PVDF sensors (accelerometers installed for preliminary tests can also be seen, but are not used in the present study).
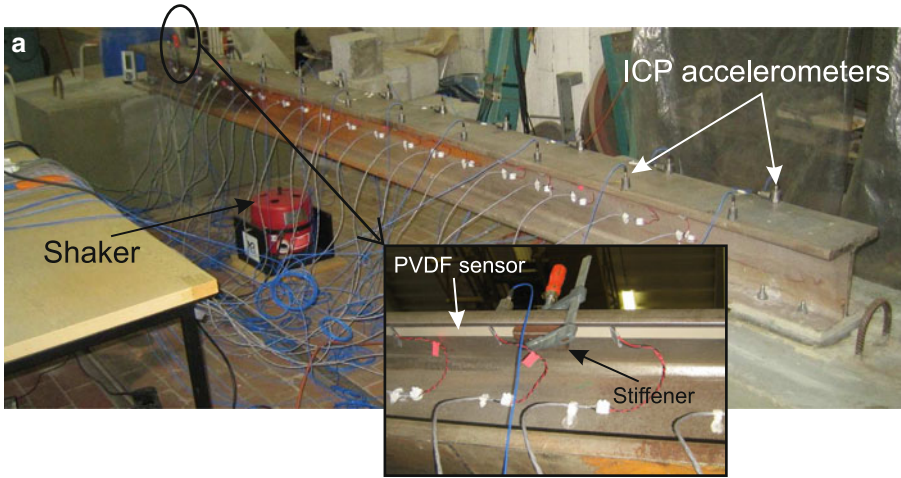
Test setup.
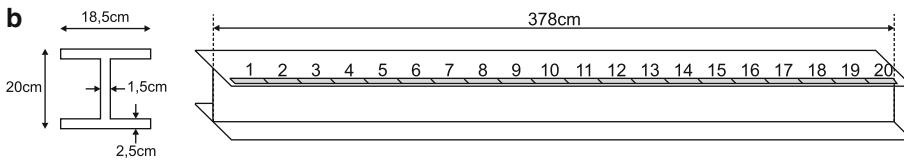
Artist front view: definition of the PVDF sensors.

Spatial Filtering for Structural Health Monitoring, Fig. 7 Experimental setup: 3.78 m steel I-beam equipped with 20 dynamic strain sensors (PVDF) for damage localization

The damage is introduced by fixing a very small steel stiffener (35 mm × 65 mm × 17 mm) directly against one of the PVDF sensors (Fig. 7a) in order to induce a local change of stiffness at that position. It has been checked that such a local change of stiffness induces a local change of strain similar to what happens with damage. The network of 20 sensors is split in five local filters of five sensors, with a small overlap: (*i*) [1:5], (*ii*) [4:8], (*iii*) [8:12], (*iv*) [12:16], and (*v*) [16:20]. The damage scenarios are described in Table 1. For the undamaged case, the measurement is performed 350 times in order to infer the statistical properties of the peak indicators. For each damaged case, 50 measurements are performed. Each measurement is referred to as a statistical *sample*.

The modal filters tuned on the two first bending mode shapes of the beam at 64 and 230 Hz are applied for each of the five local filters separately. The feature vector therefore consists in two peak indicators in each local filter (appearance of peak

Spatial Filtering for Structural Health Monitoring, Table 1 Damage scenarios

Case	Samples	Location of damage (sensor)	Location of damage (local filter)
1	1–350	None	None
2	351–400	1	[1:5]
3	401–450	2	[1:5]
4	451–500	3	[1:5]
5	501–550	4	[1:5] and [4:8]
6	551–600	5	[1:5] and [4:8]

at 64 Hz for the modal filter tuned to 230 Hz, and appearance of peak at 230 Hz for the modal filter tuned to 64 Hz). Note that the peak indicator used in this example is slightly different from the peak indicator presented in section "Spatial and Modal Filtering" but shares similar properties. As the feature vector is multivariate, the Hotelling T^2 control chart has been applied to automate the damage localization in each local filter. The first 200 undamaged samples have been considered to
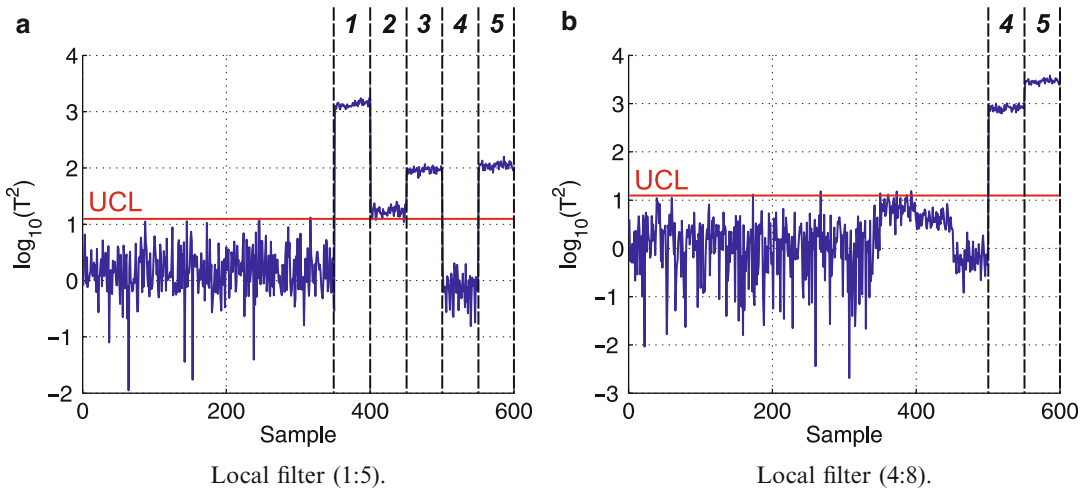
Spatial Filtering for Structural Health Monitoring, Fig. 8 Automated damage detection in local filter (1:5) and local filter (4:8)
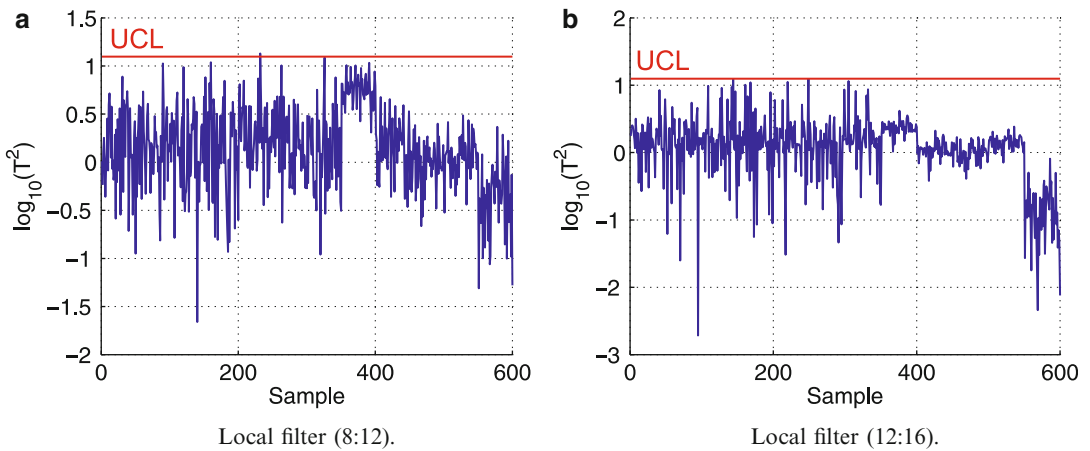
Spatial Filtering for Structural Health Monitoring, Fig. 9 Automated damage detection in local filter (8:12) and local filter (12:16)

estimate the covariance matrix, as well as to compute the control limit (γ is fixed to 0.25 %).

Figures 8, 9, and 10 show the Hotelling T^2 control chart. There is only one missing alarm in local filter [1:5] for a damage at sensor four. However, that missing alarm is compensated thanks to the overlapping of the local filters. Indeed, the damage at sensors four and five is correctly located in local filter [4:8]. The results show that the method has successfully, and automatically, localized all the damage cases.

Summary

There is a strong incentive for the development of online automated SHM techniques for large civil infrastructures. The objective of such systems is to be able to assess the structural integrity of safety critical civil infrastructure in real time. This is particularly important to detect the onset of damage due to aging or more severe damage due to accidental event such as an earthquake or a collision.
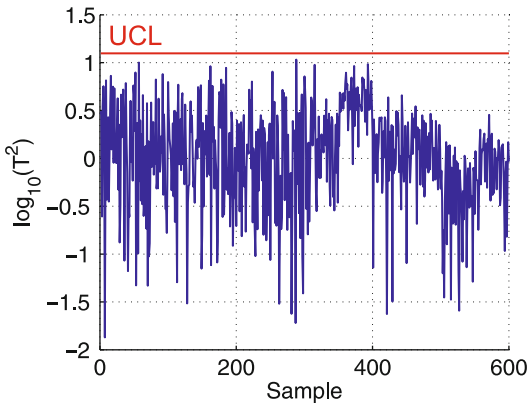
Spatial Filtering for Structural Health Monitoring, Fig. 10 Automated damage detection in local filter (16:20)

With the deployment of very large sensor networks on structures, alternatives to modal identification techniques can be interesting when the focus lies in fast and efficient damage detection and localization. This is the aim of the method presented in this chapter. The three important ingredients are (i) the use of a linear combiner to perform data reduction in the time domain through modal filtering, (ii) the transformation of the time-domain output of modal filters to the frequency domain and the subsequent feature extraction to detect the appearance of spurious peaks, and (iii) the use of control charts to automate the damage detection and localization process.

The fully integrated and automated methodology allows to process efficiently data from large sensor networks and to condense it into very limited information for diagnostics in the form of control charts. The efficiency of this technology has been illustrated on a laboratory experiment of a 3.78 m long steel I-beam.

Cross-References

▶ Operational Modal Analysis in Civil Engineering: An Overview
▶ Stochastic Structural Identification from Vibrational and Environmental Data
▶ System and Damage Identification of Civil Structures
▶ Vibration-Based Damage Identification: The Z24 Bridge Benchmark

References

De Neumann S, Andersen JE, Enckell M, Vullo E (2011) Messina bridge – structural health monitoring system. In: Proceedings IABSE-IASS conference. London

Deraemaeker A, Worden K (2010) New trends in vibration based structural health monitoring. CISM courses and lectures, vol 520. Springer Wien, New York

Deraemaeker A, Reynders E, De Roeck G, Kullaa J (2008) Vibration-based structural health monitoring using output-only measurements under changing environment. Mech Syst Signal Pr 22:34–56

Doebling SW, Farrar C, Prime MB (1998) A summary review of vibration-based damage identification methods. Shock Vib Dig 30:91–105

Ewins DJ (1984) Modal testing: theory and practice. Research Studies Press, Letchworth

Hellier C (2003) Handbook of nondestructive evaluation. McGraw-Hill, New York

Mendrok K, Uhl T (2010) Experimental verification of the damage localization procedure based on modal filtering. Struct Health Monit Int J 20:1615–1630

Montgomery DC (2009) Statistical quality control: a modern introduction. Wiley, New York

Ni YQ, Wong KY (2012) Integrating bridge structural health monitoring and condition-based maintenance management. In: Proceedings of the CSHM-4. Berlin

Reynders E, De Roeck G (2008) Reference-based combined deterministic-stochastic subspace identification for experimental and operational modal analysis. Mech Syst Signal Pr 22:617–637

Reynders E, Houbrechts J, De Roeck G (2012) Fully automated (operational) modal analysis. Mech Syst Signal Pr 29:228–250

Rofidal K (2007) U.S. coast guard responds following the collapse of the I-35W bridge in Minnesota. USCG Reservist 54:26–29

Ryan TP (2000) Statistical methods for quality improvement. Wiley, New York

Rytter A (1993) Vibration based inspection of civil engineering structures. PhD thesis, University of Aalborg

Tondreau G, Deraemaeker A (2011). Comparison of damage localization based on modal filters using strain measurements and acceleration measurements. In IWSHM2011, Standford, Sept 2011

Tondreau G, Deraemaeker A (2013) Local modal filters for automated data-based damage localization using ambient vibrations. Mech Syst Signal Pr 39: 162–180

Spatial Variability of Ground Motion: Seismic Analysis

Mounir Khaled Berrah[1], Djawad Zendagui[2] and Mustapha Djafour[2]
[1]Ecole Nationale Polytechnique, Algiers, Algeria
[2]Risam Laboratory, Faculty of Technology, University Abou Bakr Belkaïd, Tlemcen, Algeria

Synonyms

Coherency; Dynamic; Pseudo-static; Seismic analysis; Spatial variability; SVGM

Introduction

The seismic design of structures considers that all ground supports are subjected simultaneously to identical seismic excitation. This assumption is not valid for extended structures since it has been recognized that seismic motion exhibits no negligible variability. This variability, called spatial variability of ground motion (SVGM), must hence be considered in the design of extended structures.

Many questions arise: what is SVGM and how does one model it? Does SVGM increase or decrease internal forces of a structure? Is there any general conclusion that can be drawn while considering SVGM?

This entry aims at answering these questions. The first part of this entry addresses the description of SVGM through models in terms of coherency function. Subsequently, the response of structures subjected to SVGM is derived. At this stage, the concept of pseudo-static and dynamic responses is introduced. To understand the effect of the SVGM on structures, two case studies are presented: a single degree of freedom (SDOF) system and a multi-degree of freedom (MDOF) system, subjected to SVGM.

Spatial Variability of Seismic Motion (SVGM): What It Is and How to Describe It

Seismic Motion and Its Measurement

The seismic motion is regarded as one of the most important and unknown load which acts on a structure. This huge motion results from mainly two causes: (a) explosive volcanic eruptions which are very common in areas of volcanic activity and (b) tectonic activity associated with plate margins and faults. The majority of earthquakes worldwide are of this second type. Thus, the main cause of earthquakes is either the sudden movement of various plate boundaries or when plates scrape against each other. Some earthquakes are also caused by old plate boundaries or faults. The point from which starts this sudden movement is the epicenter of the earthquake. This point could be between 10 and 400 km below the surface of the earth.

Following this sudden movement, many waves radiate from the epicenter and propagate through the earth. It has been recognized that both P and S waves, known as body waves, emanate from the source and travel with a velocity that exceeds 4 km/s in earth. When they arrive at the free field, they are followed by both Love and Rayleigh waves. These latter, known as surface waves, travel only at the surface. The movements associated with body as well as surface waves are well known and are mainly decomposed into compressional and shear movement. At the free field, the seismic motion could be decomposed into two parts: (a) motion caused by body waves and (b) motion caused by surface waves.

The seismic motion is recorded by seismometers which are deployed in seismic areas. The number of seismometers is growing from year to year denoting the need for a better understanding of the seismic motion. While some of them are deployed in a single scheme, others are deployed in grouped schemes. This latter scheme forms what is called the dense seismic array. A typical dense array configuration is presented

S
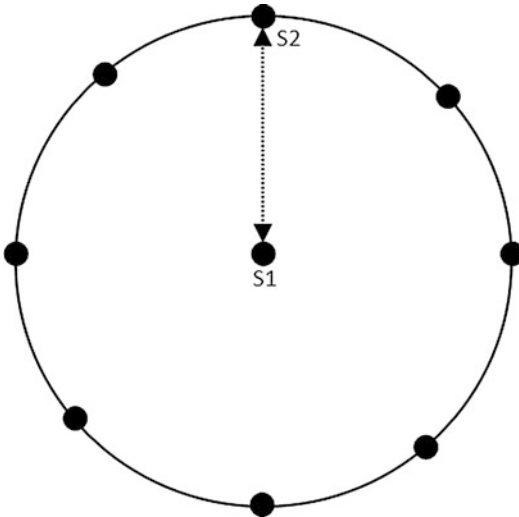
Spatial Variability of Ground Motion: Seismic Analysis, Fig. 1 Typical dense array configuration

in Fig. 1 where in general a central station S1 is surrounded by other stations such as station S2 with a separating distance.

Many arrays have been and are being deployed around the world to assess the characteristics of seismic ground motions. Among them SMART-1 array "Strong Motion ARray in Taiwan," located in Lotung, in the north-east corner of Taiwan deployed in 1980 remains one of the most important. Other dense seismic arrays have been deployed (Zerva 2009). The deployment of these arrays enhanced greatly our knowledge on how the SVGM affects structure and what the main parameters used to describe the SVGM are.

What Is SVEGM?

Let's consider that the medium is described using the Cartesian coordinate $\mathbf{x} = (x, y, z)$ and $\mathbf{u}(\mathbf{x}, t) = \left\{ u(\mathbf{x}, t) \quad v(\mathbf{x}, t) \quad w(\mathbf{x}, t) \right\}^T$ is the seismic motion measured along these axis. Hence, $v(\mathbf{x}, t)$ is the vertical motion, whereas $u(\mathbf{x}, t)$ and $w(\mathbf{x}, t)$ are the horizontal motions. At any point of the soil, both horizontal and vertical motions could be recorded. For instance, let's consider two points A and B at the free surface. Recorded earthquake along a particular axis at two points is obviously different in terms of amplitude and frequency. This difference is well accepted if

the separating distance between those points is somewhat important. However, data from dense seismic array show that even if the separating distance between two points is not very important (sometimes less than 50 m), the recorded motion at both points exhibits some differences.

The effect of this difference, called spatial variability of earthquake motion (SVEGM), on extended structures is addressed in this entry. Other sources of variability such as the result from relative surface fault motion for recording stations located on either side of a causative fault, soil liquefaction, and landslides (Zerva and Zervas 2002) are not considered herein. Finally, although SVEGM has been extensively studied for horizontal component and at free surface, some dense seismic arrays have been instrumented to record motion with respect to vertical component and/or at depth.

What causes SVEGM? It has been recognized that SVEGM is caused by several effects (Der Kiureghian 1996) (Fig. 1):

1. Incoherence effects due to scattering in the heterogeneous ground and extended source effects.
2. Traveling-wave effects, in which nonvertical waves reach different points on the ground surface at different times, producing a time shift between the motions at those points; this effect is also known as "wave passage effect."
3. Site effects due to the variation by the filtering effects of overlying soil columns; this effect is known as "site effect."

Although the attenuation of seismic motion increases the variability of seismic motion, it is not usually included since its effect on structures will not be visible below a separating distance of a dozen kilometers, which is greater than the usual separating distance for structures.

How to Measure and Describe SVEGM

For the purpose of simplicity, two points A and B located at the free surface with $z = 0$ are considered. Thus, coordinates of both points are, respectively, $\mathbf{x_A} = (x_A, 0, 0)$ and $\mathbf{x_B} = (x_B, 0, 0)$. Even if the motion could be recorded along the

**Spatial Variability of
Ground Motion: Seismic
Analysis,
Fig. 2** Propagation of
seismic waves
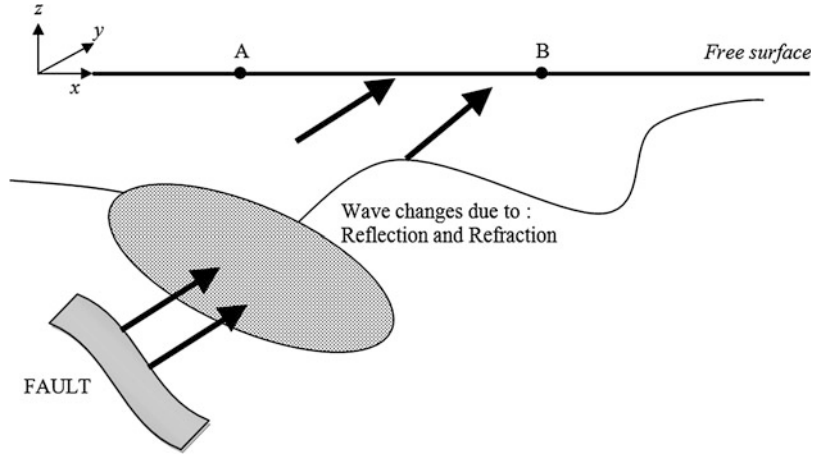
three axes, only the horizontal motion which is along the x axis, i.e., $u(\mathbf{x_A}, t)$ and $u(\mathbf{x_B}, t)$, is considered herein (Fig. 2).

The SVEGM is described by models in terms of correlation or coherence functions. This latter, which is the most used, is the ratio of the cross-spectral density function of the motion at two points separated by a horizontal offset to the square root of the product between the auto-spectral density functions at the two points:

$$\rho_{AB}(\omega) = \frac{S_{AB}(\omega)}{\sqrt{S_A(\omega)S_B(\omega)}} \qquad (1)$$

where

$$L = x_B - x_A,$$

$S_{AB}(\omega)$ is the cross-spectral density function of the motion along the x axis for the two points A and B.

$S_A(\omega)$ and $S_B(\omega)$ are the auto-spectral density functions of the motion along the x axis at, respectively, points A and B.

The coherency function is often written as

$$\rho_{AB}(\omega) = |\rho_{AB}(\omega)|\exp[i\theta_{AB}(\omega)] \qquad (2)$$

with

$$\theta_{AB}(\omega) = \tan^{-1}\left(\frac{\mathrm{Im}(S_{AB}(\omega))}{\mathrm{Re}(S_{AB}(\omega))}\right) \qquad (3)$$

where

$\mathrm{Im}(S_{AB}(\omega))$ and $\mathrm{Re}(S_{AB}(\omega))$ stands for imaginary and real parts of $S_{AB}(\omega)$.

$|\rho_{AB}(\omega)|$ is the lagged coherency and $\mathrm{Re}(S_{AB}(\omega))$ as unlagged coherency. It is worth noting that the coherency function is complex, and lagged coherency is usually called coherence function. This latter describes the similarity of the waveforms at two stations without taking into account the difference in the arrival times of the waves.

If the subscript A and B are skipped and replaced by the separating distance L, then the coherency function is rewritten as

$$\rho(L, \omega) = |\rho(L, \omega)|\exp[i\theta(L, \omega)] \qquad (4)$$

This coherency function measures the coherency along the x axis and at the free surface. It could be measured at depth by, for instance, considering that both points A and B are located at the same depth, i.e., $y = -h$, which leads to the following coherency function

$$\rho(h, L, \omega) = |\rho(h, L, \omega)|\exp[i\theta(h, L, \omega)] \qquad (5)$$

Models in terms of coherency functions have been developed during last decades (Abrahamson et al. 1991; Der Kiureghian 1996; Zendagui et al. 1999). They could be divided into three categories: empirical, semiempirical, and physically based models. An exhaustive list of these models could be found in Zerva (2009). The common feature of these models is that the

coherency function decreases with separating distance and frequency.

Generation of Variable Seismic Excitation and Response Spectra

The generation of variable seismic excitation has been extensively studied. The method developed by Shinozuka (1971) remains the most popular. This method has been extended and modified by others to take into account nonstationarity of processes (Deodatis 1996). A spectral-representation-based simulation algorithm has been developed by Deodatis (1996). The simulation uses the coherency function and allows the generation of acceleration as well as displacement time histories. This latter is necessary while performing non linear structural analysis. A recent method developed by Benmansour et al. (2012) solves the problem of integrability of seismic motion to avoid the use of baseline correction.

Alternatively, Berrah and Kausel (1992) developed an interesting method that modifies a particular response spectrum to take into account SVGM.

Structure Subjected to SVGM: General Derivation

Consider a structure with N DOFs. For the analysis of such a system, consider that the displacement vector contains two parts: (1) V includes the N DOFs of the superstructure, and (2) v_s contains the N_g components of support displacements. The equation of dynamic equilibrium is written as

$$\begin{bmatrix} m & m_s \\ m_s^T & m_{ss} \end{bmatrix} \begin{Bmatrix} \ddot{V} \\ \ddot{v}_s \end{Bmatrix} + \begin{bmatrix} c & c_s \\ c_s^T & c_{ss} \end{bmatrix} \begin{Bmatrix} \dot{V} \\ \dot{v}_s \end{Bmatrix} \\ + \begin{bmatrix} k & k_s \\ k_s^T & k_{ss} \end{bmatrix} \begin{Bmatrix} V \\ v_s \end{Bmatrix} = \begin{Bmatrix} 0 \\ R(t) \end{Bmatrix} \quad (6)$$

The mass, damping, and stiffness matrices can be determined from properties of the structure, while the support motions $\ddot{v}_s(t)$, $\dot{v}_s(t)$, and $v_s(t)$ must be specified. It is desired to determine the displacements V in the superstructure DOF and the support forces $R(t)$.

To write the governing equations in a form similar to that of the case of a single excitation, the displacements are separated into two parts:

$$\begin{Bmatrix} V \\ v_s \end{Bmatrix} = \begin{Bmatrix} u_s \\ v_s \end{Bmatrix} + \begin{Bmatrix} U \\ 0 \end{Bmatrix} \quad (7)$$

In this equation, u_s is the vector of structural displacements due to static application of prescribed support displacements v_s at each time instant. The two parts are related as

$$\begin{bmatrix} k & k_s \\ k_s^T & k_{ss} \end{bmatrix} \begin{Bmatrix} u_s \\ v_s \end{Bmatrix} = \begin{Bmatrix} 0 \\ R_s(t) \end{Bmatrix} \quad (8)$$

$R_s(t)$ are the support forces necessary to statically impose displacements v_s that vary with time; obviously, u_s varies with time and is therefore known as the vector of quasi-static displacements. Observe that $R_s(t) = 0$ in the case of identical support ground motion. The displacements of the superstructure U are known as dynamic displacements.

With the total structural displacements split into quasi-static and dynamic displacements, consider hence the first of the two partitioned equations (Eq. 6):

$$m\ddot{V} + m_s\ddot{v}_s + c\dot{V} + c_s\dot{v}_s + kV + k_sv_s = 0 \quad (9)$$

Substituting Eq. 7 and transferring all terms involving v_s and u_s to the right-hand side lead to

$$m\ddot{U} + c\dot{U} + kU = P_{eff}(t) \quad (10)$$

where the vector of effective earthquake forces is

$$P_{eff}(t) = -(m\ddot{u}_s + m_s\ddot{v}_s + c\dot{u}_s + c_s\dot{v}_s + ku_s + k_sv_s) \quad (11)$$

Equation 8 gives

$$ku_s + k_sv_s = 0 \quad (12)$$

This relation also enables us to express the quasi-static displacement u_s in terms of the specified support displacements v_s:

$$u_s = [\Lambda]v_s \quad (13)$$

where $[\Lambda] = -k^{-1}k_s$ is the influence matrix.

Equation 13 can be written in a different form as

$$u_s = \sum_{i=1}^{N_s} \Lambda_i v_{si}(t) \qquad (14)$$

where Λ_i is the ith column of the influence matrix $[\Lambda]$.

Substituting Eqs. 12 and 13 in Eq. 11, one gets

$$P_{eff}\,(t) = -(m[\Lambda] + m_s)\ddot{v}_s(t) - (c[\Lambda] + c_s)\dot{v}_s \qquad (15)$$

For many practical applications, further simplification of P_{eff} is possible in two stages. First, the damping term is usually small relative to the inertia term and can be dropped. Second, for structures with mass idealized as lumped at the DOFs, the mass matrix is diagonal, implying that m_s is a null matrix and m is diagonal. With these simplifications, P_{eff} is expressed as

$$P_{eff}\,(t) = -m[\Lambda]\ddot{v}_s(t) \qquad (16)$$

By using Eq. 14, P_{eff} can be expressed as

$$P_{eff}\,(t) = -\sum_{i=1}^{N_s} m\Lambda_i \ddot{v}_{si}(t) \qquad (17)$$

The dynamic response U can be computed by modal analysis, as the superposition of the modal contributions

$$U(t) = \sum_{n=1}^{N} \phi_n q_n(t) \qquad (18)$$

where ϕ_n are the natural modes and q_n are scalar multipliers called modal coordinates. Instead of Eq. 18, the modal equation can be expressed for the nth mode as

$$\ddot{q}_n + 2\xi_n\omega_n\dot{q}_n + \omega_n^2\,q_n = -\sum_{i=1}^{N_s} \Gamma_{ni}\ddot{v}_{si}(t) \qquad (19)$$

where

$$\Gamma_{ni} = \frac{L_{ni}}{M_n}, \;\; L_{ni} = \phi_n^T m\Lambda_i \;\text{ and }\; M_n = \phi_n^T m\phi_n \qquad (20)$$

The solution of Eq. 19 can be written as

$$q_n(t) = \sum_{i=1}^{N_s} \Gamma_{ni}D_{ni}(t) \qquad (21)$$

where $D_{ni}(t)$ is the displacement response of the nth-mode SDOF system to support acceleration $\ddot{v}_{si}(t)$.

The displacement response of the structure (Eq. 7) contains two parts:

1. The dynamic displacements are obtained by combining Eqs. 18 and 21:

$$U(t) = \sum_{i=1}^{N_s} \sum_{n=1}^{N} \Gamma_{ni}\phi_n D_{ni}(t) \qquad (22)$$

2. The quasi-static displacements u_s are given by Eq. 14.

Combining the two parts gives the total displacements of the structure

$$V(t) = \sum_{i=1}^{N_s} \Lambda_i v_{si}(t) + \sum_{i=1}^{N_s} \sum_{n=1}^{N} \Gamma_{ni}\phi_n D_{ni}(t) \qquad (23)$$

The equivalent static forces in structural DOFs are given by

$$F = kV + k_s v_s \qquad (24)$$

Substituting Eq. 7 for V and using Eq. 12 give

$$F = kU(t) \qquad (25)$$

These forces depend on the dynamic displacements only given by Eq. 22. Therefore,

$$F(t) = \sum_{i=1}^{N_s} \sum_{n=1}^{N} \Gamma_{ni}k\phi_n D_{ni}(t) \qquad (26)$$

The equivalent static forces along support DOFs are also given by the last term on the left-hand side of Eq. 6:

$$F_s = k_s^T V + k_{ss} v_s \qquad (27)$$

Spatial Variability of Ground Motion: Seismic Analysis, Fig. 3 A SDOF system subjected to SVGM
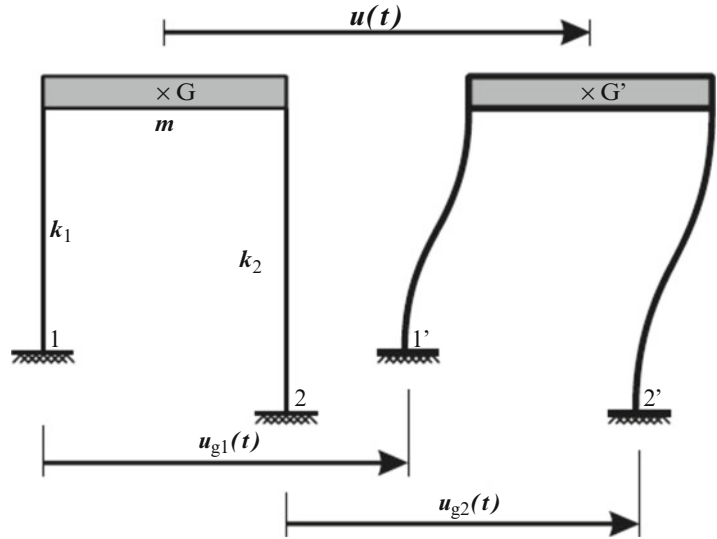
Substituting Eq. 7 and using Eq. 8 for the quasi-static support forces $R_s(t)$ give

$$F_s(t) = k_s^T U(t) + R_s(t) \qquad (28)$$

Application 1: SDOF System Subjected to SVGM

Derivation of the Relative Displacement

Consider a simple multi-support structure represented by a SDOF system with two supports 1 and 2 excited by, respectively, $u_{g1}(t)$ and $u_{g2}(t)$ (Fig. 3). The columns are assumed to be weightless and inextensible in the vertical (axial) direction, and the resistance to girder displacement provided by each column is represented by its spring constant k_1 and k_2. For the purpose of simplicity, it is assumed that the slab, which has a mass m, is infinitely rigid and that damping is neglected. The total displacement of the mass from the initial position G to the final position G' is $u(t)$

The forces acting on the slab:

- Inertial forces: $f_I = m\ddot{u}$.
- Elastic force on column (1): $f_{e1} = k_1(u - u_{g1})$
- Elastic force on column (2): $f_{e2} = k_2(u - u_{g2})$.

Hence, the equation of motion is an expression of the equilibrium of these forces given by

$$m\ddot{u} + (k_1 + k_2)u = k_1 u_{g1}(t) + k_2 u_{g2}(t) \qquad (29)$$

The natural frequency is defined by

$$\omega = \sqrt{\frac{k_1 + k_2}{m}} \qquad (30)$$

The quasi-static displacement is thus

$$u_s(t) = \frac{k_1 u_{g1}(t) + k_2 u_{g2}(t)}{k_1 + k_2} \qquad (31)$$

At this stage, it is considered that the support motions are harmonic variable both in amplitude and phase:

$$\left\{ \begin{array}{l} u_{g1}(t) = A_1 \sin \overline{\omega}_1 t \\ u_{g2}(t) = A_2 \sin \overline{\omega}_2 t \end{array} \right\} \qquad (32)$$

The total dynamic and quasi-static displacements are hence

$$u(t) = \frac{k_1}{k_1 + k_2} \frac{A_1}{\left(1 - \left(\frac{\omega_1}{\omega}\right)^2\right)} \sin \overline{\omega}_1 t$$

$$+ \frac{k_2}{k_1 + k_2} \frac{A_2}{\left(1 - \left(\frac{\omega_2}{\omega}\right)^2\right)} \sin \overline{\omega}_2 t \qquad (33)$$
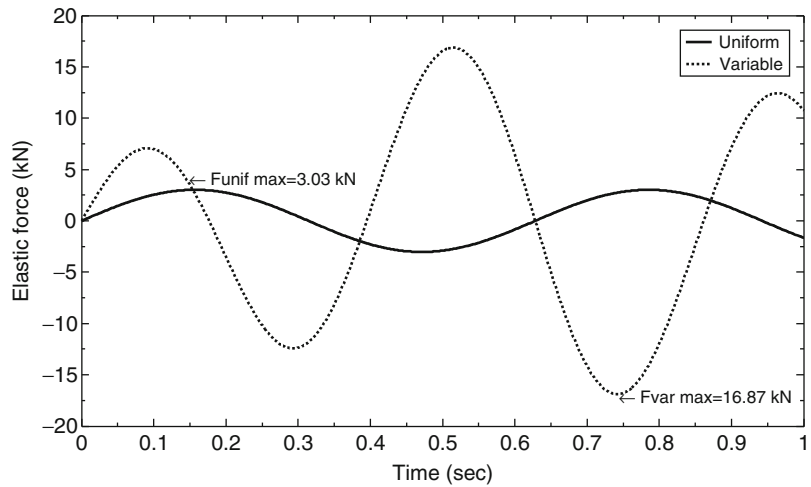
$$u_s(t) = \frac{k_1}{k_1 + k_2}\, A_1 \sin \overline{\omega}_1 t + \frac{k_2}{k_1 + k_2}\, A_2 \sin \overline{\omega}_2 t \tag{34}$$

Finally, the dynamic response could be derived as follows:

$$U(t) = u(t) - u_s(t) \tag{35}$$

Case Study

As the purpose of this entry is to see how SVGM affects a structure, it will be considered hereafter two cases: a uniform seismic motion at supports 1 and 2 equal to u_{g1} and a variable seismic motion, i.e., u_{g1} at support 1 and u_{g2} at support 2. The natural frequency is equal to 24 *rad/s*. The results are presented in terms of elastic forces of the column (1). For the case of spatially variable ground motion, two cases will be considered:

$$\text{Case S-1} \qquad \left\{ \begin{array}{l} u_{g1} = 0.1 \sin 10t \\ u_{g2} = 0.1 \sin\ 15t \end{array} \right\} (m) \tag{36}$$

$$\text{Case S-2} \qquad \left\{ \begin{array}{l} u_{g1} = 0.1 \sin 20t \\ u_{g2} = 0.1 \sin\ 25t \end{array} \right\} (m) \tag{37}$$

The frequencies of excitation have been chosen below and above the natural frequency. It is interesting to see that for both cases (Eqs. 36 and 37), SVGM (dashed line) exhibits greater value than uniform load (continuous line) (Figs. 4 and 5). Thus, for these cases, it seems that SVGM induces greater values of elastic forces than uniform seismic motion.

What about the variation of the quasi-static and dynamic components for both cases? Results show great difference on the part of both components depending on frequency of excitations. Hence, by analyzing the variations of the two components, quasi-static and dynamic (Figs. 6 and 7), it is found that for case S-1 the amplification caused by variable motion is mainly due to the importance of the quasi-static component. In this case, the frequency of the structure is larger than those of the imposed displacement. The SDOF system could be considered herein as a rigid structure. However, for case S-2, the frequency of the structure is close to the frequency of the excitation. In this case, the dynamic component is the main cause of the amplification of the response.

Application 2: MDOF System Subjected to SVGM

Derivation of the Relative Displacements

In this part, responses of a MDOF system under uniform and variable seismic input will be *computed*. It is assumed herein that the MDOF is composed of five masses connected by five springs (Fig. 8). The motion, in terms of

**Spatial Variability of
Ground Motion: Seismic
Analysis, Fig. 5** Variation
of elastic forces at column
1 for case S-2
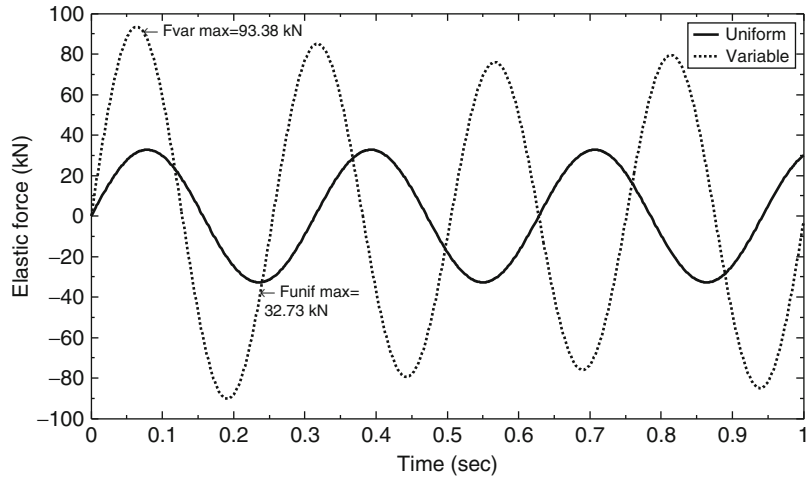
**Spatial Variability of
Ground Motion: Seismic
Analysis, Fig. 6** Variation
of the pseudo-static and
dynamic components of the
elastic forces at column
1 for case S-1
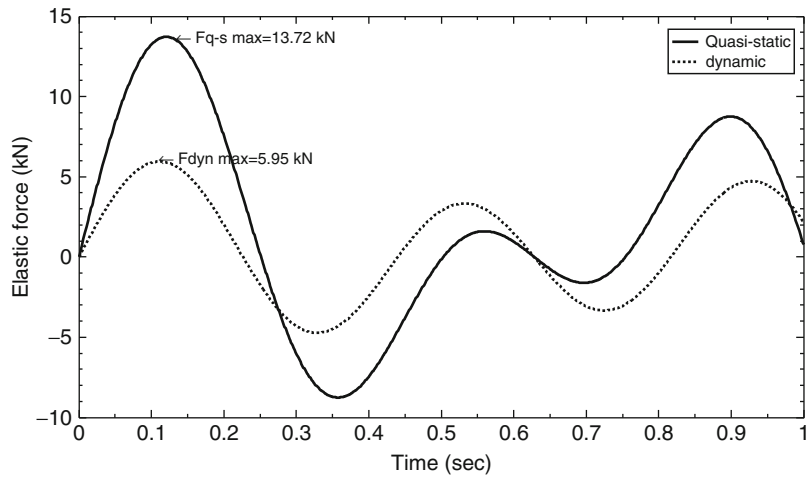
**Spatial Variability of
Ground Motion: Seismic
Analysis, Fig. 7** Variation
of the pseudo-static and
dynamic components of the
elastic forces at column
1 for case S-2
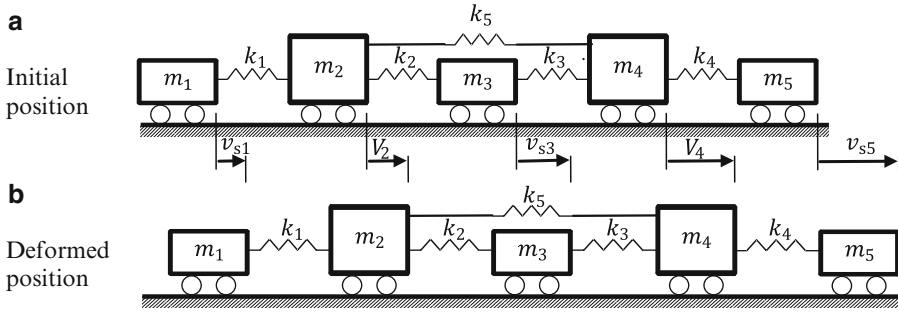
Spatial Variability of Ground Motion: Seismic Analysis, Fig. 8 MDOF system under SVGM

displacements, is imposed at the outer masses and the middle one, v_{s1}, v_{s3} et v_{s5}. Hence, only two masses do not have imposed displacements and their displacements are noted: V_2 et V_4. Thus, the displacement vector could be decomposed into two parts $\{v_s\} = \langle v_{s1} \quad v_{s3} \quad v_{s5} \rangle^T$ and $\{V\} = \langle V_2 \quad V_4 \rangle^T$.

The equation of dynamic equilibrium for all the DOFs is written as

$$\begin{bmatrix} m_1 & & & & \\ & m_2 & & & \\ & & m_3 & & \\ & & & m_4 & \\ & & & & m_5 \end{bmatrix} \begin{Bmatrix} \ddot{v}_{s1} \\ \ddot{V}_2 \\ \ddot{v}_{s3} \\ \ddot{V}_4 \\ \ddot{v}_{s5} \end{Bmatrix}$$

$$+ \begin{bmatrix} k_1 & -k_1 & & & \\ -k_1 & k_1 + k_2 + k_5 & -k_2 & & -k_5 \\ & -k_2 & k_2 + k_3 & -k_3 & \\ & -k_5 & -k_3 & k_3 + k_4 + k_5 & -k_4 \\ & & & -k_4 & -k_4 \end{bmatrix} \begin{Bmatrix} v_{s1} \\ V_2 \\ v_{s3} \\ V_4 \\ v_{s5} \end{Bmatrix} = \begin{Bmatrix} R_1 \\ 0 \\ R_3 \\ 0 \\ R_5 \end{Bmatrix} \quad (38)$$

By considering only DOFs 2 and 4,

$$\begin{bmatrix} m_2 & 0 \\ 0 & m_4 \end{bmatrix} \begin{Bmatrix} \ddot{V}_2 \\ \ddot{V}_4 \end{Bmatrix} + \begin{bmatrix} k_1 + k_2 + k_5 & -k_5 \\ -k_5 & k_3 + k_4 + k_5 \end{bmatrix} \begin{Bmatrix} V_2 \\ V_4 \end{Bmatrix}$$

$$= \begin{Bmatrix} k_1 v_{s1} + k_2 v_{s3} \\ k_3 v_{s3} + k_4 v_{s5} \end{Bmatrix} = \begin{bmatrix} k_1 & k_2 & 0 \\ 0 & k_3 & k_4 \end{bmatrix} \begin{Bmatrix} v_{s1} \\ v_{s3} \\ v_{s5} \end{Bmatrix} \quad (39)$$

The pseudo-static component is obtained by neglecting the inertial forces

$$0 + \begin{bmatrix} k_1 + k_2 + k_5 & -k_5 \\ -k_5 & k_3 + k_4 + k_5 \end{bmatrix} \begin{Bmatrix} u^s_2 \\ u^s_4 \end{Bmatrix}$$

$$= \begin{bmatrix} k_1 & k_2 & 0 \\ 0 & k_3 & k_4 \end{bmatrix} \begin{Bmatrix} v_{s1} \\ v_{s3} \\ v_{s5} \end{Bmatrix} \quad (40)$$

Thus,

$$\{u^s\} = \begin{Bmatrix} u^s_2 \\ u^s_4 \end{Bmatrix} = \begin{bmatrix} k_1 + k_2 + k_5 & -k_5 \\ -k_5 & k_3 + k_4 + k_5 \end{bmatrix}^{-1}$$

$$\begin{bmatrix} k_1 & k_2 & 0 \\ 0 & k_3 & k_4 \end{bmatrix} \begin{Bmatrix} v_{s1} \\ v_{s3} \\ v_{s5} \end{Bmatrix} \quad (41)$$

If $k_i = k$

$$\{u^s\} = \begin{Bmatrix} u^s_2 \\ u^s_4 \end{Bmatrix} = \frac{1}{8} \begin{bmatrix} 3 & 4 & 1 \\ 1 & 4 & 3 \end{bmatrix} \begin{Bmatrix} v_{s1} \\ v_{s3} \\ v_{s5} \end{Bmatrix} \quad (42)$$

The equation of movement could be rewritten in terms of the dynamic component $\{U\} = \{V\} - \{u^s\}$:

S

$$
\begin{bmatrix} m_2 & 0 \\ 0 & m_4 \end{bmatrix} \begin{Bmatrix} \ddot{U}_2 \\ \ddot{U}_4 \end{Bmatrix} + \begin{bmatrix} k_1+k_2+k_5 & -k_5 \\ -k_5 & k_3+k_4+k_5 \end{bmatrix} \begin{Bmatrix} U_2 \\ U_4 \end{Bmatrix}
$$
$$
= \begin{bmatrix} m_2 & 0 \\ 0 & m_4 \end{bmatrix} \begin{bmatrix} k_1+k_2+k_5 & -k_5 \\ -k_5 & k_3+k_4+k_5 \end{bmatrix}^{-1}
$$
$$
\begin{bmatrix} k_1 & k_2 & 0 \\ 0 & k_3 & k_4 \end{bmatrix} \begin{Bmatrix} \ddot{v}_{s1} \\ \ddot{v}_{s3} \\ \ddot{v}_{s5} \end{Bmatrix}
$$

$$(43)$$

If $m_i = m$ and $k_i = k$, then

$$
m \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} \ddot{U}_2 \\ \ddot{U}_4 \end{Bmatrix} + k \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{Bmatrix} U_2 \\ U_4 \end{Bmatrix}
$$
$$
= -\frac{m}{8} \begin{bmatrix} 3 & 4 & 1 \\ 1 & 4 & 3 \end{bmatrix} \begin{Bmatrix} \ddot{v}_{s1} \\ \ddot{v}_{s3} \\ \ddot{v}_{s5} \end{Bmatrix}
$$

$$(44)$$

The natural modes are

$$
\{\phi_1\} = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} \text{ and } \{\phi_2\} = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix} \qquad (45)
$$

Case Study

In this part, elastic forces at the spring k_1 obtained by considering uniform imposed displacements will be compared to those obtained when imposing variable displacements v_{s1}, v_{s3} et v_{s5}. For the sake of simplicity, it will be considered that the uniform displacement is set equal to v_{s1}

For the purpose of study, let's consider that

$$
m = 1\,t, \quad k = 600\,kN/m \quad \Rightarrow
$$
$$
\omega_1 = 34.64\,rad/s \text{ et } \omega_2 = 49\,rad/s
$$

For the case of spatially variable ground motion, four cases will be considered:

Case M-1 variability in terms of maximum amplitude:

$$
\begin{Bmatrix} v_{s1} = 0.10 \sin 30t \\ v_{s3} = 0.15 \sin 30t \\ v_{s5} = 0.30 \sin 30t \end{Bmatrix} (m) \qquad (46)
$$

Case M-2 variability in terms of phases:

$$
\begin{Bmatrix} v_{s1} = 0.1 \sin 30t \\ v_{s3} = 0.1 \sin 35t \\ v_{s5} = 0.1 \sin 30t \end{Bmatrix} (m) \qquad (47)
$$

Case M-3 variability in terms of temporal delay and frequency close to those of the system:

$$
\begin{Bmatrix} v_{s1} = 0.1 \sin 30t \\ v_{s3} = 0.1 \sin 30(t-0.03) \\ v_{s5} = 0.1 \sin 30(t-0.06) \end{Bmatrix} (m) \qquad (48)
$$

Case M-4 variability in terms of temporal delay and frequency greater than of the system:

$$
\begin{Bmatrix} v_{s1} = 0.1 \sin 50t \\ v_{s3} = 0.1 \sin 50(t-0.03) \\ v_{s5} = 0.1 \sin 50(t-0.06) \end{Bmatrix} (m) \qquad (49)
$$

It is worth noting that these cases model incoherence effect (Eqs. 46 and 47) and wave passage effect (Eqs. 48 and 49). Results obtained in terms of elastic forces at spring 1 for the four above cases are presented in Figs. 9, 10, 11, and 12.

It is observed that SVGM caused by incoherence effect leads to values of elastic forces greater than those obtained for uniform loading (Figs. 9 and 10). This could be readily observed while varying amplitude (Eq. 46, Fig. 9) or frequency (Eq. 47, Fig. 10). In turn, SVGM caused by wave passage effect does not yield greater values of elastic forces than uniform loading for the case where the frequency of excitation is close to the natural frequency of the structure (Eq. 48, Fig. 11). However, this conclusion is not valid for case M-4 where the excitation frequency is greater than the natural frequency of the system (Fig. 12).

Summary

This entry aims at describing the modeling of SVGM and its effects on structures. The total *force* is divided into two components: pseudo-static and dynamic forces. After deriving a general formulation for the case of MDOF systems, two cases have been considered: SDOF and MDOF systems.

It has been shown that SVGM could increase or decrease the elastic forces for both cases.
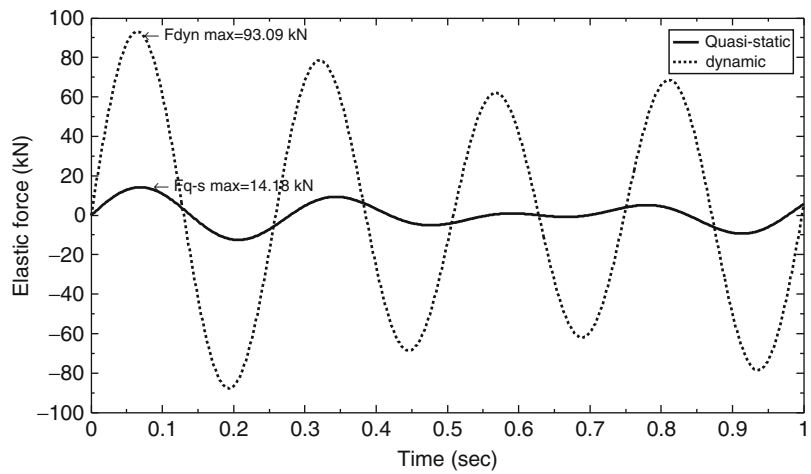
Spatial Variability of Ground Motion: Seismic Analysis, Fig. 9 Variation of elastic forces at spring 1 for case M-1
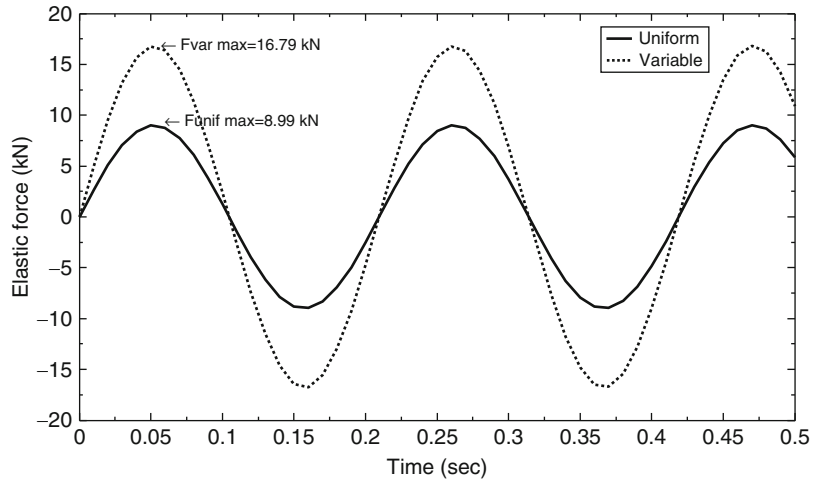
Spatial Variability of Ground Motion: Seismic Analysis, Fig. 10 Variation of elastic forces at spring 1 for case M-2
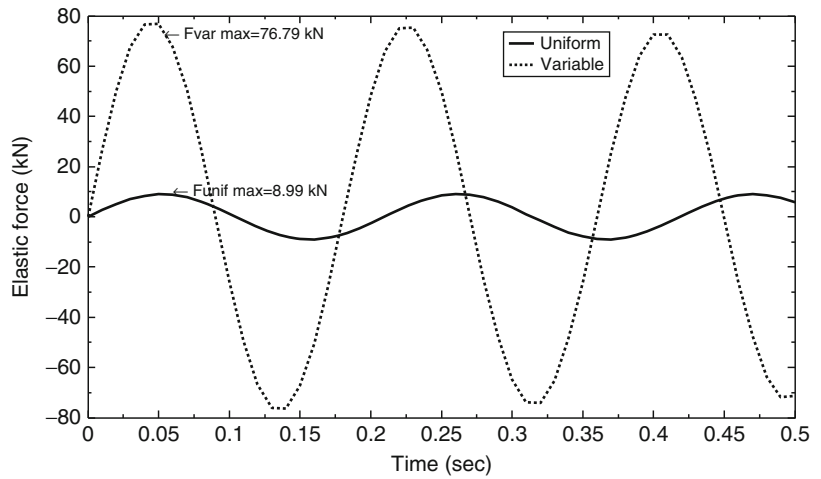
Spatial Variability of Ground Motion: Seismic Analysis, Fig. 11 Variation of elastic forces at spring 1 for case M-3

S
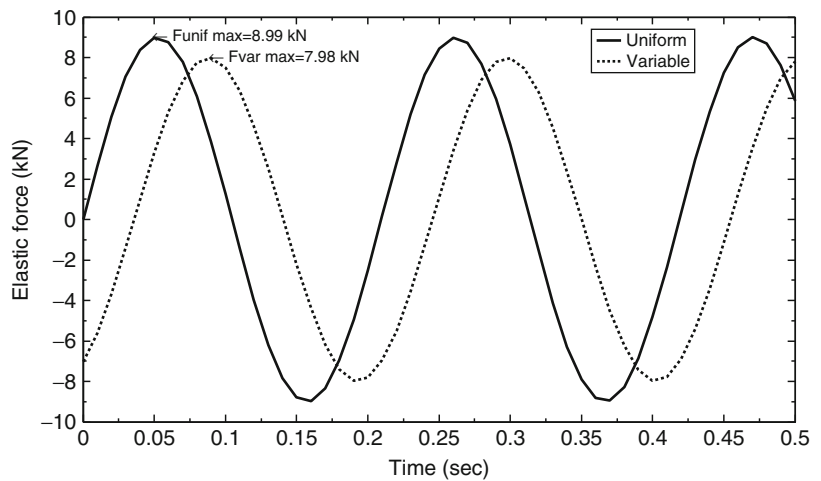
Spatial Variability of Ground Motion: Seismic Analysis,
Fig. 12 Variation of elastic forces at spring 1 for case M-4
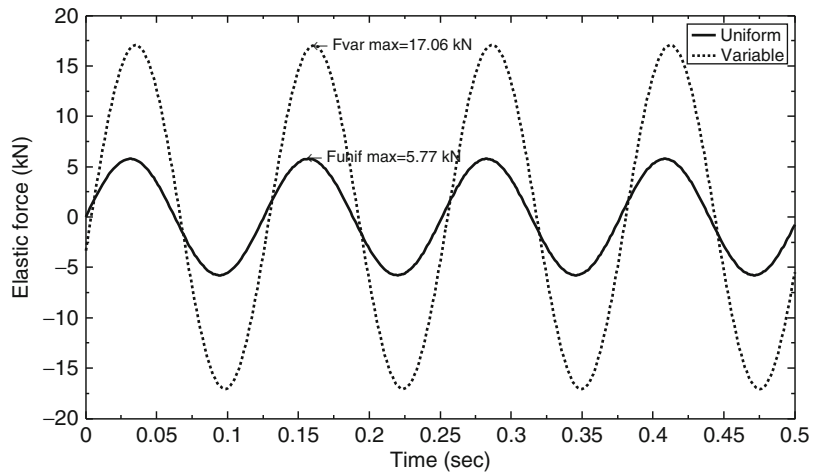
Cross-References

References

Abrahamson N, Schneider J, Stepp J (1991) Empirical spatial coherency functions for applications to soil-structure interaction analyses. Earthquake Spectra 7(1):1–27

Benmansour N, Djafour M, Bekkouche A, Zendagui D (2012) Seismic response evaluation of bridges under differential ground motion: a comparison with the new Algerian provisions. Eur J Environ Civil Eng 16(7):1–19

Berrah MK, Kausel E (1992) Response spectrum analysis of structures subjected to spatially varying motions. Earthq Eng Struct Dyn 21(6):461–470

Deodatis G (1996) Non-stationary stochastic vector processes: seismic ground motion applications. Probab Eng Mech 11(3):149–167

Der Kiureghian A (1996) Coherence model for spatially varying ground motions. Earthq Eng Struct Dyn 25(1):99–111

Shinozuka M (1971) Simulation of multivariate and multi dimensional random processes. J Acoust Soc Am 49(1):357–367

Zendagui D, Berrah M, Kausel E (1999) Stochastic de-amplification of spatially varying ground motions. Soil Dyn Earthq Eng 18(6):395–471

Zerva A (2009) Spatial variation of seismic ground motions. Modeling and engineering applications. CRC Press, Boca Raton

Zerva A, Zervas V (2002) Spatial variation of seismic ground motions: an overview. Appl Mech Rev 55(3):271–296

Spectral Finite Element Approach for Structural Dynamics

Priscilla Brandão Silva, Danilo Beli and José Roberto de França Arruda
Department of Computational Mechanics, Faculty of Mechanical Engineering, UNICAMP, Campinas, SP, Brazil

Synonyms

Forced response; Spectral element; Structural dynamics; Wave propagation

---

Introduction

Earthquake disaster in the past has motivated studies on earthquake-resistant buildings to save lives and resources (Newmark and Hall 1982). With this objective, physical models and numerical methods are used to predict the dynamic behavior of structures. The numerical methods provide quantitative analyses of physical phenomena. They can be used to design structures with geometry and materials for an adequate dynamic behavior under seismic excitation. The most frequently used in structural dynamics are the finite element method (FEM) and the boundary element method (BEM).

Based on wave propagation, the spectral finite element or spectral element method (SEM) was introduced by Beskos in 1978, organized and seemed by Doyle (1997) in the 1990s. It allows calculating relatively complex structures with different boundary conditions and discontinuities using simple theories. It combines important characteristics of FEM, dynamic stiffness method (DSM), and spectral analysis (Lee 2009).

The dynamic stiffness matrices and shape functions used in SEM are exact within the scope of the underlying physical theory, and the method allows a reduced number of degrees of freedom. The matrices are depended on frequency, but using spectral analysis, the dynamic response can be easily composed by wave superposition. Harmonic, random, or damped transient excitations can be decomposed using the discrete Fourier transform (DFT). The discrete frequencies are used to calculate the spectral matrix and discrete responses. Then, the complete dynamic response is computed by the sum of frequency components (inverse DFT). As FEM, SEM uses the assembly of a global matrix using elementary matrices and spatial discretization. However, differently from FEM, only discontinuities and locations where loads are applied need to be meshed (Ahmida and Arruda 2001).

It can be a useful tool because it joins DSM with spectral analysis to produce accurate solutions in both frequency and time domains at a low computational cost due to a dramatic system reduction, making it more adequate for uncertainty and optimization analyses. Furthermore, it can easily be associated with other numerical methods such as FEM and BEM. The limitations of SEM are general nonlinear analysis, or when exact wave solutions are not available, as it occurs in geometrically complex structures. When using SEM it is important to be careful with the time–frequency transformations to avoid leakage and aliasing errors (Lee 2009).

This chapter presents spectral element for straight and curved frame elements. First, the equations of motion are developed and physical and kinematic assumptions are presented. Then the spectrum relation is established and the spectral element is derived. Two simple examples – a two- and a three-dimensional frame structures, excited by a real earthquake signal – are analyzed. The purpose is to present the methodology and concepts of dynamic structural analysis via the spectral finite element method.

Classical Straight Frame

Equations of Motion

Consider a straight, long, and slender beam, as shown in Fig. 1. A Cartesian reference frame is placed such that the x-axis is along neutral axis of the beam. Thus, for a rectangular cross section, if h is the beam thickness, the position y will vary in the range $-h/2 \leq y \leq h/2$. Along the z-axis is the beam width.

Here, it is considered a plane frame element under flexural moments and axial forces. In straight beams, flexural and longitudinal behaviors are decoupled, and they may be modeled separately, the former via the Euler–Bernoulli beam theory and the latter via the elementary rod theory.

The Euler–Bernoulli theory assumes the vertical deflection is constant across any cross section and that cross sections remain plane after deformation, which means that the shear deformation is neglected. On the other hand, the classical rod theory considers only axial displacements and neglects lateral contractions due to Poisson's effect. Under these assumptions, the displacement field for a straight beam is given by:

Spectral Finite Element Approach for Structural Dynamics,
Fig. 1 Straight frame element model
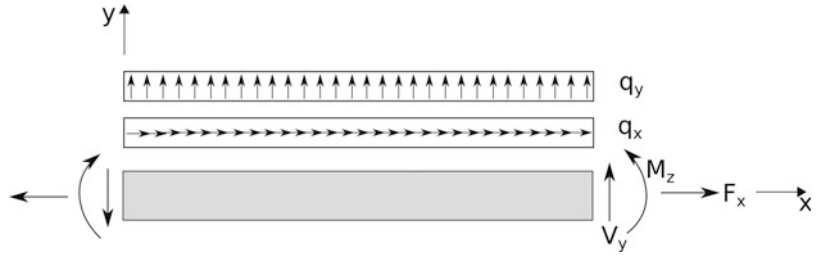
$$\overline{u}(x,y) = u(x) - y\theta_z(x)$$
$$\overline{v}(x,y) = v(x)$$

where $\bar{u}$ is the displacement in the x-direction (axial displacement), $\bar{v}$ is the displacement in the y-direction (vertical displacement), and θ_z is the rotation with respect to the z-axis.

Then, from differentiation of the displacement field, axial and transverse shear strains at a given cross section of the beam are expressed:

$$\varepsilon_{xx} = \frac{\partial \overline{u}}{\partial x} = \frac{\partial u}{\partial x} - y\frac{\partial \theta_z}{\partial x}$$
$$\varepsilon_{xy} = \frac{1}{2}\left(\frac{\partial \overline{u}}{\partial y} + \frac{\partial \overline{v}}{\partial x}\right) = \frac{1}{2}\left(-\theta_z + \frac{\partial v}{\partial x}\right).$$

The Euler–Bernoulli beam theory neglects the shear deformation, although not the shear force. This implies that the transverse shear strain must vanish, which yields:

$$\varepsilon_{xy} = 0 \rightarrow \theta_z = \frac{\partial v}{\partial x}.$$

The expression for the cross-section rotation can be used to write the axial strain, which is the only nonzero strain in the Euler–Bernoulli beam model, in terms of the vertical and the longitudinal displacements, as follows:

$$\varepsilon_{xx} = \frac{\partial u}{\partial x} - y\frac{\partial^2 v}{\partial x^2}.$$

Moreover, neglecting all stresses except for the predominant axial stress, one can write:

$$\sigma_{xx} = E\varepsilon_{xx} = E\frac{\partial u}{\partial x} - yE\frac{\partial^2 v}{\partial x^2}.$$

Once the stress and strains are determined, Newton's laws or variational formulations, such

as Hamilton's principle, may be used to determine the equations of motion for this classical beam model, which, in the absence of damping, is given by:

$$EA\frac{\partial^2 u}{\partial x^2} - \rho A\frac{\partial^2 v}{\partial t^2} = q_x(x,t)$$
$$\frac{\partial^2}{\partial x^2}\left(EI\frac{\partial^2 v}{\partial x^2}\right) + \rho A\frac{\partial^2 v}{\partial t^2} = q_y(x,t)$$

where E is the Young's modulus, I is the area moment of inertia with respect to the z-axis, ρ is the material density, A is the cross-section area, and $q_x(x,t)$ and $q_y(x,t)$ are, respectively, the longitudinal and vertical distributed loads along the beam length. The boundary conditions associated to this beam problem are:

$$u = 0 \quad \text{or} \quad F_x = EA\frac{\partial u}{\partial x}$$
$$v = 0 \quad \text{or} \quad V_y = -\frac{\partial}{\partial x}\left(EI\frac{\partial^2 v}{\partial x^2}\right)$$
$$\theta_z = 0 \quad \text{or} \quad M_z = EI\frac{\partial^2 v}{\partial x^2}.$$

The term EI is usually called the *bending stiffness*.

If some internal viscoelastic damping η is present, the equations of motion for the straight frame are modified to:

$$EA\frac{\partial^2 u}{\partial x^2} - \eta A\frac{\partial u}{\partial t} - \rho A\frac{\partial^2 u}{\partial t^2} = q_x(x,t)$$
$$\frac{\partial^2}{\partial x^2}\left(EI\frac{\partial^2 v}{\partial x^2}\right) + \eta A\frac{\partial v}{\partial t} + \rho A\frac{\partial^2 v}{\partial t^2} = q_y(x,t).$$

Spectral Analysis

Assuming a harmonic solution for the displacements of the beam (or applying the Fourier transform), i.e., by writing:
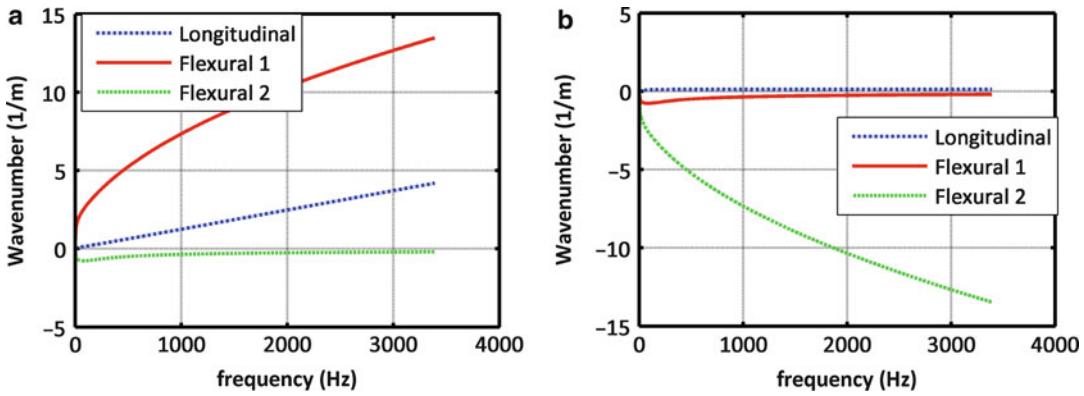
Spectral Finite Element Approach for Structural Dynamics, Fig. 2 Spectrum relations of a damped frame element: (**a**) real part of wavenumbers and (**b**) imaginary part of wavenumbers

$$u(x,t) = \sum_n \hat{u}_n(x,\omega_n)e^{i\omega_n t},$$
$$v(x,t) = \sum_n \hat{v}_n(x,\omega_n)e^{i\omega_n t}$$

in the homogenous case and considering constant properties along the beam, the equations of motion, in the frequency domain, are:

$$\frac{d^2\hat{u}}{dx^2} + k_x^2 \hat{u} = 0, \quad k_x = \pm\sqrt{\frac{\omega^2\rho A - i\omega\eta A}{EA}}$$
$$\frac{d^4\hat{v}}{dx^4} - \beta^4\hat{v} = 0, \quad \beta^2 = \pm\sqrt{\frac{\omega^2\rho A - i\omega\eta A}{EI}}.$$

The second homogeneous equation, relative to the transverse displacement, can be split into a product of two terms which must vanish. Thus, it yields:

$$\frac{d^2\hat{v}}{dx^2} + \beta^2\hat{v} = 0$$
$$\frac{d^2\hat{v}}{dx^2} - \beta^2\hat{v} = 0.$$

Then, considering a second Fourier transformation, now in the spatial domain, $\hat{u}(x,\omega) = \sum_m \tilde{u}(k_m,\omega)e^{-ik_m x}$, $\hat{v}(x,\omega) = \sum_m \tilde{v}(k_m,\omega)e^{-ik_m x}$, gives:

$$-k_1{}^2 + k_x{}^2 = 0 \quad \rightarrow \quad k_1 = \pm k_x$$
$$-k_2{}^2 + \beta^2 = 0 \quad \rightarrow \quad k_2 = \pm\beta$$
$$-k_3{}^2 - \beta^2 = 0 \quad \rightarrow \quad k_3 = \pm i\beta$$

where k_1, k_2, k_3 are the wavenumbers. They are function of the angular frequency ω, and this relationship constitutes the spectrum or dispersion relation, which is plotted in Fig. 2. Notice that when some damping is present, longitudinal and flexural modes are complex waves; thus, they exhibit a decaying behavior while propagating. In the undamped case, there are two real wavenumbers (the longitudinal and first flexural waves), which are propagating waves, and an imaginary wavenumber (the second flexural wave), which is a purely spatial decaying wave, the so-called evanescent wave. Phase and group speeds, c_0 and c_G, respectively, can be also derived (see Fig. 3), and they are expressed as:

$$c_0 = \frac{\omega}{k}, \quad c_G = \frac{d\omega}{dk}.$$

In the undamped case, unlike in the case of longitudinal motion, the phase and group speeds related to flexural waves depend on the cross-section properties, given by the area and the second moment of inertia. The group speed is twice the phase speed and both depend on frequency and, hence, are dispersive. It should be noted that, with the Euler–Bernoulli assumptions, the wave speeds tend to infinity as the frequency increases,
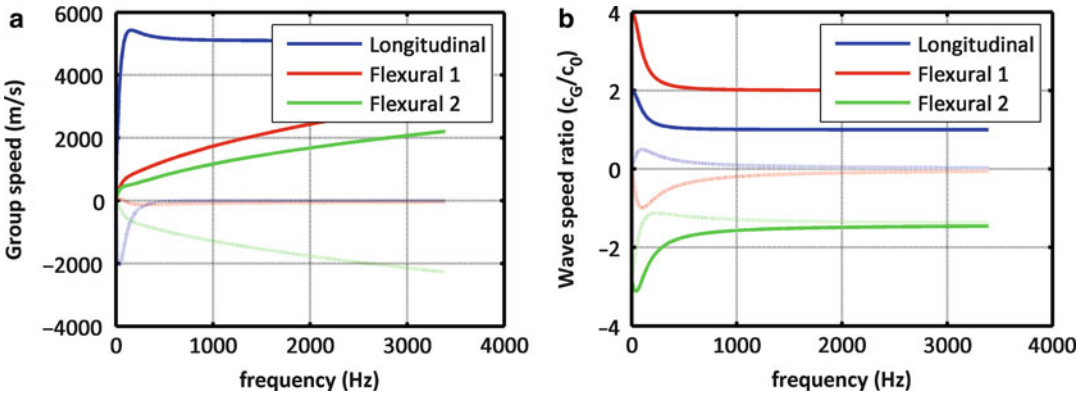
Spectral Finite Element Approach for Structural Dynamics, Fig. 3 Wave speeds of a frame element, real part in *solid line*, and imaginary part in *dotted line*: (**a**) group speeds and (**b**) speed ratios between group and phase speeds

which is not reasonable. This theory is, therefore, only suitable for low frequencies.

The displacement solution may be expressed in the spectral form as:

$$u(x,t) = \sum_n \left(A_1^+ e^{-ik_{xn}x} + A_1^- e^{ik_{xn}x} \right) e^{i\omega_n t}$$

$$v(x,t) = \sum_n \left(A_2^+ e^{-i\beta_n x} + A_3^+ e^{-\beta_n x} + A_2^- e^{i\beta_n x} + A_3^- e^{\beta_n x} \right) e^{i\omega_n t},$$

i.e., as a superposition of waves propagating in opposite directions.

It is worth recalling here that the cross-section rotation is related to the transverse displacement by the following expression:

$$\theta_z = \frac{\partial v}{\partial x}.$$

Thus, it may be also written as a superposition of waves, as follows:

$$\theta_z(x,t) = \sum_n \left(-iA_2^+ e^{-i\beta_n x} - A_3^+ e^{-\beta_n x} + iA_2^- e^{i\beta_n x} + A_3^- e^{\beta_n x} \right) \beta_n e^{i\omega_n t}.$$

From the classical theories for beams and rods, the internal loads related to the frame may be derived from the displacement field and are given by:

$$F_x = EA \frac{\partial u}{\partial x}$$

$$V_y = -\frac{\partial}{\partial x} \left(EI \frac{\partial^2 v}{\partial x^2} \right).$$

$$M_z = EI \frac{\partial^2 v}{\partial x^2}$$

Thus, they may be expressed in spectral form as:

$$F_x(x,t) = EA \sum_n k_{xn} \left(-iA_1^+ e^{-ik_{xn}x} + iA_1^- e^{ik_{xn}x} \right) e^{i\omega_n t}$$

$$V_y(x,t) = EI \sum_n \beta_n^3 \left(-iA_2^+ e^{-i\beta_n x} + A_3^+ e^{-\beta_n x} + iA_2^- e^{i\beta_n x} - A_3^- e^{\beta_n x} \right) e^{i\omega_n t}$$

$$M_z(x,t) = EI \sum_n \beta_n^2 \left(-A_2^+ e^{-i\beta_n x} + A_3^+ e^{-\beta_n x} - A_2^- e^{i\beta_n x} + A_3^- e^{\beta_n x} \right) e^{i\omega_n t}.$$

Displacements and loads evaluated at an arbitrary position along the frame may be expressed in matrix form, as follows:

$$
\hat{\mathbf{q}} = \begin{bmatrix} \hat{u} \\ \hat{v} \\ \hat{\theta}_z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & -i\beta & -\beta & 0 & i\beta & \beta \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_x^+ & 0 \\ 0 & \mathbf{\Lambda}_x^- \end{bmatrix} \begin{bmatrix} A_1^+ \\ A_2^+ \\ A_3^+ \\ B_1^- \\ B_2^- \\ B_3^- \end{bmatrix}
$$

$$
\hat{\mathbf{f}} = \begin{bmatrix} \hat{F}_x \\ \hat{V}_y \\ \hat{M}_z \end{bmatrix} = E \begin{bmatrix} -iAk_x & 0 & 0 & -iAk_x & 0 & 0 \\ 0 & -iI\beta^3 & I\beta^3 & 0 & iI\beta^3 & -I\beta^3 \\ 0 & -I\beta^2 & I\beta^2 & 0 & -I\beta^2 & I\beta^2 \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_x^+ & 0 \\ 0 & \mathbf{\Lambda}_x^- \end{bmatrix} \begin{bmatrix} A_1^+ \\ A_2^+ \\ A_3^+ \\ B_1^- \\ B_2^- \\ B_3^- \end{bmatrix}
$$

where

$$
\mathbf{\Lambda}_x^+ = \begin{bmatrix} e^{-ik_x x} & 0 & 0 \\ 0 & e^{-i\beta x} & 0 \\ 0 & 0 & e^{-\beta x} \end{bmatrix},
$$

$$
\mathbf{\Lambda}^{x-} = \begin{bmatrix} e^{ik_x(x-L)} & 0 & 0 \\ 0 & e^{i\beta(x-L)} & 0 \\ 0 & 0 & e^{\beta(x-L)} \end{bmatrix}.
$$

Here, the reference for wave amplitudes related to negative-going waves, i.e., those propagating to the left or in the negative sense of the x-axis, is placed at the end of an element of frame ($x = L$). Thus, instead of A_1^-, A_2^-, A_3^-, one has B_1^-, B_2^-, B_3^-. This is done for the purpose of numerical conditioning in the case where there are highly evanescent waves.

From Mace et al. (2005) and Mead (1973), the classification of waves into positive- or negative-going waves consists in analyzing the sense of decreasing wave amplitudes or, if the amplitude remains constant, the sense of the time-averaged power transmission. The latter can be evaluated by regarding the sign of $\Re\{i\omega\mathbf{f}_L^T\mathbf{q}_L\}$: if it is negative, the corresponding wave is a positive-going wave; otherwise, it is a negative-going wave. Alternatively, the sign of the group speed may be used, which may be shown to follow the time-averaged power transmission sign.

Spectral Element for a Classical Frame

Consider an element of classical frame (Euler–Bernoulli beam combined with an elementary rod) with length L. The displacements at both ends of the spectral element can be expressed as a function of the wave amplitudes as follows:

$$
\begin{bmatrix} \hat{u}(0) \\ \hat{v}(0) \\ \hat{\theta}_z(0) \\ \hat{u}(L) \\ \hat{v}(L) \\ \hat{\theta}_z(L) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -e^{-ik_x L} & 0 & 0 \\ 0 & 1 & 1 & 0 & e^{-i\beta L} & e^{-\beta L} \\ 0 & -i\beta & -\beta & 0 & i\beta e^{-i\beta L} & \beta e^{-\beta L} \\ e^{-ik_x L} & 0 & 0 & -1 & 0 & 0 \\ 0 & e^{-i\beta L} & e^{-\beta L} & 0 & 1 & 1 \\ 0 & -i\beta e^{-i\beta L} & -\beta e^{-\beta L} & 0 & i\beta & \beta \end{bmatrix} \begin{bmatrix} A_1^+ \\ A_2^+ \\ A_3^+ \\ B_1^- \\ B_2^- \\ B_3^- \end{bmatrix}
$$

$$\begin{bmatrix} \hat{\mathbf{q}}_0 \\ \hat{\mathbf{q}}_L \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_q^+ & \mathbf{\Phi}_q^- \mathbf{\Lambda}_L^+ \\ \mathbf{\Phi}_q^+ \mathbf{\Lambda}_L^+ & \mathbf{\Phi}_q^- \end{bmatrix} \begin{bmatrix} \mathbf{A}^+ \\ \mathbf{B}^- \end{bmatrix}.$$

Analogously, the external loads applied to the frame may be written as:

$$\begin{bmatrix} -\hat{F}_x(0) \\ -\hat{V}_y(0) \\ -\hat{M}_z(0) \\ \hat{F}_x(L) \\ \hat{V}_y(L) \\ \hat{M}_z(L) \end{bmatrix} = E \begin{bmatrix} iAk_x & 0 & 0 & iAk_x e^{-ik_x L} & 0 & 0 \\ 0 & iI\beta^3 & -I\beta^3 & 0 & -iI\beta^3 e^{-i\beta L} & I\beta^3 e^{-\beta L} \\ 0 & I\beta^2 & -I\beta^2 & 0 & I\beta^2 e^{-i\beta L} & -I\beta^2 e^{-\beta L} \\ -iAk_x e^{-ik_x L} & 0 & 0 & -iAk_x & 0 & 0 \\ 0 & -iI\beta^3 e^{-i\beta L} & I\beta^3 e^{-\beta L} & 0 & iI\beta^3 & -I\beta^3 \\ 0 & -I\beta^2 e^{-i\beta L} & I\beta^2 e^{-\beta L} & 0 & -I\beta^2 & I\beta^2 \end{bmatrix} \begin{bmatrix} A_1^+ \\ A_2^+ \\ A_3^+ \\ B_1^- \\ B_2^- \\ B_3^- \end{bmatrix}$$

$$\begin{bmatrix} \hat{\mathbf{F}}_0 \\ \hat{\mathbf{F}}_L \end{bmatrix} = \begin{bmatrix} -\mathbf{\Phi}_f^+ & -\mathbf{\Phi}_f^- \mathbf{\Lambda}_L^+ \\ \mathbf{\Phi}_f^+ \mathbf{\Lambda}_L^+ & \mathbf{\Phi}_f^- \end{bmatrix} \begin{bmatrix} \mathbf{A}^+ \\ \mathbf{B}^- \end{bmatrix}.$$

With the general method of computing $\mathbf{A}^+$ and $\mathbf{B}^-$ from the displacement equation and replacing it in the expression for external loads, one can write:

$$\begin{bmatrix} \hat{\mathbf{F}}_0 \\ \hat{\mathbf{F}}_L \end{bmatrix} = \mathbf{K}(\omega) \begin{bmatrix} \hat{\mathbf{q}}_0 \\ \hat{\mathbf{q}}_L \end{bmatrix},$$

$$\mathbf{K} = \begin{bmatrix} -\mathbf{\Phi}_f^+ & -\mathbf{\Phi}_f^- \mathbf{\Lambda}_L^+ \\ \mathbf{\Phi}_f^+ \mathbf{\Lambda}_L^+ & \mathbf{\Phi}_f^- \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_q^+ & \mathbf{\Phi}_q^- \mathbf{\Lambda}_L^+ \\ \mathbf{\Phi}_q^+ \mathbf{\Lambda}_L^+ & \mathbf{\Phi}_q^- \end{bmatrix}^{-1}.$$

Throw-Off Spectral Element for a Straight Frame
A semi-infinite element of beam can be easily modeled via SEM by assuming the wave amplitudes related to waves propagating to the left (negative-going waves) are zero, i.e., considering there is no reflection. In this case, the displacement and force vectors at the left side of the element are given by:

$$\hat{\mathbf{q}}_0 = \mathbf{\Phi}_q^+ \mathbf{A}^+$$

$$\hat{\mathbf{F}}_0 = -\mathbf{\Phi}_f^+ \mathbf{A}^+.$$

Thus, the spectral matrix for the throw-off element of frame is:

$$\hat{\mathbf{F}}_0 = \mathbf{K}_\infty(\omega) \hat{\mathbf{q}}_0, \quad \mathbf{K}_\infty = -\mathbf{\Phi}_f^+ \left(\mathbf{\Phi}_q^+ \right)^{-1}$$

$$\mathbf{K}_\infty = \frac{E(1+i)}{2} \begin{bmatrix} iAk_x & 0 & 0 \\ 0 & 2iI\beta^3 & (1+i)I\beta^2 \\ 0 & (1+i)I\beta^2 & 2I\beta \end{bmatrix}.$$

Curved Beam

Equations of Motion
Consider a slender in-plane curved beam as shown in Fig. 4. The circumferential reference frame is placed such that the s-axis is along the neutral axis of the beam, here supposed to be equidistant of the upper and lower surfaces of the beam. The beam thickness h varies in the range $-h/2 \le r \le +h/2$ along the radial r-axis and its width lies along the z-axis. Based on the Euler–Bernoulli beam theory, which neglects shear deformation, the kinematics of a curved beam can be written as:

$$\overline{u}_s(r, s) = u_s(s) + r \left(\frac{\partial u_s(s)}{\partial s} - \frac{v_r(s)}{R} \right)$$
$$\overline{v}_r(r, s) = v_r(s)$$

where $\bar{u}_s$ is the displacement in the s-direction (tangential displacement), $\bar{v}_r$ is the displacement in the y-direction (radial displacement), $\overline{\psi}$ is the rotation with respect to the z-axis, and R is the radius of the centerline.

Spectral Finite Element Approach for Structural Dynamics, Fig. 4 Curved beam element model
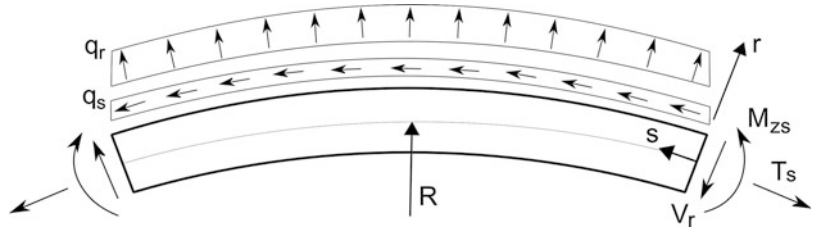
Then, from differentiation of the displacement field, considering a significant strain in s-direction, a strain at a given cross section of the curved beam is expressed:

$$\varepsilon_{ss} = \frac{\partial u_s}{\partial s} - \frac{v_r}{R} + r\left(\frac{\partial^2 v_r}{\partial s^2} + \frac{1}{R}\frac{\partial u_s}{\partial s}\right).$$

Moreover, neglecting all stresses except for the predominant circumferential stress, one can write:

$$\sigma_{ss} = E\varepsilon_{ss} = E\left[\frac{\partial u_s}{\partial s} - \frac{v_r}{R}\right] + rE\left(\frac{\partial^2 v_r}{\partial s^2} + \frac{1}{R}\frac{\partial u_s}{\partial s}\right).$$

Then, as a result of the application of Newton's laws or variational formulations, the equations of motion for the undamped curved beam are derived:

$$\frac{EI}{R}\left(\frac{1}{R}\frac{\partial^2 u_s}{\partial s^2} - \frac{\partial^3 v_r}{\partial s^3}\right) + \frac{EA}{R}\left(\frac{\partial v_r}{\partial s} + R\frac{\partial^2 u_s}{\partial s^2}\right)$$
$$- \rho A\frac{\partial^2 u_s}{\partial t^2} = q_s(s,t)$$
$$\frac{EI}{R}\left(\frac{\partial^3 u_s}{\partial s^3} - R\frac{\partial^4 v_r}{\partial s^4}\right) - \frac{EA}{R}\left(\frac{v_r}{R} + \frac{\partial u_s}{\partial s}\right)$$
$$+ \rho A\frac{\partial^2 v_r}{\partial t^2} = q_r(s,t)$$

where $q_s(s,t)$ and $q_r(s,t)$ are, respectively, tangential and radial distributed loads along the curved beam length. The boundary conditions associated to this problem are:

$$u_s = 0 \quad \text{or} \quad T_s = EA\left(\frac{\partial u_s}{\partial s} - \frac{v_r}{R}\right)$$
$$v_r = 0 \quad \text{or} \quad V_r = -EI\left(\frac{1}{R}\frac{\partial^2 u_s}{\partial s^2} + \frac{\partial^3 v_r}{\partial s^3}\right)$$
$$\psi = 0 \quad \text{or} \quad M_{zs} = EI\left(\frac{1}{R}\frac{\partial u_s}{\partial s} + \frac{\partial^2 v_r}{\partial s^2}\right).$$

If some internal viscoelastic damping η is present, the equations of motion are modified to:

$$\frac{EI}{R}\left(\frac{1}{R}\frac{\partial^2 u_s}{\partial s^2} + \frac{\partial^3 v_r}{\partial s^3}\right) + \frac{EA}{R}\left(-\frac{\partial v_r}{\partial s} + R\frac{\partial^2 u_s}{\partial s^2}\right)$$
$$- \eta A\frac{\partial u_s}{\partial t} - \rho A\frac{\partial^2 u_s}{\partial t^2} = q_s(s,t)$$
$$\frac{EI}{R}\left(\frac{\partial^3 u_s}{\partial s^3} + R\frac{\partial^4 v_r}{\partial s^4}\right) + \frac{EA}{R}\left(\frac{v_r}{R} - \frac{\partial u_s}{\partial s}\right)$$
$$+ \eta A\frac{\partial v_r}{\partial t} + \rho A\frac{\partial^2 v_r}{\partial t^2} = q_r(s,t).$$

Spectral Analysis

As before, in the straight beam case, two Fourier transforms can be successively applied to the displacement solution of a curved beam, which allows writing:

$$u_s(s,t) = \sum_n \hat{u}_{sn}(s,\omega_n)e^{i\omega_n t} = \sum_n \sum_m \tilde{u}_s(k_m,\omega)e^{-ik_m s}e^{i\omega_n t}$$
$$v_r(s,t) = \sum_n \hat{v}_{rn}(s,\omega_n)e^{i\omega_n t} = \sum_n \sum_m \check{v}_r(k_m,\omega)e^{-ik_m s}e^{i\omega_n t}.$$

In the homogenous case and considering constant properties along the beam, the displacement solution is substituted in the equations of motion, which yields, in the frequency domain,

S
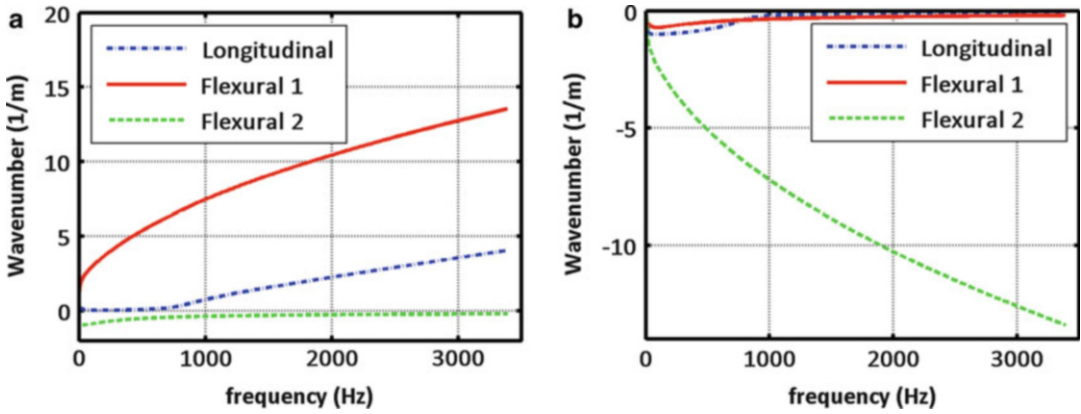
Spectral Finite Element Approach for Structural Dynamics, Fig. 5 Spectrum relations of a damped curved beam element: (**a**) real part of wavenumbers and (**b**) imaginary part of wavenumbers

$$
\begin{bmatrix}
-\dfrac{EIk^2}{R^2} - EAk^2 + i\eta A\omega + \omega^2 \rho A & \dfrac{iEIk^3}{R} + \dfrac{iEAk}{R} \\[3mm]
\dfrac{iEIk^3}{R} + \dfrac{iEAk}{R} & EIk^4 + \dfrac{EA}{R^2} - i\eta A\omega - \omega^2 \rho A
\end{bmatrix}
\begin{bmatrix} \tilde{u}_s \ \tilde{v}_r \end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \end{bmatrix}.
$$

A nontrivial solution for this system of equations is obtained by setting the determinant to vanish, which results in a sixth-order polynomial equation in k, the characteristic equation, or spectrum relation:

$$
k^6 - \left[k_0^2 + \frac{2}{R^2} \right] k^4 - \left[\beta_0^4 + k_0^2 \frac{1}{R^2} - \frac{1}{R^4} \right] k^2
$$
$$
+ \left[k_0^2 - \frac{1}{R^2} \right] \beta_0^4 = 0
$$

where $\beta_0^4 = \frac{\rho A \omega^2 + i\eta A\omega}{EI}$ and $k_0^2 = \frac{\rho A \omega^2 + i\eta A\omega}{EA}$.

This equation provides three pairs of wavenumbers $\pm k_{s1}$, $\pm k_{s2}$, $\pm k_{s3}$. One pair is linked to extensional waves and the other two pairs to flexural waves (Fig. 5).

Then, the displacement solution may be expressed in the spectral form as:

$$
u_s(s,t) = \sum_n \left(\overline{A}_{s1}^+ e^{-ik_{s1n}s} + \overline{A}_{s2}^+ e^{-ik_{s2n}s} + \overline{A}_{s3}^+ e^{-ik_{s3n}s} + \overline{A}_{s1}^- e^{ik_{s1n}s} + \overline{A}_{s2}^- e^{ik_{s2n}s} + \overline{A}_{s3}^- e^{ik_{s3n}s} \right) e^{i\omega_n t}
$$
$$
v_r(s,t) = \sum_n \left(A_{s1}^+ e^{-ik_{s1n}s} + A_{s2}^+ e^{-ik_{s2n}s} + A_{s3}^+ e^{-ik_{s3n}s} + A_{s1}^- e^{ik_{s1n}s} + A_{s2}^- e^{ik_{s2n}s} + A_{s3}^- e^{ik_{s3n}s} \right) e^{i\omega_n t},
$$
$$
\psi(s,t) = \frac{\partial v_r(s,t)}{\partial s} = \sum_n \begin{pmatrix} -ik_{s1n}A_{s1}^+ e^{-ik_{s1n}s} - ik_{s2n}A_{s2}^+ e^{-ik_{s2n}s} - ik_{s3n}A_{s3}^+ e^{-ik_{s3n}s} \\ +ik_{s1n}A_{s1}^- e^{ik_{s1n}s} + ik_{s2n}A_{s2}^- e^{ik_{s2n}s} + ik_{s3n}A_{s3}^- e^{ik_{s3n}s} \end{pmatrix} e^{i\omega_n t},
$$
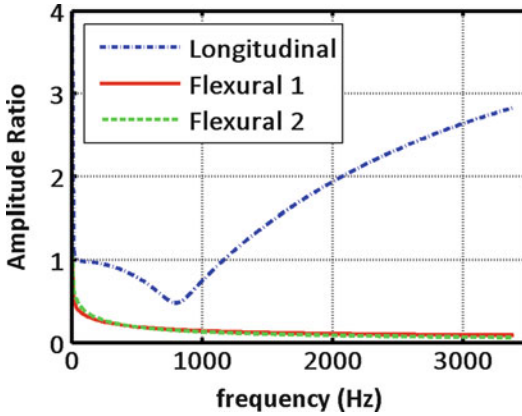
Spectral Finite Element Approach for Structural Dynamics, Fig. 6 Amplitude ratios for a damped curved beam element

$$\alpha_{sm} = \frac{\overline{A}_{sm}}{A_{sm}}$$

$$= -\frac{EIk_{smn}^4 + EAR^{-2} - i\eta A\omega - \rho A\omega^2}{i\left(EIk_{smn}^3 + EAk_{smn}\right)/R}.$$

Expressions for the internal loads in curved beams have been derived in terms of the displacement field and are rewritten below:

$$T_s = EA\left(\frac{\partial u_s}{\partial s} - \frac{v_r}{R}\right)$$

$$V_r = -EI\left(\frac{1}{R}\frac{\partial^2 u_s}{\partial s^2} + \frac{\partial^3 v_r}{\partial s^3}\right)$$

$$M_{zs} = EI\left(\frac{1}{R}\frac{\partial u_s}{\partial s} + \frac{\partial^2 v_r}{\partial s^2}\right).$$

Thus, substituting the spectral solutions for the displacement field, it yields:

i.e., as a superposition of waves propagating in opposite directions. From the homogeneous system of equations for $\tilde{u}_s$, $\tilde{v}_r$, amplitude ratios for each wave mode are obtained (Fig. 6):

$$T_s(s,t) = EA\sum_n \left(\begin{array}{c} -ik_{s1n}\alpha_{s1}A_{s1}^+ e^{-ik_{s1n}s} - ik_{s2n}\alpha_{s2}A_{s2}^+ e^{-ik_{s2n}s} - ik_{s3n}\alpha_{s3}A_{s3}^+ e^{-ik_{s3n}s} \\ +ik_{s1n}\alpha_{s1}A_{s1}^- e^{ik_{s1n}s} + ik_{s2n}\alpha_{s2}A_{s2}^- e^{ik_{s2n}s} + ik_{s3n}\alpha_{s3}A_{s3}^- e^{ik_{s3n}s} \end{array}\right) e^{i\omega_n t}$$
$$-\frac{EA}{R}\sum_n \left(\begin{array}{c} A_{s1}^+ e^{-ik_{s1n}s} + A_{s2}^+ e^{-ik_{s2n}s} + A_{s3}^+ e^{-ik_{s3n}s} \\ +A_{s1}^- e^{ik_{s1n}s} + A_{s2}^- e^{ik_{s2n}s} + A_{s3}^- e^{ik_{s3n}s} \end{array}\right) e^{i\omega_n t}$$

$$V_r(s,t) = -\frac{EI}{R}\sum_n \left(\begin{array}{c} k_{s1n}^2\alpha_{s1}A_{s1}^+ e^{-ik_{s1n}s} + k_{s2n}^2\alpha_{s2}A_{s2}^+ e^{-ik_{s2n}s} + k_{s3n}^2\alpha_{s3}A_{s3}^+ e^{-ik_{s3n}s} \\ -k_{s1n}^2\alpha_{s1}A_{s1}^- e^{ik_{s1n}s} - k_{s2n}^2\alpha_{s2}A_{s2}^- e^{ik_{s2n}s} - k_{s3n}^2\alpha_{s3}A_{s3}^- e^{ik_{s3n}s} \end{array}\right) e^{i\omega_n t}$$
$$-EI\sum_n \left(\begin{array}{c} ik_{s1n}^3 A_{s1}^+ e^{-ik_{s1n}s} + ik_{s2n}^3 A_{s2}^+ e^{-ik_{s2n}s} + ik_{s3n}^3 A_{s3}^+ e^{-ik_{s3n}s} \\ -ik_{s1n}^3 A_{s1}^- e^{ik_{s1n}s} - ik_{s2n}^3 A_{s2}^- e^{ik_{s2n}s} - ik_{s3n}^3 A_{s3}^- e^{ik_{s3n}s} \end{array}\right) e^{i\omega_n t}$$

$$M_{zs}(s,t) = \frac{EI}{R}\sum_n \left(\begin{array}{c} -ik_{s1n}\alpha_{s1}A_{s1}^+ e^{-ik_{s1n}s} - ik_{s2n}\alpha_{s2}A_{s2}^+ e^{-ik_{s2n}s} - ik_{s3n}\alpha_{s3}A_{s3}^+ e^{-ik_{s3n}s} \\ +ik_{s1n}\alpha_{s1}A_{s1}^- e^{ik_{s1n}s} + ik_{s2n}\alpha_{s2}A_{s2}^- e^{ik_{s2n}s} + ik_{s3n}\alpha_{s3}A_{s3}^- e^{ik_{s3n}s} \end{array}\right) e^{i\omega_n t}$$
$$EI\sum_n \left(\begin{array}{c} k_{s1n}^2 A_{s1}^+ e^{-ik_{s1n}s} + k_{s2n}^2 A_{s2}^+ e^{-ik_{s2n}s} + k_{s3n}^2 A_{s3}^+ e^{-ik_{s3n}s} \\ -k_{s1n}^2 A_{s1}^- e^{ik_{s1n}s} - k_{s2n}^2 A_{s2}^- e^{ik_{s2n}s} - k_{s3n}^2 A_{s3}^- e^{ik_{s3n}s} \end{array}\right) e^{i\omega_n t}$$

S

Displacements and loads evaluated at an arbitrary position along the curved beam may be expressed in matrix form, as follows:

$$\hat{\mathbf{q}}_{CB} = \begin{bmatrix} \hat{u}_s \\ \hat{v}_r \\ \hat{\psi} \end{bmatrix} = \begin{bmatrix} \alpha_{s1} & \alpha_{s2} & \alpha_{s3} & \alpha_{s1} & \alpha_{s2} & \alpha_{s3} \\ 1 & 1 & 1 & 1 & 1 & 1 \\ -ik_{s1} & -ik_{s2} & -ik_{s3} & ik_{s1} & ik_{s2} & ik_{s3} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_s^+ & 0 \\ 0 & \mathbf{\Lambda}_s^- \end{bmatrix} \begin{bmatrix} A_{s1}^+ \\ A_{s2}^+ \\ A_{s3}^+ \\ B_{s1}^- \\ B_{s2}^- \\ B_{s3}^- \end{bmatrix}$$

$$
\hat{\mathbf{f}}_{CB} = \begin{bmatrix} \hat{T}_s \\ \hat{V}_r \\ \hat{M}_{zs} \end{bmatrix} = \left(C_1 \begin{bmatrix} -ik_{s1} & -ik_{s2} & -ik_{s3} & ik_{s1} & ik_{s2} & ik_{s3} \\ k_{s1}^2 & k_{s2}^2 & k_{s3}^2 & -k_{s1}^2 & -k_{s2}^2 & -k_{s3}^2 \\ -ik_{s1} & -ik_{s2} & -ik_{s3} & ik_{s1} & ik_{s2} & ik_{s3} \end{bmatrix} \Gamma_s \\ +C_2 \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ ik_{s1}^3 & ik_{s2}^3 & ik_{s3}^3 & -ik_{s1}^3 & -ik_{s2}^3 & -ik_{s3}^3 \\ k_{s1}^2 & k_{s2}^2 & k_{s3}^2 & -k_{s1}^2 & -k_{s2}^2 & -k_{s3}^2 \end{bmatrix} \right) \begin{bmatrix} \Lambda_s^+ & 0 \\ 0 & \Lambda_s^- \end{bmatrix} \begin{bmatrix} A_{s1}^+ \\ A_{s2}^+ \\ A_{s3}^+ \\ B_{s1}^- \\ B_{s2}^- \\ B_{s3}^- \end{bmatrix}
$$

where

$$
\Lambda_s^+ = \begin{bmatrix} e^{-ik_{s1}s} & 0 & 0 \\ 0 & e^{-ik_{s2}s} & 0 \\ 0 & 0 & e^{-ik_{s3}s} \end{bmatrix}, \quad \Lambda_s^- = \begin{bmatrix} e^{ik_{s1}(s-s_0)} & 0 & 0 \\ 0 & e^{ik_{s2}(s-s_0)} & 0 \\ 0 & 0 & e^{ik_{s3}(s-s_0)} \end{bmatrix},
$$

$$
\Gamma_s = diag\begin{bmatrix} \alpha_{s1} & \alpha_{s2} & \alpha_{s3} & \alpha_{s1} & \alpha_{s2} & \alpha_{s3} \end{bmatrix},
$$

$$
C_1 = \begin{bmatrix} EA & 0 & 0 \\ 0 & -EI/R & 0 \\ 0 & 0 & EI/R \end{bmatrix}, \quad C_2 = \begin{bmatrix} -EA/R & 0 & 0 \\ 0 & -EI & 0 \\ 0 & 0 & EI \end{bmatrix}.
$$

It should be noticed that the reference for the wave amplitudes related to negative-going waves is placed at the end of the beam element, i.e., at $s = s_0$.

Spectral Element

Consider an element of curved beam with circumferential length s_0. The displacements at both ends of the spectral element can be expressed as a function of the wave amplitudes as follows:

$$
\begin{bmatrix} \hat{u}_s(0) \\ \hat{v}_r(0) \\ \hat{\psi}(0) \\ \hat{u}_s(s_0) \\ \hat{v}_r(s_0) \\ \hat{\psi}(s_0) \end{bmatrix} = \begin{bmatrix} \alpha_{s1} & \alpha_{s2} & \alpha_{s3} & \alpha_{s1}e^{-ik_{s1}s_0} & \alpha_{s2}e^{-ik_{s2}s_0} & \alpha_{s3}e^{-ik_{s3}s_0} \\ 1 & 1 & 1 & e^{-ik_{s1}s_0} & e^{-ik_{s2}s_0} & e^{-ik_{s3}s_0} \\ -ik_{s1} & -ik_{s2} & -ik_{s3} & ik_{s1}e^{-ik_{s1}s_0} & ik_{s2}e^{-ik_{s2}s_0} & ik_{s3}e^{-ik_{s3}s_0} \\ \alpha_{s1}e^{-ik_{s1}s_0} & \alpha_{s2}e^{-ik_{s2}s_0} & \alpha_{s3}e^{-ik_{s3}s_0} & \alpha_{s1} & \alpha_{s2} & \alpha_{s3} \\ e^{-ik_{s1}s_0} & e^{-ik_{s2}s_0} & e^{-ik_{s3}s_0} & 1 & 1 & 1 \\ -ik_{s1}e^{-ik_{s1}s_0} & -ik_{s2}e^{-ik_{s2}s_0} & -ik_{s3}e^{-ik_{s3}s_0} & ik_{s1} & ik_{s2} & ik_{s3} \end{bmatrix} \begin{bmatrix} A_{s1}^+ \\ A_{s2}^+ \\ A_{s3}^+ \\ B_{s1}^- \\ B_{s2}^- \\ B_{s3}^- \end{bmatrix}
$$

$$
\begin{bmatrix} \hat{\mathbf{q}}_0 \\ \hat{\mathbf{q}}_{s_0} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_{qs}^+ & \mathbf{\Phi}_{qs}^- \Lambda_{s_0}^+ \\ \mathbf{\Phi}_{qs}^+ \Lambda_{s_0}^+ & \mathbf{\Phi}_{qs}^- \end{bmatrix} \begin{bmatrix} \mathbf{A}_s^+ \\ \mathbf{B}_s^- \end{bmatrix}.
$$

Analogously, the external loads applied may be written as:

$$
\begin{bmatrix} -\hat{T}_s(0) \\ -\hat{V}_r(0) \\ \hat{M}_{zs}(0) \\ \hat{T}_s(s_0) \\ \hat{V}_r(s_0) \\ -\hat{M}_{zs}(s_0) \end{bmatrix} = \begin{bmatrix} C_1 \begin{bmatrix} ik_{s1} & ik_{s2} & ik_{s3} & -ik_{s1}e^{-ik_{s1}s_0} & -ik_{s2}e^{-ik_{s2}s_0} & -ik_{s3}e^{-ik_{s3}s_0} \\ -k_{s1}^2 & -k_{s2}^2 & -k_{s3}^2 & k_{s1}^2 e^{-ik_{s1}s_0} & k_{s2}^2 e^{-ik_{s2}s_0} & k_{s3}^2 e^{-ik_{s3}s_0} \\ -ik_{s1} & -ik_{s2} & -ik_{s3} & ik_{s1}e^{-ik_{s1}s_0} & ik_{s2}e^{-ik_{s2}s_0} & ik_{s3}e^{-ik_{s3}s_0} \\ -ik_{s1}e^{-ik_{s1}s_0} & -ik_{s2}e^{-ik_{s2}s_0} & -ik_{s3}e^{-ik_{s3}s_0} & ik_{s1} & ik_{s2} & ik_{s3} \\ k_{s1}^2 e^{-ik_{s1}s_0} & k_{s2}^2 e^{-ik_{s2}s_0} & k_{s3}^2 e^{-ik_{s3}s_0} & -k_{s1}^2 & -k_{s2}^2 & -k_{s3}^2 \\ ik_{s1}e^{-ik_{s1}s_0} & ik_{s2}e^{-ik_{s2}s_0} & ik_{s3}e^{-ik_{s3}s_0} & -ik_{s1} & -ik_{s2} & -ik_{s3} \end{bmatrix} \Gamma_s \\ +C_2 \begin{bmatrix} -1 & -1 & -1 & -e^{-ik_{s1}s_0} & -e^{-ik_{s2}s_0} & -e^{-ik_{s3}s_0} \\ -ik_{s1}^3 & -ik_{s2}^3 & -ik_{s3}^3 & ik_{s1}^3 e^{-ik_{s1}s_0} & ik_{s2}^3 e^{-ik_{s2}s_0} & ik_{s3}^3 e^{-ik_{s3}s_0} \\ k_{s1}^2 & k_{s2}^2 & k_{s3}^2 & -k_{s1}^2 e^{-ik_{s1}s_0} & -k_{s2}^2 e^{-ik_{s2}s_0} & -k_{s3}^2 e^{-ik_{s3}s_0} \\ e^{-ik_{s1}s_0} & e^{-ik_{s2}s_0} & e^{-ik_{s3}s_0} & 1 & 1 & 1 \\ ik_{s1}^3 e^{-ik_{s1}s_0} & ik_{s2}^3 e^{-ik_{s2}s_0} & ik_{s3}^3 e^{-ik_{s3}s_0} & -ik_{s1}^3 & -ik_{s2}^3 & -ik_{s3}^3 \\ -k_{s1}^2 e^{-ik_{s1}s_0} & -k_{s2}^2 e^{-ik_{s2}s_0} & -k_{s3}^2 e^{-ik_{s3}s_0} & k_{s1}^2 & k_{s2}^2 & k_{s3}^2 \end{bmatrix} \end{bmatrix} \begin{bmatrix} A_{s1}^+ \\ A_{s2}^+ \\ A_{s3}^+ \\ B_{s1}^- \\ B_{s2}^- \\ B_{s3}^- \end{bmatrix}
$$

$$
\begin{bmatrix} \hat{\mathbf{F}}_0 \\ \hat{\mathbf{F}}_{s_0} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{\Phi}_{fs}^+ & -\boldsymbol{\Phi}_{fs}^- \boldsymbol{\Lambda}_{s_0}^+ \\ \boldsymbol{\Phi}_{fs}^+ \boldsymbol{\Lambda}_{s_0}^+ & \boldsymbol{\Phi}_{fs}^- \end{bmatrix} \begin{bmatrix} \mathbf{A}_s^+ \\ \mathbf{B}_s^- \end{bmatrix}.
$$

With the general method for computing $\mathbf{A}_s^+$ and $\mathbf{B}_s^-$ from the displacement equation and replacing it in the expression for external loads, one can write:

$$
\begin{bmatrix} \hat{\mathbf{F}}_0 \\ \hat{\mathbf{F}}_{s_0} \end{bmatrix} = \mathbf{K}_{CB}(\omega) \begin{bmatrix} \hat{\mathbf{q}}_0 \\ \hat{\mathbf{q}}_{s_0} \end{bmatrix},
$$

$$
\mathbf{K}_{CB}(\omega) = \begin{bmatrix} -\boldsymbol{\Phi}_{fs}^+ & -\boldsymbol{\Phi}_{fs}^- \boldsymbol{\Lambda}_{s_0}^+ \\ \boldsymbol{\Phi}_{fs}^+ \boldsymbol{\Lambda}_{s_0}^+ & \boldsymbol{\Phi}_{fs}^- \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_{qs}^+ & \boldsymbol{\Phi}_{qs}^- \boldsymbol{\Lambda}_{s_0}^+ \\ \boldsymbol{\Phi}_{qs}^+ \boldsymbol{\Lambda}_{s_0}^+ & \boldsymbol{\Phi}_{qs}^- \end{bmatrix}^{-1}.
$$

Rod Under Torsion (Shaft)

Consider a rod element subjected to a distributed torsion moment (torque) m_x as shown in Fig. 7. Due to the applied load, at a position x, the cross section rotates about the x-axis by an angle $\theta_x(x)$. Here, it is assumed that the cross section remains plane while twisting. For wavelengths which are greater than ten times the cross-section dimensions, the Saint-Venant theory is also valid (Petyt 2010). In this case, the kinematics of a rod under torsion may be approximated as:

$$
u(x, y, z) = \psi(y, z) \frac{\partial \theta_x(x)}{\partial x}
$$
$$
v(x, y, z) = -z\theta_x(x)
$$
$$
w(x, y, z) = y\theta_x(x)
$$

where $\psi(y, z)$ is a function which represents the warping of the cross section. In this chapter, only rods of circular cross section are considered. In this case, there is no warping ($\psi = 0$) and the nonzero strain components are:

$$
\varepsilon_{xy} = \frac{1}{2}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) \rightarrow \varepsilon_{xy} = -\frac{z}{2}\frac{\partial \theta_x}{\partial x}, \; \gamma_{xy} = -z\frac{\partial \theta_x}{\partial x}
$$
$$
\varepsilon_{xz} = \frac{1}{2}\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial y}\right) \rightarrow \varepsilon_{xz} = \frac{y}{2}\frac{\partial \theta_x}{\partial x}, \; \gamma_{xz} = y\frac{\partial \theta_x}{\partial x}.
$$

The associated stresses are given by:

$$
\sigma_{xy} = G\gamma_{xy} \rightarrow \; \sigma_{xy} = -zG\frac{\partial \theta_x}{\partial x}
$$
$$
\sigma_{xz} = G\gamma_{xz} \rightarrow \; \sigma_{xz} = yG\frac{\partial \theta_x}{\partial x}
$$

where G is the shear modulus. In cylindrical coordinates, the resultant stress is:

$$
\sigma_{xr} = G\gamma_{xr} \rightarrow \; \sigma_{xr} = rG\frac{\partial \theta_x}{\partial x}.
$$

By applying Newton's laws or variational formulations, the equation of motion for this element is derived:

Spectral Finite Element Approach for Structural Dynamics, Fig. 7 Rod element model under torsion (shaft)
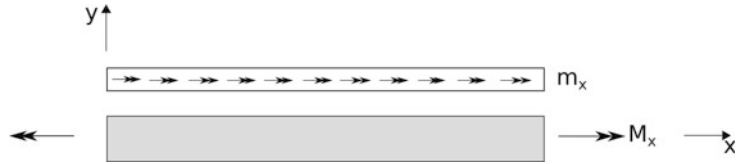
$$\frac{\partial}{\partial x}\left[GJ\frac{\partial \theta_x}{\partial x}\right] - \eta J\frac{\partial \theta_x}{\partial t} - \rho J\frac{\partial^2 \theta_x}{\partial t^2} = -m_x$$

where θ_x is the rotation with respect to the x-axis, J is the torsion constant or the polar moment of area with respect to the x-axis for concentric circular rods (round shafts), η is the damping (viscous) coefficient per unit volume, and ρ is the mass density.

The associated boundary conditions in this case are the rotation θ_x and the internal torque $M_x = GJ\frac{\partial \theta_x}{\partial x}$.

Spectral Analysis
As it was done in previous sections, by applying two Fourier transforms (time and space), the displacement field is written as spectral solutions, in the form:

$$\theta_x(s,t) = \sum_n \hat{\theta}_{xn}(x,\omega_n)e^{i\omega_n t}$$
$$= \sum_n \sum_m \widetilde{\theta}_x(k_{xxm},\omega)e^{-ik_m s}e^{i\omega_n t}.$$

Then, assuming geometric and material properties constant along the rod length and substituting this expression in the homogeneous differential equation of motion, the spectrum relation is obtained:

$$\left(-GJk_{xx}^2 - i\omega\eta J + \omega^2 \rho J\right)\theta_x = 0,$$
$$k_{xx} = \pm\sqrt{\frac{\omega^2 \rho J - i\omega\eta}{GJ}}.$$

Thus, the rotation θ_x and the internal moment M_x evaluated at an arbitrary point along the rod may be expressed as:

$$\theta_x(x,t) = \sum_n \left(A_{xx1}^+ e^{-ik_{xxn}x} + A_{xx1}^- e^{ik_{xxn}x}\right)e^{i\omega_n t}$$
$$M_x(x,t) = \sum_n ik_{xxn}GJ\left(-A_{xx1}^+ e^{-ik_{xxn}x} + A_{xx1}^- e^{ik_{xxn}x}\right)e^{i\omega_n t}.$$

Setting the reference for negative-going waves at $x = L$, the rotation and the internal load are written in matrix form:

$$\hat{q}_{xx} = \hat{\theta}_x = \begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} e^{-ik_{xx}x} & 0 \\ 0 & e^{ik_{xx}(x-L)} \end{bmatrix}\begin{bmatrix} A_{xx1}^+ \\ B_{xx1}^- \end{bmatrix}$$
$$\hat{f}_{xx} = \hat{M}_x = ik_{xx}GJ\begin{bmatrix} -1 & 1 \end{bmatrix}\begin{bmatrix} e^{-ik_{xx}x} & 0 \\ 0 & e^{ik_{xx}(x-L)} \end{bmatrix}\begin{bmatrix} A_{xx1}^+ \\ B_{xx1}^- \end{bmatrix}.$$

Spectral Element
The spectral element of a shaft element in torsion of length L is straightforwardly obtained by writing the rotations with respect to the x-axis at both ends as a function of the wave amplitudes, as follows:

$$\begin{bmatrix} \hat{\theta}_x(0) \\ \hat{\theta}_x(L) \end{bmatrix} = \begin{bmatrix} 1 & e^{-ik_{xx}L} \\ e^{-ik_{xx}L} & 1 \end{bmatrix}\begin{bmatrix} A_{xx1}^+ \\ B_{xx1}^- \end{bmatrix},$$

i.e.,

$$\begin{bmatrix} \hat{q}_{xx0} \\ \hat{q}_{xxL} \end{bmatrix} = \begin{bmatrix} \Phi_{qxx}^+ & \Phi_{qxx}^- \Lambda_L^+ \\ \Phi_{qxx}^+ \Lambda_L^+ & \Phi_{qxx}^- \end{bmatrix}\begin{bmatrix} A_{xx}^+ \\ B_{xx}^- \end{bmatrix}.$$

The applied torque may be also evaluated at both ends:

$$\begin{bmatrix} -\hat{M}_x(0) \\ \hat{M}_x(L) \end{bmatrix} = ik_{xx}GJ\begin{bmatrix} 1 & -e^{-ik_{xx}L} \\ -e^{-ik_{xx}L} & 1 \end{bmatrix}\begin{bmatrix} A_{xx1}^+ \\ B_{xx1}^- \end{bmatrix}$$

or

$$\begin{bmatrix} \hat{F}_{xx0} \\ \hat{F}_{xxL} \end{bmatrix} = \begin{bmatrix} -\Phi_{fxx}^+ & -\Phi_{fxx}^- \Lambda_L^+ \\ \Phi_{fxx}^+ \Lambda_L^+ & \Phi_{fxx}^- \end{bmatrix}\begin{bmatrix} A_{xx}^+ \\ B_{xx}^- \end{bmatrix}.$$

Spectral Finite Element Approach for Structural Dynamics, Fig. 8 3D frame member
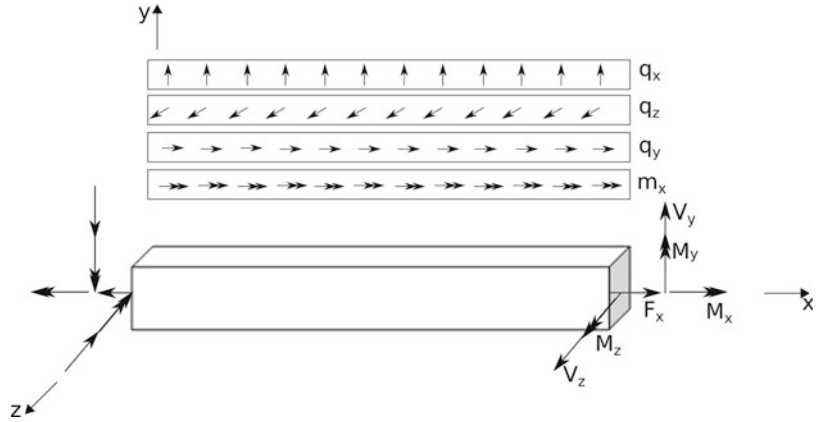
Thus, by writing the wave amplitudes as function of the rotation vector and substituting it into the expression for external applied moments, the spectral element matrix is obtained:

$$\begin{bmatrix} \hat{F}_{xx0} \\ \hat{F}_{xxL} \end{bmatrix} = \mathbf{K}_{xx}(\omega) \begin{bmatrix} \hat{q}_{xx0} \\ \hat{q}_{xxL} \end{bmatrix},$$

$$\mathbf{K}_{xx}(\omega) = \begin{bmatrix} -\boldsymbol{\Phi}_{fxx}^+ & -\boldsymbol{\Phi}_{fxx}^- \boldsymbol{\Lambda}_L^+ \\ \boldsymbol{\Phi}_{fxx}^+ \boldsymbol{\Lambda}_L^+ & \boldsymbol{\Phi}_{fxx}^- \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_{qxx}^+ & \boldsymbol{\Phi}_{qxx}^- \boldsymbol{\Lambda}_L^+ \\ \boldsymbol{\Phi}_{qxx}^+ \boldsymbol{\Lambda}_L^+ & \boldsymbol{\Phi}_{qxx}^- \end{bmatrix}^{-1},$$

$$\mathbf{K}_{xx}(\omega) = GJk_{xx} \sin(k_{xx}L) \begin{bmatrix} \cos(k_{xx}L) & -1 \\ -\cos(k_{xx}L) & 1 \end{bmatrix}.$$

3D Frame Structure

A 3D frame structure is a combination of one-dimensional elements in the 3D space that are capable to support axial, bending, and torsion efforts (Fig. 8). Thus, using the expressions derived in the previous sections, a spectral element for a 3D frame element can be easily built. It is composed of six degrees of freedom per node (three translations and three rotations), which allow an axial deformation and torsion along the main axis and bending in the two transverse planes. Its displacement field may be expressed as:

$$\overline{u}(x,y,z,t) = u(x,t) \quad \overline{\theta}_x(x,y,z,t) = \theta_x(x,t)$$
$$\overline{v}(x,y,z,t) = v(x,t) \quad \overline{\theta}_y(x,y,z,t) = \theta_y(x,t) = -\frac{\partial w(x,t)}{\partial x}$$
$$\overline{w}(x,y,z,t) = w(x,t) \quad \overline{\theta}_z(x,y,z,t) = \theta_z(x,t) = \frac{\partial v(x,t)}{\partial x}$$

It is important to emphasize at this point that the expression for the rotation w.r.t. the y-axis requires a negative sign as a result of the assumption that translations are positive in the sense of the Cartesian coordinates and also the respect of the right-hand rule for the coordinate system. By superposing the spectral solutions for a beam element under axial, flexural, and torsion deformations, the displacement solution vector is given by:

$$\hat{\mathbf{q}}_{3D}(x) = \begin{bmatrix} \hat{u}(x) \\ \hat{v}(x) \\ \hat{w}(x) \\ \hat{\theta}_x(x) \\ \hat{\theta}_y(x) \\ \hat{\theta}_z(x) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}_q^+ & \boldsymbol{\Phi}_q^- \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}_x^+ & 0 \\ 0 & \boldsymbol{\Lambda}_x^- \end{bmatrix} \begin{bmatrix} \mathbf{A}^+ \\ \mathbf{B}^- \end{bmatrix},$$

where:

$$\boldsymbol{\Phi}_q^+ = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & i\beta_z & \beta_z & 0 \\ 0 & -i\beta_y & -\beta_y & 0 & 0 & 0 \end{bmatrix},$$

$$\boldsymbol{\Phi}_q^- = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -i\beta_z & -\beta_z & 0 \\ 0 & i\beta_y & \beta_y & 0 & 0 & 0 \end{bmatrix},$$

S

$$\mathbf{\Lambda}_x^+ = \begin{bmatrix} e^{-ik_xx} & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{-i\beta_yx} & 0 & 0 & 0 & 0 \\ 0 & 0 & e^{-\beta_yx} & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{-i\beta_zx} & 0 & 0 \\ 0 & 0 & 0 & 0 & e^{-\beta_zx} & 0 \\ 0 & 0 & 0 & 0 & 0 & e^{-ik_{xx}x} \end{bmatrix},$$

$$\mathbf{\Lambda}_x^- = \begin{bmatrix} e^{ik_x(x-L)} & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{i\beta_y(x-L)} & 0 & 0 & 0 & 0 \\ 0 & 0 & e^{\beta_y(x-L)} & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{i\beta_z(x-L)} & 0 & 0 \\ 0 & 0 & 0 & 0 & e^{\beta_z(x-L)} & 0 \\ 0 & 0 & 0 & 0 & 0 & e^{ik_{xx}(x-L)} \end{bmatrix}.$$

The corresponding internal loads acting on a 3D frame member are:

$$F_x = EA\frac{\partial u}{\partial x} \qquad M_x = GJ\frac{\partial \theta_x}{\partial x}$$
$$V_y = -\frac{\partial}{\partial x}\left(EI_z\frac{\partial^2 v}{\partial x^2}\right) \quad M_y = -EI_y\frac{\partial^2 w}{\partial x^2}$$
$$V_z = -\frac{\partial}{\partial x}\left(EI_y\frac{\partial^2 w}{\partial x^2}\right) \quad M_z = EI_z\frac{\partial^2 v}{\partial x^2}$$

which, in vector form, yields:

$$\hat{\mathbf{f}}_{3D}(x) = \begin{bmatrix} \hat{F}_x(x) \\ \hat{V}_y(x) \\ \hat{V}_z(x) \\ \hat{M}_x(x) \\ \hat{M}_y(x) \\ \hat{M}_z(x) \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{\Phi}_f^+ & \mathbf{\Phi}_f^- \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_x^+ & 0 \\ 0 & \mathbf{\Lambda}_x^- \end{bmatrix} \begin{bmatrix} \mathbf{A}^+ \\ \mathbf{B}^- \end{bmatrix},$$

with:

$$\mathbf{\Phi}_f^+ = \begin{bmatrix} -ik_xEA & 0 & 0 & 0 & 0 & 0 \\ 0 & -iEI_z\beta_y^3 & EI_z\beta_y^3 & 0 & 0 & 0 \\ 0 & 0 & 0 & -iEI_y\beta_z^3 & EI_Y\beta_z^3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -ik_{xx}GJ \\ 0 & 0 & 0 & EI_y\beta_z^2 & -EI_y\beta_z^2 & 0 \\ 0 & -EI_z\beta_y^2 & EI_z\beta_y^2 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{\Phi}_f^- = \begin{bmatrix} -ik_xEA & 0 & 0 & 0 & 0 & 0 \\ 0 & iEI_z\beta_y^3 & -EI_z\beta_y^3 & 0 & 0 & 0 \\ 0 & 0 & 0 & iEI_y\beta_z^3 & -EI_Y\beta_z^3 & 0 \\ 0 & 0 & 0 & 0 & 0 & ik_{xx}GJ \\ 0 & 0 & 0 & EI_y\beta_z^2 & -EI_y\beta_z^2 & 0 \\ 0 & -EI_z\beta_y^2 & EI_z\beta_y^2 & 0 & 0 & 0 \end{bmatrix}.$$

A spectral element of frame with length L is straightforwardly obtained by writing displacement and external loads at both ends, as follows:

$$\begin{bmatrix} \hat{\mathbf{q}}_{3D0} \\ \hat{\mathbf{q}}_{3DL} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}_q^+ & \boldsymbol{\Phi}_q^- \boldsymbol{\Lambda}_L^+ \\ \boldsymbol{\Phi}_q^+ \boldsymbol{\Lambda}_L^+ & \boldsymbol{\Phi}_q^- \end{bmatrix} \begin{bmatrix} \mathbf{A}^+ \\ \mathbf{B}^- \end{bmatrix}$$

$$\begin{bmatrix} \hat{\mathbf{F}}_{3D0} \\ \hat{\mathbf{F}}_{3DL} \end{bmatrix} = \begin{bmatrix} -\hat{\mathbf{f}}_{3D}(x=0) \\ \hat{\mathbf{f}}_{3D}(x=L) \end{bmatrix}$$
$$= \begin{bmatrix} -\boldsymbol{\Phi}_f^+ & -\boldsymbol{\Phi}_f^- \boldsymbol{\Lambda}_L^+ \\ \boldsymbol{\Phi}_f^+ \boldsymbol{\Lambda}_L^+ & \boldsymbol{\Phi}_f^- \end{bmatrix} \begin{bmatrix} \mathbf{A}^+ \\ \mathbf{B}^- \end{bmatrix}.$$

Then, the spectral element matrix of a 3D frame member is readily obtained:

$$\begin{bmatrix} \hat{\mathbf{F}}_{3D0} \\ \hat{\mathbf{F}}_{3DL} \end{bmatrix} = \mathbf{K}_{3D}(\omega) \begin{bmatrix} \hat{\mathbf{q}}_{3D0} \\ \hat{\mathbf{q}}_{3DL} \end{bmatrix},$$
$$\mathbf{K}_{3D}(\omega) = \begin{bmatrix} -\boldsymbol{\Phi}_f^+ & -\boldsymbol{\Phi}_f^- \boldsymbol{\Lambda}_L^+ \\ \boldsymbol{\Phi}_f^+ \boldsymbol{\Lambda}_L^+ & \boldsymbol{\Phi}_f^- \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_q^+ & \boldsymbol{\Phi}_q^- \boldsymbol{\Lambda}_L^+ \\ \boldsymbol{\Phi}_q^+ \boldsymbol{\Lambda}_L^+ & \boldsymbol{\Phi}_q^- \end{bmatrix}^{-1}.$$

Spectral Element Modeling

In previous sections, it has been shown how to formulate spectral element matrices for straight and curved frame members. However, in real engineering systems, those structures do not appear isolated; they are usually connected to other basic structures involving or not discontinuities. In this section, it will be shown that the spectral element method can be efficiently used to compute the forced response of coupled systems in the frequency and in the time domains. This is particularly useful in earthquake engineering, as it allows simulating the effect of seismic ground motion at a very low computational cost compared to conventional computational analyses via FEM or BEM. This is possible as SEM allows building exact element matrices at each frequency. Therefore, mesh refinement is not required with the increase in frequency, and it is possible to build spectral models with very few degrees of freedom. On the other hand, as it makes use of matrix formulations as FEM does. The assemblage of spectral elements is quite simple via the direct stiffness method (Doyle 1997). It is also important to note that the global problem is solved in the frequency domain. For every discrete frequency, an assemblage process identical to that of the static case is performed and the global response vector is obtained. The discrete Fourier transform is used to obtain the spectral representation of the input data, and the inverse DFT allows one to recover the response of the coupled system in the time domain.

Global Dynamic Stiffness Matrix Assembling Procedure

Consider a frame element as shown in Fig. 9. Two types of coordinate system are defined: a global one, defined by the orthogonal X, Y, and Z directions, and a local one, defined by x, y, and z axes with the x-axis parallel to the main axis (neutral axis) of the frame member.

The spectral element matrix, as formulated previously for frame member, is written with respect to the local reference frame, as follows:

$$\hat{\mathbf{K}}_{xyz}(\omega)\hat{\mathbf{q}}_{xyz}(\omega) = \hat{\mathbf{F}}_{xyz}(\omega).$$

Load and displacement vectors in the global reference frame may be related to those defined in the local coordinate system by application of successive rotations from the local axes to the global axes:

$$\hat{\mathbf{q}}_{XYZ} = \mathbf{T}\hat{\mathbf{q}}_{xyz}$$
$$\hat{\mathbf{F}}_{XYZ} = \mathbf{T}\hat{\mathbf{F}}_{xyz}$$

with:

$$\mathbf{T} = \begin{bmatrix} \mathbf{R}_x(\psi)\mathbf{R}_y(\varphi)\mathbf{R}_z(\theta) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_x(\psi)\mathbf{R}_y(\varphi)\mathbf{R}_z(\theta) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{R}_x(\psi)\mathbf{R}_y(\varphi)\mathbf{R}_z(\theta) \end{bmatrix},$$

Spectral Finite Element Approach for Structural Dynamics, Fig. 9 Frame element relative to local and global coordinate systems
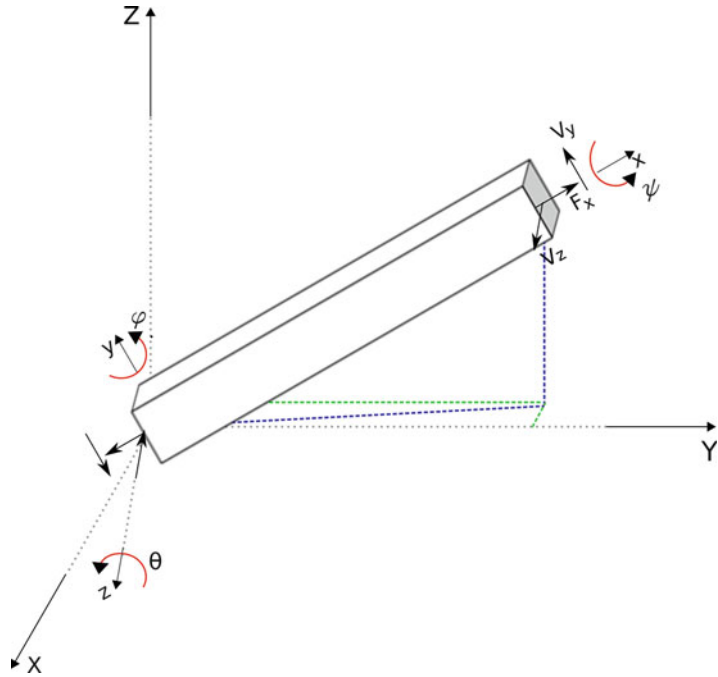
where $\mathbf{R}_x$, $\mathbf{R}_y$, and $\mathbf{R}_z$ are rotation matrices, therefore unitary $\left(\mathbf{R}^T\mathbf{R} = \mathbf{I}\right)$, defined as:

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{bmatrix},$$

$$\mathbf{R}_y = \begin{bmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{bmatrix},$$

$$\mathbf{R}_z = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Applying the transformation matrix to the local system, the equilibrium system of equations for a frame member in the global reference frame is written:

$$\hat{\mathbf{K}}_{XYZ}(\omega)\hat{\mathbf{q}}_{XYZ}(\omega) = \hat{\mathbf{F}}_{XYZ}(\omega),$$

with:

$$\hat{\mathbf{K}}_{XYZ}(\omega) = \mathbf{T}\hat{\mathbf{K}}_{xyz}(\omega)\mathbf{T}^T.$$

The transformation of coordinate system is performed for each structural member. This allows one to enforce kinematic compatibility and equilibrium at all nodes connecting structural members. Then, the global dynamic system is constructed following classic direct stiffness procedure:

$$\hat{\mathbf{K}}_G(\omega) = \sum_i \mathbf{L}^{(i)T}\hat{\mathbf{K}}^{(i)}_{XYZ}(\omega)\mathbf{L}^{(i)},$$

where $\mathbf{L}^{(i)}$ is the incidence matrix linking the local vector of degrees of freedom of a structural member i to the global vector of degrees of freedom.

For each discrete frequency, the global problem to be solved is:

$$\hat{\mathbf{K}}_G\hat{\mathbf{q}}_G = \hat{\mathbf{F}}_G,$$
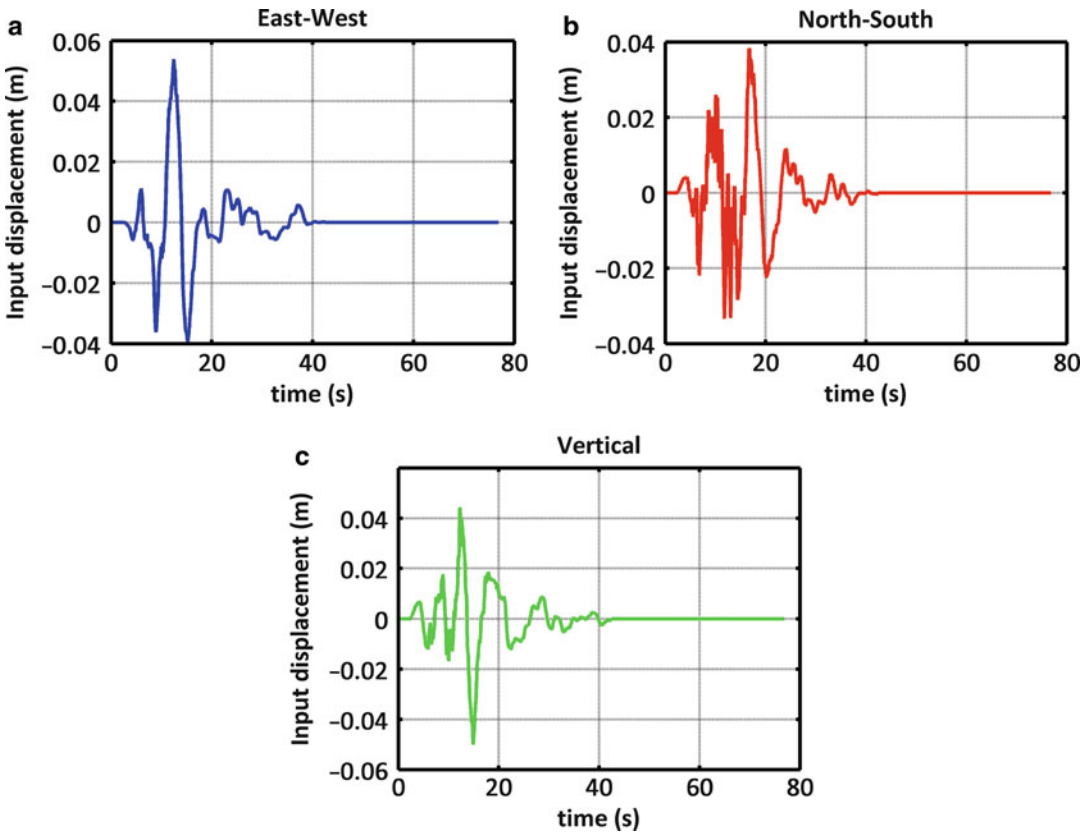
Spectral Finite Element Approach for Structural Dynamics, Fig. 10 Seismic records using as input displacements in numerical experiments
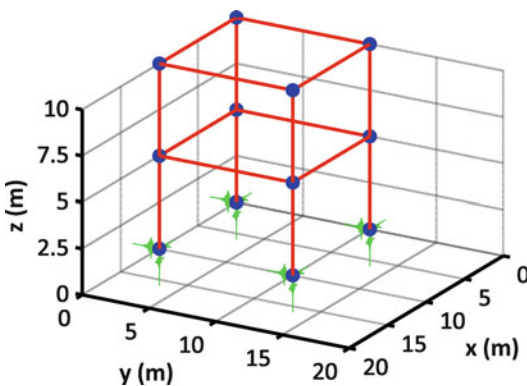
Spectral Finite Element Approach for Structural Dynamics, Fig. 11 3D frame model

subjected to:

$$\hat{\mathbf{F}}_F = \hat{\mathbf{F}}_0,$$
$$\hat{\mathbf{q}}_P = \hat{\mathbf{q}}_0,$$

where the subscripts F and P are used, respectively, to denote sets of degrees of freedom where Neumann (prescribed loads) and Dirichlet (prescribed displacements/rotations) boundary conditions are applied.

The solution of the global system is possible by rewriting it as:

$$\begin{bmatrix} \hat{\mathbf{K}}_{FF} & \hat{\mathbf{K}}_{FP} \\ \hat{\mathbf{K}}_{PF} & \hat{\mathbf{K}}_{PP} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{q}}_F \\ \hat{\mathbf{q}}_0 \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{F}}_0 \\ \hat{\mathbf{F}}_P \end{bmatrix},$$

S
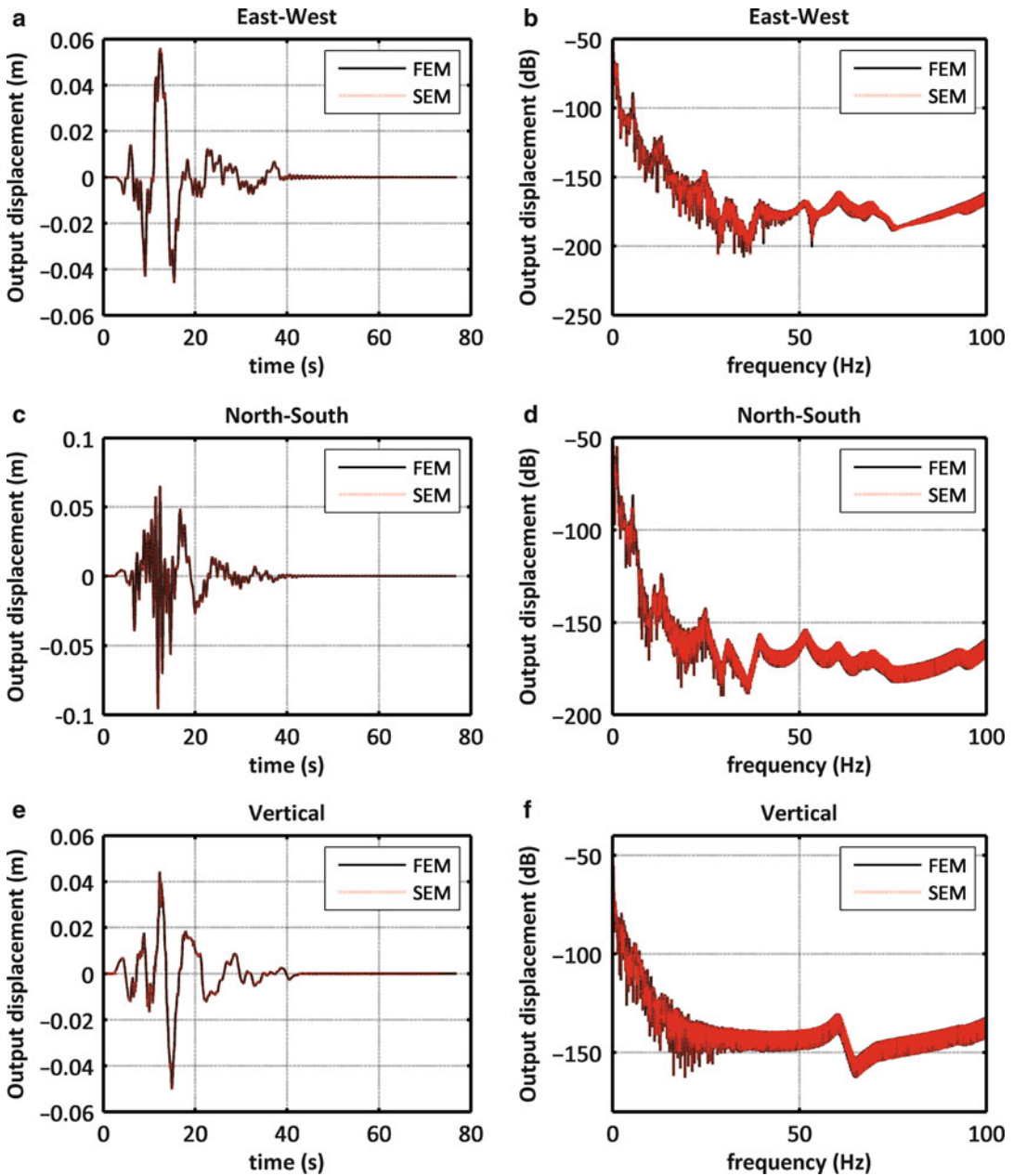
Spectral Finite Element Approach for Structural Dynamics, Fig. 12 Response of the 3D frame structure to the earthquake measured at (x = 10, y = 10) m

which allows one to get:

$$\hat{\mathbf{q}}_F = \hat{\mathbf{K}}_{FF}^{-1}\left(\hat{\mathbf{F}}_0 - \hat{\mathbf{K}}_{FP}\hat{\mathbf{q}}_0\right)$$
$$\hat{\mathbf{F}}_P = \left(\hat{\mathbf{K}}_{PP} - \hat{\mathbf{K}}_{PF}\hat{\mathbf{K}}_{FF}^{-1}\hat{\mathbf{K}}_{FP}\right)\hat{\mathbf{q}}_0 + \hat{\mathbf{K}}_{PF}\hat{\mathbf{K}}_{FF}^{-1}\hat{\mathbf{F}}_0.$$

Numerical Examples

Earthquake Signal

Seismic records from the Loma Prieta earthquake which struck the San Francisco Bay area of

California in 1989 are used as input displacements applied to the bottom nodes of the framework. The input signal in time domain has been obtained from an online database available in Silva et al. (2000) and it is shown in Fig. 10.

3D Frame

Here, the dynamic response of a 3D framework subjected to seismic ground motion is analyzed. The model is presented in Fig. 11. Each frame member is made of steel with mass density $\rho = 7,800 \, \text{Kg.m}^{-3}$, Young's modulus $E = 210 \, \text{GPa}$, structural loss factor $\eta = 0.04$, and Poisson's ratio $v = 0.3$. All members have a circular cross section of radius 0.15 m.

The response of the structure at $(x = 10, y = 10)$ m along the three orthogonal directions is shown in Fig. 12, in the time and in the frequency domains. It is worth pointing out that the SEM has yielded a great reduction of the problem size compared to classic FEM. The response obtained from the SEM using 16 elements is in good agreement with the one obtained by solving a classic FE problem composed of 1200 FE elements. An animation of the framework dynamics during the passage of seismic waves is simulated via the SEM and it is shown in Video 1.

2D Frame

A 2D frame structure modeled using straight and curved beam elements is shown in Fig. 13. The frame has a square cross section 0.20 m × 0.20 m and steel as material, with Young's modulus $E = 210$ GPa, mass density $\rho = 7,800 \, \text{kg/m}^3$, and structural damping $\eta = 0.02$.

As in the previous section, the structure is excited at the base nodes by the Loma Prieta earthquake signal. Because of the in-plane configuration, only the east–west and vertical displacement data are used.

The output displacements at the node (x = 5, y = 0) m are shown in Fig. 14. The results are validated with FEM results obtained using
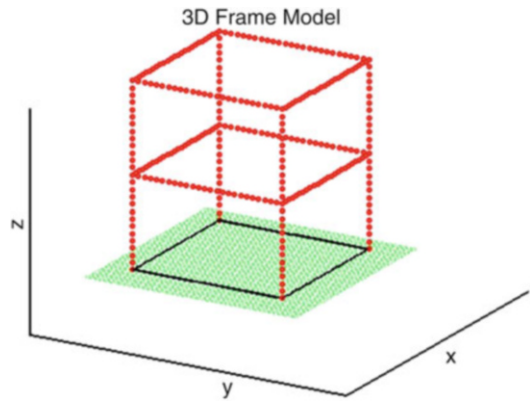
Spectral Finite Element Approach for Structural Dynamics, Video 1 Animation of a 3D frame under seismic ground motion
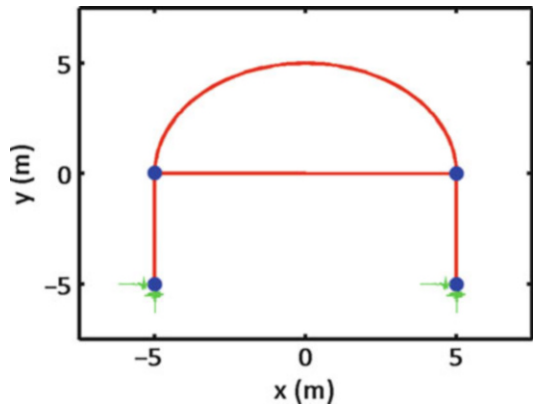
Spectral Finite Element Approach for Structural Dynamics, Fig. 13 2D frame model

ANSYS®. As in the previous numerical problem, fewer degrees of freedom were required to solve the problem via the spectral element method compared to the classic FEM, which results in CPU time savings. The 2D frame motion caused by the earthquake passage was computed via the SEM and it is available in Video 2. The methodology presented in this chapter can be used to derive other elements, such as plates and shells (Doyle 1997), and to study more complex structures under dynamic excitation.

S
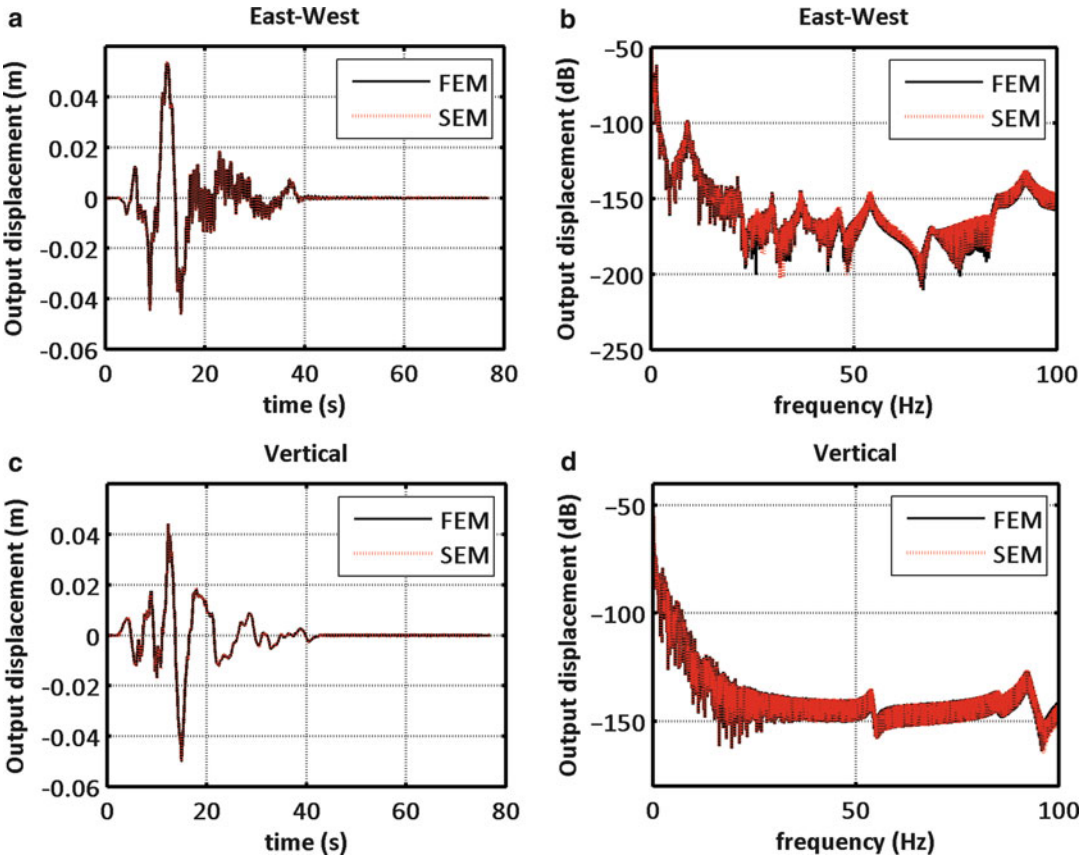
Spectral Finite Element Approach for Structural Dynamics, Fig. 14 Response of the 2D frame structure to the earthquake measured at $(x = 5, y = 0)$ m
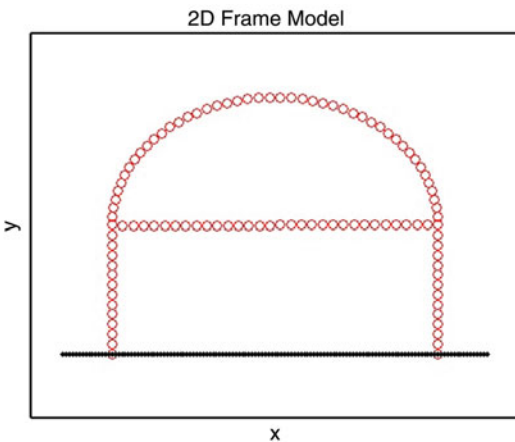
Spectral Finite Element Approach for Structural Dynamics, Video 2 Animation of a 2D frame under seismic ground motion

Summary

Elementary matrix of straight and curved beams, rod under axial force, and torsion were developed using spectral finite element approach. This method can be used to analyze the structural dynamic behavior. From the perspective of earthquake engineering, SEM allows to design structures subject to seismic excitation. The approach has advantages such as low computational cost, accuracy in frequency/time response, and simple modeling of complex structures. The equations of motion, spectral relation, and the spectral elementary matrix were derived. The simple elements were combined into a 3D frame element and a global matrix was assembled. Time and frequency responses of 3D and 2D frame structures under a

known seismic ground motion were computed with SEM and results were validated using well-known FEM commercial software.

Cross-References

References

Ahmida KM, Arruda JRF (2001) Spectral element-based prediction of active power flow in Timoshenko beams. Int J Solids Struct 38:1669–1679

Doyle JF (1997) Wave propagation in structures, 2nd edn. Springer, New York

Lee U (2009) Spectral element method in structural dynamic. Wiley, Singapore

Mace BR, Duhamel D, Brennan MJ, Hinke L (2005) Finite element prediction of wave motion in structural waveguides. J Acoust Soc Am 117(5):2835–2843

Mead DJ (1973) A general theory of harmonic wave propagation in linear periodic systems with multiple coupling. J Sound Vib 27(2):235–260

Newmark NM, Hall WJ (1982) Earthquake spectra and design. Earthquake Engineering Research Institute, Berkeley

Petyt M (2010) Introduction to finite element vibration analysis, 2nd edn. Cambridge University Press, Cambridge

Silva W, Bartling D, Spitzer C (2000) PEER strong motion database. Regents of the University of California. http://peer.berkeley.edu/smcat/search.html. Accessed 9 July 2014

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures

George Vasdravellis[1] and
Theodore L. Karavasilis[2]
[1]Institute for Infrastructure and Environment, Heriot-Watt University, Edinburgh, UK
[2]School of Engineering, University of Warwick, Coventry, UK

Synonyms

Mitigation of residual drifts; Optimized seismic dampers; Seismic resilience; Steel self-centering frames

Introduction

Conventional steel moment-resisting frames (MRFs) are currently designed to form a global plastic mechanism under the design basis earthquake (DBE) by developing plastic hinges at the ends of the beams and the bases of the columns. This design methodology offers many advantages, including collapse prevention and initial economy; however, plastic hinges in structural members involve difficulty to inspect and repair damage and local buckling as well as residual drifts. The socioeconomic losses associated with damage and residual drifts are repair costs, increased downtime, and possibly demolition due to the complications associated with large residual drifts (McCormick et al. 2008).

A challenge of modern earthquake engineering is the development, standardization, and practical implementation of resilient minimal-damage structures with the inherent potential to overcome the socioeconomic losses related to earthquake damage by minimizing or avoiding inelastic deformations and residual drifts. Examples of minimal-damage steel structures include frames equipped with rate-dependent passive dampers (Karavasilis et al. 2011; Whittle et al. 2012),

S

frames with steel energy dissipation devices (Karavasilis et al. 2012), and self-centering moment-resisting frames (SC-MRFs) with posttensioned (PT) connections.

SC-MRFs exhibit softening force–drift behavior while eliminating inelastic deformations and residual drifts under strong earthquakes as the result of gap openings developed in beam-to-column interfaces and self-centering capability due to elastic pretensioning elements (e.g., high-strength steel bars) which clamp beams to the columns. PT connections use carefully designed energy dissipation devices which are activated when gaps open and can be classified into yielding devices which dissipate energy through inelastic deformations and devices which dissipate energy through friction. Yielding devices were proposed as angles bolted on the top and bottom flanges of the beam and on the column flanges, dissipating energy through inelastic bending (Ricles et al. 2001, 2002; Garlock et al. 2005); buckling-restrained steel bars placed between the beam flanges and welded on the beam and column, dissipating energy through axial deformations (Christopoulos et al. 2002); reduced flange plates welded around a square hollow section column and bolted on the beam flanges (Chou et al. 2006); and reduced-section or cross-shaped steel plates placed below the bottom flange of the beam (Chou and Lai 2009). Friction-based devices were proposed as friction-bolted surfaces placed on the top and bottom flanges of the beam (Rojas et al. 2004; Kim and Christopoulos 2008a), on the web of the beam (Tsai et al. 2008), or on the bottom flange of the beam (Wolski et al. 2009).

A new steel PT beam–column connection using web hourglass shape steel pins (WHPs) as EDs has been recently developed by the authors. The development phases included (a) large-scale experimental evaluation (Vasdravellis et al. 2013a) of the proposed connection, (b) detailed finite element models (FEM) and associated parametric studies (Vasdravellis et al. 2013b), and (c) numerical simulations and seismic analysis of SC-MRFs equipped with PT connections with WHPs (Dimopoulos et al. 2013). A more recent study evaluated the

hysteresis and fracture capacity of WHPs made of high performance steel, i.e., high-strength steel and stainless steel (Vasdravellis et al. 2014). It was found that the connection isolates inelastic deformations in WHPs, avoids damage in other connection parts as well as in beams and columns, and eliminates residual drifts. The WHPs do not interfere with the composite slab and are very easy to replace without bolting or welding, and so, the connection enables nondisruptive repair and rapid return to building occupancy in the aftermath of a strong earthquake. This chapter presents the development of the PT connection through the experimental evaluation and the main results of the numerical analyses.

Steel Posttensioned Connection with Web Hourglass Pins

Figure 1 shows a schematic representation of an SC-MRF incorporating the proposed PT connection and the details of an exterior PT connection. Two high-strength steel bars located at the mid-depth of the beam, one at each side of the web, pass through holes drilled on the column flanges. The bars are posttensioned and anchored to the exterior column flange and hence clamp the beam to the column. Four cylindrical steel pins (WHPs) are inserted in aligned holes drilled on the web of the beam and on strong supporting plates. The supporting plates are welded to the column flanges and are of greater thickness to provide fixed support boundary conditions to the WHPs. Energy dissipation is provided by the inelastic bending of the WHPs, which are symmetrically placed (close to the top and bottom beam flange) to provide increased lever arm and hence increased internal moment resistance. As shown in Fig. 2, the WHPs are designed to have an hourglass shape to provide enhanced energy dissipation and fracture capacity (Kobori et al. 1992). Both sides of the beam web are reinforced with steel plates to increase the contact surface of the WHPs with the web. In that way, possible ovalization of the holes drilled on the web and the reinforcing plates under the WHP-bearing forces will be negligible, and

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures,
Fig. 1 (**a**) SC-MRF incorporating the proposed PT connection; (**b**) exterior PT connection details
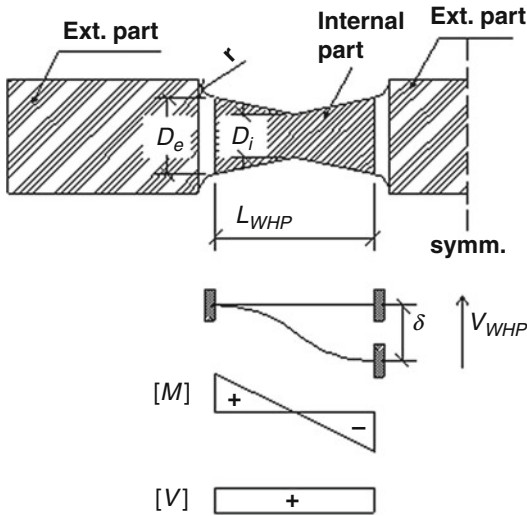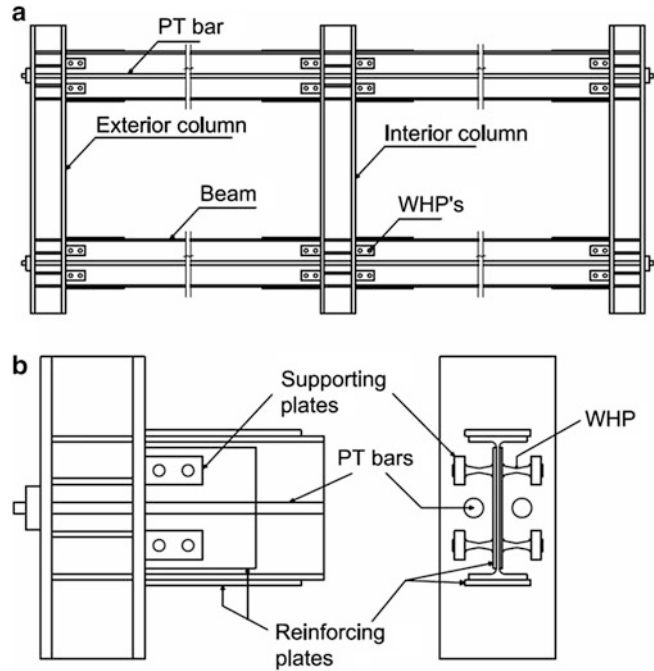
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 2 Geometry of half a WHP, assumed static system, and internal force diagrams

pinching behavior under cyclic deformations can be avoided. The connection includes beam flange-reinforcing plates to avoid excessive early yielding in the beam flanges under high

PT bar forces. In addition, the panel zone is strengthened with horizontal stiffeners and continuity plates along the web of the column.

Connection Behavior and Simplified Analysis

Flexural Behavior

The connection behavior is characterized by gap opening and closing in the beam–column interface as a result of the re-centering force in the PT bars. Figure 3 shows the gap-opening mechanism in the connection where d_{1u} and d_{1l} are the distances of the upper and lower WHPs from the center of rotation (COR), respectively; d_2 is the distance of the PT bars from the COR; F_{PT} is the total force in both PT bars; $F_{WHP,u}$ and $F_{WHP,l}$ are the forces in the upper and lower WHPs, respectively; and C_F is the compressive force on the beam–column bearing surface. It is assumed that the COR is located at the inner edge of the beam flange-reinforcing plate. This assumption was verified during the connection experiments described later.

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures,
Fig. 3 Gap-opening mechanism in the proposed PT connection
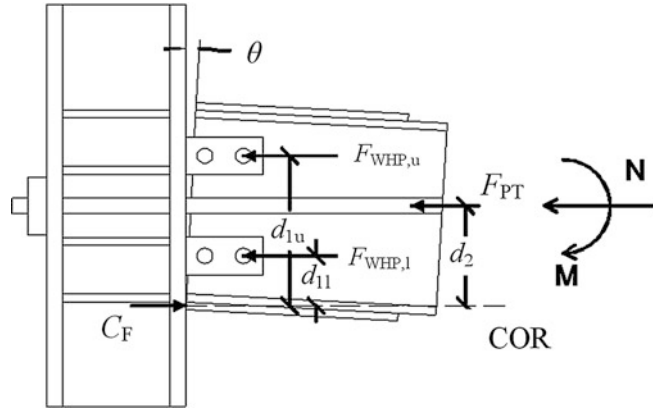
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures,
Fig. 4 Force versus displacement cyclic behavior of one WHP
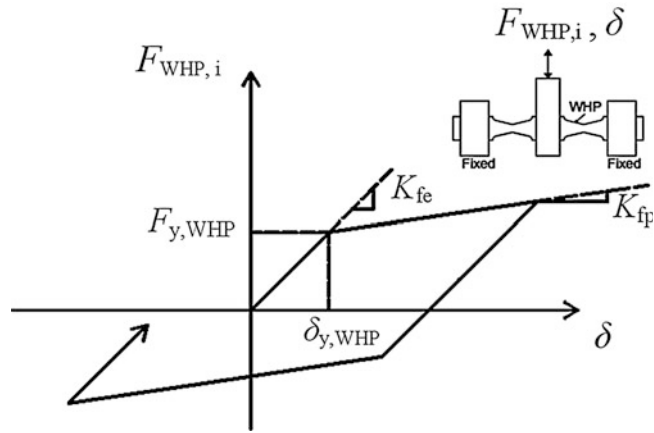
Figure 4 shows the assumed bilinear elastoplastic force–displacement ($F_{\mathrm{WHP},i}$-δ) behavior of one of the four (i = 1 to 4) WHPs. Figure 5a plots the assumed cyclic moment–rotation (M-θ) behavior of the connection. M is the sum of the moment contributions from the PT bars, M_{PT}, and the WHPs, M_{WHP}. Figure 5b shows the nonlinear elastic moment contribution of both PT bars. In Fig. 5b, M_{d} is the decompression moment, i.e., equal to $F_{\mathrm{PT},i}d_2$, where $F_{\mathrm{PT},i}$ is the total initial posttensioning force in both PT bars. Figure 5c shows the moment contribution of the four WHPs. The M_{WHP}-θ curve changes slope two times since the upper WHPs yield at a rotation θ_2 and the lower WHPs yield afterwards at rotation θ_3. Point 1 in Fig. 5a corresponds to the decompression moment, M_{d}, of the connection. After decompression, gap

opens, and the connection behavior becomes nonlinear elastic. Following the procedure described in Garlock et al. (2007), the stiffness K_1 between points 1 and 2 is equal to the sum of the rotational stiffness contribution of the PT bars, K_{PT}, and the rotational stiffness contribution of the WHPs, K_{e}:

$$K_1 = K_{\mathrm{PT}} + K_{\mathrm{e}} = \frac{K_{\mathrm{PT},a}K_b}{K_{\mathrm{PT},a} + K_b}d_2^2 + 2K_{\mathrm{fe}}\left(d_{1l}^2 + d_{1u}^2\right)$$

(1)

where K_{fe} is the elastic stiffness of the force–displacement relationship of one WHP shown in Fig. 4, $K_{\mathrm{pt},a}$ is the total axial elastic stiffness of both PT bars, and K_b is the axial elastic stiffness of the beam. At point 2, the upper WHPs yield and M continues to increase with slope K_2:
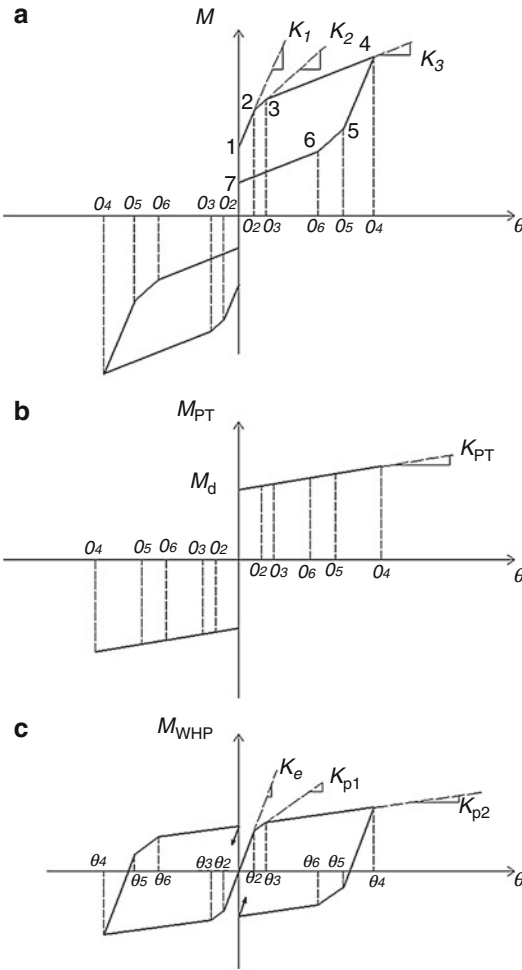
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 5 Conceptual cyclic moment–rotation behavior of the proposed PT connection

$$K_2 = K_{PT} + K_{p1}$$
$$= \frac{K_{PT,a}K_b}{K_{PT,a} + K_b} d_2^2 + 2(K_{fe}d_{1l}^2 + K_{fp}d_{1u}^2) \quad (2)$$

where K_{fp} is the post-elastic stiffness of the force–displacement relationship of one WHP shown in Fig. 4. At point 3, the lower WHPs yield and M continues to increase with slope K_3:

$$K_3 = K_{PT} + K_{p2}$$
$$= \frac{K_{PT,a}K_b}{K_{PT,a} + K_b} d_2^2 + 2K_{fp}(d_{1l}^2 + d_{1u}^2) \quad (3)$$

When loading is reversed, the connection begins to unload, from point 4 to point 5, ideally with the same elastic stiffness, K_1. After point 5, the connection unloads with stiffness equal to K_2 between points 5 and 6 and with stiffness K_3 between points 6 and 7 until the gap closes. The behavior of the connection is symmetric, since the WHPs are placed symmetrically to the beam centerline. The moment in the connection, M, is calculated as:

$$M = M_{PT} + M_{WHP} \quad (4)$$

Following the procedure described by Garlock et al. (2007), MPT is given by:

$$M_{PT} = M_d + \frac{K_{PT,a}K_b}{K_{PT,a} + K_b} d_2^2 \theta \quad (5)$$

According to Fig. 3, MWHP can be obtained as:

$$M_{WHP} = F_{WHP,u}d_{1u} + F_{WHP,l}d_{1l} \quad (6)$$

Plastic Design of WHPs

Plastic analysis is used to calculate the force in the WHPs as they deform due to gap opening and closing. The assumed static system for half a WHP is shown in Fig. 2. The yield strength, V_{WHP}, of half a WHP is controlled either by the plastic moment of resistance, M_{pl}, or the plastic shear resistance, V_{pl}, (Eurocode 3 2003), i.e.:

$$M_{pl} = \frac{D_e^3}{6} f_y \quad (7)$$

$$V_{pl} = \frac{0.9 \frac{\pi D_i^2}{4} f_y}{\sqrt{3}} \quad (8)$$

where f_y is the yield strength of the WHP material, D_e is the equivalent external diameter (to be defined later), and D_i is the diameter at the mid-length of half a WHP, as indicated in Fig. 2. The factor 0.9 in Eq. 8 accounts for the relation between the average shear stress and the maximum shear stress in a circular section. The internal WHP part is connected to the external WHP part using a radius of 5 mm to avoid stress

Fig. 6 Prototype steel building: (**a**) plan view and (**b**) elevation and cross-sections of the interior steel MRF used in this study
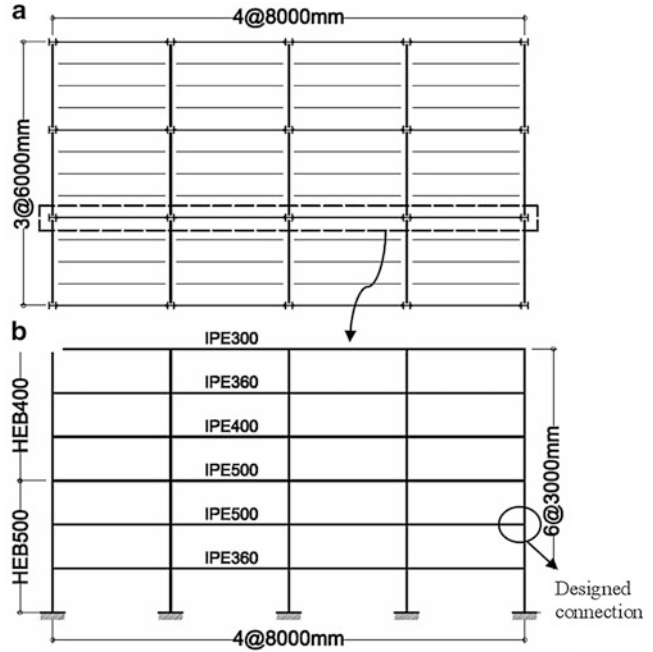
concentration and early fracture. It is assumed that D_e, which controls the WHP bending resistance, is equal to the diameter at the start of the round-shaped part with radius r. Plastic analysis assumes that the plastic moment of resistance should be reached at the ends before the plastic shear resistance is reached at the mid-length of half a WHP. By equilibrium, the aforementioned condition can be written as:

$$V_{\text{WHP}} = \frac{2M_{\text{Pl}}}{L_{\text{WHP}}} < V_{\text{pl}} \qquad (9)$$

where L_{WHP} is the clear length of the bending parts of half a WHP. From Eqs. 7 to 9, any combination of f_y, D_e, D_i, and L_{WHP} can be chosen to provide the desired V_{WHP}. The yield force of a WHP, $F_{y,\text{WHP}}$, is then calculated as:

$$F_{y,\text{WHP}} = 2V_{\text{WHP}} \qquad (10)$$

The virtual work method along with analytical integration was applied to derive the elastic stiffness K_{fe} of a WHP:

$$K_{\text{fe}} = 2\frac{9\pi D_e^3 D_i EG}{\left(40ED_e^2 L_{\text{WHP}} + 48GL_{\text{WHP}}^3\right)} \qquad (11)$$

where E is the modulus of elasticity and G the shear modulus of the material. The WHP post-yield stiffness K_{fp} is assumed to be 2 % the initial stiffness, i.e., $K_{\text{fp}} = 0.02K_{\text{fe}}$.

Design of Prototype SC PT Connection with WHPs

Prototype Building

Figure 6a shows the plan view of the prototype building used for this study. The study focuses on the interior steel MRF shown in Fig. 6b. The yield stress of structural steel is equal to 300 MPa. The design seismic action, referred to herein as DBE, has a return period equal to 475 years and is expressed by the elastic design spectrum of Eurocode 8 (2004) with peak ground acceleration equal to 0.3 g and soil B. The interior MRF is first designed as a conventional MRF according to EC3 (2003) and EC8. The behavior (or strength reduction) factor q is equal to 6.5. Eurocode 8 imposes a serviceability limit on the peak story drift, θ_{max}, under the frequent earthquake (FOE) with a return period equal to 95 years. The FOE has intensity equal to 40 % the intensity of the DBE, and the limit on θ_{max} is 0.5 % and

0.75 % for non-ductile and ductile nonstructural components, respectively. A strength-based design under the DBE was first performed. Beams and columns from this design had to be increased to satisfy the limit on θ_{max} under the FOE. The final sections are shown in Fig. 6b and were found iteratively, i.e., by decreasing the value of q, designing the MRF under the DBE, and then checking drifts under the FOE. The conventional MRF has story drifts equal to 0.64 % under the FOE, 1.6 % under the DBE, and 2.4 % under the maximum considered earthquake (MCE). The MCE has intensity equal to 150 % the intensity of the DBE.

The sections of the conventional MRF are used for designing the proposed PT connection (see next section). In that way, a SC-MRF using the proposed PT connections would have the same initial stiffness and period of vibration with the conventional MRF.

Performance-Based Design of PT Connection with WHPs

The design focuses on the exterior connection of the 2nd floor of the prototype steel MRF shown in Fig. 6b. This connection is designed as a PT connection with WHPs using the methodology presented by Garlock et al. (2007) that adopts two performance objectives and associated structural limit states, namely, (1) immediate occupancy under the DBE by avoiding damage in beams and columns while permitting gap opening and (2) collapse prevention under the MCE by avoiding PT bar yielding and beam local buckling while permitting minor yielding in beams and columns.

The design methodology proposed by Garlock et al. (2007) uses beam flange-reinforcing plates to avoid early beam flange yielding and buckling. However, the connection can experience sudden loss of strength and stiffness due to beam web local buckling when deformed beyond the MCE drift. In this work, the use of the structural details (drilled holes on the beam flanges and beam web stiffeners) proposed by Kim and Christopoulos (2009) in a connection designed according to the procedure described by Garlock is also examined with the aim to further delay beam local buckling

for drifts higher than the MCE drifts. Such a design approach is slightly different than the one proposed by Kim and Christopoulos (2008b) where the connection is designed to form a ductile plastic hinge for drifts approximately equal to the DBE drifts.

The design of the PT connection is based on drifts equal to 0.64 % under the FOE, 1.6 % under the DBE, and 2.4 % under the MCE according to the prototype building design. These drifts are significantly lower than those used in previous works which are based on the IBC 2 % drift limit under the DBE (Rojas et al. 2004; Garlock et al. 2007). To initiate the design procedure, the moment corresponding to point 2 in Fig. 5a, referred to herein as M_{IGO}, is set equal to 66 % of the nominal plastic moment of resistance of the beam section. In addition, M_d is set equal to 69 % of M_{IGO}. These M_{IGO} and M_d values provide an energy dissipation factor, β (Garlock et al. 2007), equal to 0.31. The posttensioning force and the length of the reinforcing plates are then designed. PT bars are designed to avoid yielding for $\theta \leq 0.07$ rad.

Hysteretic Characterization of WHPs

Component tests on the WHPs were conducted to assess their energy dissipation capacity and ductility under cyclic loading. Figure 7 shows the experimental setup for the tests. A pair of WHPs was tested to reproduce their behavior in the proposed PT connection. Two supporting plates were welded on a fixed plate. A plate simulating the web of the beam was bolted to a thick plate attached to the actuator. The thicknesses of the plates were equal to those of the plates designed for the large-scale PT connection tests discussed later.

Four coupon tests were conducted to identify the material properties of the WHPs. The results of the coupon tests are summarized in Table 1. The WHPs were made of 1020-Grade carbon steel, with an average yield strength 557 MPa, ultimate strength 598 MPa, and elongation at fracture equal to 20 %. This type of steel was chosen to achieve the required strength while keeping the sizes of the WHPs relatively small.
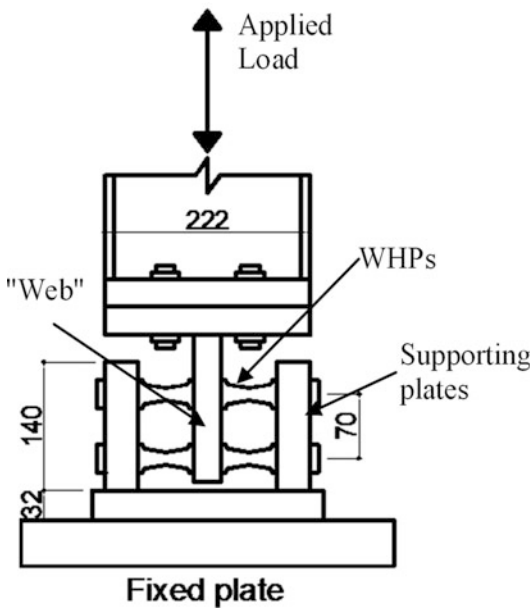
S

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 7 Component test setup for WHPs (dimensions in mm)

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Table 1 Material properties of WHPs from coupon tests

Coupon no.	Yield stress (MPa)	Tensile stress (MPa)	E (GPa)	% Elongation at fracture
1	551	595	207	17
2	550	594	209	20
3	564	599	210	23
4	563	604	208	20
Average	*557*	*598*	*208*	*20*

changes sign due to slight ovalization of the holes of the supporting plates under the bearing forces induced by the WHPs.

Large-Scale Experimental Program

Test Setup

Experiments on PT connections equipped with WHPs were conducted in the test setup shown in Fig. 9. The test specimens were based on the prototype second-floor exterior beam-to-column connection (see Fig. 6) at 0.6 scale. A strong *310UC158* column was used to minimize the column deformations since they do not influence the behavior of the main connection components (beam, PT bars, and WHPs). Two additional steel members were welded to the column to form a truss system which increases the horizontal stiffness of the test setup (*310UC158* horizontal member and *200UC52* diagonal member in Fig. 9). The whole system was bolted on the strong floor. The displacement history was applied vertically by a hydraulic actuator positioned at a distance of 1,800 mm from the inner face of the column.

Test Specimens

Two connection specimens (SC-WHP1 and SC-WHP2) were designed according to the procedure described earlier. The resulting beam and column sizes and details of the connection specimens are summarized in Table 2. The specimens are identical except for the beam web stiffeners and beam flange drilled holes used in specimen SC-WHP2.

The hourglass shape of the WHPs was designed to be consistent with their bending moment diagram shown in Fig. 2. Such a design results in almost uniform distribution of inelastic deformations along the length of the WHPs and, hence, in enhanced energy dissipation and fracture capacity. For the tested configuration, the design resulted in WHPs with external diameter $D_e = 20$ mm, internal diameter $D_i = 14$ mm, and $L_{whp,i} = 40$ mm, while the external part had a diameter equal to 30 mm.

Figure 8a shows the deformed WHPs during the component test. WHPs were capable of sustaining repeated large inelastic cycles without fracture up to displacements associated with connection rotation of 0.07 rad. Figure 8b plots the force–displacement hysteresis of the WHPs. WHPs provide a yielding force equal to 150 kN which is in excellent agreement with the 147 kN WHP yielding force calculated by the plastic analysis procedure presented previously. The ultimate strength of the WHPs at the initiation of fracture is 160 kN. Slight pinching (short flat region) is observed at the points where the force

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures,
Fig. 8 WHPs test results: (**a**) deformed WHPs and (**b**) WHPs hysteresis

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures,
Fig. 9 PT connection test setup (dimensions in mm)
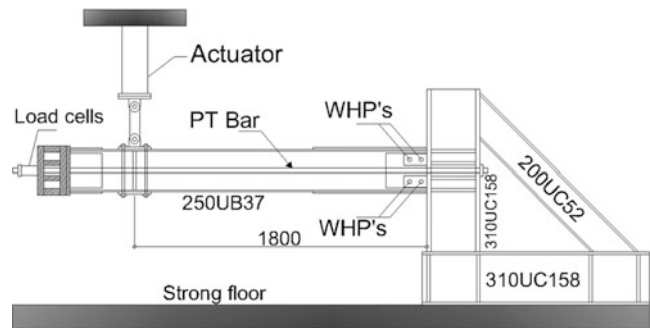
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Table 2 Specimen details

	SC-WHP1	SC-WHP2
Beam section	250UB37	250UB37
Column section	310UC158	310UC158
$F_{PT,i}$ (kN)[a]	518	504
L_{rp} (mm)[b]	700	700
Web stiffeners length (mm)	n/a	540
Drilled holes diameter (mm)	n/a	27

[a]$F_{PT,i}$: Initial posttensioning force
[b]L_{rp}: Length of reinforcing plates on the beam flanges

Figure 10a shows the SC-WHP1 specimen, while Fig. 10b shows a close-up view of its connection region. Beam flanges are reinforced with 700 mm long and 10 mm thick plates, while the web of the beam is reinforced with 10 mm thick plates as shown in Fig. 10b. The thickness of the WHP-supporting plates is 36 mm, their width is 80 mm, and a weld radius of 6 mm was used for the welding of all the plates in the connection.

Figure 11 shows the SC-WHP2 specimen which is identical to the SC-WHP1 apart from four 540 mm long longitudinal stiffeners which are welded on the web and four 27 mm holes drilled on the flanges immediately after the end of the reinforcing plate. The longitudinal stiffeners have thickness 10 mm, start 270 mm before and are extended 270 mm beyond the end of the reinforcing plates. The longitudinal stiffeners aim to resist the web buckling in the beam which typically occurs after the end of the reinforcing plates due to the increased bending moment and compressive force developed in this section as the gap opening increases. The drilled holes aim to reduce the moment

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 10 SC-WHP1 specimen: (**a**) overview and (**b**) close-up view of the connection region

capacity of the beam in this region in order to allow for the formation of a plastic hinge under large drifts. A dog-bone detailing could be used instead.

Material Properties

The beam and column sections were made of steel with nominal yield strength equal to 300 MPa. The reinforcing plates, WHPs supporting plates, and stiffeners were made of steel with nominal yield strength equal to 350 MPa. Tensile tests were carried out on coupons cut from the flanges and webs of the beams. Table 3 summarizes the mean actual yield stress, tensile stress, and modulus of elasticity resulting from the material tests. The coupon tests showed that the mean yield stresses are 361 and 388 MPa for the flanges and the web of the beam, respectively, indicating considerable overstrength of the steel material. The material of the PT bars has nominal yield strength equal to 930 MPa, tensile stress 1,050 MPa, and elongation capacity 6 %, according to the specifications of the supplier.
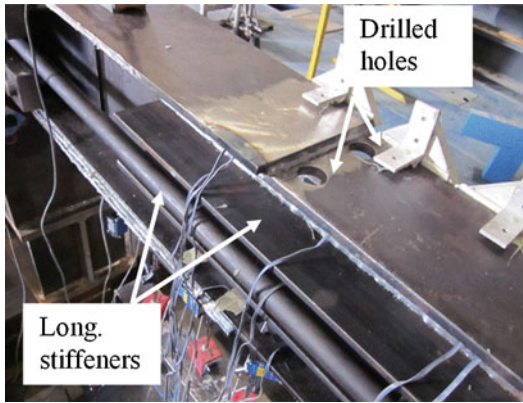
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 11 Details of SC-WHP2 specimen

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Table 3 Material test results

Coupon	Sample no.	Yield stress (MPa)	Tensile stress (MPa)	Modulus of elasticity (GPa)
Flange	1	354	535	208
	2	370	536	207
	3	359	530	208
	Average	*361*	*534*	*208*
Web	1	385	547	205
	2	392	548	199
	3	387	545	176
	Average	*388*	*547*	*193*

Test Procedure

The loading protocol for the two specimens consists of cyclic vertical displacements of increasing amplitude imposed in a quasi-static fashion with a speed approximately equal to 21 mm/min. The AISC (2005) loading protocol was used. This protocol consists of three initial sets of six cycles at 6.75, 9, and 13.5 mm displacements, four subsequent cycles at 18 mm, and six sets of two cycles at 27, 36, and 54 mm. These displacements correspond to drifts equal to 0.00375, 0.005, 0.0075, 0.01, 0.015, 0.02, and 0.03. The specimens were further imposed to drifts equal to 0.06 and, then, up to 0.09–0.10 to identify the connection failure mode.

Experimental Results

Specimen SC-WHP1

Figure 12a shows the force–drift hysteresis of the SC-WHP1 specimen, while Fig. 12b shows the associated moment–rotation hysteresis. In Fig. 12a, the drift is calculated as the ratio of the vertical displacement at the tip of the actuator to the horizontal distance of the tip of the actuator to the inner face of the column. The rotation is calculated as the gap-opening angle at the beam–column interface measured during the tests. In Fig. 12b, the moment is calculated as the product of the actuator force and the distance

of the actuator tip to the inner face of the column. Figure 12a shows that the connection maintains self-centering capability for drifts up to 6 %. The initial stiffness of the connection is 5 kN/mm and is equal to the theoretical stiffness of a welded (fully restrained) connection. During the last cycle of 9 % drift, the connection lost its self-centering capability and experienced a residual drift equal to 3 %.

To assess the reparability of the proposed connection after a strong earthquake, the SC-WHP1 specimen was initially imposed to displacements up to the DBE level (i.e., 1.6 % drift). The test was then interrupted, and damaged WHPs were substituted by new ones. The substitution was easily accomplished according to the following steps: (a) a temporary support was placed below the beam close to the beam–column interface to accommodate the self-weight of the beam; (b) the PT force was relaxed; (c) the damaged WHPs were taken out with the aid of a wooden hammer; and (d) new WHPs were placed in the connection. Since no welding or bolting is needed for the WHPs replacement, the proposed connection can be repaired with minimal disturbance to building use or occupation in the aftermath of a major earthquake. After the substitution of the WHPs, the test was restarted. Figure 12a plots superimposed loops from the tests before and after the replacement of WHPs. During the first cycles, the connection loses part of its initial
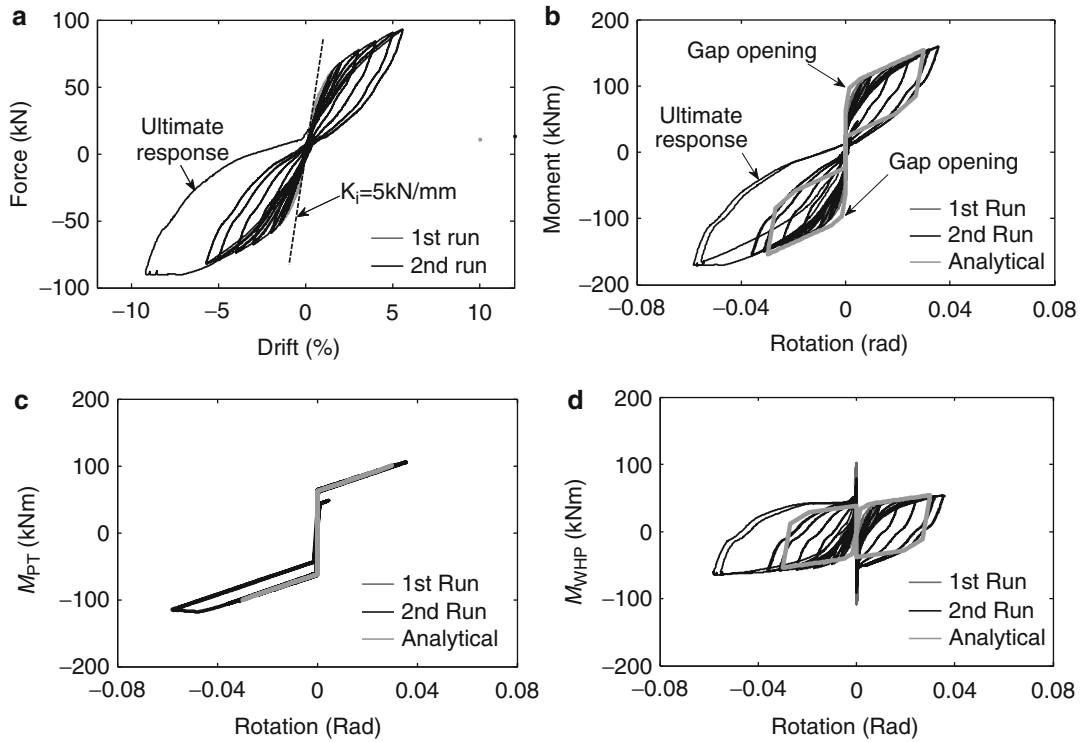
S

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures,
Fig. 12 Hysteretic response of SC-WHP1 specimen and comparison with analytical predictions

stiffness as a result of the unavoidable local yielding of the beam edges at the beam–column interface. Nevertheless, Fig. 12a shows that the connection gains its stiffness after drifts of 1 %. In fact, the two loops are perfectly superimposed to each other, demonstrating that the proposed connection can be repaired after a strong earthquake without compromising its stiffness, strength, and energy dissipation capacity.

Figure 12b plots superimposed moment–rotation loops from the tests before and after replacement of WHPs along with the analytical moment–rotation prediction using the analytical expressions of the simplified procedure described earlier. The moment contribution of the PT bars, M_{PT}, is calculated from the posttension force recorded by the load cells and is plotted in Fig. 12c, while the contribution of the WHPs, M_{WHP}, is found by subtracting M_{PT} from the total moment of the connection and is plotted in Fig. 12d.

Figure 12b shows that the connection achieves rotations larger than 0.035 rad with full self-centering capability. The experimental moments at decompression and gap opening are $M_d = 62$ kNm and $M_{IGO} = 95$ kNm, respectively. These values are slightly smaller than the design values of $M_d = 66$ kNm and $M_{IGO} = 100$ kNm. However, the predicted analytical curve captures well the cyclic behavior in terms of both stiffness and strength. The analytical prediction is also in good agreement with both the M_{PT}-θ and the M_{WHP}-θ loops, as shown in Fig.12c, d, respectively.

Displacement measurements at the beam–column interface indicated a vertical translation of the beam axis with respect to its initial position of about 4 mm. This translation took place gradually during the testing procedure and is attributed to a gradual loss of friction between the beam and the column interface. Shear tabs could be welded on the column flange and bolted on the beam web using slotted holes to

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 13 Gap opening at 6 % imposed drift of the SC-WHP1 specimen
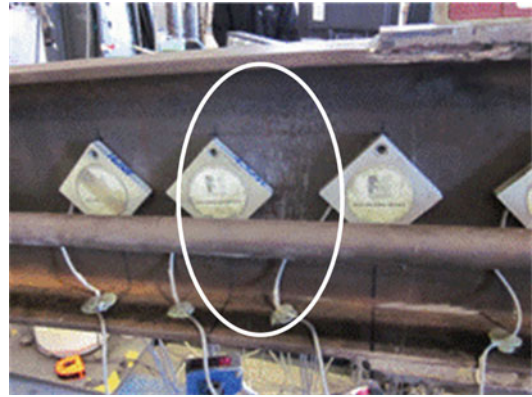
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 14 Web buckling of the SC-WHP1 specimen

accommodate gravity loads and eliminate the vertical beam displacement without influencing the connection moment behavior. Another alternative to accommodate gravity loads could be the use of a seat angle positioned below the bottom flange of the beam.

Figure 13 shows a close-up view of the gap-opening angle corresponding to 6 % drift. Apart from the intended damage in the WHPs and local yielding in the beam edges at the beam–column interface, no other evidence of damage was observed for drifts lower or equal to 6 %. Yielding in the beam flanges at the beam–column bearing surface was evident from the first cycles of the test. In fact, it was visually verified that the extreme edges of the beam at the beam–column interface became rounded during the test. As a result, the location of the COR was not constant during the experimental process. At the initial cycles, the COR was located approximately at the center of the reinforcing plate, while at larger drifts, it was stabilized at the inner side of the reinforcing plates.

During the last cycle of 9 % drift, the specimen failed due to local buckling of the web of the beam immediately after the end of the reinforcing plates. Figure 14 shows a close-up view of the buckled region. Web buckling took place at 7 % drift, while local buckling of the beam flanges initiated at the same time. Strain gauge measurements did not indicate

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 15 PT force versus drift for the SC-WHP1 specimen

yielding at any point throughout the beam section when local buckling initiated.

Figure 15 plots the variation of the force in the PT bars with respect to drift. The PT force increases with gap opening and returns to its initial value of 504 kN when gap closes. The loss in posttensioning during the test was about 7 kN. The maximum force in the PT bars has values of 850 kN and 970 kN at 6 % and 9 % drift, respectively. These values show that the PT bars remained elastic during the test since their nominal yield limit is 1,227 kN. A drop in the PT force due to beam local buckling is observed during the last cycle.
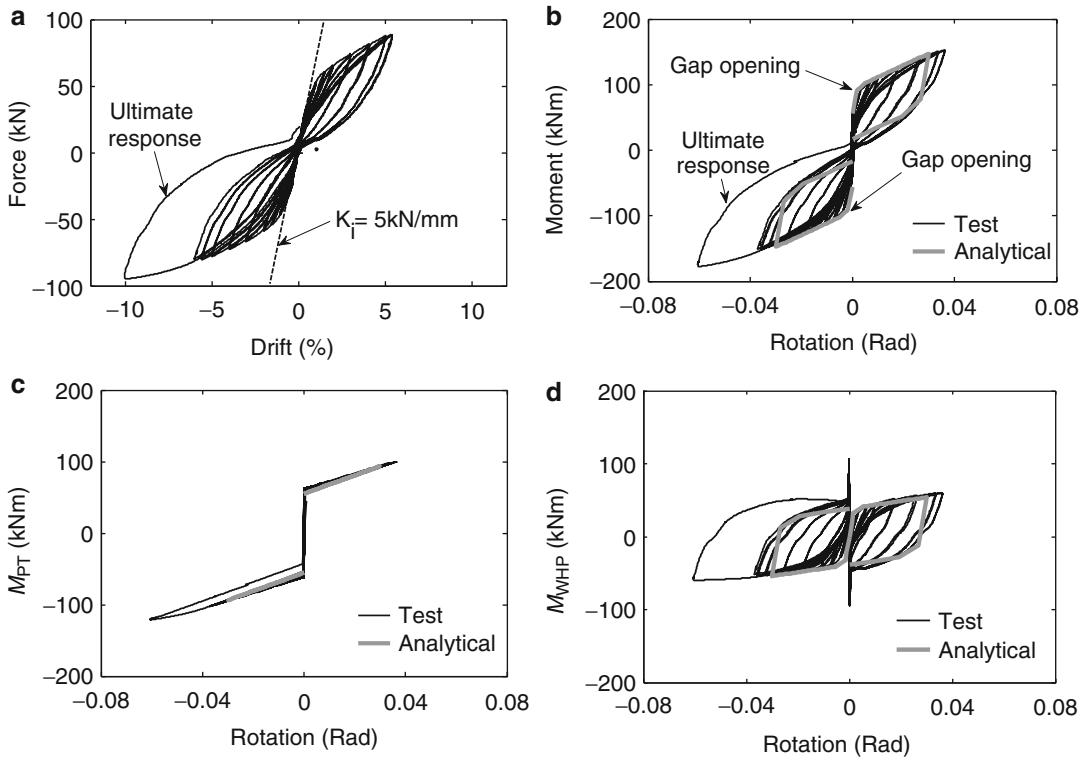
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 16 Hysteretic response of specimen SC-WHP2 and comparison with analytical predictions

Specimen SC-WHP2

Figure 16a shows the cyclic force–drift behavior of the SC-WHP2 specimen. The behavior is stable with full self-centering capacity up to 6 % drift similar to specimen SC-WHP1. A final experimental loop of 10 % drift caused the connection to fail, and a residual drift of 4 % was present at the end of the test.

The moment versus relative rotation hysteresis is plotted in Fig. 16b. The experimental M_d value was 60 kNm, while M_{IGO} was 92 kNm. These values are somewhat smaller than the corresponding values of the SC-WHP1 specimen due to the slightly smaller initial posttensioning force (504 instead of 518 kN). The simplified design procedure predicts well the cyclic envelope with the predicted values for M_d and M_{IGO} being 64 and 98 kNm, respectively. Figure 16c–d shows that the analytical predictions are in good

agreement with the individual experimental moment contributions M_{PT} and M_{WHP}.

The detailing employed for specimen SC-WHP2 resulted in a different failure mode. No web or flange buckling was observed in the beam section after the end of the beam flange-reinforcing plates. In addition, no plastic hinge was developed in this region. The strain gauges indicated that the yield strain was not reached in any point throughout the beam section at the end of the beam flange-reinforcing plate. Figure 17 shows that the SC-WHP2 specimen failed due to excessive yielding in the beam flanges at the beam–column interface under large bearing forces. Local yielding at the bearing interface of specimen SC-WHP2 was more pronounced than in specimen SC-WHP1 and resulted in a gradual reduction in posttensioning force. Figure 18 shows that the loss of force in the PT bars was
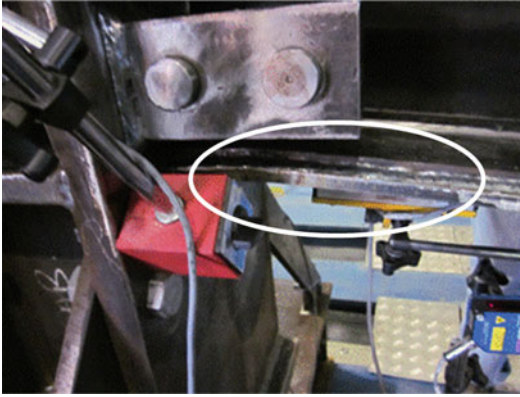
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 17 Local yielding in the beam flange at the bearing interface of the SC-WHP2 specimen

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 18 PT force versus drift for the SC-WHP2 specimen

50 kN during the test, with the largest amount attributed to loading cycles associated with drifts larger than the MCE drift. A reduction of 15 kN in the force of the PT bars was observed as a result of the local failure at the final loading cycle.

Finite Element Simulations and Parametric Study

This section presents FEM models which can be used to reliably assess the design and behavior of the connection. A detailed nonlinear FEM model was developed in Abaqus (Dassault Systems 2010). The FEM model was calibrated against experimental results and found capable to trace the nonlinear cyclic behavior of the connection and capture all possible local failure modes. The calibrated FEM model was used to conduct a series of simulations to study the effect of different parameters on the connection behavior.

Models for the WHPs

The cyclic behavior of the WHPs was first simulated. The FEM model geometry reproduced the actual geometry of the test setup of the component tests that were carried out to characterize the behavior of the WHPs. Figure 19 shows the setup used for the WHPs characterization tests along

with the FEM model and its boundary conditions. Since the actual geometry was symmetric, only one fourth of the specimens was modeled to decrease computational time.

Figure 20a shows the deformed shape of the WHP and the equivalent plastic strain (*PEEQ*) distribution at a displacement of 12 mm. Figure 20a shows that the hourglass shape geometry results in uniform distribution of plastic deformation along the length of the internal part of the WHP, while the external parts of the WHP and the supporting plates are essentially elastic. The *PEEQ* distribution in the WHP's supporting plates is plotted separately in Fig. 20b which shows minor plastic concentrations. These results agree with the experimental test observations where a negligible ovalization of the holes of the supporting plates was evidenced.

Figure 21 compares the force–displacement hysteresis from FEM analysis with the experimental hysteresis from the WHPs characterization tests. The FEM model can trace the nonlinear cyclic behavior of the WHPs with good accuracy. The model captures the pinching behavior at zero force which is observed in the experimental hysteretic curve as the result of the negligible ovalization of the holes of the supporting plates. Figure 21 shows that the experimental hysteresis of the WHPs deteriorates after a displacement amplitude of 12 mm. The WHP failed due to
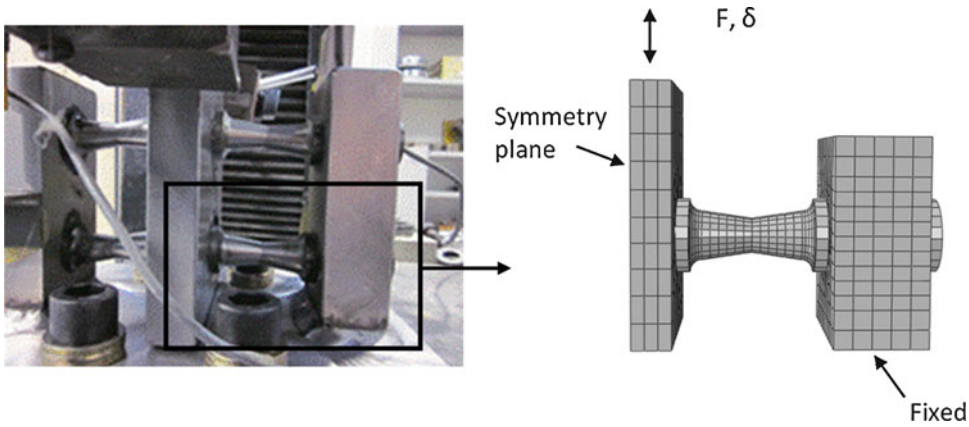
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 19 The FEM model used for the WHPs
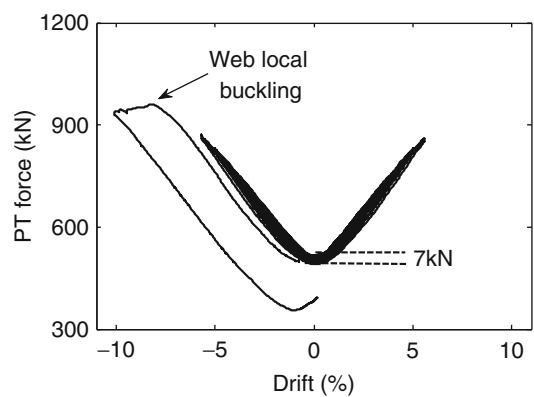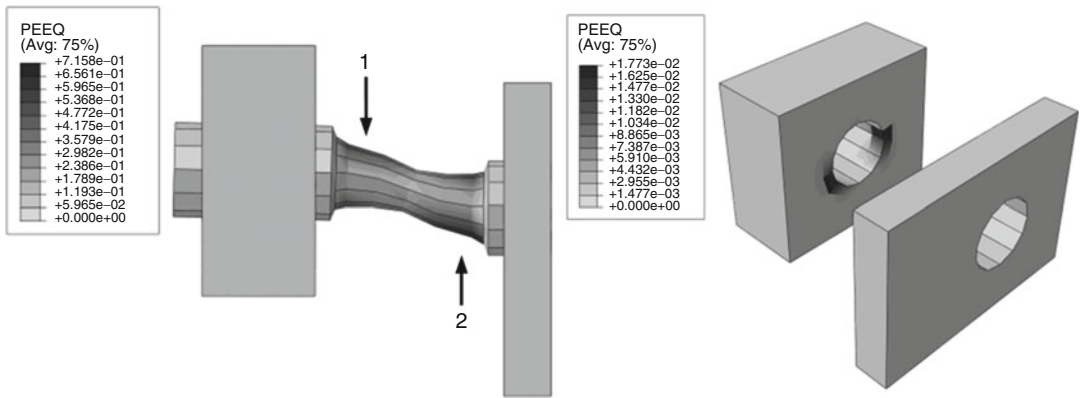

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 20 Deformed shape of the WHP and equivalent plastic strain (*PEEQ*) contour plot

ductile fracture at the section close to the internal support plate under a displacement amplitude of 12 mm.

Models for the Connection

A three-dimensional FEM model was developed to simulate the behavior of the connections as shown in Fig. 22. The geometry of the tests was reproduced in full detail. The column, the PT bars, the WHPs, and the plates were modeled using C3D8R solid elements. The beam was modeled using solid elements with incompatible modes (C3D8I). C3D8I are first-order elements that are enhanced by incompatible modes to

improve their bending behavior. The mesh was refined in regions where severe plastic deformations or buckling phenomena were expected to occur, i.e., close to the beam–column interface and at the end of the beam flange-reinforcing plates. A coarser mesh was used for regions that were expected to remain essentially elastic, i.e., the anchor block, the tip of the beam, and the top and bottom parts of the column. The WHPs and their supporting plates were modeled using the verified model for the WHPs described in the previous section.

The vertical load was applied as an imposed displacement, U_2, at a distance equal to 1,800 mm

from the connection face as shown in Fig. 22a. The analysis consisted of several steps. In the first step, the contact interactions were established to ensure that numerical problems due to contact formulation will not be encountered during the next steps. The posttensioning force was applied during the second analysis step by imposing an axial displacement at the free ends of the PT bars

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 21 Comparison of the experimental and the numerical force–displacement hysteresis of the WHPs

capable of producing the desired $F_{PT,i}$ in the beam. In the subsequent steps, the cyclically displacement history was applied. Displacement-controlled nonlinear analysis was performed.

Assessment of the Finite Element Model

Figure 23 plots the force–drift hysteresis from FEM analyses along with the experimental hysteresis of the two connection specimens. The FEM model can capture well the overall cyclic behavior of the connections. The predicted values for the M_d and M_{IGO} are in very good agreement with the corresponding experimental ones, while the initial and post-elastic stiffness are almost identical. The variation of F_{PT} from experiments is compared with the FEM prediction in Fig. 24. The FEM simulations and experiments are in reasonably good agreement. The numerical model predicts a slightly smaller loss in F_{PT} than that measured one in the experiments for specimen SC-WHP1, while the loss in F_{PT} is more accurately predicted for specimen SC-WHP2.

Following the loading protocol of the experimental program, the FEM models were pushed to large drifts to investigate all possible global and local failure modes. Figure 23 shows that the two

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 22 FEM model details: (**a**) FEM discretization, (**b**) model for specimen SC-WHP1, and (**c**) model for specimen SC-WHP2

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 23 Force–drift hysteresis of the two connection specimens tested in Kim and Christopoulos (2008b) and comparison with FEM results: (**a**) SC-WHP1 and (**b**) SC-WHP2

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 24 F_{PT}-drift plots for the two connection specimens tested in Kim and Christopoulos (2008b) and comparison with FEM results: (**a**) SC-WHP1 and (**b**) SC-WHP2

large last cycles of 9 % drift for specimen SC-WHP1 and 10 % drift for specimen SC-WHP2 resulted in residual inelastic drifts and loss of self-centering capability. These large drifts resulted in different failure modes of the two connection specimens, consistent with the experimental observations. The FEM model for specimen SC-WHP1 failed due to web buckling immediately after the end of the reinforcing plates. The FEM model for specimen SC-WHP2 failed due to yielding in the region of the drilled holes on the beam flange and in the bearing interface, while the web of the beam remained elastic until the end of the analysis. This can be verified by the *PEEQ* value distribution in the deformed configurations of the connections (Fig. 25) for

three drift levels: DBE (i.e., Design Basis Earthquake (Charney and Downs 2004)) = 1.6 %, MCE = 2.4 %, and Ultimate = 9–10 %. In this figure, darker areas indicate larger plastic strain concentration. The FEM models captured the failure modes observed in the experimental program. Figure 25 shows that under the DBE drift, the connection damage is isolated in the WHPs. Under the MCE drift, damage is concentrated in the WHPs and in the beam flanges in the beam–column interface. Damage in the beam–column interface is more evident in the connection SC-WHP2. Under the ultimate drift, damage in the connection is spread along the beam web immediately after the end of the reinforcing plates and in the beam flanges at the

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 25 Contour plots of the equivalent plastic strain (*PEEQ*) in the connection at different drift levels: SC-WHP1 (*left*) and SC-WHP2 (*right*)
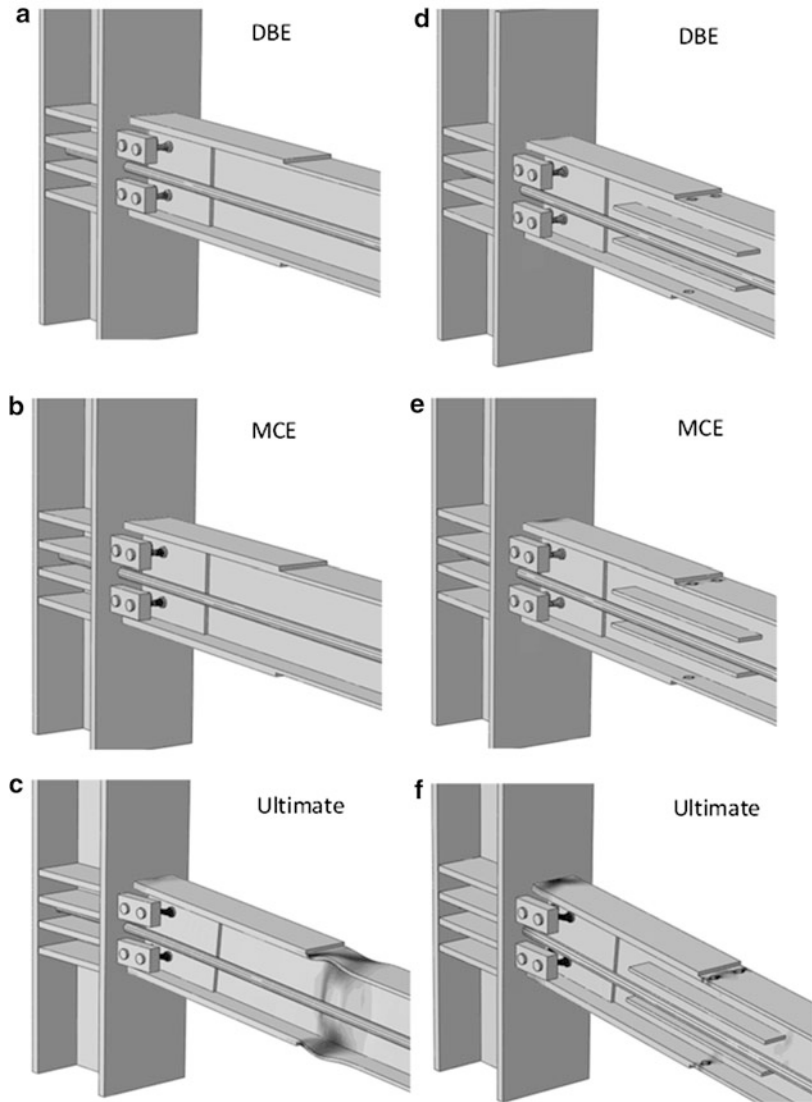
bearing interface. Figure 25f shows that damage in the beam flanges is more evident in the SC-WHP2 model.

The comparisons between the FEM analyses and experiments show that the proposed FEM model is capable of reproducing the inelastic response of the tested PT connections up to the ultimate deformation levels and to capture accurately all possible failure modes. Therefore, it is a reliable tool for the simulation of the hysteretic behavior of PT steel connections (Figs. 23 and 24).

Seismic Analysis of SC-MRFs Using PT Connections with WHPs

Modeling of PT Connection with WHPs for Nonlinear Analysis in OpenSees

A model for the PT connection with WHPs and the associated beams and columns was developed in OpenSees (Mazzoni et al. 2006) as shown in Fig. 26. The beams and columns were modeled using the nonlinear force-based beam–column fiber element which can strictly satisfy equilibrium and accurately capture the distribution of
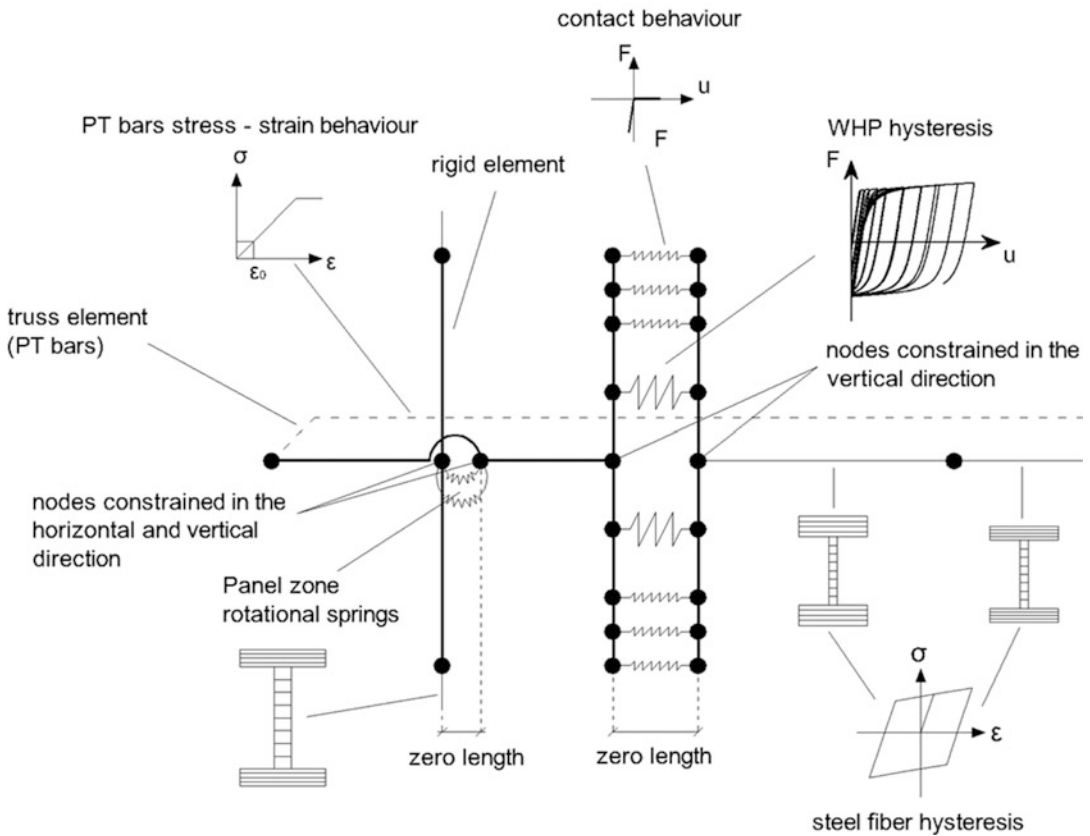
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures,
Fig. 26   Model for an exterior PT connection with WHPs and associated columns and beam

inelasticity along the depth of the section and along the length of the physical member. For the beam, two fiber elements with cross-sections having different flange thickness were used to account for the beam flange-reinforcing plates. Each fiber was associated with uniaxial bilinear elastoplastic stress–strain behavior (Steel01 in OpenSees) with a post-yield stiffness ratio equal to 0.01.

Rigid elastic beam–column elements were used to model the beam–column interface where gap opening and closing take place. To accurately capture the gap-opening mechanism in the beam–column interface, three zero-length contact spring elements were placed at equal spaces along the beam flange thickness. These contact springs were associated with an elastic compression – no tension force–displacement behavior (ENT material in Openness). A value

of the compression stiffness equal to 20 times the axial stiffness of the beam Kb was assigned to these contact springs. Larger values for this stiffness were found to produce practically the same results but with higher computational cost, i.e., more iterations to achieve convergence and equilibrium in nonlinear analysis. To capture the hysteretic energy dissipation capacity of the connection, two zero-length hysteretic springs were placed at the exact locations of WHPs along the depth of the beam web. These springs were associated with a smooth hysteretic Giuffre–Menegotto–Pinto model with isotropic hardening (Steel 02 material in OpenSees). To account for panel zone shear deformations and possible yielding, the panel zone was modeled using the Scissors model which introduces four additional rigid elastic beam–column elements and two nodes in the center of the panel zone
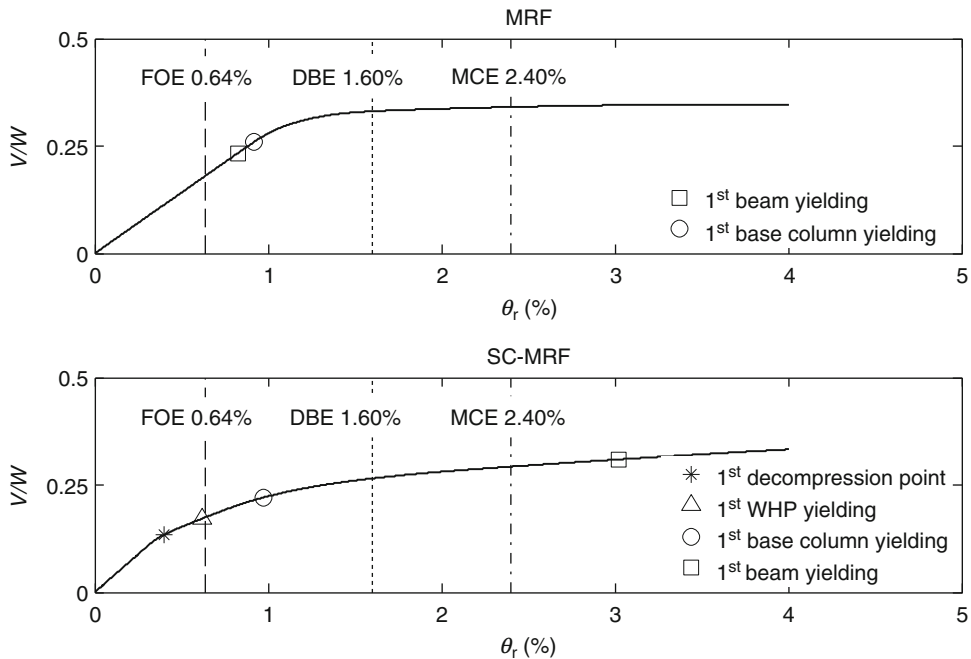
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures,
Fig. 27 Base shear coefficient – roof drift behavior from nonlinear monotonic (pushover) static analysis

connected with two zero-length rotational springs. These springs are associated with bilinear elastoplastic hysteretic rules (Steel01 material in OpenSees) with properties calculated to reflect the contribution of the column web (including doubler plates) and the column flanges in the force–shear deformation panel zone behavior. This simple panel zone model has been found to produce identical results to those of the more computationally expensive Krawinkler panel zone model (Charney and Downs 2004). PT bars were modeled using a truss element running parallel to beam centerline axis and connected to the exterior nodes of the panel zones of the exterior columns of the SC-MRF. The truss element has a cross-section area APT equal to that of both PT bars. To account for posttensioning, an initial strain equal to FPT,i/(APT · EPT) was first assigned to the truss element where EPT is the modulus of elasticity of the PT bar material. However, posttensioning results in shortening of the beams which in turn decreases the posttensioning force. To account for this decrease, the initial strain in the truss element

was increased to ensure that the posttensioning force in the PT bars will be equal to FPT,i after the beam shortening.

The accuracy of the developed model in OpenSees was assessed by comparing the force–deformation and moment–rotation responses of the connection with the experimental results. Detailed comparisons can be found in (Dimopoulos et al. 2013), where it is shown that the numerical model is in acceptable agreement with the experimental response and can be considered as reliable.

Nonlinear Monotonic and Cyclic Static Analysis

Figure 27 shows the base shear coefficient (V/W) – roof drift (θ_r) behavior of the conventional MRF and the SC-MRF from nonlinear monotonic (pushover) static analysis. V is the base shear and W is the seismic weight. An inverted triangular force distribution along with roof displacement control was used in these analyses. The MRF and the SC-MRF have comparable base shear strengths and comparable
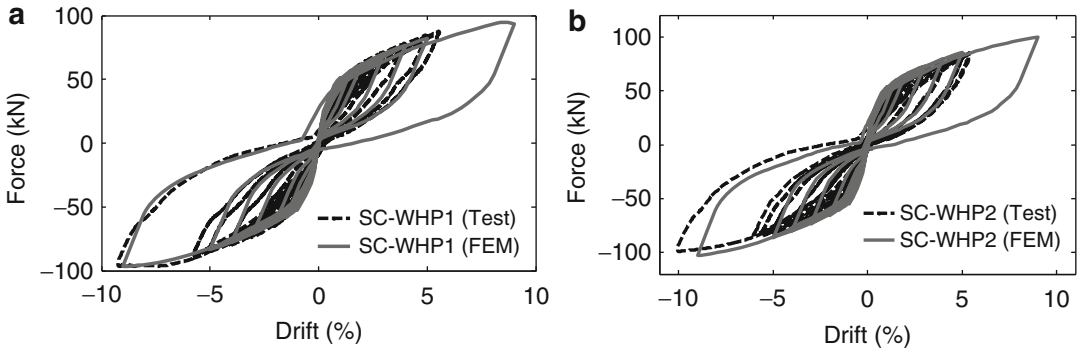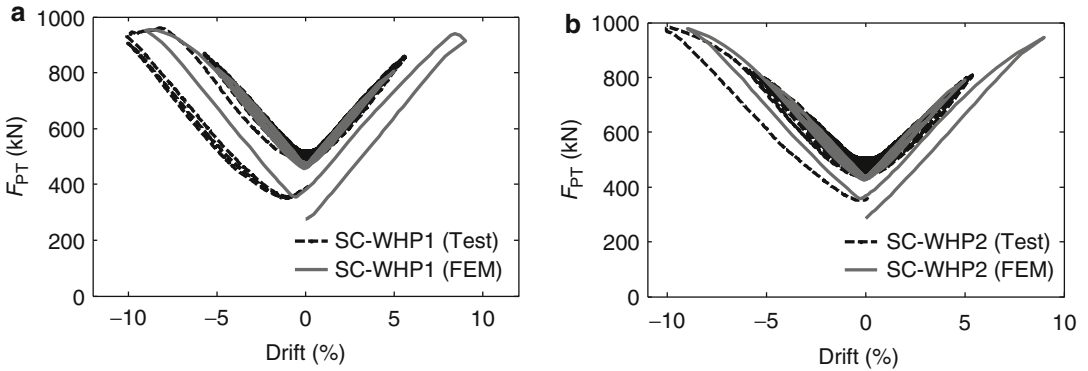
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 28 Base shear coefficient – roof drift behavior from nonlinear cyclic (push–pull) static analysis
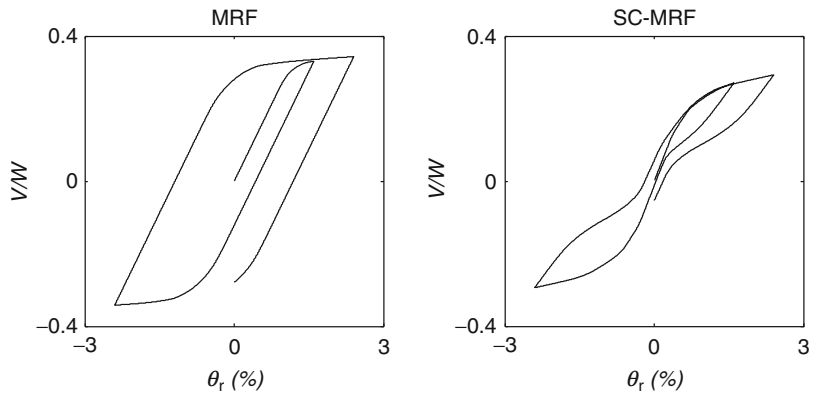
initial stiffness. The pushover curves are plotted along with points associated with structural limit states and vertical lines corresponding to roof drifts expected under the FOE, DBE, and MC-E. The structural limit states for the conventional MRF are beam yielding and base column yielding and occur at θ_r equal to 0.82 % and 0.92 %, respectively. The conventional MRF avoids damage under the FOE but experiences significant damage under the DBE. The structural limit states for the SC-MRF are decompression in a PT connection, WHP yielding, base column yielding, and beam yielding. Figure 27 shows that the beams of the SC-MRF are damage-free for θ_r equal or lower than 3 %, i.e., drifts higher than the MCE. Damage in the SC-MRF is experienced at the column bases that yield at θ_r equal to 0.97 %. No PT bar yielding is observed. The first decompression occurs at θ_r equal to 0.4 % while WHPs yield at θ_r equal to 0.62 % which is almost equal to the FOE drift. Decompression does not involve damage, while yielding of the WHPs is acceptable under low drifts since WHPs can be easily replaced without bolting or welding. The conventional MRF experiences softening at θ_r equal to 1.25 %, while the SC-MRF shows a more gradual softening behavior. In particular, the SC-MRF shows softening due to decompression in the PT connections at low drifts and further softening due to plastic deformations at the column bases and yielding of a large number of WHPs at θ_r equal to 1 %. Figure 28 shows the V/W – θ_r behavior of the MRF and the SC-MRF from nonlinear cyclic (push–pull) static analysis. The first cycle of the analysis is performed up to the DBE drift while the next cycle up to the MCE drift. The SC-MRF shows full re-centering capability under the DBE, adequate energy dissipation, and a small residual drift under the MCE due to plastic deformations at the column bases. The conventional MRF shows large energy dissipation capacity due to plastic deformations at the beam ends and at the column bases and the possibility of experiencing large residual drifts under the DBE and MCE.

Nonlinear Dynamic Time History Analyses

Ground Motions
A set of 20 earthquake ground motions recorded on ground type B were used in 2D nonlinear dynamic time history analyses to evaluate the performance of the SC-MRF and the performance of the conventional MRF. None of the ground motions exhibit near-fault forward directivity effects. The ground motions were scaled to the DBE level using the scaling procedure of Somerville (1997). Table 4 provides the scale factors and information on the 20 earthquake ground motions.

Modeling for Nonlinear Dynamic Analysis
Two-dimensional nonlinear analytical models of the conventional MRF and the SC-MRF were developed for nonlinear dynamic analyses in OpenSees. Nonlinear beam–column fiber elements were used for the beams and columns

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures,
Table 4 Properties of the ground motions used for nonlinear dynamic analyses

Earthquake	Station	Component	Magnitude (M_w)	Distance (km)	Scale factor		
					FOE	DBE	MCE
Imperial Valley 1979	Cerro Prieto	H-CPE237	6.53	15.19	0.82	2.05	3.08
Loma Prieta 1989	Hollister – S & P	HSP000	6.93	27.67	0.29	0.72	1.08
Loma Prieta 1989	Woodside	WDS000	6.93	33.87	1.40	3.49	5.24
Loma Prieta 1989	WAHO	WAH090	6.93	17.47	0.48	1.20	1.80
Manjil 1990	Abbar	ABBAR-T	7.37	12.56	0.28	0.70	1.05
Cape Mendocino 1992	Fortuna – Fortuna Blvd	FOR000	7.01	15.97	0.99	2.47	3.71
Cape Mendocino 1992	Rio Del Overpass – FF	RIO360	7.01	14.33	0.50	1.25	1.88
Landers 1992	Desert – Hot Springs	LD-DSP000	7.30	21.78	0.95	2.37	3.56
Northridge 1994	LA – W 15th St	W15090	6.69	25.60	1.14	2.86	4.29
Northridge 1994	Moorpark – Fire Sta	MRP180	6.69	16.92	0.78	1.94	2.91
Northridge 1994	N Hollywood – Cw	CWC270	6.69	7.89	0.53	1.33	2.00
Northridge 1994	Santa Susana Ground	5108-360	6.69	1.69	0.78	1.95	2.93
Northridge 1994	LA – Brentwood VA	0638-285	6.69	12.92	0.85	2.12	3.18
Northridge 1994	LA – Wadsworth VA	5082-235	6.69	14.55	0.62	1.54	2.31
Kobe 1995	Nishi-Akashi	NIS090	6.90	7.08	0.48	1.19	1.79
Kobe 1995	Abeno	ABN090	6.90	24.85	1.00	2.49	3.74
ChiChi 1999	TCU105	TCU105-E	7.62	17.18	0.96	2.39	3.59
ChiChi 1999	CHY029	CHY029-N	7.62	10.97	0.53	1.32	1.98
ChiChi 1999	CHY029	CHY041-N	7.62	19.83	0.56	1.40	2.10
Hector 1999	Hector	HEC090	7.13	10.35	0.42	1.04	1.56

and the Scissors model (Charney and Downs 2004) for the panel zones of the conventional MRF and the SC-MRF. A diaphragm constraint is imposed on each floor level of the conventional MRF. Stiff truss elements were used to connect the internal nodes of each bay of the SC-MRF to allow beam shortening due to posttensioning, gap opening in PT connections, and to consider the use of a discontinuous slab proposed by Chou et al. (2009). The model of Fig. 26 was used to model the PT connections and the associated beams and columns of the SC-MRF. Each dynamic analysis was extended well beyond the actual earthquake time (the ground motions were padded with zeros) to allow for damped free vibration decay and accurate residual drifts calculation.

Seismic Response Results

Figure 29 compares the roof drift time histories of the conventional MRF and the SC-MRF under the 5082-235 ground motion scaled to the DBE and MCE. Near the end of the time histories, the SC-MRF oscillates around the origin, indicating negligible residual drift, while the conventional MRF experiences residual drifts. The peak roof displacements of the MRF and the SC-MRF are similar. Figure 30 shows the stress–strain hysteresis at points A and B (extreme column base flange fibers), and C and D (extreme beam flange fibers) of the first story of the conventional MRF under the 5082-235 ground motion scaled to the DBE and MCE. Figure 31 presents similar information for the SC-MRF. The stress–strain hysteresis immediately after the end of the beam

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures,
Fig. 29 Comparison of the roof drift time histories under the 5082-235 ground motion scaled to the DBE and MCE
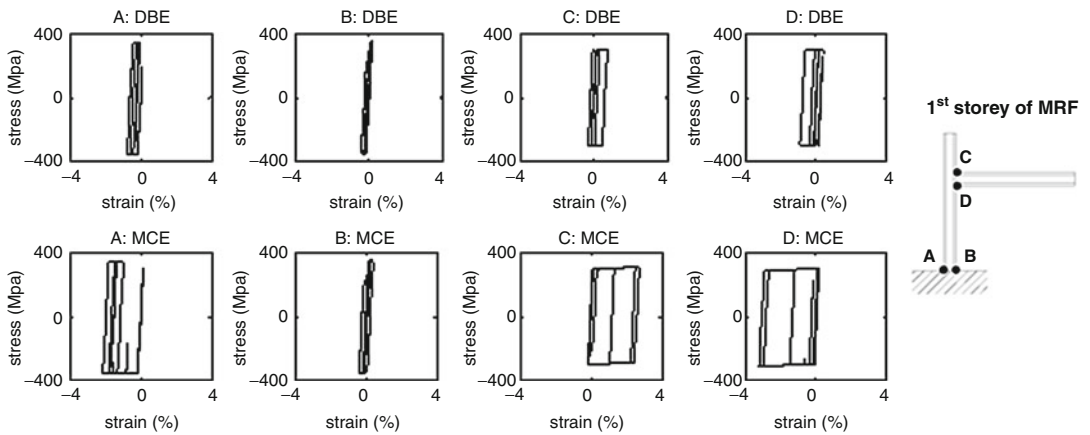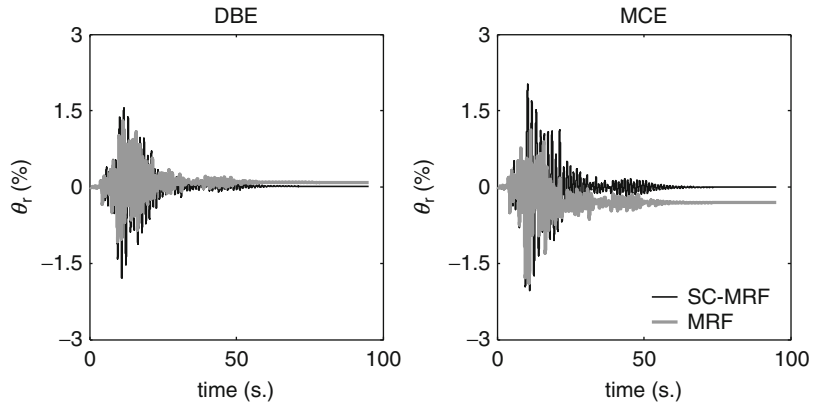
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 30 Stress–strain hysteresis at points A, B, C, and D of the conventional MRF under the 5082-235 ground motion scaled at the DBE and MCE
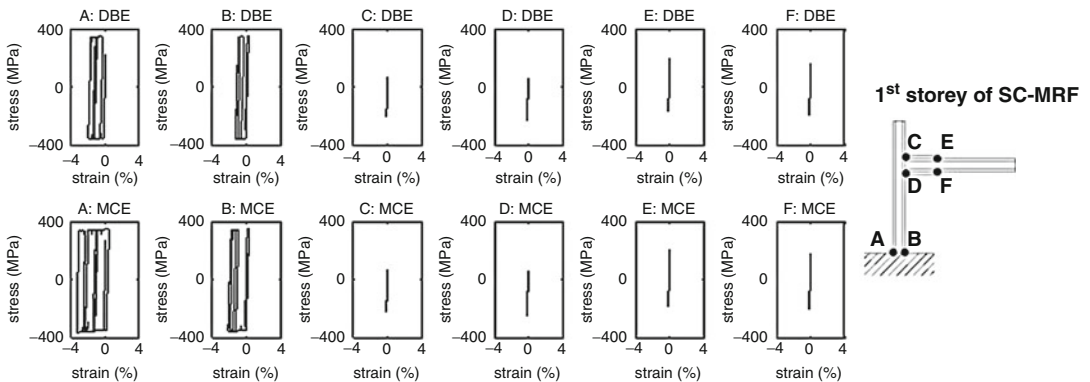
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 31 Stress–strain hysteresis at points A, B, C, D, E, and F of the SC-MRF under the 5082-235 ground motion scaled at the DBE and MCE
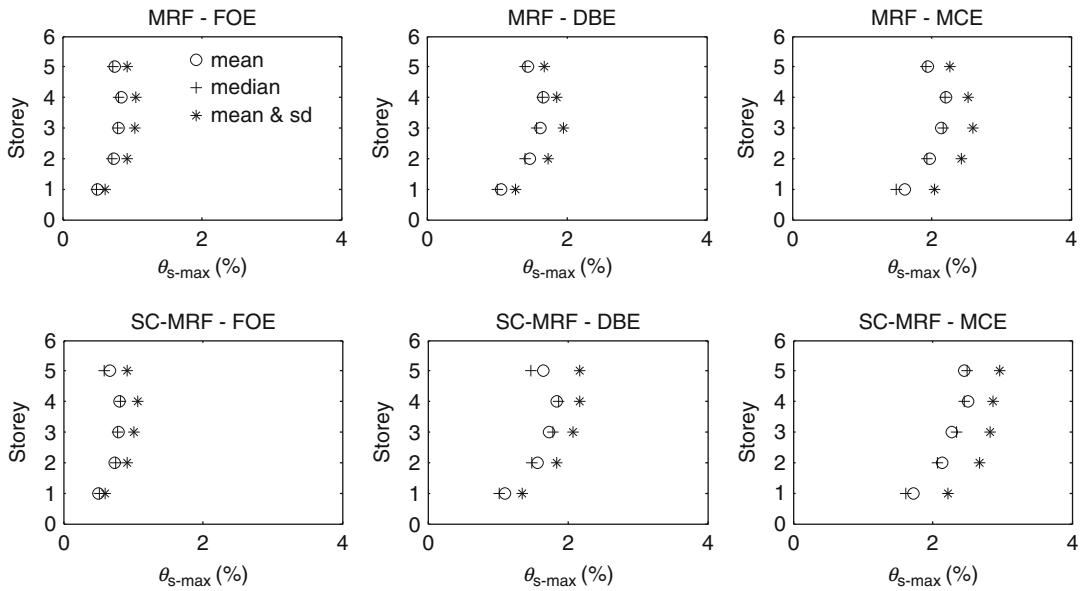
Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 32 Statistics of peak story drifts of the conventional MRF and the SC-MRF under 20 earthquake ground motions scaled to the FOE, DBE, and MCE

flange-reinforcing plate (points E and F) of the SC-MRF is also shown. The SC-MRF eliminates beam plastic deformations under both the DBE and MCE, while the conventional MRF experiences appreciable beam damage. Figures 28 and 29 show that the column bases of the SC-MRF experience larger plastic deformations than those of the column bases of the conventional MRF.

Figure 32 shows μ, $\mu + \sigma$, and median $\theta_{s\text{-max}}$ values under the earthquake ground motions of Table 4 scaled to the FOE, DBE, and MCE. The μ, $\mu + \sigma$ and median height-wise $\theta_{s\text{-max}}$ distributions show identical shapes. The MRF has the largest $\theta_{s\text{-max}}$ in the fourth story with μ values equal to 0.75 % under the FOE, 1.65 % under the DBE, and 2.2 % under the MCE, i.e., close to the design values of 0.64 % under the FOE and 1.6 % under the DBE and smaller than the design value of 2.4 % under the MCE. The SC-MRF has the largest $\theta_{s\text{-max}}$ in the fourth story with mean values equal to 0.75 % under the FOE, 1.8 % under the DBE, and 2.5 % under the MCE, i.e., slightly larger than the DBE and MCE design ones. Figure 33 shows μ, $\mu + \sigma$, and median values of the residual story drifts, $\theta_{s\text{-res}}$. $\theta_{s\text{-res}}$ values show a

uniform height-wise distribution for the conventional MRF and large dispersion compared to that of $\theta_{s\text{-max}}$. The largest $\theta_{s\text{-res}}$ of the conventional MRF occurs in the first story with mean values equal to 0.1 % under the DBE and 0.3 % under the MCE. The associated $\mu + \sigma$ $\theta_{s\text{-res}}$ values are equal to 0.25 % under the DBE and 0.6 % under the MCE. The latter $\theta_{s\text{-res}}$ values indicate that repair of damage in the conventional MRF would be costly and disruptive after the DBE and not financially viable after the MCE (Mc Cormick et al. 2008). These results highlight the need for Eurocode 8 to include residual deformations as an additional seismic performance parameter. The SC-MRF practically eliminates residual story drifts apart from the first story that has μ and $\mu + \sigma$ $\theta_{s\text{-res}}$ values equal to 0.1 % and 0.15 % under both the DBE and MCE. The latter $\theta_{s\text{-res}}$ values are lower than the global sway imperfections defined in EC3 (2003) and so it can be assumed that there will be no need for these residual drifts to be straightened out. Figure 33 shows small $\mu + \sigma$ $\theta_{s\text{-res}}$ values in the third story of the SC-MRF due to modest yielding in the beam ends.
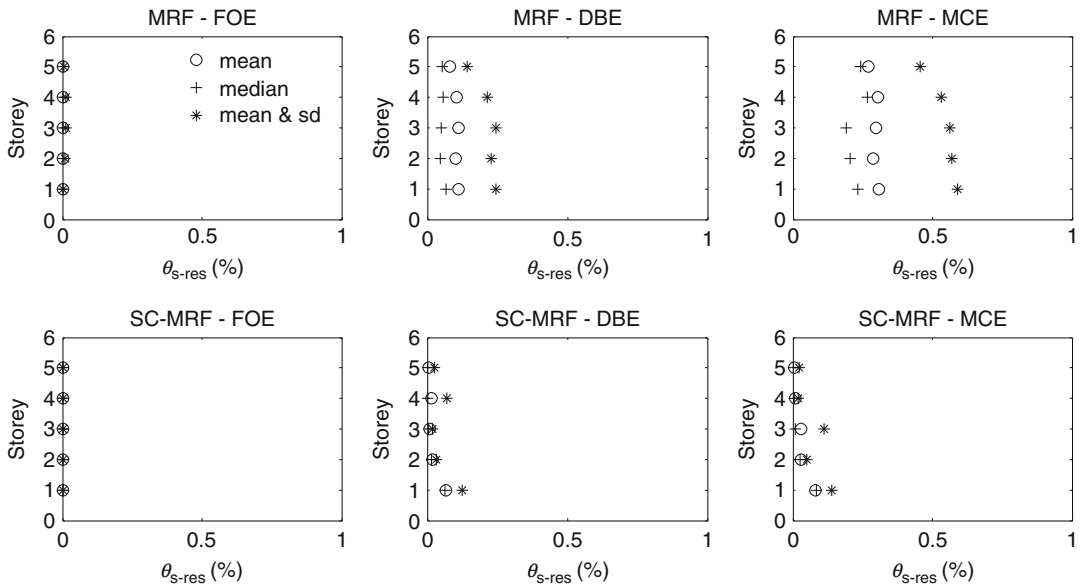
S

Steel Posttensioned Connections with Web Hourglass Pins: Toward Earthquake Resilient Steel Structures, Fig. 33 Statistics of residual story drifts of the conventional MRF and the SC-MRF under 20 earthquake ground motions scaled to the FOE, DBE, and MCE

Conclusions

A new steel posttensioned connection was developed for use in self-centering moment-resisting frames. The connection performance was validated through a large-scale experimental program and extensive numerical analyses, including seismic analysis of steel frames equipped with the proposed connection. The main conclusions of this research are:

(a) *Based on the experimental results:*
  • The proposed connection has stable self-centering behavior, enhanced energy dissipation capacity, and strength and stiffness comparable to those of a welded connection.
  • The proposed connection eliminates residual deformations and avoids beam damage for drifts lower or equal to 6 %.
  • A simplified analytical procedure using plastic analysis and simple mechanics was found to accurately predict the connection behavior.
  • Repeatable tests on a connection specimen were conducted along with replacing

damaged WHPs. These tests showed that WHPs can be easily replaced without welding or bolting, and hence, the proposed connection can be repaired with minimal disturbance to building use or occupation in the aftermath of a major earthquake.
  • Web local buckling at the connection region after the beam flange-reinforcing plates can be avoided by using web stiffeners. However, this detailing results in excessive local yielding at the beam–column interface due to high bearing forces.
  • To accommodate the gravity loads of the frame and to facilitate the replacement procedure of the damaged WHPs, it is necessary to include a vertical loading transfer system, e.g., a shear tab with slotted holes or a seat angle below the beam.
  • It should be emphasized that the tests described herein were performed on a 300 mm deep beam section. For a given rotation, the demand on the WHPs and PT bars is expected to increase for deeper sections. The response of the proposed

connection should be verified by using deeper beam sections.

(b) *Based on the finite element simulations:*
- The developed nonlinear FEM models can be reliably used to assess the design of the proposed connection as they are capable to trace the hysteretic behavior and predict the local failure modes both of the individual WHPs and the connection when subjected to either monotonic or cyclic loading.
- The FEM analyses confirm that for drifts equal or lower than those expected under the design basis earthquake (DBE), damage in a carefully designed connection is concentrated in the WHPs which are components that can be very easily replaced without welding or bolting.

(c) *Based on the seismic analyses of steel frames with the proposed connection:*
- The proposed model for the PT connection with WHPs and the associated beams and columns has been calibrated against experimental results and found to accurately simulate the hysteretic behavior of the PT connection.
- Nonlinear static monotonic (pushover) analysis shows that the conventional MRF and the SC-MRF have comparable base shear strength and initial stiffness. The conventional MRF experiences significant damage in beams at the DBE drift. On the other hand, the SC-MRF has damage-free beams for drifts even higher than the MCE drift.
- Nonlinear static cyclic (push–pull) analysis shows that the SC-MRF has full re-centering capability and adequate energy dissipation capacity under the DBE.
- Seismic analyses show that the conventional MRF and the SC-MRF have comparable peak story drifts. In particular, the conventional MRF has slightly lower peak story drifts than the SC-MRF. For both frames through the mean peak story drifts are close to the design values.
- Seismic analyses show that the SC-MRF practically eliminates residual story drifts

apart from the first story which sustains small residual drifts due to plastic deformations at the column bases. The mean plus one standard deviation value of the first story residual drift of the SC-MRF is equal to 0.15 % under both the DBE and MCE which is considered small and does not need to be straightened out.
- Seismic analyses show that the mean plus one standard deviation value of the maximum residual story drift of the conventional MRF is 0.25 % under the DBE and 0.6 % under the MCE. These values indicate that repair of damage in the conventional MRF would be costly and disruptive after the DBE and not financially viable after the MCE.
- Seismic analyses show that the beams of the SC-MRF do not exhibit any yielding even under the MCE, while significant inelastic deformations are developed in the beams of the MRF under both the DBE and MCE. On the other hand, the column bases of the SC-MRF experience larger inelastic deformations than those of the conventional MRF.

Summary

This contribution presents the experimental and analytical validation on a new self-centering steel posttensioned connection using web hourglass shape steel pins (WHPs) as energy dissipation devices. The connection isolates inelastic deformations in WHPs, avoids damage in other connection parts as well as in beams and columns, and eliminates residual drifts. WHPs do not interfere with the composite slab and can be very easily replaced without bolting or welding, and so, the connection enables nondisruptive repair and rapid return to building occupancy in the aftermath of a strong earthquake. The experimental results are first presented followed by parametric studies using nonlinear finite element models. A simplified nonlinear model for the connection and the associated beams and columns that consist of nonlinear beam–column

elements and hysteretic and contact zero-length spring elements appropriately placed in the beam–column interface was also developed in OpenSees. The model was calibrated against experimental results and found to accurately simulate the connection behavior. A prototype building was selected and designed as a conventional steel moment-resisting frame (MRF) according to Eurocode 8 or as a self-centering steel MRF (SC-MRF) using the connection with WHPs. Seismic analyses results show that the conventional MRF and the SC-MRF have comparable peak story drifts and highlight the inherent potential of the SC-MRF to eliminate damage in beams and residual drifts. It is shown that repair of damage in the conventional MRF will be costly and disruptive after the design basis earthquake and not financially viable after the maximum considered earthquake due to large residual drifts.

Cross-References

▶ Seismic Analysis of Steel Buildings: Numerical Modeling
▶ Seismic Collapse Assessment
▶ Seismic Strengthening Strategies for Existing (Code-Deficient) Ordinary Structures

References

Charney FA, Downs WM (2004) Connections in steel structures V. ESSC/AISC workshop, Amsterdam, 3–4 June

Chou CC, Lai YJ (2009) Post-tensioned self-centering moment connections with beam bottom flange energy dissipators. J Constr Steel Res 65(10–11):1931–1941

Chou CC, Chen JH, Chen YC, Tsai KC (2006) Evaluating performance of post-tensioned steel connections with strands and reduced flange plates. Earthq Eng Struct Dyn 35(9):1167–1185

Chou C, Tsai K, Yang W (2009) Self-centering steel connections with steel bars and a discontinuous composite slab. Earthq Eng Struct Dyn 38:403–422. doi:10.1002/eqe.

Christopoulos C, Filiatrault A, Uang CM, Folz B (2002) Posttensioned energy dissipating connections for moment-resisting steel frames. J Struct Eng 128(9):1111–1120

Dassault Systems (2010) Abaqus theory manual. Dassault Systèmes Simulia, Providence

Dimopoulos AI, Karavasilis TL, Vasdravellis G, Uy B (2013) Seismic design, modelling and assessment of self-centering steel frames using post-tensioned connections with web hourglass shape pins. Bull Earthq Eng 11(5):1797–1816. doi:10.1007/s10518-013-9437-4

EC3 Eurocode 3 (2003) Design of steel structures

EC8 Eurocode 8 (2004) Design of structures for earthquake resistance

Garlock M, Ricles JM, Sause R (2005) Experimental studies of full-scale posttensioned steel connections. J Struct Eng 131(3):438–448

Garlock M, Sause R, Ricles JM (2007) Behavior and design of posttensioned steel frame systems. J Struct Eng 133(3):389–399

Karavasilis TL, Ricles JM, Sause R, Chen C (2011) Experimental evaluation of the seismic performance of steel MRFs with compressed elastomer dampers using large-scale real-time hybrid simulation. Eng Struct 33(6):1859–1869

Karavasilis TL, Kerawala S, Hale E (2012) Model for hysteretic behaviour of steel energy dissipation devices and evaluation of a minimal-damage seismic design approach for steel frames. J Constr Steel Res 70:358–367

Kim HJ, Christopoulos C (2008a) Friction damped posttensioned self-centering steel moment-resisting frames. J Struct Eng 134(11):1768–1779

Kim HJ, Christopoulos C (2008b) Seismic design procedure and seismic response of post-tensioned self-centering steel frames. Earthq Eng Struct Dyn 38(3):355–376

Kim H, Christopoulos C (2009) Numerical models and ductile ultimate deformation response of post-tensioned self-centering moment connections. Earthq Eng Struct Dyn 38:1–21. doi:10.1002/eqe.

Kobori T, Miura Y, Fukuzawa E, Yamada T, Arita T, Takenaka Y, Miyagawa N, Tanaka N, Fukumoto T (1992) Development and application of hysteresis steel dampers. In: Earthquake engineering, tenth world conference. Balkema, Rotterdam

Mazzoni S, McKenna F, Scott M, Fenves G (2006) Open system for earthquake engineering simulation (OpenSees). User Command Language Manual, Pacific Earthquake Engineering Research Center, University of California, Berkeley

Mc Cormick J, Aburano H, Ikenaga M, Nakashima M (2008) Permissible residual deformation levels for building structures considering both safety and human elements. 14th WCEE, Beijing

Ricles J, Sause R, Garlock M, Zhao C (2001) Posttensioned seismic-resistant connections for steel frames. J Struct Eng 127(2):113–121

Ricles J, Sause R, Peng SW, Lu LW (2002) Experimental evaluation of earthquake resistant posttensioned steel connections. J Struct Eng 128(7):850–859

Rojas P, Ricles JM, Sause R (2004) Seismic performance of post-tensioned steel moment resisting frames with friction devices. J Struct Eng 131(4):529–540

Somerville P (1997) Development of ground motion time histories for phase 2 of the FEMA/SAC steel project, Report No. SAC/DB-97/04, Sacramento

Tsai KC, Chou CC, Lin CL, Chen PC, Jhang SJ (2008) Seismic self-centering steel beam-to-column moment connections using bolted friction devices. Earthq Eng Struct Dyn 37:627–645

Vasdravellis G, Karavasilis TL, Uy B (2013a) Large-scale experimental validation of steel post-tensioned connections with web hourglass pins. J Struct Eng 139(6):1033–1042

Vasdravellis G, Karavasilis TL, Uy B (2013b) Finite element models and cyclic behavior of self-centering steel post-tensioned connections with web hourglass pins. Eng Struct 52:1–16

Vasdravellis G, Karavasilis TL, Uy B (2014) Design rules, experimental evaluation and fracture models for high-strength and stainless-steel hourglass shape energy dissipation devices. J Struct Eng (ASCE). doi:10.1061/(ASCE)ST.1943-541X.00010

Whittle J, Williams MS, Karavasilis TL, Blakeborough T (2012) A comparison of viscous damper placement methods for improving seismic building design. J Earthq Eng 16:540–560

Wolski M, Ricles JM, Sause R (2009) Experimental study of a self-centering beam-column connection with bottom flange friction device. J Struct Eng 135(5):479–488

Steel Structures

Roberto Leon
The Charles Edward Via, Jr. Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, VA, USA

Synonyms

Metal structures; Steel connections; Steel construction

Introduction

The use of metals as the principal construction material for modern buildings and bridges dates back to the beginning of the Industrial Revolution (Bodsworth 2001; Ashton 1968). Three metallic materials have been commonly used in the construction industry: cast and wrought iron, steel, and aluminum. Cast iron, which was first manufactured as early as fifth century BC, is typically considered to be too brittle and difficult to join but received widespread application in construction through the late eighteenth century. The use of cast iron is uncommon in today's construction industry, but some ductile cast irons continue to be used in specialized applications such as water pipes. The use of cast iron was rapidly replaced by wrought iron, which was easier to work with, leading to the construction of the first large metal structure, the Iron Bridge at Coalbrookdale by A. Darby in 1780 (www. greatbuildings.com/buildings/Iron_Bridge_at_ Coalbrookdale.html). Extensive use of iron continued through the first half of the nineteenth century, culminating in iconic structures such as the Crystal Palace in London (http://en. wikipedia.org/wiki/The_Crystal_Palace).

Wrought iron began to be replaced by steel in civil structures when the latter became commercially available in large quantities in the late nineteenth century AD through the development of the Siemens and Bessemer processes and the development of rolling technologies (Bodsworth 2001; Ashton 1968; Halmos 2000). Initially steel was rolled into flat plates and small sections (angles and channels or L- and C-sections) that were assembled into larger structural sections through riveting (Fig. 1), just as it had been done for wrought iron. Riveted, or built-up sections, were used through the first half of the twentieth century.

Steel rolling, which began to produce large sections at the turn of the twentieth century (Fig. 2), led to the development of optimal W- (or H-) sections for carrying flexure and axial loads. Advanced fabrication techniques, recycling, and improvements in energy consumption during production have resulted in extremely efficient and sustainable construction practices in the steel industry in the twenty-first century. Much of today's structural steel is manufactured in minimills that utilize scrap steel as their source materials; some estimates are that up to 95 % of steel is recycled in the USA.

The use of aluminum, which became common in the 1940s when the need to build light

Steel Structures,
Fig. 1 Riveted built-up
steel sections in a bridge
truss (thebridgehunter.
areavoices.com,
© RTLeon)
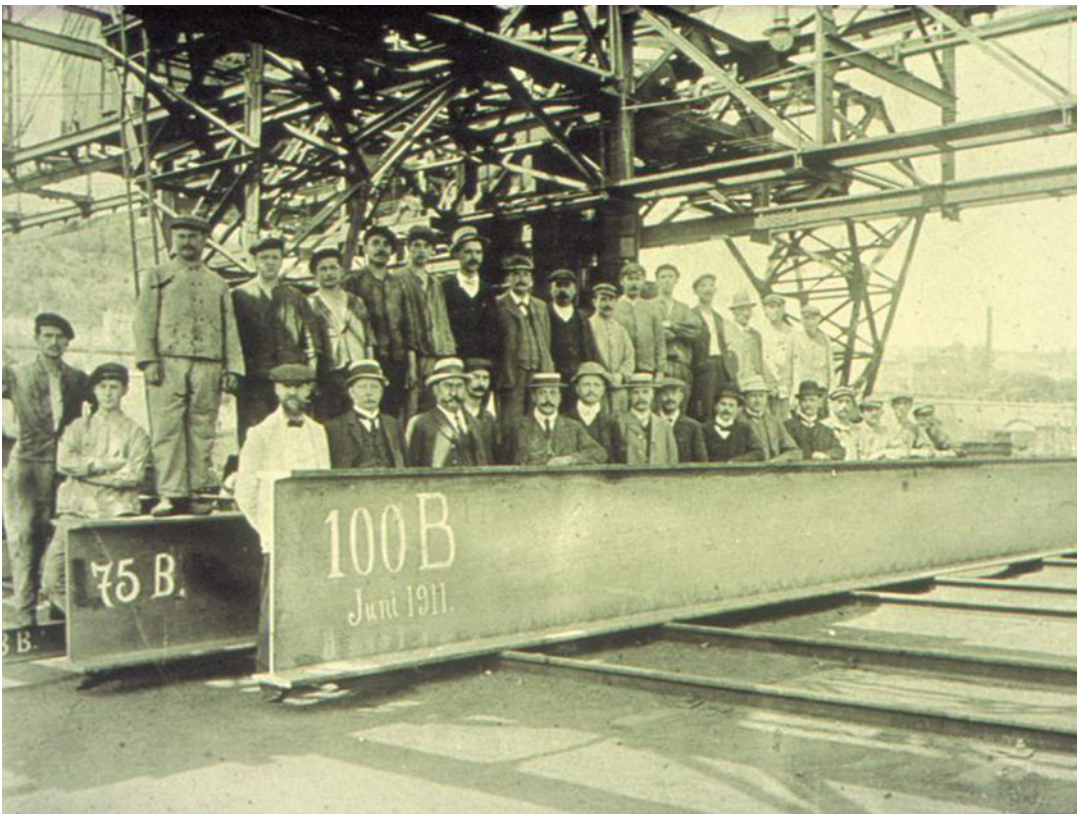
Steel Structures, Fig. 2 Rolling of deep sections (75 and 100 cm) in 1911. Note that flanges are not as wide as modern sections (© J.C. Gerardy, ArcelorMittal)

Fig. 3 Modern rolling of
large steel wide-flange
section showing selective
cooling (*darker* area) to
improve metal performance
(© J.C. Gerardy,
ArcelorMittal)
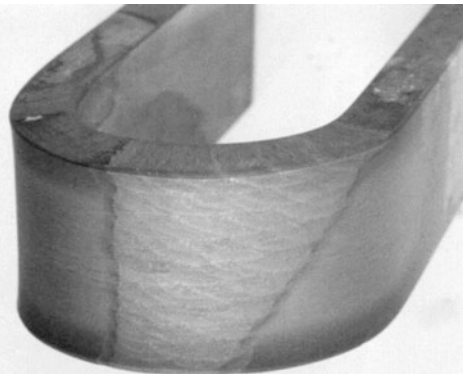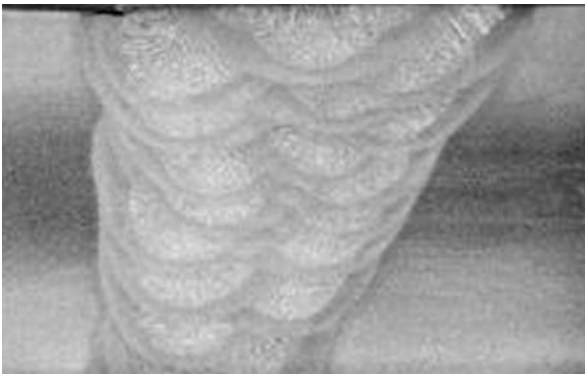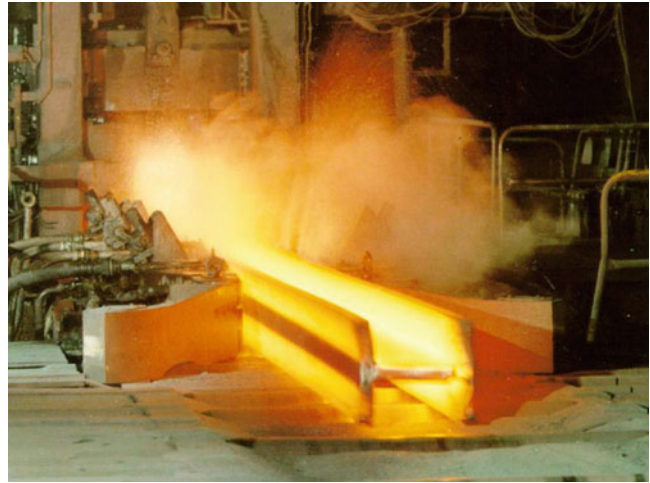
Steel Structures, Fig. 4 Carefully executed weld (*left*) showing numerous passes and good workmanship; bend test from this weld showing excellent deformation capacity as the weld does not fracture even if bent at 180° (© J.C. Gerardy, ArcelorMittal)

airplanes made its strength-to-weight ratio very attractive, is also limited in today's construction industry due to cost and welding issues. Its use is limited mostly to facades, finishes, and nonstructural elements.

Steel for Seismic Applications

A large number of improvements through better metallurgy (heat treatments (Tylecote 1976) and the development of alloy steels with large deformation capacity, Fig. 3) and joinery (welding and high-strength bolts, Fig. 4) in the twentieth century have made steel the preferred construction material to withstand the large cyclic loads imposed by earthquakes.

The use of steel in multistory framed buildings began in the 1880s in Chicago, notably with the Home Insurance Building (Condit 1968) and similar skyscraper buildings. The use of steel frames was still in its maturing stage when the excellent behavior of steel structures, when subjected to large ground motions, became evident after the San Francisco earthquake and fire of 1906 (Kurzman 2001). These events highlighted the lateral resistance (strength) and deformation capacity (ductility) of steel structures (Hamburger and Meyer 2006), but also their susceptibility to fire unless suitably protected (Fig. 5).

Steel Structures, Fig. 5 Damaged buildings from the 1906 San Francisco earthquake and fire. The two tall steel frames were under construction and were not damaged. The structure of the burned building in the center of the photograph was undamaged (en.wikipedia.org)

Most structural members were made up of riveted plates and connections for larger structures (Fig. 6) and riveted smaller rolled sections for buildings (Fig. 7). Structures that contained steel columns and beams encased in masonry or concrete survived the earthquake and fire relatively unscathed (Fig. 5). The superior performance of steel frames was enshrined in the report by ASCE (1906), which stated: *The well-designed steel frame offers the best solution of the question of an earthquake proof building, as all the stresses can be cared for.*

This and similar observations by many others led to the widespread utilization of such systems in the earthquake-prone areas of the Western USA in the first half of the twentieth century. These structures consisted of rolled steel sections joined by connections made by riveted steel

angles or plates. Initially, the encasement was made from masonry rubble inside masonry facades, but this was later replaced by encasement with lightly reinforced concrete (Fig. 7). Today, the synergistic use of steel in tension and concrete in compression has led to the development of composite steel–concrete construction particularly for high-rise construction (Viest et al. 1997).

Three important developments have occurred in the steel construction industry since the 1950s. The first development is the change of joining methods, as rivets have been replaced by both high-strength bolts and welds. The use of industrialized welding, which was developed in the early twentieth century for the shipping industry, led to much simpler, stronger, and stiffer connections (Fig. 8). These connections were considered

Steel Structures,
Fig. 6 Typical built-up
riveted members and
connection for an elevated
railway structure
(Engineering News, 1914)
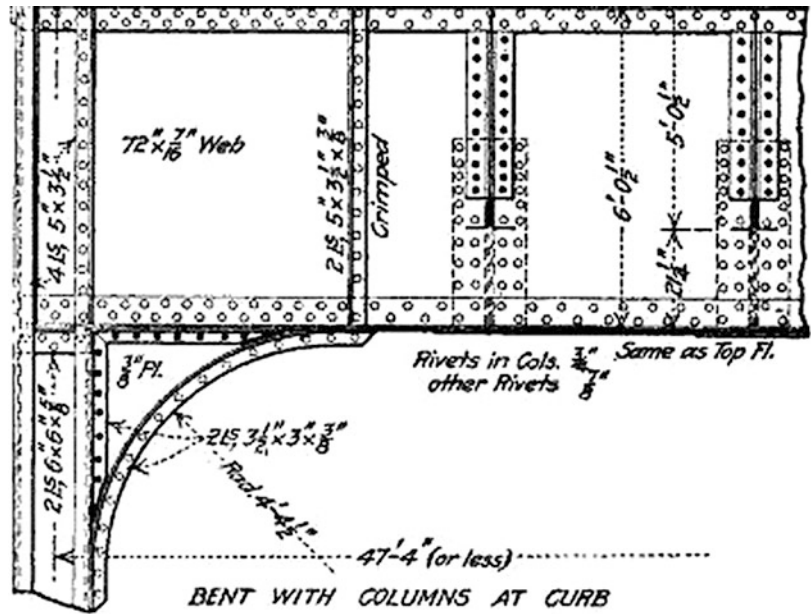
Steel Structures,
Fig. 7 Riveted building
beam-to-column
connection before
encasement (© RTLeon)
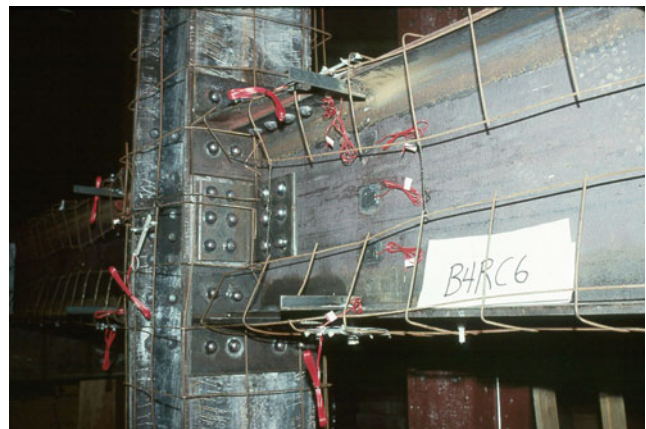
for many years to be the most reliable structural system in seismic areas. Failures in welded connections observed after the 1994 Northridge (Fig. 9) and 1995 Kobe earthquakes (FEMA 355E 2000a) have led to radical changes in the specifications for such joints. While both joining methods can be used in seismic construction, the use of welding requires extensive planning, inspection, and QA/QC procedures to ensure desirable ductile performance (FEMA 355B 2000b). In the aftermath of those failures, design requirements for bolted connections have also been increased even though few bolted connection failures have been observed after earthquakes, except for the case of under-designed brace connections. Welded connections are still preferred by many designers because of their perceived superior strength and stiffness characteristics and ease of design when compared to bolted connections.

The second development is that most floors in steel structures have become composite ones, with the use of shear studs to connect floor beams and girders to concrete slab cast in metal deck (Viest et al. 1997). The use of the metal deck as both formwork and reinforcement, the
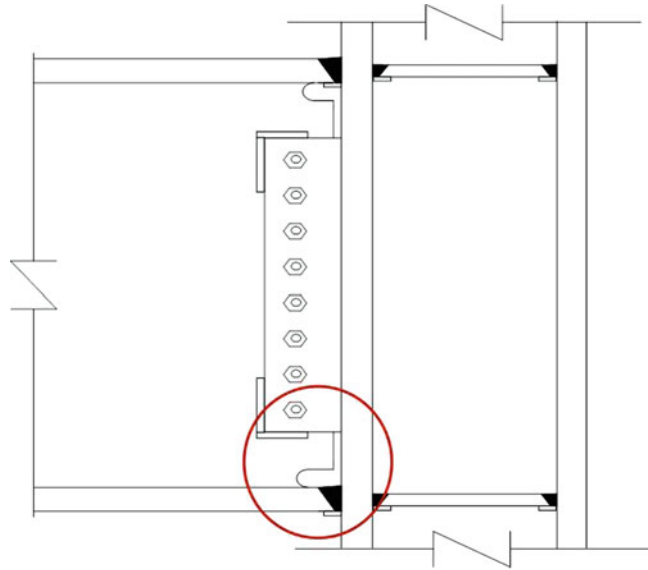
ease of welding the shear studs that act as connectors between the steel beam and the concrete floor, and the tremendous increases in strength and stiffness have made composite floor construction a most economical system (Fig. 10). Floors of this type provide sufficient in-plane stiffness, such that all parts of the lateral force-resisting system can be assumed to work together (rigid floor diaphragm assumption), so that all lateral load-resisting systems can act concurrently. This characteristic is important as many frames in the USA utilize lateral load systems engaging only partially the perimeter frames.

The third important change has been the development of numerous new structural systems and proprietary connections that have given designers a very wide choice of technologies for use in seismic design. The development of such systems has been made possible by both extensive experimental testing and advanced simulation tools brought forth by the use of computers and improvements in analytical methods. Some of these systems are discussed under *Structural Systems* later in this article. Striking evidence of the advantages of these new systems can be seen in the excellent performance of steel *eccentrically braced frames* (EBF) in the recent

Steel Structures,
Fig. 10 Typical composite
floor system (© RTLeon)
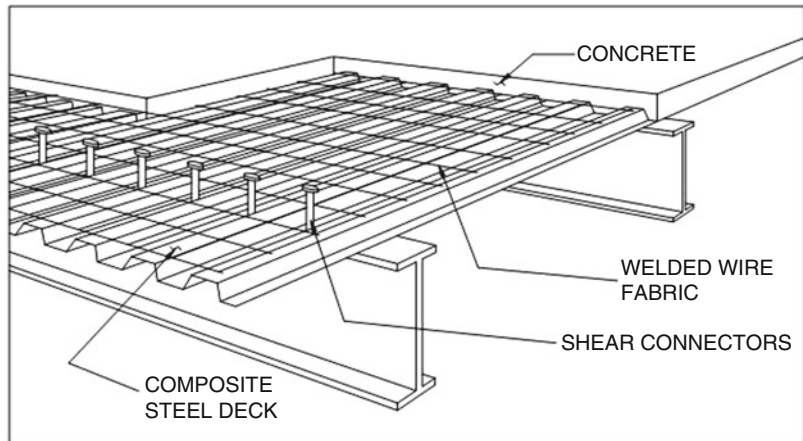
Steel Structures,
Fig. 11 Ductility of steel
as shown by the >25 %
elongation of the original
coupon (*bottom*) to the
fractured one (*top*)
(© RTLeon)
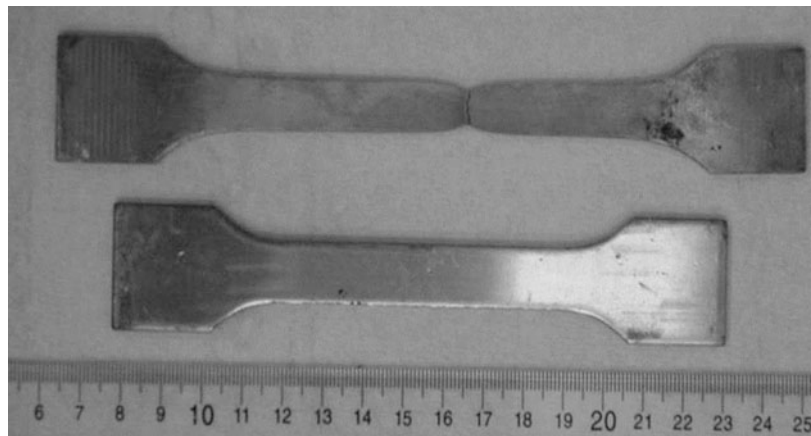
Christchurch earthquakes (Clifton et al. 2011) as compared to that of modern reinforced *special moment-resisting reinforced concrete frame* (SMRF-CR) and shear wall structures (Leon et al. 2014).

Steel Properties

For structural design purposes, steel is an iron-carbon alloy that contains many additional alloying elements such as Mn, Mo, Cr, Ni, and V. A large variety of mechanical properties can be achieved with steels by either varying its alloys (primarily carbon which constitutes 0.2–0.6 % of steel) and/or subjecting the material to different heat treatments that change its grain microstructure.

Steel is an ideal material for seismic design because both of its isotropic and homogeneous properties and its ability to undergo large local plastic deformations before failure. Figure 11 shows a typical mild steel test coupon elongating in excess of 15 % over a 2 in. length and more than 50 % locally. The stress–strain (or load-deformation) behavior for this type of steel is shown by the red line in Fig. 12. After an initial stiff elastic response, the steel reaches its yield point (knee in the curve), the steel deforms plastically (i.e., undergoing large deformations with little small increments of load, shown by the flat portion of the curve), and picks up additional resistance as the deformation becomes large (last part of the curve, termed the strain-hardening region). The curve for a steel coupon subjected to large deformation reversals such as those experienced

Steel Structures,
Fig. 12 Cyclic
stress–strain behavior of a
mild structural steel
(© Alan Pense, Lehigh
University)
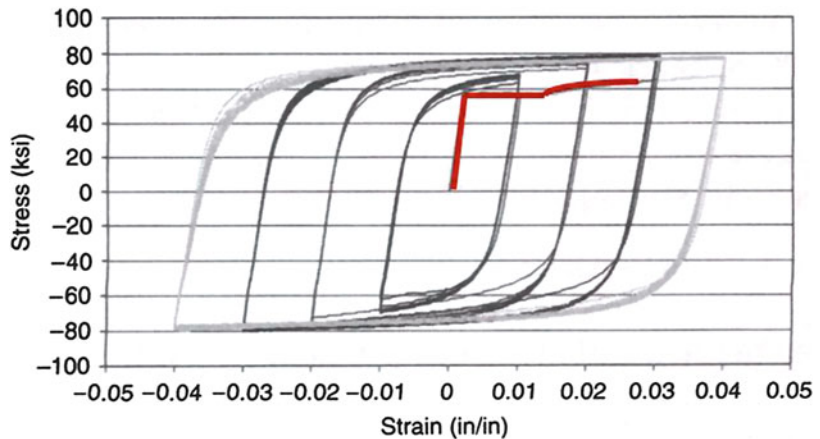
locally during a large seismic event is shown by the rest of the curves in Fig. 12. If loaded cyclically, the steel will show fat, stable, and hardening hysteresis (stress–strain) loops, evidence of good energy dissipation capacity. Failure by ductile fracture will ideally occur after many cycles of deformation. For seismic design, steel will perform best when deformed in shear.

Steels for structural use are classified as carbon steels, high-strength low-alloy steels, and alloy steels. For design purposes, these steels can be assumed to have a density of 7.85 g/cm^3, a modulus of elasticity of 210 GPa, and a Poisson's ratio of 0.3. Carbon steels are classified based on the percentage of carbon. Mild carbon steels (0.15–0.29 % C) with yield points in the range of 220–250 MPa and tensile strengths of 400–500 MPa are the most common structural carbon steels. Typically, an increase in carbon percent raises the yield point and increases hardness, but reduces ductility and makes welding more difficult. These drawbacks can be minimized by heat treatments.

The stress–strain behavior of a mild carbon steel is characterized by a sharp yield point and large ultimate strains (percent elongation >20 %). High-strength low-alloy steels exhibit higher yield points (345–480 MPa) and somewhat larger tensile strengths (450–600 MPa) without an appreciable drop in ductility. Alloy steels are heat-treated steels (mostly quenched and tempered) with yield points of 550–760 MPa, but lower tensile to yield ratios and ductility than for other structural steels.

Other important engineering properties needed for seismic design are notch toughness and weldability. Notch toughness refers to the ability of the material to resist fracture propagation from an existing defect when subjected to dynamic loads. Fracture propagation is resisted through local plastic deformations and is primarily a function of temperature and heat treatment. Weldability is primarily a matter of obtaining a structural joint free of any undesirable defects by utilizing appropriate electrode materials and weld procedure specifications (WPS).

Structural Systems

As noted earlier, one of the great developments in seismic construction has been the promotion of a large variety of steel structural systems. These systems are classified as *special, intermediate*, and *ordinary* based on the amount of seismic detailing present. A *special* system will require great care in design and detailing to obtain the large deformation capacity needed to activate the plastic deformations and energy dissipation characteristic of these systems. The benefit of a special system is lower design lateral forces, which results in savings in both material and construction costs.

At its most basic, lateral force-resisting systems can be visualized as a continuum, with a pure frame system (Fig. 13a) at one extreme and a wall system at the other end (Fig. 13e). In these
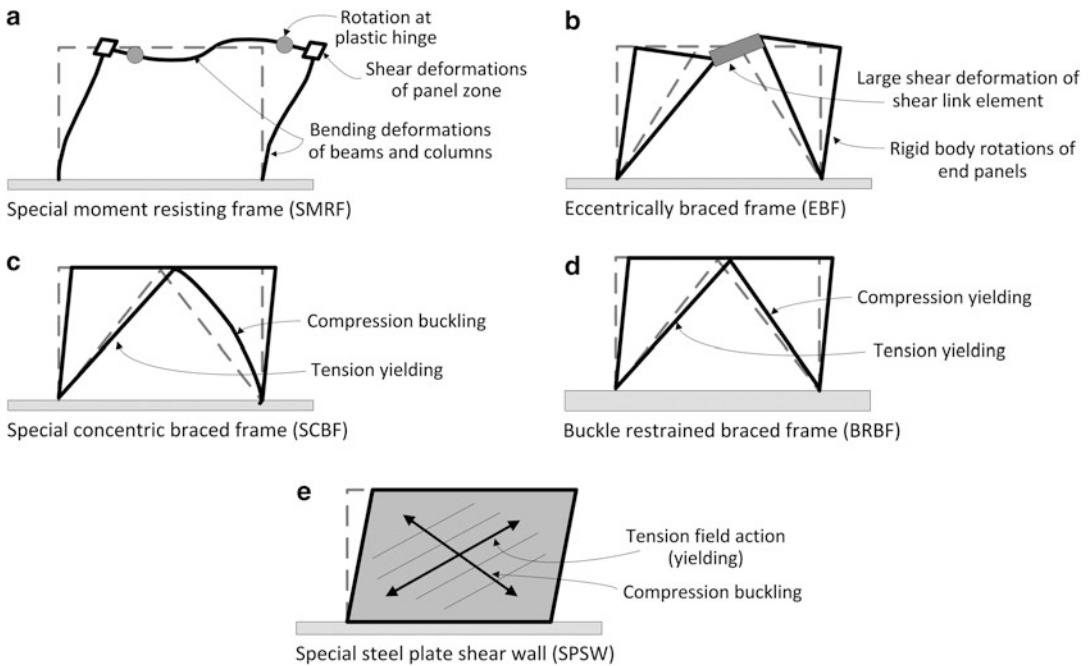
Steel Structures, Fig. 13 Deformation mechanisms for typical steel seismic force-resisting systems (© RTLeon)

figures, only the primary structural elements are shown; secondary structural elements, such as out-of-plane braces, floor slabs, chord, collectors, and diaphragms, are not shown. Also not shown are *nonstructural elements*, such as partition walls, which are assumed not to contribute significantly to lateral resistance and should be designed so as not to interact with the primary structural members. If not properly isolated, nonstructural elements can have deleterious effects on frame behavior.

The *moment-resisting frame* (MRF), a combination of slender beams and columns, was the first type of steel structural system used and which showed excellent seismic behavior (Fig. 5). In a pure moment-resisting frame (MRF, Fig. 13a), the lateral deformations are to be accommodated primarily by bending of the beams and columns, shear deformation of the panel zone, and the formation of *plastic hinges*, or areas of concentrated plasticity, in the beams. The yielding of the beam steel and the formation of a plastic hinge are shown by the flaking of the whitewash in the critical section of the beam in Fig. 14. All of these deformation mechanisms can

be very ductile when properly designed, but pure frame structures tend to be rather flexible. To eliminate or limit damage to nonstructural elements and contents under both small earthquakes and large wind loads, special moment-resisting frames need to meet some maximum deformation criteria (typically ≈3 % drift under earthquake loads and 0.25 % under wind loads) and thus are said to be *drift controlled*. It is clear that rigid and strong connections that limit the shear deformations of the panel zones and promote beam yielding are desirable (note thick end plate and large bolts in the connection in Fig. 14), as long as the local buckling of the beam flanges and web can be delayed. In addition, mechanisms that can reduce the strength demand at the welds at the joint without substantially decreasing stiffness, such as reduced beam sections (RBS, Fig. 15), are desirable. In an RBS, the steel in the beam flanges is cut to intentionally weaken the beam and promote the formation of a plastic hinge away from the welded connection.

At the other extreme, one can think of a wall structure composed of a steel frame infilled with a steel plate (Fig. 13e) as the analogue to a
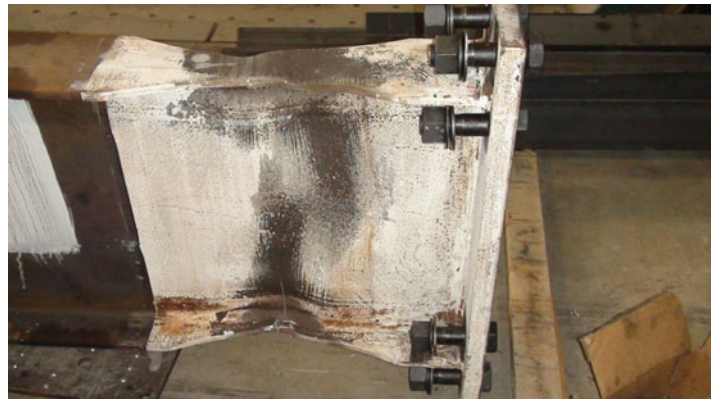
reinforced concrete shear wall system. For economy reasons, in steel structures it is desirable to use a rather thin plate. When displaced laterally, the plate will buckle in compression along one diagonal, allowing a tension field to develop in the other diagonal (Fig. 16). This creates a large tensile strut mechanism that transfers the horizontal forces similarly to a braced frame (Fig. 13c). This force transfer is exactly the inverse of a concrete shear wall structure, where a compression strut will form, as the concrete is strong in compression but weak in tension. The great advantage of a steel shear wall over a concrete one is that whereas the behavior of a concrete shear wall will degrade rapidly due to an X-crack pattern that develops with cycling, the steel shear wall performance will not degrade substantially due to the

post-buckling strength of metal plates. Steel-plated shear walls are relatively modern structures, with a few implementations as far back as the late 1970s, but have become more accepted only in the last decade or so.

In between these two limits, one can visualize a large number of alternatives. Starting at the bottom of Fig. 13, one can visualize a system where the struts in the walls are replaced by discrete, strong, and rigid braces (or diagonal elements, Fig. 17). This system is known as a *buckling-restrained brace frames* (BRBF, Fig. 13d), because the braces in compression are so stout that they will not buckle but yield. The buckling in these systems is limited to an interior core of the brace, with the exterior portion being disconnected and acting only to restrict the

Steel Structures, Fig. 16 Thin plate under lateral loads showing tension-field action from *top right* to *bottom left* and buckling in the opposite direction (© M. Kurata)
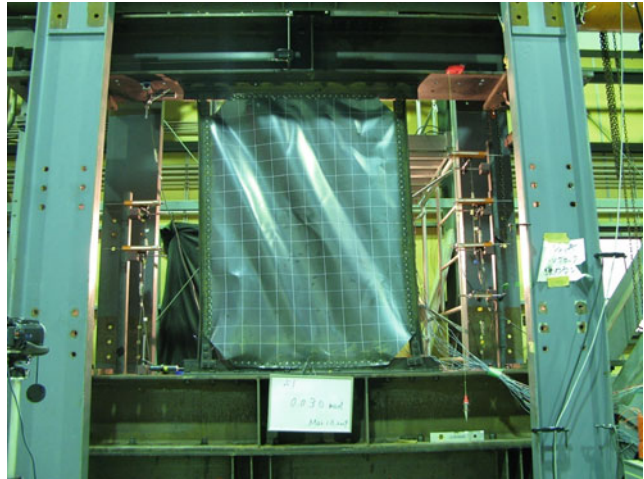
Steel Structures, Fig. 17 Use of large BRBs in a modern structure (Berkeley Animal Shelter) (© ci. berkeley.ca.usa)
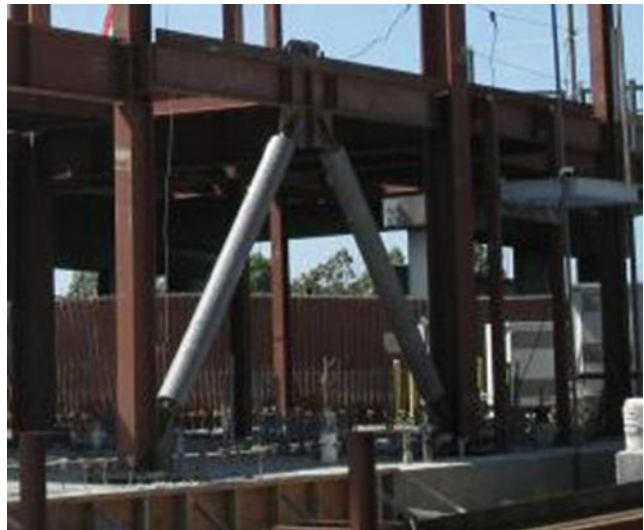
buckling of the brace. Many of these buckling-restrained braces are proprietary devices. BRBF systems originated in Japan in the early 1990s and are a very common type of structural system in the Western USA today.

The next system is similar, except that the braces are not so stiff and they will buckle slightly in compression. This system is known as a *special concentrically braced frame* (SCBF, Fig. 13c). In this system, the brace in tension is assumed to carry most of the force (generally about 70 %) as the compression brace will buckle and be able to provide only its post-buckling resistance. *Concentrically braced*

frames (CBF) originated when engineers began to stiffen MRF for wind loads in the 1910s using light diagonals, as in the Sather Tower in Berkeley, CA (Fig. 18). Modern CBF are much stiffer than SMRF, but generally less ductile as the buckling of the braces can lead to a concentration of deformations in a few floors. It is not clear when CBF began to be used in seismic systems (i.e., going from CBF to SCBF), but seismic design recommendations for braced frames began to appear in the 1960s. The original CBF had slender X-shaped braces in which only the brace in tension provided resistance. There are a large number of variations of this

S
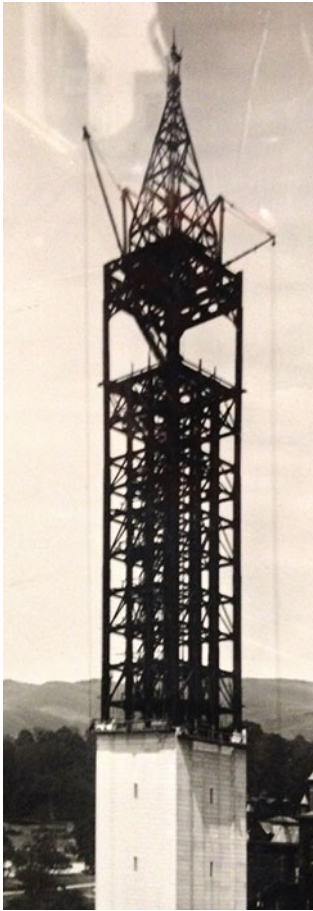
Steel Structures, Fig. 18 Sather Tower under construction (1914)

system (Fig. 19), each with its own design requirements and behavior characteristics.

Finally, one can think of beginning to move the brace in each bay horizontally, such that a hybrid structure between a pure moment frame and a braced frame is achieved. This is known as an *eccentrically braced frame* (Fig. 13b). In these structures, the central portion of the beam, known as the link, will deform and yield in shear if the link is short and in a combination of flexure and shear if the link is longer. Eccentrically braced frames have performed extremely well in recent earthquakes (Fig. 20), with numerous such structures providing superior performance during the 2011 Christchurch earthquake (Clifton et al. 2011).

Basis of Design

The seismic design of steel structures is predicated on having little or no structural damage under a service-level earthquake, moderate but rapidly repairable damage under a moderate earthquake, and not to collapse and cause loss of lives in the event of the design-level earthquake. Until recently, design specifications typically addressed only the design-level earthquake through a series of prescriptive design provisions intended to address ultimate strength design (ULD) criteria (limit states). However, the costs and time related to both loss of contents and functionality under smaller events and the desire to provide some freedom from prescriptive building codes have begun to drive the design process toward a performance-based design (PBD) approach. A pure PBD approach will be completely non-prescriptive, giving the structural engineer complete freedom in determining both the limit states and how to comply with them. In the short term (next 5–15 years), it is likely that a hybrid model for seismic design with elements of both ULS and PBD will emerge.

Whether a ULD or a PBD approach is used, the basis for design of steel structures remains the need to provide ductility, redundancy, robustness, and resiliency. Ductility comes from a synergistic use of the very ductile behavior of steel at the material level to create elements, connections, and structural systems capable of sustaining large inelastic deformations without appreciable stiffness or strength degradation. From the design standpoint, most of our ductility is ensured through a series of prescriptive requirements that have been developed through observed performance in past earthquakes (Fig. 19), advanced analytical studies (Fig. 20), and complex laboratory testing (Figs. 14, 15, and 16).

Redundancy is the ability of the structure to redistribute forces as inelastic action occurs in order to efficiently activate all major lateral load resistance systems. In its simplest form, redundancy can be defined as providing multiple load paths in a structure. Redundancy is primarily a function of the number and ductility of the lateral
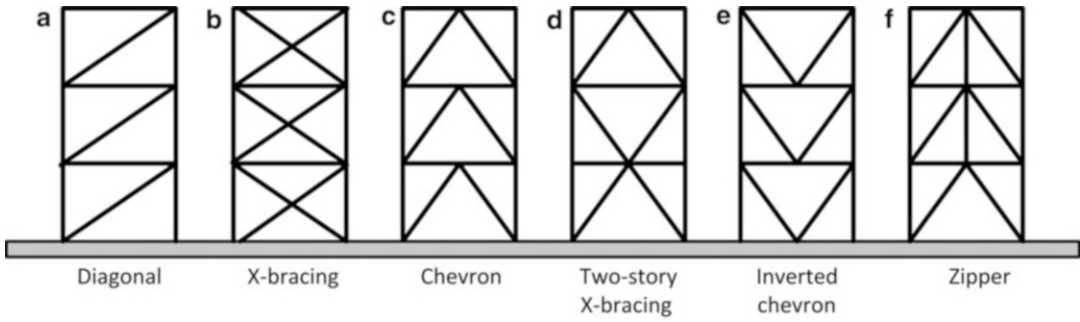
Steel Structures, Fig. 19 Different braced frame configurations (© RTLeon)

Steel Structures, Fig. 20 Damaged eccentrically brace frame, showing yielding of the shear link element (© RTLeon)

load-resisting systems employed, the system's rational 3D configuration, the ductility of connections, the strength of diaphragms and collector elements, and the detailing of the structure. Importantly, redundancy implies activating not only the lateral load-resisting system but also the gravity load system and other nonstructural elements that can contribute to the strength and stiffness of the system.

Robustness refers to the ability of the structural system to limit any failure to a relatively small part of the structure so that the consequences of the failure are not disproportionate to the initial failure. For example, a common way of determining the robustness of a moment frame is to remove an interior column in a lower floor (something that may happen during an earthquake) and assess the stability of the system under gravity loads (to make sure that the structures does not "pancake" and crush its

occupants). If the structure does not collapse, regardless of its damaged state, it will be considered robust. Clearly, a client can establish more stringent criteria for robustness in her/his structure if deemed necessary.

Resiliency refers to the ability of the structure to fulfill its intended functions after a large event. For example, a level-four trauma hospital would be resilient if it can continue to function uninterruptedly after the design-level earthquake. While it is possible today to provide that level of performance for the structural system, it is rather more difficult to do it for all the other systems (from electricity, water, oxygen, and the myriad of IT systems) present in such hospitals. For a large store, it may mean being able to get back in business within a few days and at a minimum of cost. Resiliency can be quantified in many ways, including the ability to resist ground motions significantly larger than the design one, the

costs to repair the structure, the time needed to carry out such repairs, and the indirect costs associated with business interruption.

Connections

In order to satisfy ductility, redundancy, robustness, and resiliency criteria, steel structures in seismic areas need to be designed primarily to avoid any type of brittle failure modes, which can result in sudden losses of local or global strength, stiffness, and stability. In general, it is possible to design and fabricate steel members (beams, columns, braces, and walls) to avoid these type of failures through close attention to proper material selection, using compact sections (i.e., keeping a small enough ratio of the width of an element to its length.) and providing lateral bracing to prevent out-of-plane buckling. The potential weak links in steel structures are the connections, which must transfer very large forces through complex force paths. The design of connections for seismic loads is a complex design issue as it is difficult to predict failure modes and their interactions.

From the standpoint of steel as a material, brittle failures can occur because of a combination of poor selection of or unmatched materials, large stress concentrations due to poor design or execution, and large triaxial state of stresses. These conditions can often arise in connections. Problems with welded connections were evidenced by many failures in the 1994 Northridge earthquake (Fig. 9). In the welds to the beam flanges in these connections, large triaxial stresses arise from a combination of multidirectional forces, residual stresses from welding, and the inevitable imperfections in the welds. These conditions lead to brittle material behavior because shear forces, the source of most deformations in steel, are small when compared to normal forces. Even if only minor overall damage to the structures was observed in structures with failed connections after the Northridge earthquake, a major effort was launched in the USA to better understand connection behavior.

As part of that effort, known as the SAC Project (FEMA 355B 2000b), extensive experimental studies were conducted on bolted and welded connections. The probability of brittle failure in connections can be minimized by (1) proper attention to selecting appropriate base materials and welding consumables with high toughness; (2) careful consideration of surface preparation, preheating, and welding sequences; (3) inspection; and (4) nondestructive testing. This amounts to the need for a very comprehensive quality assurance/quality control (QA/QC) plan throughout all phases of the construction projects (materials, fabrication, and erection). In the USA, these requirements are now embodied in the required weld procedure specifications (WPS).

Connections with bolts are usually less susceptible to material-type problems, but require more care in the fabrication as slight misalignments and exceeding tolerances can lead to severe problems during erection. In addition, the force transfer mechanisms in connections with bolts are less obvious and require more detailed computations than those with welds. Finally, in general, bolted connections will be less rigid than welded ones, requiring that this semirigid behavior be included in the analyses, substantially complicating the design process. Bolted connections, such as the end plates shown in Figs. 14 and 15, when properly detailed, can provide excellent performance.

From the structural standpoint, ductility should arise in "plastic hinges" or zones of concentrated plasticity primarily in beams. The behavior of a plastic hinge can be visualized as that of a nonlinear rotational spring. The ability of these hinges to rotate depends primarily on delaying the onset of any local or global buckling of the section. Local buckling (Fig. 14) is mostly dependent on the slenderness (width/thickness) ratio of the flanges and webs, and codes prescribe strict limits to ensure that this type of buckling does not occur until large rotations are achieved. Global buckling of a beam or column, which in this case implies large out-of-plane displacements, is also controlled by slenderness criteria,
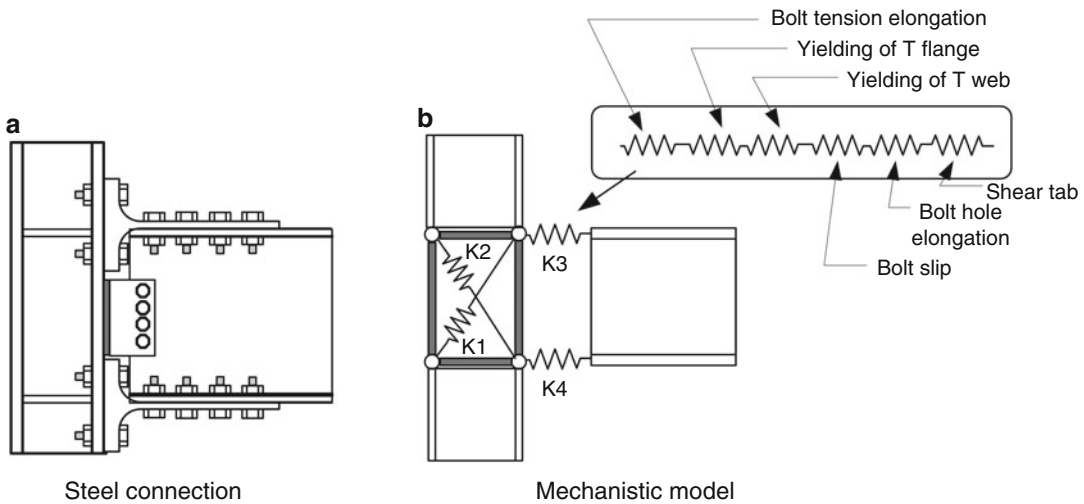
Steel Structures, Fig. 21 (**a**) T-stub bolted connection and (**b**) corresponding component model (© RTLeon)

in the form of length/moment of inertia ratios. Global buckling can be minimized by the addition of bracing along the member length, but this leads to additional costs and complications in the construction process. Global buckling or collapse of the structure can occur when large lateral deformations arise and the gravity forces result in large additional second-order effects. Global buckling is addressed in design by specifying strict drift (lateral deformation) limits and conducting advanced analyses that ensure overall stability.

As the results of the research began to filter into practice, a number of new design technologies became popular. For example, for the successful design of a T-stub connection (Fig. 21a), it is necessary to first determine all the possible ductile and brittle failure modes and prioritize them from most brittle to most ductile. One approach to this task is the component approach, in which each deformation mechanism in a joint is identified and individually quantified through a series of small laboratory component tests and associated analytical studies. These tests are carefully designed to measure one deformation component at a time. Each of these components is then represented by a spring with either linear or nonlinear characteristics (Fig. 21b). These springs are arranged in series or in parallel and the overall moment-rotation (M-θ) curve derived with the aid of simple computer programs that conduct the analysis of the spring system. In this example, the K1 and K2 springs model the panel zone deformation due to shear, while springs K3 and K4 model the bending deformations of the T-stubs. Springs K3 and K4 are made up of the contributions of several other springs that model different deformation components. With the aid of this approach, it is possible to achieve designs that meet and exceed current performance requirements. Care in the design of these connections can lead to very successful designs (Fig. 22).

Because the design of connections is a difficult and time-consuming task, in the USA, the idea of "prequalifying" connections has arisen (AISC 358 2010). In this case, an extensive experimental campaign is conducted to test a range of beam and column sizes with a particular type of connection. If the results indicate reliable ductile cyclic behavior to rotations around 4 %, after careful examination of all the results, the connection will become "prequalified." This implies that a series of simple, predefined design steps is all that is necessary to design that type of connection.

Steel Structures,
Fig. 22 Successful design
of a T-stub connection
(© RTLeon)

Steel Structures,
Fig. 23 Innovative,
proprietary connection
(ConXTech, www.
conxtech.com)
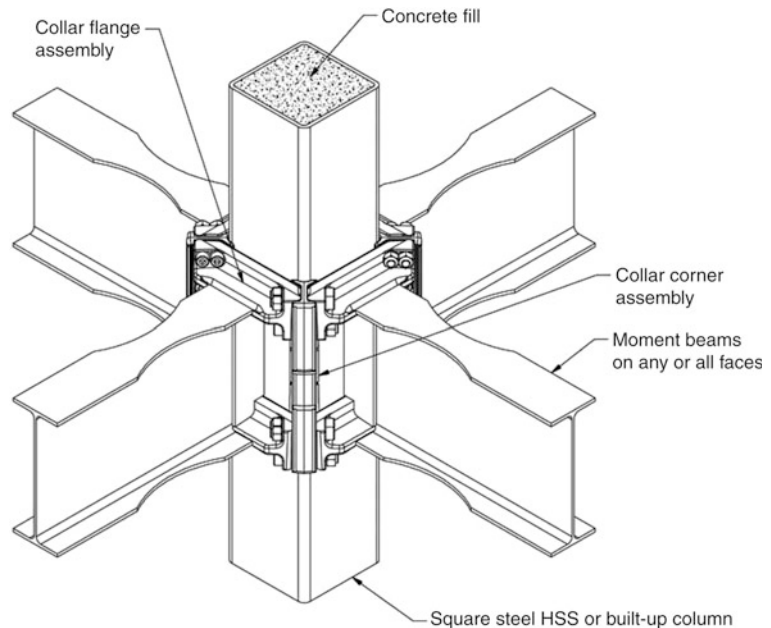
An interesting and important consequence of the "prequalification" process is that a number of proprietary connections have entered the market. In these cases, the manufacturer of the connection will provide all the necessary connection design data and manufacture all proprietary pieces. For example, the ConXL connection (Fig. 23, ConXtech Inc.) consists of a concrete-filled tube HSS column, RBS beams, and a series of steel castings that make the field assemblage of the connection a fast and simple operation.

Summary

Properly designed modern steel structures can provide superior performance when subjected to very large earthquake motions. The key to good performance is in strict attention to detail in every phase of the construction process: (a) initial selection of the structural form; (b) ductile design of members, particularly with respect to plastic hinge formation; (c) careful attention to connection strength and ductility; (d) frequent

interaction between the designer and fabricator to ensure proper fabrication procedures and tolerances; and (e) proper erection, along with extensive QA/QC for any field welding and bolting needed.

Cross-References

References

AISC 358 (2010) Prequalified connections for special and intermediate moment frames for seismic applications. American Institute of Steel Construction, Chicago

ASCE (1906) The effects of the San Francisco earthquake of April 18th, 1906, on Engineering Constructions: report of Committee on Fire and Earthquake Damage To Buildings. Trans ASCE 57(2):208–263

Ashton TS (1968) Iron and steel in the industrial revolution. Augustus M. Kelley, New York

Bodsworth C (ed) (2001) British iron and steel AD1800-2000 and beyond, Book 472. IOM Communications, London

Clifton C et al (2011) Steel structures damage from the Christchurch earthquake of February 22, 2011. Bull N Z Soc Earthq Eng 44(4):297–318

Condit CW (1968) American building. The University of Chicago Press, Chicago

FEMA 355E (2000a) State of the art report on past performance of steel moment-frame buildings. FEMA, Washington, DC

FEMA 355B (2000b) State of the art report on welding and inspection. FEMA, Washington, DC

Halmos GT (2000) Roll forming handbook. CRC Press, Boca Raton

Hamburger R, Meyer JD (2006) The performance of steel-frame buildings with infill masonry walls in the 1906 San Francisco earthquake. Earthq Spectra, EERI 22 (2):S43–S68

Kurzman D (2001) Disaster! The Great San Francisco earthquake and fire of 1906. Harper Perennial, New York

Leon et al (2014) Performance of RC beam-column joints in the Christchurch 2010–2011 earthquakes. In: SP-296: symposium honoring James O. Jirsa's contributions in structural concrete: a time to reflect, CD. ACI, Farmington Hills

Tylecote RF (1976) A history of metallurgy. Metals Society, London

Viest et al (1997) Composite construction design for buildings. McGraw-Hill/ASCE, New York, 416 pp

Stochastic Analysis of Linear Systems

Giuseppe Muscolino
Dipartimento di Ingegnria Civile, Informatica, Edile, Ambientale e Matematica Applicata, Università degli Studi di Messina, Messina, Italy

Synonyms

Evolutionary frequency response function; Evolutionary power spectral density function; Gaussian zero-mean random models of seismic accelerations; Non-geometric spectral moments; Stochastic analysis

Introduction

The stochastic analysis of structural vibrations deals with the description and characterization of structural loads and responses that are modeled as stochastic processes. The probabilistic characterization of the input process could be extremely complex in time domain where the probability density functions depend on the autocorrelation functions which experimentally have to be specified over given set points. Since this approach is difficult to be used in applications, stochastic vibration analysis of structural linear systems subjected to Gaussian input processes is quite often performed in the frequency domain by means of the spectral analysis. This analysis is a very powerful tool for the analytical and experimental treatment of a large class of physical as well as structural problems subjected to random excitations. The main reasons are (a) the spectrum has an immediate physical interpretation as a power-frequency distribution; (b) the spectrum provides information on the stochastic structure of the process; and (c) the spectrum may be estimated by fairly simple numerical techniques which do not require any specific assumption of the structure of the process.

S

In the framework of earthquake engineering, the stationary non-white input models were suggested first. These models, which account for site properties and for the dominant frequency in ground motion, fail to reproduce the time-varying intensity typical of real earthquakes ground-motion accelerograms. In order to overcome this drawback, the so-called quasi-stationary (or uniformly modulated) random processes have been introduced (see, e.g., Shinozuka and Sato 1967; Jennings et al. 1969; Hsu and Bernard 1978). These processes are constructed as the product of a stationary zero-mean Gaussian random process by a deterministic function of time; for this reason they are also called separable nonstationary stochastic processes.

Furthermore, a time-varying frequency content is observed in actual accelerogram records. This nonstationary frequency is prevalently due to different arrival times of the primary, secondary, and surface waves that propagate at different velocities through the Earth's crust. To take into account both the simultaneous amplitude and frequency non-stationarity, Spanos and Solomos (1983) proposed a non-separable model introducing a particular *evolutionary power spectral density* (EPSD) function; Fan and Ahmadi (1990) proposed a generalization of the Kanai-Tajimi filter model with time-dependent coefficients; Conte and Peng (1997) defined the ground-motion accelerations as the sum of a finite number of pairwise independent uniformly modulated zero-mean Gaussian stochastic process, the so-called sigma-oscillatory process.

Once the problem is formulated from a mathematical point of view, the further step deals with the evaluation of the structural response to perform the prediction of the safety of structural systems. In this framework, the maximum absolute peak of stationary or nonstationary stochastic responses may be useful in design information of several engineering situations (see, e.g., Lin 1976; Lutes and Sarkani 2004; Muscolino and Palmeri 2005; Li and Chen 2009). Approximate procedures to calculate the statistics of the maximum absolute peak of the response have been proposed. These procedures lead to the probabilistic assessment of structural failure as a function of barrier crossing rates, distribution of peaks, and extreme values. The latter quantities can be evaluated, for stationary input process, as a function of the well-known *geometric spectral moments* (GSMs) introduced by Vanmarcke (1972). For stationary stochastic response processes, the GSMs are defined as the geometric moments of the one-sided *power spectral density* (PSD) of the response process. Application of spectral methods to nonstationary random processes is more difficult than for the stationary ones; indeed for nonstationary processes the geometric approach fails (Di Paola 1985; Muscolino 1991). To perform the structural reliability in the latter cases, the so-called *nongeometric spectral moments* (NGSMs) have been introduced (Michaelov et al. 1999a, b).

In this study the before outlined topics will be addressed in order to evaluate the spectral characteristics of the structural response that are useful to perform the reliability assessment of linear systems subjected to stationary or nonstationary mono-/multi-correlated excitations.

Spectral Representation of Stochastic Processes

To carry out the spectral analysis, it is necessary to determine the spectral properties of the involved functions. These properties may be determined through the *Fourier-Stieltjes transform* (Priestley 1999).

Let us consider now a zero-mean stationary stochastic process, $F(t)$. This process is characterized by the feature that its *statistical moments* do not change over time and generally arise from any "stable" system which has achieved a "steady state." Moreover, the probabilistic structure of a stationary process is invariant under a shift of the time origin. This is a consequence of the fact that a sample of the process will almost certainly not "decay" to zero at infinity. Then the stationary processes possess infinite energy. Since a stationary process possesses infinite energy, its kth sample, $F^{(k)}(t)$, cannot be represented by the Fourier transform. In fact in this case, the Dirichlet condition is not satisfied. It follows that the spectral representation of a sample of a stationary

stochastic process can be performed only by the *Fourier-Stieltjes integral* (Priestley 1999):

$$F^{(k)}(t) = \int\limits_{-\infty}^{+\infty} \exp(\mathrm{i}\omega t)\,\mathrm{d}N^{(k)}(\omega) \qquad (1)$$

where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit and $N^{(k)}(\omega)$ is the kth sample of the complex stochastic process $N(\omega)$, satisfying the condition

$$\mathrm{E}\langle \mathrm{d}N(\omega_1)\,\mathrm{d}N^*(\omega_2)\rangle = \delta(\omega_1 - \omega_2)S_{NN}(\omega_1)\mathrm{d}\omega_1\mathrm{d}\omega_2 \tag{2}$$

where $\delta(\bullet)$ is the Dirac delta, the symbol $\mathrm{E}\langle\bullet\rangle$ means stochastic average, and the asterisk indicates the complex conjugate quantity. Notice that in Eq. 2 sometimes the Kronecker delta is introduced instead of the Dirac delta; this is to avoid the inconsistence of this relationship for $\omega_1 = \omega_2$ (Spanos and Solomos 1983). The relationship (2) shows that the stochastic process $N(\omega)$ is a process with orthogonal increments, in the sense that its increments $\mathrm{d}N(\omega_1)$ and $\mathrm{d}N(\omega_2)$ at any two distinct points ω_1 and ω_2 are uncorrelated random variables. Furthermore, in Eq. 2 $S_{NN}(\omega)$, which is a real and symmetric function, $S_{NN}(-\omega) = S_{NN}(\omega)$, is the *power spectral density* (PSD) function of the process $N(\omega)$. According to the theory of stationary stochastic process, the autocorrelation function, $R_{FF}(\tau)$, of the zero-mean stationary stochastic process $F(t)$ is a real function given as (Lin 1976; Lutes and Sarkani 2004; Li and Chen 2009)

$$\begin{aligned} R_{FF}(\tau) &= \mathrm{E}\langle F(t+\tau)F(t)\rangle \\ &= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} \exp[\mathrm{i}(\omega_1(t+\tau) - \omega_2 t)] \quad (3) \\ &\quad \mathrm{E}\langle \mathrm{d}N(\omega_1)\,\mathrm{d}N^*(\omega_2)\rangle \end{aligned}$$

which in virtue of Eq. 2 leads to

$$R_{FF}(\tau) = \mathrm{E}\langle F(t+\tau)F(t)\rangle = \int\limits_{-\infty}^{\infty} \exp(\mathrm{i}\omega\tau)S_{FF}(\omega)\mathrm{d}\omega \tag{4}$$

In this equation $S_{FF}(\omega)$ is the PSD function of the stationary process $F(t)$.

In postulating the stationarity of the stochastic process, very strong assumptions regarding the structure of the process are made. Once these assumptions are dropped, the process can become nonstationary in many different ways. In the framework of the spectral analysis of nonstationary processes, Priestley (see, e.g., Priestley 1999) introduced the *evolutionary power spectral density* (EPDS) function. The EPSD function has essentially the same type of physical interpretation of the PSD function of stationary processes. The main difference is that whereas the PSD function describes the power-frequency distribution for the whole stationary process, the EPSD function is time dependent and describes the local power-frequency distribution at each instant time. The theory of EPSD function is the only one which preserves this physical interpretation for the nonstationary processes. Moreover, since the spectrum may be estimated by fairly simple numerical techniques, which do not require any specific assumption of the structure of the process, this model, based on the EPSD function, is nowadays the most adopted model for the analysis of structures subjected to nonstationary processes as the seismic motion due to earthquakes.

In the Priestley spectral representation of nonstationary processes, a sample of the nonstationary stochastic process is defined by the *Fourier-Stieltjes integral* as follows:

$$F^{(k)}(t) = \int\limits_{-\infty}^{+\infty} \exp(i\omega t)\,a(\omega, t)\,\mathrm{d}N^{(k)}(\omega) \tag{5}$$

where $a(\omega, t)$ is a slowly varying complex deterministic time-frequency modulating function which has to satisfy the condition $a(\omega, t) \equiv a^*(-\omega, t)$ and $N(\omega)$ is an orthogonal process satisfying the condition (2). In Eq. 2 $S_{NN}(\omega)$ is the PSD function of the so-called "embedded" stationary counterpart process, $N(\omega)$ (Michaelov et al. 1999a). It follows that the autocorrelation function of the zero-mean Gaussian nonstationary random process $F(t)$ can be obtained as

S

$$
\begin{aligned}
R_{FF}(t_1,t_2) &= \mathrm{E}\langle F(t_1)F(t_2)\rangle \\
&= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} \exp[\mathrm{i}(\omega_1 t_1 - \omega_2 t_2)]\,a(\omega_1,t) \\
&\quad a^*(\omega_2,t_2)\,\mathrm{E}\langle \mathrm{d}N(\omega_1)\ \mathrm{d}N^*(\omega_2)\rangle \mathrm{d}\omega_1\,\mathrm{d}\omega_2
\end{aligned}
\tag{6}
$$

It is a real function which, in virtue of Eq. 2, leads to

$$
\begin{aligned}
R_{FF}(t_1,t_2) &= \int\limits_{-\infty}^{\infty} \exp[\mathrm{i}\omega(t_1 - t_2)]a(\omega,t_1)a^*(\omega,t_2) \\
S_{NN}(\omega)\mathrm{d}\omega &= \int\limits_{-\infty}^{\infty} \exp[\mathrm{i}\omega(t_1 - t_2)]S_{FF}(\omega,t_1,t_2)\mathrm{d}\omega
\end{aligned}
\tag{7}
$$

where

$$
S_{FF}(\omega,t_1,t_2) = a(\omega,t_1)\,a^*(\omega,t_2)\,S_{NN}(\omega) \tag{8}
$$

According to the Priestley evolutionary process model (Priestley 1999), the function

$$
S_{FF}(\omega,t) = |a(\omega,t)|^2 S_{NN}(\omega) \tag{9}
$$

is the so-called EPSD function of the nonstationary process $F(t)$. In the previous equations the symbol $|\bullet|$ denotes the modulus of the function in brackets. The processes characterized by the EPSD function $S_{FF}(\omega, t)$ are called *fully nonstationary* or *non-separable* random process, since both time and frequency content change, and they cannot be decoupled. If the modulating function is a time-dependent function, $a(\omega, t) \equiv a(t)$, the nonstationary process is called *quasi-stationary* or *uniformly modulated* or s*eparable* random process. In this case the time content change is independent by the frequency content change; indeed, the EPSD function assumes the following expression:

$$
S_{FF}(\omega,t) = a^2(t)\,S_{NN}(\omega) \tag{10}
$$

In the stochastic analysis the one-sided PSD is generally used; the latter can be suitably defined in the Priestley representation by the following equation:

$$
G_{FF}(\omega,t_1,t_2) = \begin{cases} a(\omega,t_1)\,a^*(\omega,t_2)\,G_{NN}(\omega) \equiv 2S_{FF}(\omega,t_1,t_2), \omega \ge 0; \\ 0, \omega < 0 \end{cases}
\tag{11}
$$

where $G_{NN}(\omega)$ $(G_{NN}(\omega) = 2S_{NN}(\omega),\ \omega \ge 0;$ $G_{NN}(\omega) = 0, \omega < 0)$ is the one-sided PSD function of the stationary counterpart of the input process $F(t)$. In this case the autocorrelation function of the process $F(t)$ is given by the following relationship:

$$
\begin{aligned}
\overline{R}_{FF}(t_1,t_2) &= \int\limits_{-\infty}^{\infty} \exp[\mathrm{i}\omega(t_1 - t_2)]a(\omega,t_1) \\
& a^*(\omega,t_2)\,G_{NN}(\omega)\mathrm{d}\omega
\end{aligned}
\tag{12}
$$

Note that since the one-sided PSD function $G_{NN}(\omega)$ is not symmetric, the corresponding autocorrelation function, $\overline{R}_{FF}(t_1,t_2)$, is a complex function (Di Paola 1985), whose real part coincides with the function defined in Eq. 6:

$\mathrm{Re}\{\overline{R}_{FF}(t_1,t_2)\} \equiv R_{FF}(t_1,t_2)$. It can be easily proved that the complex function (12) is the autocorrelation function of a complex process $\overline{F}(t)$ defined as (Di Paola and Petrucci 1990)

$$
\overline{F}(t) = \sqrt{2} \int\limits_{0}^{+\infty} \exp(\mathrm{i}\omega t)\,a(\omega,t)\,\mathrm{d}N(\omega) \tag{13}
$$

The real part of $\overline{F}(t)$ is proportional to the process $F(t)$, while the imaginary part of $\overline{F}(t)$ is a nonstationary process having stationary counterpart proportional to Hilbert transform of the real part of the stationary counterpart of the process itself (Di Paola 1985; Di Paola and Petrucci 1990. Muscolino 1991). The complex process, $\overline{F}(t)$, which generates the complex

autocorrelation function (12), has been called *pre-envelope process* by Di Paola (1985).

Reliability of Linear Structural Systems Subjected to Stochastic Excitations

The structural systems are conceived and designed to survive to natural actions. If the excitations are modeled as random processes, the dynamic responses are random processes too, and the structural safety needs to be evaluated in a probabilistic sense. Among the models of failure, the simplest one, which is also the most widely used in practical analyses, is based on the assumption that a structure fails as soon as the response at a critical location exits a prescribed safe domain for the first time. The probability of failure, in this case, coincides with the first passage probability, i.e., the probability that the absolute value of the random response process $X(t)$ of a selected structural response (e.g., strain or stress at a critical point) will exceed a specified safety bound, b, within a specified time interval (Lin 1976). In random vibration theory, the problem of probabilistically predicting this event is termed *first passage problem*. Unfortunately, this is one of the most complicated problems in computational stochastic mechanics. The solution of this problem has not been derived in exact form, even in the simplest case of the stationary response of a single-degree-of-freedom (SDoF) linear oscillator under zero-mean Gaussian white noise (Lin 1976; Lutes and Sarkani 2004; Muscolino and Palmeri 2005). Hence, a large number of approximated techniques have been proposed in literature, which differ in generality, complexity, and accuracy. In the framework of approximate methods, the time-dependent reliability of the structure, based on the first passage failure criterion, can be expressed, for a symmetric barrier, as (Lutes and Sarkani 2004)

$$L_{|X|}(b,t) = L_{|X|}(b,0)\exp\left[-\int_0^t \eta_X(b,\rho)\,\mathrm{d}\rho\right] \tag{14}$$

where $\eta_X(b,t)$ is the so-called *hazard function* and $L_{|X|}(b,0)$ is the reliability at time $t = 0$, which for the nonstationary case can be assumed as unity.

For narrow-band zero-mean stationary Gaussian process, the hazard function has been derived by Vanmarcke (1972) as

$$\eta_X(b) = \frac{1}{\pi}\sqrt{\frac{\lambda_{2,X}}{\lambda_{0,X}}}\left[\frac{1 - \exp\left(-b\,\delta_X^{1.2}\sqrt{\frac{\pi}{2\,\lambda_{0,X}}}\right)}{\exp\left(\frac{b^2}{2\,\lambda_{0,X}}\right) - 1}\right] \tag{15}$$

with

$$\delta_X = \sqrt{1 - \frac{\lambda_{1,X}^2}{\lambda_{0,X}\,\lambda_{2,X}}} \tag{16}$$

In this equation δ_X is the so-called *bandwidth parameter* of the process $X(t)$ (Vanmarcke 1972, 1975). This parameter measures the variation of the narrowness of the stochastic process $X(t)$. Usually, a stochastic process with $0 \leq \delta_X \leq 0.35$ is called a narrow-band stochastic process. Finally, in the previous equations $\lambda_{i,X}$ ($i = 0, 1, 2$) are the so-called *geometric spectral moments* (GSMs), introduced by Vanmarcke (1972) as the geometric moments of the one-sided PSD of the response process:

$$\lambda_{i,X} \equiv \lambda_{i,X}^G = \int_0^\infty \omega^i\,G_{XX}(\omega)\,\mathrm{d}\omega \tag{17}$$

In this equation the apex G emphasizes the geometric evaluation of the GSMs. Notice that $\lambda_{0,X}$ coincides with the variance of the zero-mean process $X(t)$, $\lambda_{0,X} \equiv \sigma_X^2 = \mathrm{E}\langle X^2(t)\rangle$; $\lambda_{2,X}$ coincides with the variance of the zero-mean process $\dot{X}(t)$, $\lambda_{2,X} \equiv \sigma_{\dot{X}}^2 = \mathrm{E}\langle \dot{X}^2(t)\rangle$; while $\lambda_{1,X}$ does not coincide with the cross-covariance of the processes $X(t)$ and $\dot{X}(t)$.

For narrow-band zero-mean nonstationary Gaussian process, the hazard function has been derived by Corotis et al. (1972) as

$$
\eta_X(b,t) = \frac{1}{\pi}\sqrt{\frac{\lambda_{2,X}(t)}{\lambda_{0,X}(t)}}\left[\frac{1-\exp\left(-b\delta_X^{1+d}(t)\sqrt{\frac{\pi}{2\lambda_{0,X}(t)}}\right)}{\exp\left(\frac{b^2}{2\lambda_{0,X}(t)}\right)-1}\right]
$$
(18)

where d $(d=0$ or $d=0.2)$ is an empirical parameter and $\delta_X(t)$ is the bandwidth parameter which in the nonstationary case is time dependent. Because of the non-stationarity of random process, this parameter involves complex functions, and it is defined by Michaelov et al. (1999a, b) as

$$
\delta_X(t) = \sqrt{1 - \frac{\mathrm{Re}\{\lambda_{1,X}(t)\}^2}{\lambda_{0,X}(t)\lambda_{2,X}(t)}}
$$
(19)

The evaluation of the time-dependent quantities $\lambda_{i,X}(t)$ is conceptually more complicated than for stationary processes; indeed for these processes the geometric approach fails for $i=1$ and $i=2$ (Di Paola 1985; Di Paola and Petrucci 1990; Muscolino 1991; Michaelov et al. 1999a, b), that is

$$
\lambda_{i,X}(t) \neq \lambda_{i,X}^G(t)
$$
$$
= \int_0^\infty \omega^i \, G_{XX}(\omega,t) \, \mathrm{d}\omega, \quad i=1,2 \quad (20)
$$

The physical inconsistency on the evaluation of the geometric GSMs, in the nonstationary case, as the moments of the one-sided EPSD function was pointed out by Corotis et al. (1972). In fact they discovered that for the case of the transient response of an oscillator subjected to stationary Gaussian white noise processes, the second GSM does not exist because it is unbounded. At same time in the stationary case, this GSM, which is the limit of the transient as the time approaches infinity, is finite. The first that considered the problem of spectral characteristics from a nongeometric point of view was Di Paola (1985). The basic idea

of this approach is to establish a time-domain interpretation of the SM. In order to do this, the pre-envelope covariances as the covariances of structural systems subjected to a complex-valued random process (the so-called pre-envelope process) have been introduced (Di Paola and Petrucci 1990). The real part of this process is proportional to the original nonstationary process, while the imaginary part is an auxiliary random process related to the real part in such a way that the complex process exhibits power in the positive frequency range only. Since the use of complex pre-envelope process is not very intuitive, Michaelov et al. (1999a, b) evaluated the pre-envelope covariances as a function of the EPSD of the response and recalled them as *nongeometric spectral moments* (NGSMs). It has to be emphasized that the NGSMs contain more information than the "conventional" covariances. Indeed, the NGSMs have been proved to be more appropriate for describing nonstationary process and can be effectively employed in structural reliability applications (Di Paola 1985; Di Paola and Petrucci 1990; Muscolino 1991; Michaelov et al. 1999a, b).

It has to be emphasized that in the framework of nonstationary analysis of structures, other time-dependent parameters, very useful in describing the time-variant spectral properties of the stochastic process, are (i) the mean frequency, $v_X^+(t)$, which evaluate the variation in time of the mean up-crossing rate of the time axis, and (ii) the central frequency, $\omega_{C,X}(t)$, which scrutinizes the variation of the frequency content of the stochastic process with respect to time. The two functions introduced before can be evaluated as a function of NGSMs and have been defined, respectively, as (Michaelov et al. 1999a, b)

$$
v_X^+(t) = \frac{1}{2\pi}\sqrt{\frac{\lambda_{2,X}(t)}{\lambda_{0,X}(t)}}; \quad \omega_{C,X}(t) = \frac{\mathrm{Re}\{\lambda_{1,X}(t)\}}{\lambda_{0,X}(t)}
$$
(21)

In stationary case the mean frequency, v_X^+, and the central frequency, $\omega_{C,X}$, are not time dependent. Moreover, in the latter case, because of the $\mathrm{Im}\{\lambda_{1,X}\}=0$, it follows $\mathrm{Re}\{\lambda_{1,X}\}\equiv\lambda_{1,X}$.

Response of Single-Degree-of-Freedom (SDoF) Oscillators

Fundamental of Deterministic Analysis

The theory of deterministic linear systems plays a fundamental role in the dynamic analysis of structures subjected to stochastic excitations. For this reason in this section the fundamental of deterministic analysis of SDoF subjected to deterministic excitation is synthetically reviewed. Particular care has been devoted to the state-space approach. This approach is the best suited for the development of formulations in the framework of random vibrations. In fact, its adaptability to numerical method of solution of differential equations and its extension to multi-degree-of-freedom (MDoF) systems are very straightforward.

The equation of motion of a linear oscillator with mass, m; viscous damping, c; stiffness, k; and subjected to the excitation $f(t)$ and at rest at initial time of motion, $t = t_0$, can be written as

$$
\begin{aligned}
& m \; \ddot{u}(t) + c \; \dot{u}(t) + k \; u(t) = f(t); \\
& u(t_0) = 0, \; \dot{u}(t_0) = 0
\end{aligned}
\quad (22)
$$

or in canonical form as follows:

$$
\begin{aligned}
& \ddot{u}(t) + 2\,\xi_0\,\omega_0 \; \dot{u}(t) + \omega_0^2 \; u(t) = F(t); \\
& u(t_0) = 0, \; \dot{u}(t_0) = 0
\end{aligned}
\quad (23)
$$

where $u(t)$ is the displacement response of the mass, $\omega_0 = \sqrt{k/m}$ is the natural circular frequency, $\xi_0 = c/2\sqrt{mk}$ is the damping ratio, and $F(t) = f(t)/m$; a dot over a variable denotes differentiation with respect to time t. In state variables the equation of motion of the oscillator, in canonical form, can be written as a set of two first-order differential equations:

$$
\dot{\mathbf{y}}(t) = \mathbf{D}_0 \; \mathbf{y}(t) + \mathbf{v}_0 \; F(t); \quad \mathbf{y}(t_0) = \mathbf{0} \quad (24)
$$

where

$$
\mathbf{y}(t) = \begin{bmatrix} u(t) \\ \dot{u}(t) \end{bmatrix}, \quad \mathbf{D}_0 = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -2\,\xi_0\,\omega_0 \end{bmatrix}, \quad \mathbf{v}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (25)
$$

Denoting by $\mathbf{y}_p(t)$ the particular solution vector, the solution of Eq. 24 can be evaluated as (Borino and Muscolino 1986)

$$
\mathbf{y}(t) = \mathbf{y}_p(t) + \Theta_0(t - t_0)\left[\mathbf{y}(t_0) - \mathbf{y}_p(t_0)\right] \quad (26)
$$

where $\Theta_0(t)$ is the so-called transition matrix:

$$
\Theta_0(t) = \exp(\mathbf{D}_0 t) = \begin{bmatrix} -\omega_0^2 g_0(t) & h_0(t) \\ -\omega_0^2 h_0(t) & \dot{h}_0(t) \end{bmatrix} \quad (27)
$$

with

$$
\begin{aligned}
g_0(t) &= -\frac{1}{\omega_0^2}\exp(-\xi_0\omega_0 t)\left[\cos(\overline{\omega}_0 t) + \frac{\xi_0\omega_0}{\overline{\omega}_0}\sin(\overline{\omega}_0 t)\right]; \\
h_0(t) &= \dot{g}_0(t) = \frac{1}{\overline{\omega}_0}\exp(-\xi_0\omega_0 t)\sin(\overline{\omega}_0 t); \\
\dot{h}_0(t) &= \exp(-\xi_0\omega_0 t)\left[\cos(\overline{\omega}_0 t) - \frac{\xi_0\omega_0}{\overline{\omega}_0}\sin(\overline{\omega}_0 t)\right]
\end{aligned}
\quad (28)
$$

and $\overline{\omega}_0 = \omega_0\sqrt{1 - \xi_0}$ is the damped natural circular frequency. Notice that the contribution of the last term in the right member of Eq. 26 decreases in the time because the transition matrix satisfies the following condition:

$$
\lim_{t \to \infty} \Theta_0(t) = \mathbf{0} \quad (29)
$$

Alternatively, by applying the so-called *parameter variation method*, the solution of Eq. 24 can be written in integral form as follows:

$$
\mathbf{y}(t) = \Theta_0(t - t_0)\mathbf{y}(t_0) + \int_{t_0}^{t} \Theta_0(t - \tau) \; \mathbf{v} \; F(\tau)\mathrm{d}\tau \quad (30)
$$

For quiescent systems, because the relationship $\mathbf{y}(t_0) = \mathbf{0}$ is satisfied, Eqs. 26 and 30 become

$$
\begin{aligned}
\mathbf{y}(t) &= \mathbf{y}_p(t) - \Theta_0(t - t_0)\mathbf{y}_p(t_0) \\
&= \int_{t_0}^{t} \Theta_0(t - \tau) \; \mathbf{v} \; F(\tau)\mathrm{d}\tau
\end{aligned}
\quad (31)
$$

Notice that in this case the first element of vector $\mathbf{y}(t)$, written in integral form, coincides with the well-known Duhamel integral. Starting by the integral form solution in state variables,

S

it is possible to derive a very powerful unconditionally stable numerical procedure for the evaluation of the structural response (see, e.g., Borino and Muscolino 1986).

Stochastic Response

Mathematically strictly speaking, as a consequence of the introduction of the one-sided PSD, the input process is a complex one. It follows that the response processes, $u(t)$, is a complex function too. After some algebra, for the quiescent oscillator (Eq. 23), the NGSMs can be evaluated, in time domain, as (Di Paola 1985; Di Paola and Petrucci 1990, Muscolino 1991)

$$\lambda_{0,uu}(t) \equiv \mathrm{E}\langle u(t)\,u^*(t)\rangle = \int_0^t\!\!\int_0^t h(t-\tau_1)h(t-\tau_2)\overline{R}_{FF}(\tau_1,\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2;$$

$$\lambda_{1,uu}(t) = -\mathrm{i}\int_0^t\!\!\int_0^t h(t-\tau_1)\dot{h}(t-\tau_2)\overline{R}_{FF}(\tau_1,\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2; \qquad (32)$$

$$\lambda_{2,uu}(t) \equiv \mathrm{E}\langle \dot{u}(t)\,\dot{u}^*(t)\rangle = \int_0^t\!\!\int_0^t \dot{h}(t-\tau_1)\dot{h}(t-\tau_2)\overline{R}_{FF}(\tau_1,\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2$$

where $\overline{R}_{FF}(t_1,t_2)$ is the complex autocorrelation function defined in Eq. 12. Moreover, the presence of the imaginary unit in the second of Eq. 32 inverts the roles of the real and imaginary parts of $\lambda_{1,uu}(t)$ with respect to the variances $\lambda_{0,uu}(t)$ and $\lambda_{2,uu\ell}(t)$; furthermore, while $\lambda_{0,uu}(t)$ and $\lambda_{2,uu}(t)$ are real functions, $\lambda_{1,uu}(t)$ is a complex one. Substituting Eq. 12 into Eq. 32, the following relationships are obtained:

$$\lambda_{0,uu}(t) = \int_0^\infty Z_0^*(\omega,t)Z_0(\omega,t)G_{NN}(\omega)\mathrm{d}\omega;$$

$$\lambda_{1,uu}(t) = -\mathrm{i}\int_0^\infty Z_0^*(\omega,t)\dot{Z}_0(\omega,t)G_{NN}(\omega)\mathrm{d}\omega; \quad (33)$$

$$\lambda_{2,uu}(t) = \int_0^\infty \dot{Z}_0^*(\omega,t)\dot{Z}_0(\omega,t)G_{NN}(\omega)\mathrm{d}\omega$$

where

$$Z_0(\omega,t) = \int_0^t h(t-\tau)\exp(\mathrm{i}\omega\tau)a(\omega,\tau)\,\mathrm{d}\tau;$$

$$\dot{Z}_0(\omega,t) = \int_0^t \dot{h}(t-\tau)\exp(\mathrm{i}\omega\tau)a(\omega,\tau)\,\mathrm{d}\tau \qquad (34)$$

Notice that the function $Z_0(\omega,t)$ is the so-called *evolutionary frequency response function* of the oscillator (Li and Chen 2009). Remarkably, since the integrals (Eq. 34) are convolution integrals of Duhamel's type, they can be interpreted as the response, in terms of state variables, of the quiescent oscillator, at time $t = 0$, subjected to the deterministic complex function $f(\omega,t) = \exp(\mathrm{i}\omega t)\,a(\omega,t)$. By introducing the state variables, the *evolutionary frequency response vector function* can be defined as

$$\mathbf{Y}_0(\omega,t) = \begin{bmatrix} Z_0(\omega,t) \\ \dot{Z}_0(\omega,t) \end{bmatrix} \qquad (35)$$

It follows that relationships (33) can be rewritten in compact form as follows:

$$\begin{aligned}\boldsymbol{\Sigma}_{uu}(t) &= \begin{bmatrix} \lambda_{0,uu}(t) & \mathrm{i}\lambda_{1,uu}(t) \\ -\mathrm{i}\lambda_{1,uu}^*(t) & \lambda_{2,uu}(t) \end{bmatrix} \\ &\equiv \int_0^\infty G_{NN}(\omega)\,\mathbf{Y}_0^*(\omega,t)\mathbf{Y}_0^T(\omega,t)\mathrm{d}\omega \end{aligned} \quad (36)$$

This matrix coincides with the so-called *pre-envelope covariance matrix* introduced by Di Paola and Petrucci (1990) which is a complex

matrix. As a conclusion, for input processes characterized by one-sided EPSD function, the cross-covariance matrix (Eq. 36) of an oscillator is a complex matrix whose elements are the NGSMs.

Generalizing Eq. 32, it can be easily proved that the complex cross-correlation function matrix of the zero-mean response process, which collects the cross-correlation of the pre-envelope response processes, can be evaluated as

$$\mathbf{R}_{uu}(t_1, t_2) = \int_0^\infty G_{NN}(\omega) \ \mathbf{Y}_0^*(\omega, t_1)\mathbf{Y}_0^T(\omega, t_2)\mathrm{d}\omega$$

(37)

Finally, according to the Priestley evolutionary process model (Priestley 1999), this complex function matrix $\mathbf{R}_{uu}(t_1, t_2)$ can be also rewritten as

$$\mathbf{R}_{uu}(t_1, t_2) = \int_0^\infty \exp[\mathrm{i}\omega(t_1 - t_2)]\mathbf{G}_{uu}(\omega, t_1, t_2)\mathrm{d}\omega$$

(38)

where

$$\mathbf{G}_{uu}(\omega, t_1, t_2) = G_{NN}(\omega) \ \mathbf{Y}_0^*(\omega, t_1)\mathbf{Y}_0^T(\omega, t_2)$$

(39)

Consequently the NGSMs of the oscillator response are the elements of the pre-envelope covariance matrix, given as

$$\Sigma_{uu}(t) \equiv \mathbf{R}_{uu}(t, t) \equiv \int_0^\infty \mathbf{G}_{uu}(\omega, t) \ \mathrm{d}\omega$$

(40)

Let us assume now the modulating function $a(\omega, t) = U(t)$. Starting from the nonstationary formulation, it is possible to deduce the formulation in the case of stationary input. This result is obtained by performing the limit as $t \to \infty$ into Eq. 34:

$$\lim_{t\to\infty} Z_0(\omega, t) = \int_0^\infty h_0(t-\tau)\exp(\mathrm{i}\omega\tau)\mathrm{d}\tau = \exp(\mathrm{i}\omega t) \ H_0(\omega);$$

$$\lim_{t\to\infty} \dot{Z}_0(\omega, t) = \int_0^\infty \dot{h}_0(t-\tau)\exp(\mathrm{i}\omega\tau)\mathrm{d}\tau = \mathrm{i}\omega\exp(\mathrm{i}\omega t) \ H_0(\omega)$$

(41)

where $h_0(t)$ and $\dot{h}_0(t)$ are the functions defined in Eq. 28 and $H_0(\omega)$ is the *frequency response function* of the oscillator defined as

$$H_0(\omega) = \int_0^\infty h_0(\tau)\exp(-\mathrm{i}\omega\tau)\mathrm{d}\tau$$

$$= \frac{1}{\omega_0^2 - \omega^2 + \mathrm{i}\,2\xi_0\omega_0\omega}$$

(42)

By substituting Eq. 41 into Eq. 35 and the results into Eq. 36, the following relationship is obtained:

$$\Sigma_{uu} = \int_0^\infty G_{NN}(\omega) \ \mathbf{H}_0^*(\omega)\mathbf{H}_0^T(\omega) \, \mathrm{d}\omega$$

$$\equiv \begin{bmatrix} \lambda_{0,\,uu} & \mathrm{i}\lambda_{1,\,uu} \\ -\mathrm{i}\lambda_{1,\,uu}^* & \lambda_{2,\,uu} \end{bmatrix}$$

(43)

where

$$\mathbf{H}_0(\omega) = \begin{bmatrix} H_0(\omega) \\ \mathrm{i}\omega H_0(\omega) \end{bmatrix}$$

(44)

The elements of this matrix Σ_{uu} are the GSMs, introduced by Vanmarcke (1972), that is

$$\lambda_{i,\,uu} \equiv \lambda_{i,\,uu}^G$$

$$= \int_0^\infty \omega^i |H_0(\omega)|^2 G_{NN}(\omega)\mathrm{d}\omega; \quad i = 0, 1, 2$$

(45)

Response of Multi-degree-of-Freedom (MDoF) Systems

Fundamental of Deterministic Analysis

Let us consider the equation of motion of a linear quiescent n-degree-of-freedom (n-DoF) classically

damped structural system whose dynamic behavior is governed by the equation of motion:

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \mathbf{f}(t) \qquad (46)$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are the $(n \times n)$ mass, damping, and stiffness matrices of the structure; $\mathbf{u}(t)$ is the $(n \times 1)$ vector of displacements, having for ith element $u_i(t)$; and $\mathbf{f}(t)$ denotes the external load vector. Under the assumption of classically damped system, the equation of motion can be decoupled by applying the modal analysis. To this aim let us introduce the modal coordinate transformation:

$$\mathbf{u}(t) = \mathbf{\Phi}\mathbf{q}(t) = \sum_{j=1}^{m} \mathbf{\phi}_j q_j(t) \Rightarrow u_i(t)$$
$$= \sum_{j=1}^{m} \phi_{ij} q_j(t) \qquad (47)$$

In this equation, $\mathbf{\Phi} = [\,\mathbf{\phi}_1 \quad \mathbf{\phi}_2 \quad \ldots \quad \mathbf{\phi}_m\,]$ is the modal matrix, of order $n \times m$, collecting the m eigenvectors $\mathbf{\phi}_j$, normalized with respect to the mass matrix $\mathbf{M}$, solutions of the following eigenproblem:

$$\mathbf{K}^{-1}\mathbf{M}\mathbf{\Phi} = \mathbf{\Phi}\mathbf{\Omega}^{-2}; \quad \mathbf{\Phi}^T\mathbf{M}\mathbf{\Phi} = \mathbf{I}_m \qquad (48)$$

In this equation $\mathbf{\Omega}$ is a diagonal matrix listing the undamped natural circular frequency ω_j, $\mathbf{I}_m$ is the identity matrix of order m, and the apex T means transpose operator. Once the modal matrix $\mathbf{\Phi}$ is evaluated, by applying the coordinate transformations (47) to Eq. 46, the following set of decoupled second-order differential equations is obtained:

$$\ddot{\mathbf{q}}(t) + \mathbf{\Xi}\dot{\mathbf{q}}(t) + \mathbf{\Omega}^2\mathbf{q}(t) = \mathbf{\Phi}^T\mathbf{f}(t) \qquad (49)$$

in which $\mathbf{\Xi}$ is a generalized damping matrix given by

$$\mathbf{\Xi} = \mathbf{\Phi}^T\mathbf{C}\mathbf{\Phi} \qquad (50)$$

For classically damped structures the modal damping matrix $\mathbf{\Xi}$ is a diagonal matrix listing the

quantities $2\xi_j\omega_j$, ξ_j being the modal damping ratio. It follows that the jth differential Eq. 49 can be written as

$$\ddot{q}_j(t) + 2\xi_j\omega_j\,\dot{q}_j(t) + \omega_j^2 q_j(t) = \mathbf{\phi}_j^T\mathbf{f}(t) \quad (51)$$

In the state space, Eq. 51 can be written in the first-order form as

$$\dot{\mathbf{y}}_j(t) = \mathbf{D}_j\mathbf{y}_j(t) + \mathbf{V}_j\mathbf{f}(t) \qquad (52)$$

where

$$\mathbf{y}_j(t) = \begin{bmatrix} q_j(t) \\ \dot{q}_j(t) \end{bmatrix}; \quad \mathbf{D}_j = \begin{bmatrix} 0 & 1 \\ -\omega_j^2 & -2\xi_j\omega_j \end{bmatrix};$$
$$\mathbf{V}_j = \begin{bmatrix} \mathbf{0} \\ \mathbf{\phi}_j^T \end{bmatrix}$$
$$(53)$$

Stochastic Response for Mono-correlated Stochastic Input Processes

Let us assume now that the forcing term is a mono-correlated zero-mean Gaussian random process vector given by the relationship:

$$\mathbf{f}(t) = \mathbf{b}\overline{F}(t) \qquad (54)$$

where $\mathbf{b}$ is the $(n \times 1)$ vector of spatial distribution of loads and $\overline{F}(t)$ is a zero-mean Gaussian nonstationary random process. It follows that the jth differential Eq. 51 can be written as

$$\ddot{q}_j(t) + 2\xi_j\omega_j\,\dot{q}_j(t) + \omega_j^2 q_j(t) = p_j\overline{F}(t), \quad j = 1, 2, \ldots, m$$
$$(55)$$

where

$$p_j = \mathbf{\phi}_j^T\mathbf{b} \qquad (56)$$

After very simple algebra it can be shown that the jth NGSMs, $\lambda_{j,u_iu_i}(t)$ $(i = 0, 1, 2)$, of the ith nodal displacement response, $u_i(t)$, are given as a function of modal NGSMs, $p_j\lambda_{j,k\ell}(t)$ $(\ell, k = 1 \ldots, m)$, by the following relationships:

$$\lambda_{0,u_i u_i}(t) = \sum_{k=1}^{m}\sum_{\ell=1}^{m} p_k p_\ell \phi_{ik} \phi_{i\ell} \lambda_{0,k\ell}(t);$$

$$\lambda_{1,u_i u_i}(t) = \sum_{k=1}^{m}\sum_{\ell=1}^{m} p_k p_\ell \phi_{ik} \phi_{i\ell} \lambda_{1,k\ell}(t); \quad (57)$$

$$\lambda_{2,u_i u_i}(t) = \sum_{k=1}^{m}\sum_{\ell=1}^{m} p_k p_\ell \phi_{ik} \phi_{i\ell} \lambda_{2,k\ell}(t)$$

It has to be emphasized that the zeroth NGSM, $\lambda_{0,u_i u_i}(t)$, and the second-order NGSM, $\lambda_{2,u_i u_i}(t)$, are real functions that coincide with the variances of the response in terms of displacement and velocity, respectively, while the first-order NGSM, $\lambda_{1,u_i u_i}(t)$, is a complex quantity whose real part can be evaluated as the cross-covariance

between the response process and the response velocity process of the same linear system subjected to a nonstationary input whose stationary counterpart is proportional to its Hilbert transform (Di Paola 1985; Langley 1986; Di Paola and Petrucci 1990; Muscolino 1991). As shown in Eq. 57 the nodal NGSMs can be evaluated as a function of $\lambda_{r,k\ell}(t)$, $r = 0, 1, 2$, which are the so-called time-dependent modal NGSMs "purged" by p_j factors. After some algebra, these quantities, which are complex ones, can be evaluated, in time domain, for quiescent structural systems as (Di Paola 1985; Di Paola and Petrucci 1990; Muscolino 1991)

$$\lambda_{0,k\ell}(t) = \int_0^t\int_0^t h_k(t-\tau_1)h_\ell(t-\tau_2)\overline{R}_{FF}(\tau_1,\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2;$$

$$\lambda_{1,k\ell}(t) = -\mathrm{i}\int_0^t\int_0^t h_k(t-\tau_1)\dot{h}_\ell(t-\tau_2)\overline{R}_{FF}(\tau_1,\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2; \qquad k = 1, 2, \ldots, m; \ \ \ell = 1, 2, \ldots, m;$$

$$\lambda_{2,k\ell}(t) = \int_0^t\int_0^t \dot{h}_k(t-\tau_1)\dot{h}_\ell(t-\tau_2)\overline{R}_{FF}(\tau_1,\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2$$

$$(58)$$

where $\overline{R}_{FF}(t_1,t_2)$ is the complex autocorrelation function defined in Eq. 12. Notice that the presence of the imaginary unit in the second term of the second of Eq. 58 inverts the roles of the real and imaginary parts of $\lambda_{1,k\ell}(t)$ with respect to the cross-covariance $\lambda_{0,k\ell}(t)$ and $\lambda_{2,k\ell}(t)$; furthermore for $k = \ell$ while $\lambda_{0,k\ell}(t)$ and $\lambda_{2,k\ell}(t)$ become real quantities, $\lambda_{1,k\ell}(t)$ remains a complex one.

It can be easily proved that the "purged" NGSMs, given in Eq. 58, can be evaluated as a function of the statistics of the response of a dummy oscillator whose motion is governed by

$$\ddot{z}_j(t) + 2\xi_j\omega_j\,\dot{z}_j(t) + \omega_j^2 z_j(t) = \overline{F}(t) \qquad (59)$$

where $\overline{F}(t)$ is a complex process whose imaginary part is a process having stationary counterpart proportional to Hilbert transform of the real part of the stationary counterpart of the complex

process itself; it follows that $z_j(t) = q_j(t)/p_j$ is a complex process too. Notice that, with the position $p_j = 1$, Eq. 59 coincides with Eq. 55. Then by substituting the complex function $\overline{R}_{FF}(t_1,t_2)$, defined in Eq. 12, into Eq. 58, after very simple algebra, it is possible to evaluate the "purged" NGSMs as

$$\lambda_{0,k\ell}(t) = \int_0^\infty Z_k^*(\omega,t)Z_\ell(\omega,t)G_{NN}(\omega)\mathrm{d}\omega;$$

$$\lambda_{1,k\ell}(t) = -\mathrm{i}\int_0^\infty Z_k^*(\omega,t)\dot{Z}_\ell(\omega,t)G_{NN}(\omega)\mathrm{d}\omega;$$

$$\lambda_{2,k\ell}(t) = \int_0^\infty \dot{Z}_k^*(\omega,t)\dot{Z}_\ell(\omega,t)G_{NN}(\omega)\mathrm{d}\omega$$

$$(60)$$

where, for $j = k, \ell$, the following positions have been made:

$$Z_j(\omega, t) = \int_0^t h_j(t - \tau) \exp(i\omega\tau) a(\omega, \tau) d\tau;$$

$$\dot{Z}_j(\omega, t) = \int_0^t \dot{h}_j(t - \tau) \exp(i\omega\tau) a(\omega, \tau) d\tau \quad (61)$$

with

$$h_j(t) = \frac{1}{\overline{\omega}_j} \exp\left(-\xi_j \omega_j t\right) \sin\left(\overline{\omega}_j t\right) \quad (62)$$

and $\overline{\omega}_j = \omega_j \sqrt{1 - \xi_j^2}$ ($j = k, \ell$) is the damped circular frequency of the jth dummy oscillator. Moreover, introducing the state variable, the *modal evolutionary frequency response function vector* is defined as

$$\mathbf{Y}_j(\omega, t) = \begin{bmatrix} Z_j(\omega, t) \\ \dot{Z}_j(\omega, t) \end{bmatrix}; \quad j = k, \ell \quad (63)$$

Then, relationships (33) can be rewritten in compact form as follows:

$$\begin{aligned} \boldsymbol{\Sigma}_{k\ell}(t) &= \begin{bmatrix} \lambda_{0,k\ell}(t) & i\lambda_{1,k\ell}(t) \\ -i\lambda_{1,k\ell}^*(t) & \lambda_{2,k\ell}(t) \end{bmatrix} \\ &\equiv \int_0^\infty G_{NN}(\omega) \ \mathbf{Y}_k^*(\omega, t) \mathbf{Y}_\ell^T(\omega, t) d\omega \end{aligned} \quad (64)$$

which coincides with the cross-modal *pre-envelope covariance matrix* which is a complex matrix (Di Paola and Petrucci 1990). As a conclusion for input processes characterized by one-sided PSD function, the cross-covariance matrix (Eq. 64) between the k, ℓth dummy oscillators is a complex matrix whose elements are the "purged" NGSMs. It can be easily proved that the complex cross-correlation function matrix of the zero-mean modal "purged" response process, which collects the cross-correlation of the

pre-envelope response processes, according to the Priestley evolutionary process model (Priestley 1999), can be evaluated as

$$\begin{aligned} \mathbf{R}_{k\ell}(t_1, t_2) &= \int_0^\infty G_{NN}(\omega) \ \mathbf{Y}_k^*(\omega, t_1) \mathbf{Y}_\ell^T(\omega, t_2) d\omega \\ &\equiv \int_0^\infty \exp[i\omega(t_1 - t_2)] \mathbf{G}_{k\ell}(\omega, t_1, t_2) d\omega \end{aligned}$$

$$(65)$$

where

$$\mathbf{G}_{k\ell}(\omega, t_1, t_2) = G_{NN}(\omega) \ \mathbf{Y}_k^*(\omega, t_1) \mathbf{Y}_\ell^T(\omega, t_2) \quad (66)$$

is the cross-modal EPSD function matrix between the k, ℓth dummy oscillators. By means of the modal transformation (Eq. 47), the nodal autocorrelation and the EPSD function matrices of the displacement response, $u_i(t)$, can be evaluated, after very simple algebra, respectively, as follows:

$$\begin{aligned} \mathbf{R}_{u_i u_i}(t_1, t_2) &= \sum_{k=1}^m \sum_{\ell=1}^m p_k p_\ell \phi_{ik} \phi_{i\ell} \mathbf{R}_{k\ell}(t_1, t_2), \\ \mathbf{G}_{u_i u_i}(\omega, t_1, t_2) &= \sum_{k=1}^m \sum_{\ell=1}^m p_k p_\ell \phi_{ik} \phi_{i\ell} \mathbf{G}_{k\ell}(\omega, t_1, t_2) \end{aligned}$$

$$(67)$$

Consequently the nodal displacement NGSMs are the elements of the nodal pre-envelope covariance matrix, given as

$$\begin{aligned} \boldsymbol{\Sigma}_{u_i u_i}(t) &\equiv \mathbf{R}_{u_i u_i}(t, t) \equiv \int_0^\infty \mathbf{G}_{u_i u_i}(\omega, t) \ d\omega \\ &= \sum_{k=1}^m \sum_{\ell=1}^m p_k p_\ell \phi_{ik} \phi_{i\ell} \boldsymbol{\Sigma}_{k\ell}(t) \end{aligned} \quad (68)$$

Let us assume now the modulating function $a(\omega, t) = U(t)$. Starting from the nonstationary formulation, it is possible to deduce the formulation in the case of stationary input. This result

is obtained by performing the limit as $t \to \infty$ into Eq. 61:

$$\lim_{t \to \infty} Z_j(\omega, t) = \int_0^\infty h_j(t - \tau)\exp(\mathrm{i}\omega\tau)\mathrm{d}\tau$$
$$= \exp(\mathrm{i}\omega t)\ H_j(\omega); \qquad (69)$$
$$\lim_{t \to \infty} \dot{Z}_j(\omega, t) = \int_0^\infty \dot{h}_j(t - \tau)\exp(\mathrm{i}\omega\tau)\mathrm{d}\tau$$
$$= \mathrm{i}\omega\ \exp(\mathrm{i}\omega t)\ H_j(\omega)$$

where $h_j(t)$ is the functions defined in Eq. 62 and $H_j(\omega)$ is the *frequency response function* of the jth modal oscillator:

$$H_j(\omega) = \int_0^\infty h_j(\tau)\exp(-\mathrm{i}\omega\tau)\mathrm{d}\tau$$
$$= \frac{1}{\omega_j^2 - \omega^2 + \mathrm{i}2\xi_j\omega_j\omega} \qquad (70)$$

By substituting Eq. 69 into Eq. 63 and the results into Eq. 64, the following relationship is obtained:

$$\Sigma_{k\ell} = \int_0^\infty G_{NN}(\omega)\ \mathbf{H}_k^*(\omega)\mathbf{H}_\ell^T(\omega)\,\mathrm{d}\omega$$
$$\equiv \begin{bmatrix} \lambda_{0,k\ell} & \mathrm{i}\lambda_{1,k\ell} \\ -\mathrm{i}\lambda_{1,k\ell}^* & \lambda_{2,k\ell} \end{bmatrix} \qquad (71)$$

where

$$\mathbf{H}_j(\omega) = \begin{bmatrix} H_j(\omega) \\ \mathrm{i}\omega H_j(\omega) \end{bmatrix} \quad j = k, \ell \qquad (72)$$

The elements of the matrix $\Sigma_{k\ell}$ are the cross GSMs, introduced by Vanmarcke (1972), that is,

$$\lambda_{i,k\ell} \equiv \lambda_{i,k\ell}^G$$
$$= \int_0^\infty \omega^i\ H_k^*(\omega)H_\ell(\omega)G_{NN}(\omega)\mathrm{d}\omega; \quad i = 0, 1, 2$$
$$\qquad (73)$$

Stochastic Response for Multi-correlated Stochastic Input Processes

In earthquake engineering it is common to assume that the entire base of a structure is subjected to a uniform ground motion. This hypothesis inherently implies that the ground motion is a result of spatially uniform motion. Thus it is certainly true when the base dimensions of the structure are small relative to the seismic vibration wavelengths. From an analytical point of view, this hypothesis leads to mono-correlated zero-mean Gaussian models of the ground-motion acceleration. It follows that this model of earthquake excitations is undoubtedly advantageous, because it substantially facilitates the stochastic analysis of dynamic problem. Indeed, in this case the stochastic analysis requires the knowledge of only one PSD of the input process.

The assumption of uniform ground motion is inappropriate for structures, such as long span bridges, which require accounting for spatial variability of the support motions. The main sources of ground-motion spatial variability are the loss of coherency of seismic waves with distance, the difference in the arrival times of waves at separate supports, and the difference in soil conditions underneath the supports. In these cases the multi-correlated model of the seismic excitation is more appropriate.

In order to take into account of the spatial variability of earthquake-induced ground motions, let us consider an n-DoF structural system subjected to an N support motion. It follows that the stochastic forcing vector process has to be modeled as a multi-correlated zero-mean Gaussian random vector process:

$$\mathbf{f}(t) = \mathbf{B}\,\overline{\mathbf{F}}(t) \qquad (74)$$

where $\mathbf{B}$ is the $(n \times N)$ matrix of spatial distribution of loads and $\overline{\mathbf{F}}(t)$ is a zero-mean Gaussian nonstationary multi-correlated random vector process of order N. Following the Priestley spectral representation of nonstationary processes (Priestley 1999), this vector can be defined by the *Fourier-Stieltjes integral* as follows (Di Paola and Petrucci 1990):

$$\overline{\mathbf{F}}(t) = \sqrt{2} \int\limits_{-\infty}^{+\infty} \exp(i\omega t) \mathbf{A}(\omega, t) \, d\mathbf{N}(\omega) \quad (75)$$

where $\mathbf{A}(\omega, t)$ is a slowly varying complex deterministic time-frequency modulating diagonal function matrix, of order $N \times N$, which has to satisfy the condition $\mathbf{A}(\omega, t) \equiv \mathbf{A}^*(-\omega, t)$, while $\mathbf{N}(\omega)$ is an orthogonal vector process satisfying the following conditions:

$$\mathrm{E}\langle d\mathbf{N}(\omega_1) \, d\mathbf{N}^{*T}(\omega_2)\rangle$$
$$= \delta(\omega_1 - \omega_2) \mathbf{G_{NN}}(\omega_1) d\omega_1 d\omega_2 \quad (76)$$

This relationship shows that the stationary counterpart of the multi-correlated stochastic process is a vector process, $\mathbf{N}(\omega)$ (of order N), with orthogonal increments. Furthermore, $\mathbf{G_{NN}}(\omega)$ is an $(N \times N)$ Hermitian matrix function which describes the one-sided PSD function matrix of the so-called "embedded" stationary counterpart vector process, $\mathbf{N}(\omega)$. After some algebra it can be proved that the autocorrelation function matrix of the zero-mean Gaussian nonstationary random vector process $\overline{\mathbf{F}}(t)$ can be obtained as

$$\overline{\mathbf{R}}_{\mathbf{FF}}(t_1, t_2) = \mathrm{E}\left\langle \overline{\mathbf{F}}(t_1) \overline{\mathbf{F}}^{*T}(t_2)\right\rangle = \int\limits_{0}^{\infty} \exp[i\omega(t_1 - t_2)] \mathbf{A}(\omega, t_1) \mathbf{G_{NN}}(\omega) \mathbf{A}^*(\omega, t_2) d\omega$$

$$= \int\limits_{-\infty}^{\infty} \exp[i\omega(t_1 - t_2)] \mathbf{G_{FF}}(\omega, t_1, t_2) d\omega \quad (77)$$

where

$$\mathbf{G_{FF}}(\omega, t_1, t_2) = \mathbf{A}(\omega, t_1) \mathbf{G_{NN}}(\omega) \mathbf{A}^*(\omega, t_2) \quad (78)$$

According to the Priestley evolutionary process model (Priestley 1999), the EPSD function matrix of the nonstationary multi-correlated process $\overline{\mathbf{F}}(t)$ is given as

$$\mathbf{G_{FF}}(\omega, t) = \mathbf{A}(\omega, t) \mathbf{G_{NN}}(\omega) \mathbf{A}^*(\omega, t) \quad (79)$$

In the framework of stochastic seismic analysis, the main distinct phenomena that give rise to the spatial variability of earthquake-induced ground motions are (i) the *incoherence effect* associated to loss of coherency of seismic waves due to the differential superpositioning of waves arriving from an extended source, (ii) the *wave-passage effect* due to difference in the arrival times of waves at separate supports, (iii) the *attenuation effect* due to gradual decay of wave amplitudes with distance due to energy

dissipation in the ground medium, and (iv) the *site-response effect* associated to spatially varying of the local soil profiles. These effects are taken into account by the so-called coherency function, $\gamma_{\mathbf{NN},s\,r}(\omega)$, defined as

$$\gamma_{\mathbf{NN},sr}(\omega) = \frac{G_{\mathbf{NN},sr}(\omega)}{\sqrt{G_{\mathbf{NN},rr}(\omega) \; G_{\mathbf{NN},ss}(\omega)}} \quad (80)$$

where $G_{\mathbf{NN},s\,r}(\omega)$ is the sth and rth element of the matrix $\mathbf{G_{NN}}(\omega)$ that is the one-sided PSD at the sth and the rth supports. The coherency function is usually written as

$$\gamma_{\mathbf{NN},sr}(\omega) = \rho_{sr}(\omega) \exp(-i\omega \; d_{sr}/c) \quad (81)$$

in which $\exp(-i \omega \; d_{s\,r}/c)$ is a measure of the wave-passage delay due to the apparent velocity of waves, c represents the velocity of the seismic waves through the ground (c decreases as the distance between the support points increases or the soil is softer), d_{sr} is the distance between the rth and sth support points, and $\rho_{sr}(\omega)$ is the real frequency-dependent spatial correlation function

between the two support points. Several models of the coherency function have been proposed in literature (see, e.g., Harichandran and Vanmarcke 1986; Zerva 1991). Moreover, the one-sided cross-PSD of ground accelerations in a particular direction between surface points r and s, $G_{\mathrm{NN},sr}(\omega)$, may be written as the product of the coherence function by the target one-sided PSD of ground acceleration $G_0(\omega)$ as follows:

$$G_{\mathbf{NN},sr}(\omega) = \begin{cases} \gamma_{\mathbf{NN},sr}(\omega)\,G_0(\omega), & \text{for } r \neq s \\ G_0(\omega), & \text{for } r = s \end{cases}$$
$$(82)$$

Let us evaluate now the NGSMs of the structural response. In the case of multi-correlated excitations, the forcing vector $\mathbf{f}(t)$ assumes the expression (Eq. 74); it follows that the differential equation (49), governing the motion in modal subspace, can be written as

$$\ddot{\mathbf{q}}(t) + \mathbf{\Xi}\dot{\mathbf{q}}(t) + \mathbf{\Omega}^2\mathbf{q}(t) = \mathbf{P}\overline{\mathbf{F}}\,(t) \qquad (83)$$

where

$$\mathbf{P} = \mathbf{\Phi}^T\mathbf{B} \qquad (84)$$

After very tedious algebra it can be shown that the jth NGSMs, $\lambda_{j,u_iu_i}(t)$ ($j = 0, 1, 2$), of the jth nodal displacement response, $u_i(t)$, are given as a function of modal NGSMs, $p_{j\,v}\lambda_{j,k\ell rs}(t)$ ($\ell, k = 1, \ldots, m$; $r, s = 1, \cdots, N$), by the following relationships:

$$\lambda_{0,u_iu_i}(t) = \sum_{k=1}^{m}\sum_{\ell=1}^{m}\sum_{r=1}^{N}\sum_{s=1}^{N}p_{kr}p_{\ell s}\phi_{ik}\phi_{i\ell}\lambda_{0,k\ell rs}(t);$$
$$\lambda_{1,u_iu_i}(t) = \sum_{k=1}^{m}\sum_{\ell=1}^{m}\sum_{r=1}^{N}\sum_{s=1}^{N}p_{kr}p_{\ell s}\phi_{ik}\phi_{i\ell}\lambda_{1,k\ell rs}(t);$$
$$\lambda_{2,u_iu_i}(t) = \sum_{k=1}^{m}\sum_{\ell=1}^{m}\sum_{r=1}^{N}\sum_{s=1}^{N}p_{kr}p_{\ell s}\phi_{ik}\phi_{i\ell}\lambda_{2,k\ell rs}(t)$$
$$(85)$$

where p_{jv} is the jth and vth element of the matrix $\mathbf{P}(j = k, \ell; v = r, s)$. Once again it has to be emphasized that the zeroth NGSM, $\lambda_{0,u_iu_i}(t)$, and the second-order NGSM, $\lambda_{2,u_iu_i}(t)$, are real functions that coincide with the variances of the

response in terms of displacement and velocity, respectively, while the first-order NGSM, $\lambda_{1,u_iu_i}(t)$, is a complex quantity (Di Paola 1985; Langley 1986; Di Paola and Petrucci 1990; Muscolino 1991).

As shown in Eq. 85 the nodal NGSMs can be evaluated as a function of $\lambda_{j,k\ell rs}(t)$, $j = 0, 1, 2$, which are the so-called time-dependent modal NGSMs "purged" by p_{jv} factors. After some algebra, these quantities, which are complex ones, can be evaluated, in frequency domain, for quiescent structural systems, at time $t = 0$, as (Di Paola 1985; Di Paola and Petrucci 1990, Muscolino1991):

$$\lambda_{0,k\ell rs}(t) = \int_0^\infty Z^*_{kr}(\omega,t)Z_{\ell s}(\omega,t)G_{\mathbf{NN},sr}(\omega)\,\mathrm{d}\omega;$$
$$\lambda_{1,k\ell rs}(t) = -\mathrm{i}\int_0^\infty Z^*_{kr}(\omega,t)\dot{Z}_{\ell s}(\omega,t)G_{\mathbf{NN},sr}(\omega)\,\mathrm{d}\omega;$$
$$\lambda_{2,k\ell rs}(t) = \int_0^\infty \dot{Z}^*_{kr}(\omega,t)\dot{Z}_{\ell s}(\omega,t)G_{\mathbf{NN},sr}(\omega)\,\mathrm{d}\omega$$
$$(86)$$

where for $j = k, \ell$, the following positions have been made:

$$Z_{jv}(\omega,t) = \int_0^t h_j(t-\tau)\exp(\mathrm{i}\omega\tau)a_{vv}(\omega,\tau)\;\mathrm{d}\tau;$$
$$\dot{Z}_{jv}(\omega,t) = \int_0^t \dot{h}_j(t-\tau)\exp(\mathrm{i}\omega\tau)a_{vv}(\omega,\tau)\;\mathrm{d}\tau;$$
$$j = k, \ell; \quad v = r, s$$
$$(87)$$

with $h_j(t)$ the function defined in Eq. 62 and $a_{vv}(\omega, t)$ the vth element of the diagonal matrix $\mathbf{A}(\omega, t)$. Moreover, introducing the state variable, the *modal evolutionary frequency response function vector* is defined as

$$\mathbf{Y}_{jv}(\omega,t) = \begin{bmatrix} Z_{jv}(\omega,t) \\ \dot{Z}_{jv}(\omega,t) \end{bmatrix}; \quad j = k, \ell; \quad v = r, s$$
$$(88)$$
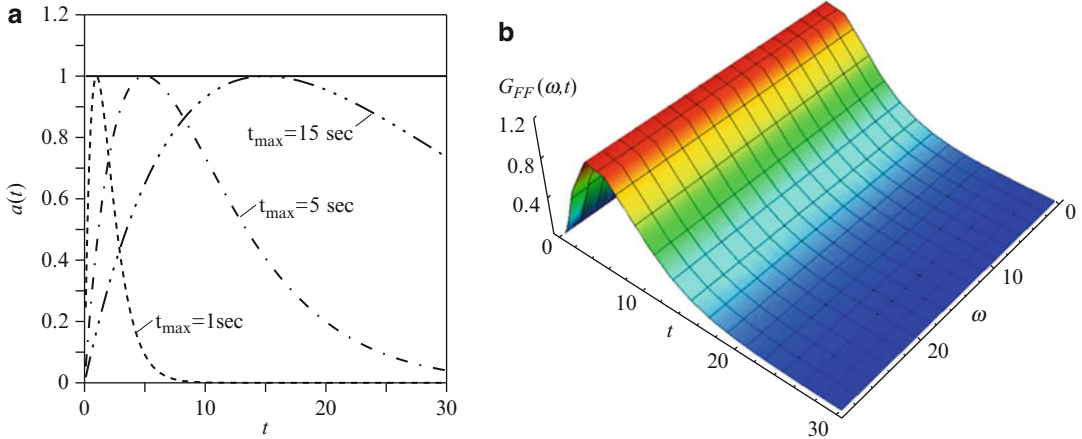
Stochastic Analysis of Linear Systems, Fig. 1 Hsu and Bernard (1978) model of the uniformly modulated nonstationary excitation: (**a**) modulating function for different value of the time instant, t_{max}, in which it takes the maximum value; (**b**) EPSD function $G_{FF}(\omega, t) = |a(t)|^2 G_{NN}(\omega)$ with $G_{NN}(\omega) = 1$ cm^2/sec^3 and $\alpha = 1/5$

Then relationships (33) can be rewritten in compact form as follows:

$$\Sigma_{k\ell rs}(t) = \begin{bmatrix} \lambda_{0,k\ell rs}(t) & i\lambda_{1,k\ell rs}(t) \\ -i\lambda_{1,k\ell rs}^*(t) & \lambda_{2,k\ell rs}(t) \end{bmatrix}$$
$$\equiv \int_0^\infty \mathbf{Y}_{kr}^*(\omega,t) G_{\mathbf{NN},sr}(\omega) \mathbf{Y}_{\ell s}^T(\omega,t) d\omega \tag{89}$$

where $\Sigma_{k\ell rs}(t)$ coincides with the cross-modal *pre-envelope covariance matrix* (Di Paola and Petrucci 1990). The one-sided PSD $G_{\mathbf{NN},sr}(\omega)$ can be evaluated as a function of the coherence function introduced in Eq. 81. Notice that according to the Priestley evolutionary process model (Priestley 1999), the complex cross-correlation function matrix of the zero-mean modal "purged" response process can be evaluated as

$$\mathbf{R}_{k\ell rs}(t_1,t_2) = \int_0^\infty \mathbf{Y}_{kr}^*(\omega,t_1) G_{\mathbf{NN},sr}(\omega) \mathbf{Y}_{\ell s}^T(\omega,t_2) d\omega$$
$$\equiv \int_0^\infty \exp[i\omega(t_1-t_2)] \mathbf{G}_{k\ell rs}(\omega,t_1,t_2) d\omega \tag{90}$$

where

$$\mathbf{G}_{k\ell rs}(\omega,t_1,t_2) = \mathbf{Y}_{kr}^*(\omega,t_1) G_{\mathbf{NN},sr}(\omega) \mathbf{Y}_{\ell s}^T(\omega,t_2) \tag{91}$$

is the cross-modal EPSD function matrix between the k, ℓth dummy oscillators at the rth and sth support points. By means of the modal transformation (47), the nodal autocorrelation and the EPSD function matrices of the displacement response, $u_i(t)$, can be evaluated, after very simple algebra, respectively, as follows:

$$\mathbf{R}_{u_i u_i}(t_1,t_2) = \sum_{k=1}^m \sum_{\ell=1}^m \sum_{r=1}^N \sum_{s=1}^N p_{kr} p_{\ell s} \phi_{ik} \phi_{i\ell} \mathbf{R}_{k\ell rs}(t_1,t_2),$$
$$\mathbf{G}_{u_i u_i}(\omega,t_1,t_2) = \sum_{k=1}^m \sum_{\ell=1}^m \sum_{r=1}^N \sum_{s=1}^N p_{kr} p_{\ell s} \phi_{ik} \phi_{i\ell} \mathbf{G}_{k\ell rs}(\omega,t_1,t_2) \tag{92}$$
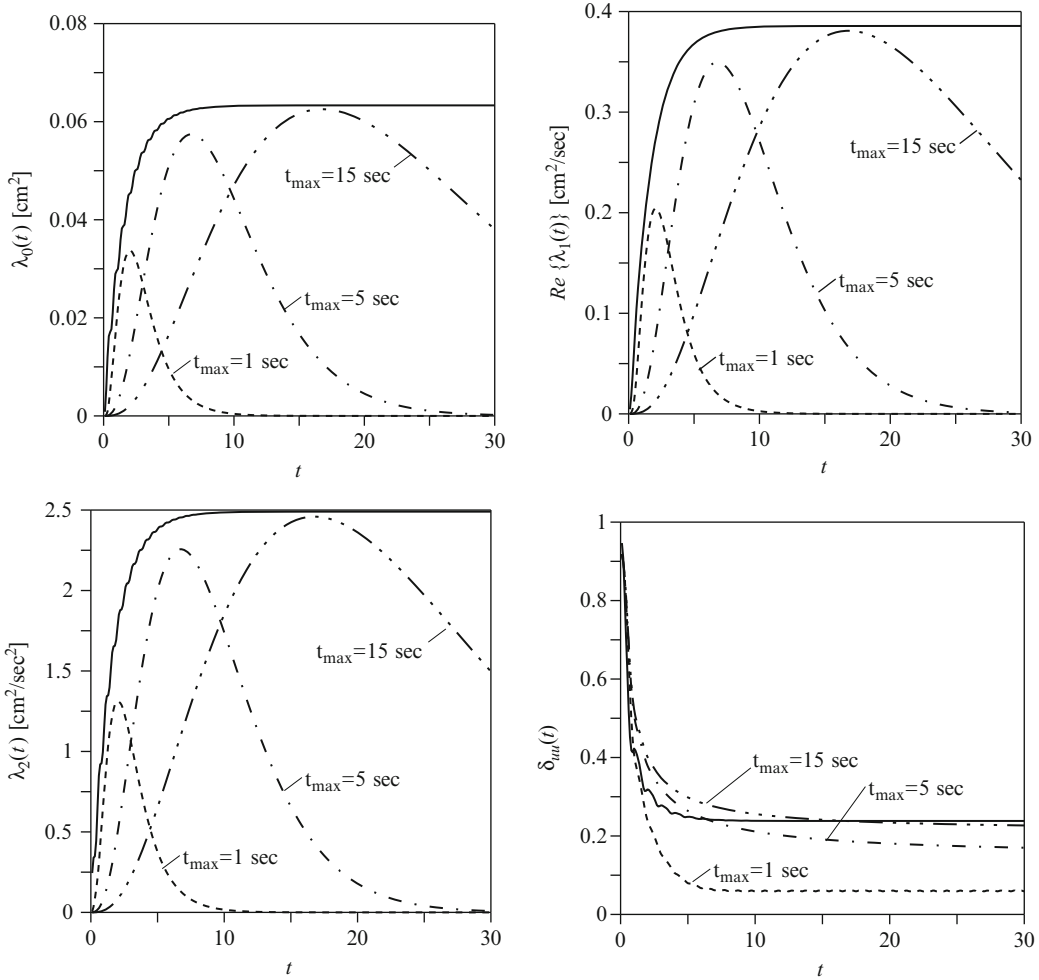
Stochastic Analysis of Linear Systems, Fig. 2 NGSM functions, $\lambda_{j,uu}(t)$ ($j = 0, 1, 2$) and bandwidth parameter, $\delta_{uu}(t)$, of the transient response of an oscillator ($\omega_0 = 2\pi$ rad/sec, $\xi_0 = 0.05$) in the stationary case (*solid line*) and adopting the Hsu and Bernard (1978) model, for different value of the time instant, t_{max}, in which the modulating $a(t)$ function take the maximum value (*dashed lines*)

Consequently the nodal displacement NGSMs are the elements of the nodal pre-envelope covariance matrix, given as

$$\Sigma_{u_i u_i}(t) \equiv \mathbf{R}_{u_i u_i}(t,t) \equiv \int_0^\infty \mathbf{G}_{u_i u_i}(\omega,t) \, d\omega$$

$$= \sum_{k=1}^m \sum_{\ell=1}^m \sum_{r=1}^N \sum_{s=1}^N p_{kr} p_{\ell s} \phi_{ik} \phi_{i\ell} \Sigma_{k\ell rs}(t)$$

$$(93)$$

Calculating Nonstationary Stochastic Responses

In this section, in order to evidence the main differences between the stochastic responses of structural systems subjected to both uniformly and fully nonstationary models of seismic excitations, the spectral characteristics of the response of two SDoF oscillators are evaluated.
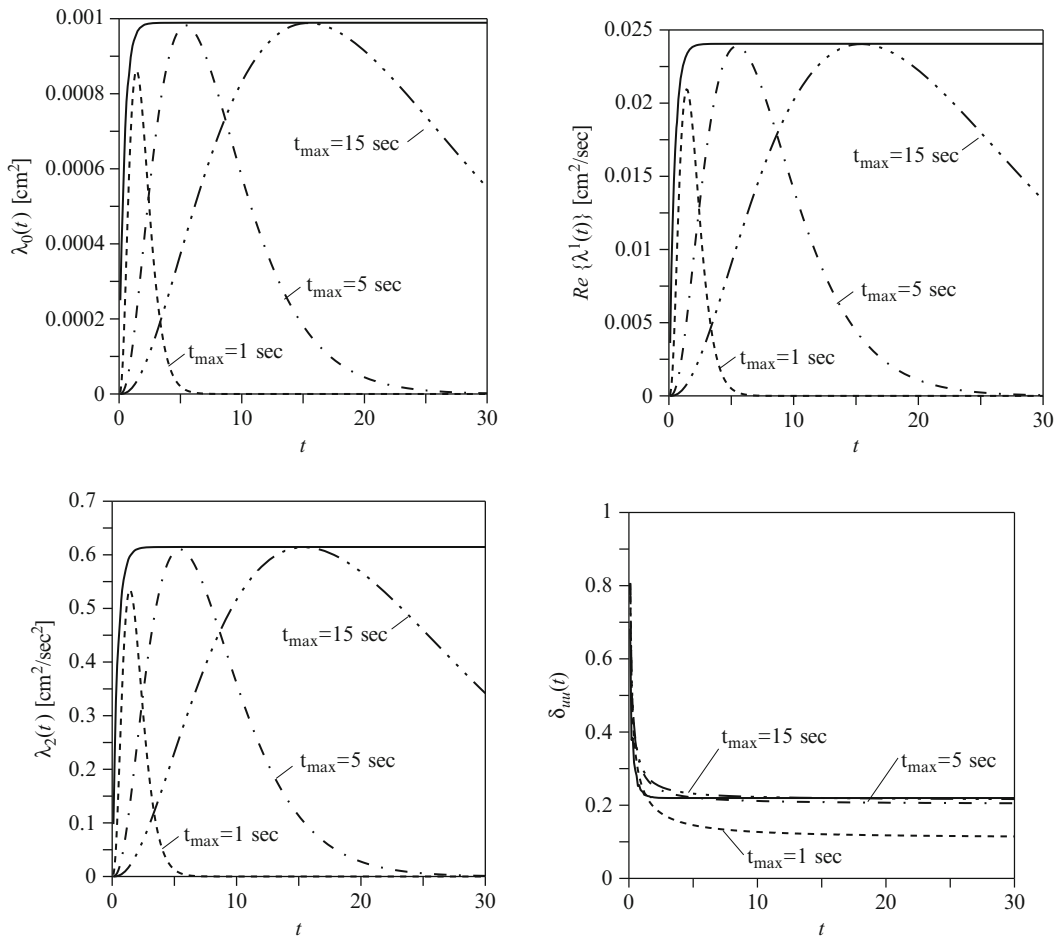
Stochastic Analysis of Linear Systems, Fig. 3 NGSMs, $\lambda_{j,uu}(t)$ ($j = 0, 1, 2$) and bandwidth parameter, $\delta_{uu}(t)$, of the transient response of an oscillator ($\omega_0 = 8\pi$ rad/sec, $\xi_0 = 0.05$) in the stationary case (*solid line*) and adopting the Hsu and Bernard (1978) model, for different value of the time instant, t_{max}, in which the modulating $a(t)$ function take the maximum value (*dashed lines*)

Comparison Between Steady-State and Nonstationary Responses of SDoF Systems

Let us consider an oscillator, whose differential equation governing the motion is written in canonical form in Eq. 23, forced by the uniformly modulate Gaussian zero-mean nonstationary process $F(t)$, defined as

$$F(t) = a(t) \ N(t) \qquad (94)$$

where $N(t)$ is a Gaussian white noise process, with one-sided PSD $G_{NN}(\omega) = 1$ cm^2/sec^3, and $a(t)$ is the normalized to one modulating function proposed by Hsu and Bernard (1978):

$$a(t) = \alpha \ \exp(1) t \ \exp(-\alpha t) \ U(t) \qquad (95)$$

which takes its maximum value at time, $t_{max} = 1/\alpha$, and $U(t)$ is the *unit step function* defined as

$$U(t - t_0) = \begin{cases} 0, & t \le t_0 ; \\ 1, & t > t_0 \end{cases} \qquad (96)$$

In Fig. 1 the modulating function (95) is depicted for different value of $\alpha = 1/t_{max}$.
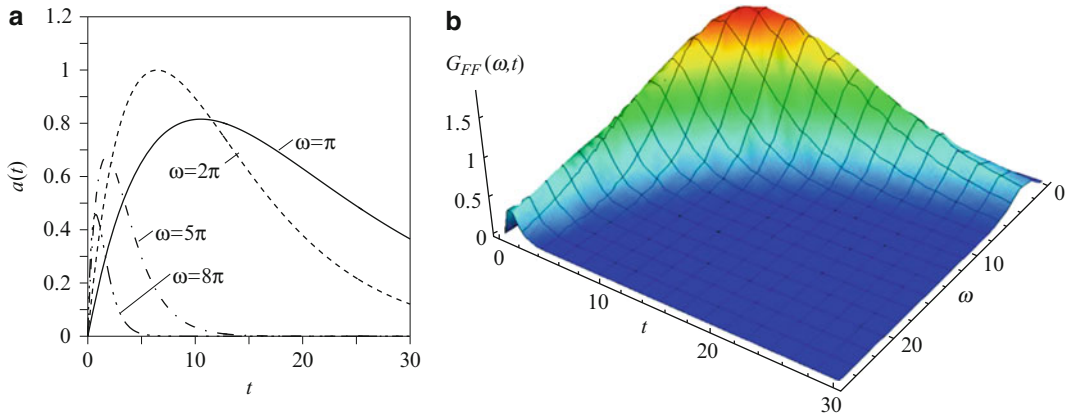
Stochastic Analysis of Linear Systems, Fig. 4 Spanos and Solomos (1983) model of the fully nonstationary excitation model: (**a**) sections of the normalized to one modulating function $a(\omega, t)$ at different values of abscissa ω; (**b**) EPSD function $G_{FF}(\omega, t) = |a(\omega, t)|^2\, G_{NN}(\omega)$ with $G_{NN}(\omega) = 1\ \text{cm}^2/\text{sec}^3$

In the same figure the EPSD function $G_{FF}(\omega, t)$ $= |a(t)|^2\, G_{NN}(\omega)$ of the quasi-stationary (separable) Hsu and Bernard (1978) model is also shown assuming $\alpha = 1/5$ and $G_{NN}(\omega) = 1\ \text{cm}^2/\text{sec}^3$.

In Fig. 2 the NGSM, $\lambda_{j,uu}(t)$ ($j = 0, 1, 2$), and bandwidth parameter, $\delta_{uu}(t)$, functions of the response of an oscillator with natural circular frequency $\omega_0 = 2\pi$ rad/sec and damping ratio $\xi_0 = 0.05$ for the Hsu and Bernard (1978) model of the uniformly modulated nonstationary excitation are depicted. These NGSMs are compared to the transient NGSMs of the stationary case ($a(t) = U(t)$), for different value of the time instant, $t_{\max}$. In Fig. 3 the comparisons are performed for an oscillator with natural circular frequency $\omega_0 = 8\pi$ rad/sec and damping ratio $\xi_0 = 0.05$. These figures shows that when the response approaches to its steady-state condition, defined by the steady-state time $t_{SS} \geq 3/(\xi_0\, \omega_0)$, the maximum values of the NGSM functions coincide with the corresponding values of the NGSM for stationary input processes. If this condition is not satisfied, the stationary approximation of input process leads to overestimated results. In fact, for the first oscillator for which the steady-state time is $t_{SS} \approx 9.55$ sec, the NGSMs with $t_{\max} < t_{SS}$ possess maximum values lesser than the corresponding stationary case. Similar results

are obtained for the second oscillator for which $t_{SS} \approx 2.39$ sec. Moreover, the analysis of the bandwidth parameters shows that if $t_{\max} < t_{SS}$ the parameter $\delta_{uu}(t)$ is narrower than the value obtained in the relative stationary case.

Comparison Between Uniformly and Fully Nonstationary Responses of SDoF Systems

Let us consider now the Spanos and Solomos (1983) model of the fully nonstationary Gaussian zero-mean process. For this model the normalized to one evolutionary modulating function can be written as

$$a(\omega, t) = \varepsilon(\omega)\, t \exp(-\alpha(\omega)\, t)\, U(t) \qquad (97)$$

Selecting the following parameters $\varepsilon(\omega) = \omega\sqrt{2}/5\pi a_{\max}$ and $\alpha(\omega) = 0.15/2 + \varepsilon^2(\omega)/4$, the normalizing to one coefficient is $a_{\max} = 1.34$. The unitary maximum is reached at $\omega = 1.937\,\pi$ rad/sec and at $t = 6.667$ sec. In Fig. 4, the section of the modulating function at different values of abscissa ω is depicted together with the one-sided EPSD function $G_{FF}(\omega, t) = |a(\omega, t)|^2\, G_{NN}(\omega)$ with $G_{NN}(\omega) = 1\ \text{cm}^2/\text{sec}^3$. This figure evidences the frequency dependence of this model of nonstationary input process.

S
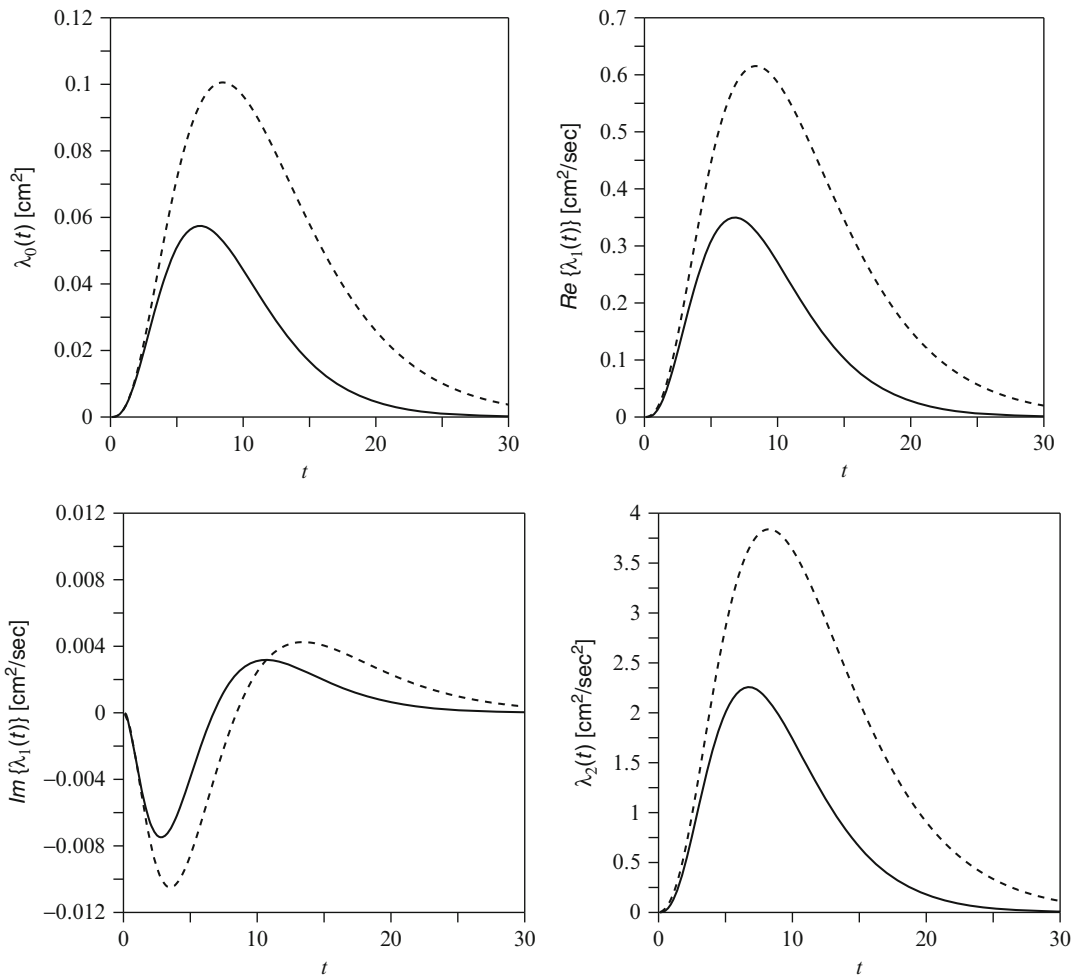
Stochastic Analysis of Linear Systems, Fig. 5 NGSMs of the response of an oscillator with $\omega_0 = 2\pi$ rad/sec, $\xi_0 = 0.05$ for the Hsu and Bernard (1978) model (*solid line*) and for the Spanos and Solomos (1983) model (*dashed line*)

In Figs. 5 and 6 the NGSMs, $\lambda_{j,uu}(t)$ ($j = 0, 1, 2$), of the two oscillators with natural circular frequency $\omega_0 = 2\pi$ rad/sec and $\omega_0 = 8\pi$ rad/sec and damping ratio $\xi_0 = 0.05$, obtained by means of the two modulating functions before described, are depicted and compared. Analyzing these figures it can be evidenced that the temporal variation of the frequency content of the EPSD function, often neglected for mathematic convenience, has substantial effects on the structural response. In fact the maximum values of the NGSM functions depend on the form of modulating function as well as the dynamic characteristics of the structural systems. Moreover, in some cases, the quasi-stationary modellation of the modulating function can lead to very unconservative results (see Fig. 5) of the stochastic structural response.

In Figs. 7 and 8 the mean frequencies, $v_{uu}^+(t)$, and the normalized time-varying central frequencies, $\omega_{C,uu}(t)/\omega_0$, defined in Eq. 21, of the response of the two oscillators before analyzed for the Hsu and Bernard (1978) and Spanos and Solomos (1983) modulating functions of the nonstationary zero-mean Gaussian input process, are depicted.
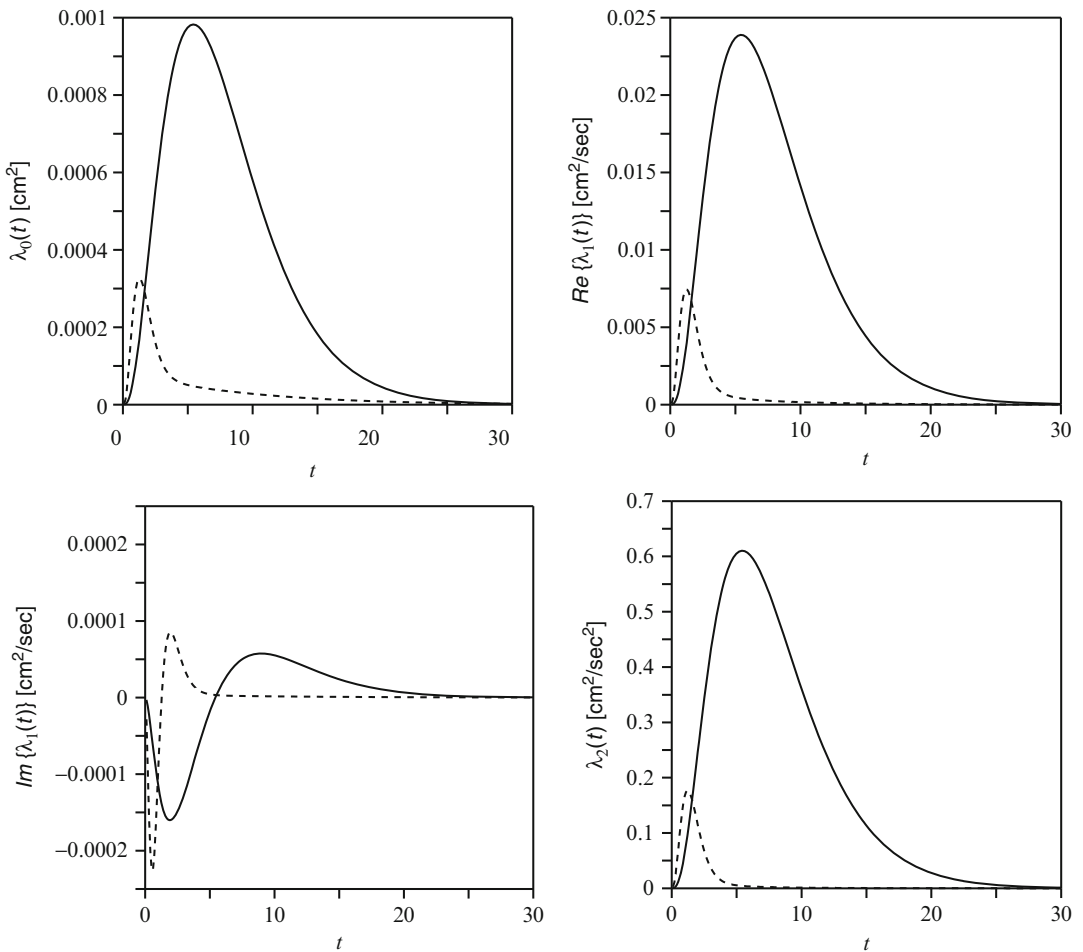
Stochastic Analysis of Linear Systems, Fig. 6 NGSMs of the response of the oscillator with $\omega_0 = 8\pi$ rad/sec, $\xi_0 = 0.05$ for the Hsu and Bernard (1978) model (*solid line*) and for the Spanos and Solomos (1983) model (*dashed line*)

Figure 7 shows that for the Hsu and Bernard (1978) modulating function, the mean frequencies, $v_{uu}^+(t)$, get to an asymptotic value very close to the natural frequency of the oscillators. This behavior is not verified for the oscillator with higher natural circular frequency subjected to the Spanos and Solomos (1983) model of the nonstationary input process, where the mean frequency decreases evidencing the frequency dependence of the structural response. Similar results are obtained for the normalized time-varying central frequencies, $\omega_{C,uu}(t)/\omega_0$, depicted in Fig. 8.

Summary

The dynamic behavior of structural systems subjected to uncertain dynamic excitations can be performed through the stochastic analysis, which requires the probabilistic characterization of both input and output processes. The characterization of output processes can be extremely complex, when nonstationary and/or non-Gaussian input processes are involved. However, in several cases the approximate description of the dynamic structural response based on its spectral characteristics may be sufficient.

S

Stochastic Analysis of Linear Systems, Fig. 7 Mean frequencies of the response of the oscillator for the Hsu and Bernard (1978) model (*solid line*) and for the Spanos and Solomos (1983) model (*dashed line*); (**a**) oscillator with $\omega_0 = 2\pi$ rad/sec and $\xi_0 = 0.05$; (**b**) oscillator with $\omega_0 = 8\pi$ rad/sec and $\xi_0 = 0.05$
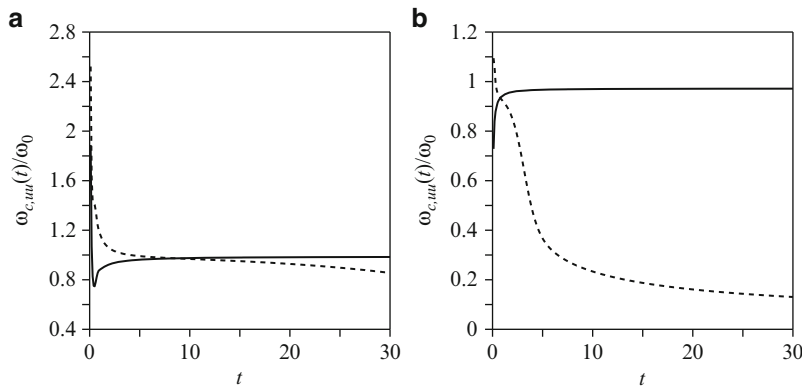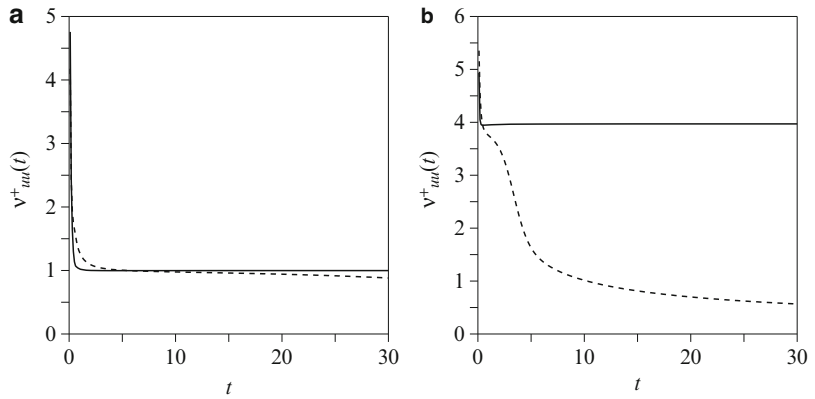
Stochastic Analysis of Linear Systems, Fig. 8 Time-varying central frequency, normalized by the natural circular frequency, of the oscillator for the Hsu and Bernard (1978) model (*solid line*) and for the Spanos and Solomos (1983) model (*dashed line*); (a) oscillator with $\omega_0 = 2\pi$ rad/sec and $\xi_0 = 0.05$; (**b**) oscillator with $\omega_0 = 8\pi$ rad/sec and $\xi_0 = 0.05$

In this study a unitary approach to evaluate the spectral characteristics of the structural response, to perform the reliability assessment, of classically damped linear systems subjected to stationary or nonstationary mono-/multi-correlated zero-mean Gaussian excitations, is described.

The main steps of the described approach are (i) the use of modal analysis to decouple the equation of motion; (ii) the determination, in state variable, of the evolutionary frequency response vector functions and of the evolutionary power spectral density function matrix of the structural response; and (iii) the evaluation of the nongeometric spectral moments as well as the spectral characteristics of the stochastic response of linear systems subjected to stationary or nonstationary mono-/multi-correlated zero-mean Gaussian seismic excitations.

Cross-References

▶ Probability Density Evolution Method in Stochastic Dynamics
▶ Stochastic Analysis of Nonlinear Systems
▶ Stochastic Ground Motion Simulation

References

Borino G, Muscolino G (1986) Mode-superposition methods in dynamic analysis of classically and non-classically damped linear systems. Earthq Eng Struct Dyn 14:705–717

Conte JP, Peng B-F (1997) Fully nonstationary analytical earthquake ground-motion model. J Eng Mech (ASCE) 123:15–24

Corotis RB, Vanmarcke EH, Cornell CA (1972) First passage of nonstationary random processes. J Eng Mech (ASCE) 98:401–414

Di Paola M (1985) Transient spectral moments of linear systems. SM Arch 10:225–243

Di Paola M, Petrucci G (1990) Spectral moments and pre-envelope covariances of nonseparable processes. J Appl Mech (ASME) 57:218–224

Fan FG, Ahmadi G (1990) Nonstationary Kanai-Tajimi models for El Cento 1940 and Mexico City 1985 earthquakes. Probab Eng Mech 5:171–181

Harichandran RS, Vanmarcke EH (1986) Stochastic variation of earthquake ground motion in space and time. J Eng Mech (ASCE) 112:154–174

Hsu T-I, Bernard MC (1978) A random process for earthquake simulation. Earthq Eng Struct Dyn 6:347–362

Jennings PC, Housner GW, Tsai C (1969) Simulated earthquake motions for design purpose. In: Proceedings of the 4th world conference earth engineering, Santiago, A-1, pp 145–160

Langley RS (1986) Structural response to non-stationary non-white stochastic ground motion. Earthq Eng Struct Dyn 14:909–924

Li J, Chen JB (2009) Stochastic dynamics of structures. Wiley, Singapore

Lin YK (1976) Probabilistic theory of structural dynamics. Krieger, Huntington

Lutes LD, Sarkani S (2004) Random vibrations – analysis of structural and mechanical vibrations. Elsevier, Boston

Michaelov G, Sarkani S, Lutes LD (1999a) Spectral characteristics of nonstationary random processes – a critical review. Struct Saf 21:223–244

Michaelov G, Sarkani S, Lutes LD (1999b) Spectral characteristics of nonstationary random processes – response of a simple oscillator. Struct Saf 21:245–267

Muscolino G (1991) Nonstationary pre-envelope covariances of nonclassically damped systems. J Sound Vib 149:107–123

Muscolino G, Palmeri A (2005) Maximum response statistics of MDOF linear structures excited by non-stationary random processes. Comput Method Appl Mech Eng 194:1711–1737

Priestley MB (1999) Spectral analysis and time series. Academic, London

Shinozuka M, Sato Y (1967) Simulation of nonstationary random process. J Eng Mech (ASCE) 93:11–40

Spanos P, Solomos GP (1983) Markov approximation to transient vibration. J Eng Mech (ASCE) 109:1134–1150

Vanmarcke EH (1972) Properties of spectral moments with applications to random vibrations. J Eng Mech (ASCE) 98:425–446

Vanmarcke EH (1975) On the distribution of the first-passage time for normal stationary random processes. J Appl Mech (ASME) 42:215–220

Zerva A (1991) Effect of spatial variability and propagation of seismic ground motions on the response of multiply supported structures. Probab Eng Mech 6:212–221

Stochastic Analysis of Nonlinear Systems

Annibale Luigi Materazzi and Marco Breccolotti
Department of Civil and Environmental Engineering, University of Perugia, Perugia, Italy

Synonyms

Earthquake engineering; Hysteresis; Nonlinearity; Random vibrations

Introduction

The present chapter is devoted to the probabilistic analysis of the random response of nonlinear structural systems exposed to random excitation with special attention to earthquake action. The random system response may be due to random excitation, to random system properties, to random boundary conditions. The nonlinear character of the response is mainly due to the nonlinear materials properties and to the effect of large displacements (the so-called P-Δ effect).

The first pioneering studies on this topic took place in the 1960s and 1970s (VanMarcke et al. 1970; Iwan 1973; Atalik and Utku 1976; Spanos 1976), when equivalent linearization techniques were used, taking advantage of the availability of first digital computers.

The attention is limited herein to systems having deterministic properties, including deterministic boundary conditions, considering those types of nonlinearity that can really occur during earthquakes.

After a brief review of the types of nonlinear behavior that can be expected, the available methods for the analysis of the response are discussed.

They can be classified into analytical and simulation methods. Among the first ones are considered in the following: the Fokker-Planck-Kolmogorov (FPK) equation, the equivalent linearization, the perturbation method, and the

S

stochastic averaging. The simulation methods based on Monte Carlo techniques are then presented.

Types of Nonlinearity in Earthquake Engineering

During severe earthquakes most structures undergo great amplitude time-varying displacements which can induce inelastic behavior in the structural members.

In the case of reinforced concrete structures, several phenomena take place, like cracking and crushing, together with yielding and strain hardening of steel. Also the bond between steel and concrete may be imperfect, as great shear forces must be exchanged and slippage can occur. Moreover, the reversal of the sign of displacements, due to the structural vibration, leads to the phenomenon of the hysteresis and to the subsequent dissipation of energy.

This apparently undesirable behavior is deliberately taken into account by the modern strategies for aseismic design of reinforced concrete structures, based on the concept of the capacity design, as it allows the reduction of the inertia forces acting upon the structures by means of energy dissipation.

The modeling of the interaction of concrete and steel under severe earthquakes was studied in depth in the past (Park et al. 1972, Takeda et al. 1970, Popov et al. 1972). Among the other aspects, the hysteretic behavior after steel yielding, the stiffness deterioration due to cracking of concrete, the strength deterioration and softening consequent to cumulative severe large deformations, and the pinching behavior due to shear, along with bond deterioration, were investigated.

Simple hysteretic models that can be used in the practice are presented in Fig. 1. Model (a) represents an elastic-perfectly plastic behavior, while model (b) is a variant allowing for strain hardening. Model (c) (Masing 1926) takes into account the stiffness deterioration, using smooth curves, and model (d) is the degrading stiffness model due to Takeda.

The Masing model is also suitable to represent the friction dissipation which takes place in masonry structures under earthquakes.

An interesting and powerful nonlinear model is represented by the Bouc-Wen equation (Bouc 1967; Wen 1976).

The restoring force F is given by

$$F = \alpha k_0 x + (1 - \alpha)QZ \qquad (1)$$

where k_0 is the initial system stiffness, α is the ratio between the post- and the pre-yielding stiffness, x is the displacement, Q is the yield strength, and Z is a nondimensional parameter which takes into account the hysteresis, satisfying the following nonlinear first-order differential equation:

$$q\frac{dZ}{dt} = A\dot{x} - \beta|\dot{x}|Z|Z|^{n-1} - \tau\dot{x}|Z|^n \qquad (2)$$

where $\beta, \tau, A,$ and n are the dimensionless parameters which control the shape of the hysteresis loop, q is the yield displacement, and $\dot{x}$ is the velocity.

Typical force-deformation hysteresis loops generated using Eqs. 1 and 2 are shown in Fig. 2. In the examples the dynamical system is acted upon by a sinusoidal motion having amplitude 0.1 m and frequency 1.0 Hz. The other parameters of the model are $q = 0.025$ m, $\beta = \tau = 0.5$, $k_0 = 1$, and $A = 1$. Thus, by changing the different parameters of the Wen's model, many hysteretic behaviors can be obtained such as the Masing type, the elastic-perfectly plastic, and the elastic plastic with hardening.

Analytical Techniques

The solution of the stochastic differential equations that govern the motion of nonlinear dynamical systems has been first achieved by means of analytical techniques. Exact and approximate solutions can be obtained by these techniques depending on the complexity of the problem. The stochastic response of nonlinear single degree of freedom system with elastic behavior and Gaussian external excitation can be generally
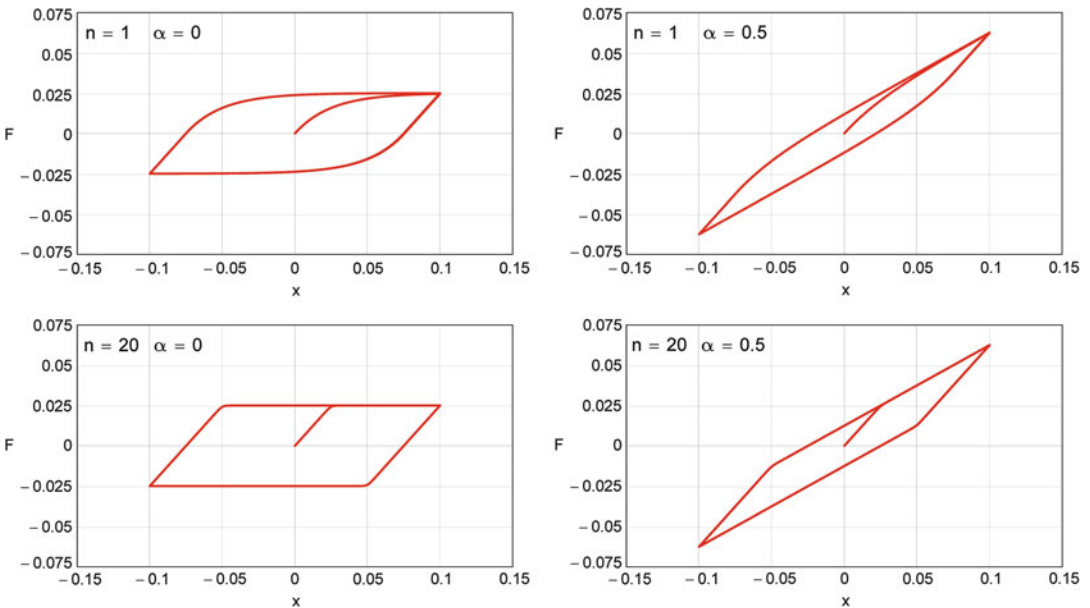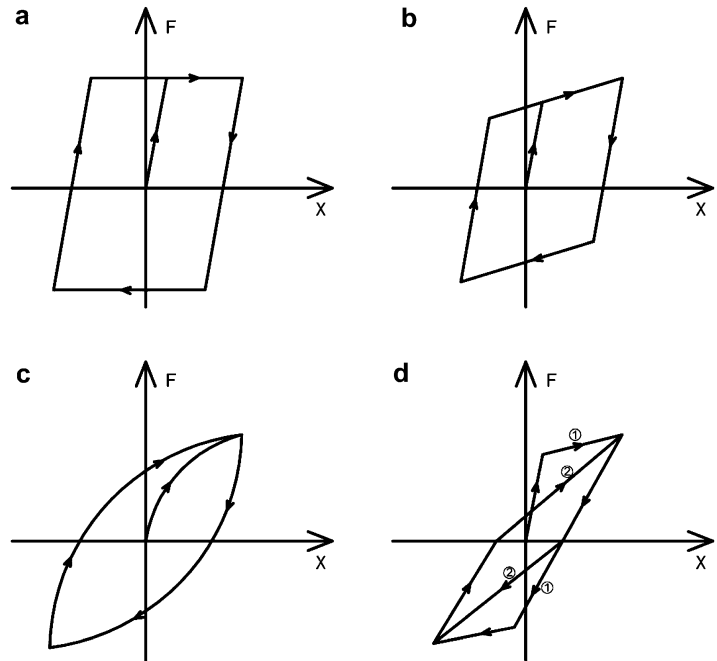
Stochastic Analysis of Nonlinear Systems, Fig. 2 Different hysteresis loops from the Bouc-Wen equation

solved exactly through the FPK equation.
Nonlinear elastic multi-degree of freedom can
be solved exactly via FPK equation only for
very restrictive conditions on the stochastic exter-
nal excitation. Thus, approximate solutions have
been searched for nonlinear multi-degree of free-
dom systems and for inelastic systems. In the next
paragraphs a brief description of the FPK equa-
tion and of some approximate methods is
provided.

Fokker-Planck-Kolmogorov Equation

Introduction

The theory of stochastic processes began in the nineteenth century when physicists were trying to show that heat in a medium is essentially a random motion of the constituent molecules. At the end of that century, some researches began to adopt more direct mathematical models of random disturbances instead of considering random motion as due to collisions between objects having a random distribution of initial positions and velocities. In this context several physicists, among which Fokker (1914) and Planck (1915), developed partial differential equations, which were versions of what was subsequently called the Fokker-Planck equation, to study the theory of Brownian motion.

The theory of the Fokker-Planck equation was made considerably more general by Kolmogorov (1931). He assumed the process to be continuous with respect to time and Markovian, i.e., a process for which the future probability density conditional on the present and past is actually independent of the past. On introducing further assumptions, Kolmogorov was able to show that the probability density of the process obeys a partial differential equation of the Fokker-Planck type. He also gave another partial differential equation which the future probability density, conditional on the present state, obeys with respect to the present state. The latter equation is called Kolmogorov's first equation, and the Fokker-Planck-type equation is called Kolmogorov's second equation. Sometimes the Fokker-Planck equation is called the Fokker-Planck-Kolmogorov (FPK) equation.

Solution of the Fokker-Planck-Kolmogorov Equation for an SDOF Elastic Nonlinear Second-Order System

Let us consider the nonlinear nonconservative stochastic system

$$m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + \frac{\partial G(x)}{dx} = \xi(t) \qquad (3)$$

that describes the displacement x of a body with mass m, damping c, and potential energy $G(x)$

subjected to a (scalar) white-noise forcing function $\xi(t)$ with second moment rate b.

By placing $y = m\frac{dx}{dt}$ the system may be rewritten in the state space as a system of first-order differential equations:

$$\begin{aligned} \frac{dx}{dt} &= \frac{y}{m} \\ \frac{dy}{dt} &= -\frac{\partial G(x)}{dx} - \frac{c}{m}y + \xi(t) \end{aligned} \qquad (4)$$

This is a special case of the system:

$$\begin{aligned} \frac{dx}{dt} &= \frac{\partial H}{\partial y} \\ \frac{dy}{dt} &= -\frac{\partial H}{\partial x} - c\frac{\partial H}{\partial y} + \xi(t) \end{aligned} \qquad (5)$$

where the symbol H represents the total energy (potential plus kinetic):

$$H = G(x) + \frac{1}{2m}y^2 \qquad (6)$$

The steady-state probability density $p(x,y)$ for the system of Eq. 5 can be determined by solving the appropriate stationary FPK equation:

$$\left[-\frac{\partial}{\partial x}\left(\frac{\partial H}{\partial y}p \right) + \frac{\partial}{\partial y}\left(\frac{\partial H}{\partial x}p \right) \right] + \frac{\partial}{\partial y}\left(c\frac{\partial H}{\partial y}p \right)$$
$$+ \frac{1}{2}b\frac{\partial^2 p}{\partial y^2} = 0$$
$$\qquad (7)$$

If p is any function of the energy H, the terms in the square brackets vanish, and therefore if it can be found a particular function $p(H)$ which makes the remaining two terms vanish, $p(H)$ is a solution of Eq. 7.

By substituting $p(x, y) = p(H(x, y))$ and assuming null the terms in the square bracket, Eq. 7 becomes

$$\frac{\partial}{\partial y}\left(c\frac{\partial H}{\partial y}p(H) \right) + \frac{1}{2}b\frac{\partial^2}{\partial y^2}p(H) = 0 \quad (8)$$

Stochastic Analysis of Nonlinear Systems, Fig. 3 Parameters describing a continuous piecewise-linear restoring force
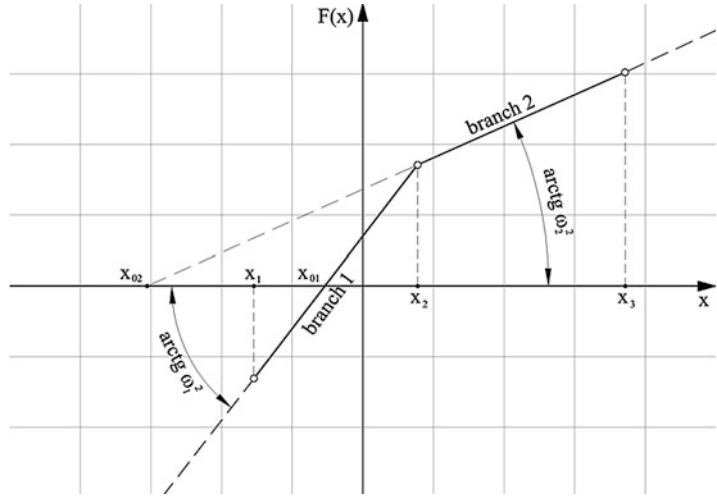
The integration with respect to y yields

$$c\frac{\partial H}{\partial y}p(H) + \frac{1}{2}b\frac{\partial}{\partial y}p(H) = L(x) \qquad (9)$$

where $L(x)$ is an arbitrary function.

As boundary conditions for the stationary FPK equation can be assumed, that p is zero at infinity. Thus, the left side of Eq. 9 vanishes when $y \to \infty$ and hence $L(x) = 0$. Equation 9 simplifies to

$$c\frac{\partial H}{\partial y}p(H) + \frac{1}{2}b\frac{\partial p}{\partial H}\frac{\partial H}{\partial y} = 0 \qquad (10)$$

Since $\frac{\partial H}{\partial y}$ cannot be identically null (otherwise H would be independent of y and hence so will be p), then Eq. 10 yields

$$\frac{\partial p}{\partial H} + \frac{2c}{b}p = 0 \qquad (11)$$

whose general solution representing the steady-state probability density is

$$p(x,y) = C\exp\left[-\frac{2c}{b}H(x,y)\right] \qquad (12)$$

being C a normalizing constant.

For the case of an SDOF system with unit mass and continuous linear piecewise restoring force, a closed form solution of the FPK equation can, thus, be determined. This model can be used, for instance, to determine the response of cracked reinforced concrete beams with cracks that alternatively open and close (Breccolotti et al. 2008).

In this case the motion of the vibrating system can be described by the equation

$$\ddot{x} + \beta_1\dot{x} + F_i(x) = f(t) \qquad (13)$$

where $F_i(x) = \omega_i^2(x - x_{0i})$ for $x_i \le x \le x_{i+1}$, $i = 1, 2, \ldots$ (see Fig. 3), ω_i and x_{0i} are positive constants, and $f(t)$ is a Gaussian stochastic process having null mean and uniform power spectrum S_0 over the entire frequency domain.

The following boundary conditions have to be satisfied:

$$w_i(x_{i+1},\dot{x},\ddot{x}) = w_{i+1}(x_{i+1},\dot{x},\ddot{x}) \qquad (14)$$

$$\sum_i \int_x \iint_{\dot{x}\,\ddot{x}} w_i(x,\dot{x},\ddot{x})\ dx\ d\dot{x}\ d\ddot{x} = 1 \qquad (15)$$

$$\omega_i^2(x_{i+1} - x_{0i}) = \omega_{i+1}^2(x_{i+1} - x_{0i+1}) \qquad (16)$$

that represent the continuity of the probability density function of the response, the normalizing condition, and the continuity of the piecewise-linear restoring force characteristics, respectively.

It can be demonstrated that the solution of the FPK equation corresponding to this system, representing the stationary probability density function (PDF) of the response, can be written in the form

$$
\begin{aligned}
w_i(x,\dot{x}) &= C_i \cdot e^{-\left(\frac{\omega_i^2}{S_0}\beta_1 x^2 + \frac{\beta_1}{S_0}\dot{x}^2 - \frac{2x_{0i}\omega_i^2}{S_0}\beta_1 x\right)} \\
&= C_i \cdot e^{-\frac{\omega_i^2}{S_0}\beta_1 x^2 + \frac{2x_{0i}\omega_i^2}{S_0}\beta_1 x} \cdot e^{-\frac{\beta_1}{S_0}\dot{x}^2} \\
&= w_{i,x}(x) \cdot w_{\dot{x}}(\dot{x})
\end{aligned}
\tag{17}
$$

It can be noted, as demonstrated by Lin and Cai (1995), that for an SDOF system having nonlinear stiffness, linear damping, and exposed to a Gaussian white-noise excitation, the stationary displacement (x) and the velocity $(\dot{x})$ are independent random variables.

Solution of the Fokker-Planck-Kolmogorov Equation for an nDOF Elastic Nonlinear Second-Order System

Several solutions of the FPK equation for nonlinear elastic systems with n degrees of freedom (nDOF) subjected to random excitation have been developed in the past (Piszczec and Nizioł 1986). Let us consider the following set of n equations for an nDOF system, characterized by elastic nonlinear properties:

$$
\ddot{x}_i + \beta_i \dot{x}_i + \frac{1}{m_i}\frac{\partial U}{\partial x_i} = f_i(t)
\tag{18}
$$

where $f_i(t)$ are independent and uncorrelated white noises with null mean and spectral density S_i and β_i are positive constants.

The term $U(x_1, x_2, \ldots)$ represents the potential energy.

This system of differential equations can be solved assuming the validity of the following conditions:

$$
\frac{S_i m_i}{2\beta_i} = k \quad i = 1, 2, \ldots
\tag{19}
$$

being k a positive constant. In this case the solution of the FPK equation is

$$
w(x_1, x_2, \ldots, \dot{x}_1, \dot{x}_2, \ldots)
$$
$$
= C\exp\left[-\frac{1}{k}\left(\frac{1}{2}\sum_{i=1}^{n} m_i \dot{x}_i^2 + U\right)\right]
\tag{20}
$$

This form of the probability density function is known as the Maxwell-Boltzmann distribution, being the expression inside the parentheses the total mechanical energy of the system.

Equivalent Linearization

The very restrictive conditions that make available exact solutions for the stochastic dynamical systems motivated the development of approximate solution techniques. Methods such as the equivalent linearization have been developed to generate first-order approximate solutions. These techniques can also be applied in some cases to single-degree-of-freedom nonlinear oscillator with hysteretic behavior.

The adaptation of the classical equivalent linearization technique of the deterministic theory to systems subjected to random excitations was independently developed by Booton (1953) and Caughey (1953). Later on other researchers have extended the method to encompass approximate solutions of the stationary random response of multi-degree-of-freedom nonlinear oscillators. In this case the mathematical equation that describes the response of the system is

$$
M\underline{\ddot{x}} + C\underline{\dot{x}} + K\underline{x} + \underline{f}\left(\underline{x},\underline{\dot{x}}\right) = \underline{G}
\tag{21}
$$

where M, C, and K are the constant mass, damping, and stiffness square matrices with dimension n, $\underline{f}\left(\underline{x},\underline{\dot{x}}\right)$ is an n-vector function of the dependent variable $\underline{x}$ and its derivative, and

$$
\underline{G}^T = (g_1, \ldots, g_n)
\tag{22}
$$

with n stationary Gaussian processes g_i.

The principle of the method is the replacement of the nonlinear dynamical system described by Eq. 21 by an auxiliary linear system for which the exact analytic solution is known. The replacement is made so as to be optimum with respect to some measure of the difference between the original and the auxiliary system.

Stochastic Analysis of Nonlinear Systems, Fig. 4 Bilinear hysteretic restoring force
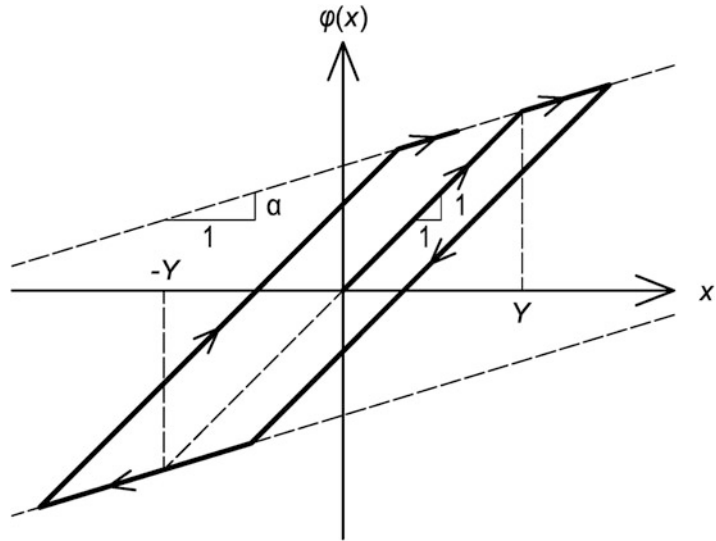
This optimum auxiliary linear system, which will be called *equivalent*, is defined as

$$M\underline{\ddot{x}} + (C + C_e)\underline{\dot{x}} + (K + K_e)\underline{x} = \underline{G} \qquad (23)$$

where C_e and K_e are t-independent matrices. These matrices must be such that the difference $\underline{d}$ between systems of Eqs. 21 and 23, defined by

$$\underline{d} = \underline{f}(\underline{x}, \underline{\dot{x}}) - C_e\underline{\dot{x}} - K_e\underline{x}, \qquad (24)$$

is minimized for every $\underline{x}$ which belongs to the class of solutions of the system (23).

Since the excitation $\underline{G}$ of the linear system is Gaussian, it is well known that the response $\underline{x}$ will be Gaussian as well. Therefore, the matrices C_e and K_e must be such that the difference $\underline{d}$ is minimized for every stationary Gaussian random vector $\underline{x}$.

The equivalent linearization technique is thus composed of the following steps:

1. Identification of a class of approximate solution functions and the parameters defining each member of the class.
2. Selection of the norm of the difference vector $\underline{d}$.
3. Selection of the averaging operator G.
4. Determination of the matrices C_e and K_e of the equivalent linear system in terms of the identification parameters of $\underline{x}$.

5. Solution of the equivalent linear system to obtain equations for the specification of the identification parameters of $\underline{x}$.

The application of the method can be conveniently shown by analyzing the hysteretic stochastic dynamical single DOF system described by the following differential equation (Iwan and Lutes 1968):

$$\ddot{x} + 2\beta_0\omega_0\dot{x} + \omega_0^2\varphi(x) = \frac{n(t)}{m} \qquad (25)$$

where m denotes the mass, ω_0 is the undamped natural circular frequency for small-amplitude vibrations, β_0 is the fraction of critical damping for small-amplitude vibrations, and $\varphi(x)$ is the bilinear hysteretic restoring force, shown in Fig. 4, having a unit slope for small amplitudes and a second slope α. The excitation $n(t)$ is a stationary random function with a uniform power spectral density S_0 and a Gaussian probability distribution.

The equivalent linear system

$$\ddot{x} + 2\beta_{eq}\omega_{eq}\dot{x} + \omega_{eq}^2 x = \frac{n(t)}{m} \qquad (26)$$

may be identified by choosing the parameters ω_{eq} and β_{eq} so as to minimize the mean-squared

difference between the nonlinear inelastic equation of motion (25) and the linear equation (26).

For the case of bilinear hysteretic oscillator with "small" nonlinearity, a solution for the parameters ω_{eq} and β_{eq} has been found by Caughey (1960) under the following assumptions:

1. The response of the nonlinear system is contained within a narrow frequency band.
2. The probability density of the amplitude of this narrowband response follows the Rayleigh distribution.

These assumptions lead to the following expressions for ω_{eq} and β_{eq}:

$$\left(\frac{\omega_{\text{eq}}}{\omega_0}\right)^2 = 1 - \left[\frac{8(1-\alpha)}{\pi}\right] \quad (27)$$
$$\int_1^\infty \left(2^{-3} + \frac{1}{\lambda z}\right)(z-1)^{\frac{1}{2}} e^{-\frac{z}{\lambda}} dz$$

and

$$\beta_{\text{eq}} = \beta_0\left(\frac{\omega_0}{\omega_{eq}}\right) + \left(\frac{\omega_0}{\omega_{eq}}\right)^2 \frac{(1-\alpha)}{\sqrt{\pi\lambda}} erfc\left(\lambda^{-\frac{1}{2}}\right) \quad (28)$$

where

$$\lambda = \frac{2\sigma_x^2}{Y^2}$$

In general Eq. 27 must be evaluated numerically. However, when $\lambda \gg 1$, the asymptotic expansion shown in Eq. 29 can be used:

$$\left(\frac{\omega_{\text{eq}}}{\omega_0}\right)^2 = \alpha + \left[\frac{8(1-\alpha)}{\pi}\right]$$
$$\left(0.6043\lambda^{-\frac{3}{4}} - 0.2451\lambda^{-\frac{5}{4}} - 0.1295\lambda^{-\frac{7}{4}}\right) \quad (29)$$

After finding ω_{eq} and β_{eq}, the RMS levels of response of the equivalent linear system can be obtained from

$$\sigma_{\dot{x}}^2 = \omega_{\text{eq}}^2 \sigma_x^2 = \frac{\pi S_0}{4m^2 \beta_{\text{eq}}\omega_{\text{eq}}} \quad (30)$$

Perturbation Method

The random response of slightly nonlinear vibrating systems can also be obtained by applying the classical perturbation method (Crandall 1963). The method is based on the assumption that the nonlinearity is small enough to allow the solution of the stochastic differential equation of motion to be expressed as a power series. If the following SDOF system is considered

$$\ddot{x} + 2\varsigma\omega_0\dot{x} + \omega_0^2 x + \varepsilon\eta(x,\dot{x}) - f(t) = 0 \quad (31)$$

where η is a nonlinear function and ε is a sufficiently small parameter, the solution can be expressed as a power series in ε:

$$x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \ldots \quad (32)$$

Substituting Eq. 32 into Eq. 31 and grouping the different terms having the same power of ε, a set of linear equation in $x_0, x_1, x_2, \ldots$ can be obtained. If the external excitation $f(t)$ is weakly stationary, the steady-state response to Eq. 31 can be constructed from the following solutions of the set of linear equations:

$$x_0(t) = \int_{-\infty}^\infty f(t-\tau)h(\tau)d\tau$$
$$x_1(t) = \int_{-\infty}^\infty \eta[x_0(\tau),\dot{x}_0(\tau)]h(\tau)d\tau \quad (33)$$
$$x_2(t) = \int_{-\infty}^\infty \eta[x_1(\tau),\dot{x}_1(\tau)]h(\tau)d\tau$$
$$\ldots$$

being h the impulse response function corresponding to $\varepsilon = 0$.

Stochastic Averaging

The method of stochastic averaging, firstly introduced by Landau and Stratonovich (1962), has proved to be a very useful tool for deriving approximate solutions to problems involving the dynamical response of lightly damped systems to broadband random excitation. It is based on the principle that the rate of change with respect to time of the oscillator's total energy is equal to the power input due to random excitation, minus the power dissipation due to the damping mechanism.

In lightly damped structures the slowly varying energy can be treated as a constant over an appropriate period of oscillation, and oscillatory terms can be approximated by their time averages over one period of oscillation. Furthermore, under broadband random excitation, the relaxation time of the oscillator response is much greater than the correlation time of the excitation. Thus, it is possible to model the power input due to the excitation as a nonzero mean component plus an additional, fluctuating component with the character of white noise.

Thanks to these properties, the method can also be applied to strongly nonlinear stiffness oscillators and to certain nonlinearities of the hysteretic kind.

The essence of the standard stochastic averaging method is embodied in a limit theorem due to Stratonovich (1963) and Khasminskii (1966):

$$\dot{X} = \varepsilon^2 f(X, t) + \varepsilon g(X, t, Y(t)) \qquad (34)$$

where $X(t)$ is an n-vector stochastic process, usually representing the response, and $Y(t)$ is an m-vector stochastic excitation process. If the elements of $Y(t)$ are broadband processes, with zero means, and the vectors f and g satisfy certain requirements (which are almost invariably met in practice), then it can be shown that $X(t)$ may be uniformly approximated over a time interval of order $O(\varepsilon^{-1})$ by an n-dimensional Markov process, which satisfies the Ito equation:

$$dX = \varepsilon^2 m(X)dt + \varepsilon \sigma(X)dW \qquad (35)$$

The symbol $W(t)$ denotes an n-vector of independent Wiener (or Brownian) processes, with unit variance, and m and σ are, respectively, the "drift vector" and the "diffusion matrix."

If ε is small, then the elements of $X(t)$ must be slowly varying, with respect to time. The equations of motion which occur in random vibration problems can be written in state space form, involving displacement and velocity response variables. However, these variables are usually rapidly fluctuating with respect to time. Evidently a transformation of variables is required to cast the equations of motion into the form of Eq. 34.

It has been shown that the stochastic averaging method is applicable to an oscillator with bilinear restoring force-displacement characteristic when the energy dissipation due to hysteresis is relatively low. In this case, after a preliminary averaging operation, the equation of motion could be cast into the standard form of Eq. 34, enabling the stationary response distribution to be determined together with statistics such as the standard deviation of the displacement and the average yielding rate.

Nevertheless, the standard stochastic averaging technique cannot be used for examining the effect of strongly nonlinear restoring forces since, to $O(\varepsilon^2)$, this effect vanishes. In these cases it is possible to combine the equivalent linearization method with stochastic averaging and treat the equivalent frequency as amplitude dependent.

An alternative approach to the analysis of nonlinear oscillators based on a consideration of the energy envelope has been developed by Roberts and Spanos (1986).

Consider an oscillator with the following equation of motion:

$$\ddot{x} + \varepsilon^2 h(x, \dot{x}) + G(x) = \varepsilon z(t) \qquad (36)$$

with ε assumed small enough to ensure that $x(t)$ is $O(\varepsilon^0)$. The symbol $G(x)$ denotes an arbitrary nonlinear stiffness term, while the energy envelope $V(t)$ may be defined as the sum of the kinetic and the potential energy:

$$V(t) = \frac{\dot{x}^2}{2} + U(x) \qquad (37)$$

and

$$U(x) = \int_0^x G(\xi)d\xi \qquad (38)$$

For the special case where the damping is linear $\varepsilon^2 h(x, \dot{x}) = 2\varsigma\omega_0\dot{x}$ and the nonlinear stiffness has the power-law form

$$G(x) = k|x|^v \text{sgn}(x), \qquad (39)$$

it is possible to obtain analytical solution for the transition density function $p(V, t|V_1, t_1)$ of $V(t)$. By introducing the nondimensional energy variable

$$\chi(t) = \frac{V(t)}{\gamma k \sigma^{\upsilon+1}}, \qquad (40)$$

the following transition density $p(\chi, t|\chi_1, t_1) = p(\chi, \tau|\chi_1)$ can be found:

$$p(\chi, \tau|\chi_1) = \frac{1}{(1-q)}\left(\frac{\chi}{\chi_1 q}\right)^{\frac{\rho}{2}}$$
$$\exp\left\{-\frac{\chi + q\chi_1}{1-q}\right\} I_p\left\{\frac{2\sqrt{\chi\chi_1 q}}{1-q}\right\} \qquad (41)$$

being

$$\rho = \frac{1}{\alpha} - 1 \qquad (42)$$

$$q = e^{-2\alpha\zeta\omega_0 t} \qquad (43)$$

and with $I_p(.)$ as the modified ρ – order Bessel function of the first kind.

Path Integral Solution

The FPK equation that describes the evolution of the response's probability density (PD) of a nonlinear system excited by an external white noise can be solved numerically by path integral (PI) solution procedures. In essence the PI method is a stepwise calculation of the joint probability density function of a set of state space variables describing a white-noise-excited nonlinear dynamic system. Among the first efforts to develop the PI method into numerical tools are those of Wehner and Wolfer (1983), Sun and Hsu (1990), and Naess and Johnsen (1993).

The PI method has been proved to provide extremely accurate results for the tail behavior of the joint probability density function of the state space vector and thus for the estimation of extreme responses of nonlinear dynamical systems excited by forces, external or parametric, that can be approximated as white noise, filtered

white-noise processes, and combined normal and Poisson white noise (Pirrotta and Santoro 2011).

The PI method is based on the fact that the state space vector, Y_t say, obtained as a solution of a stochastic differential equation is a Markov vector process. This makes it possible to use a time stepping procedure to obtain the joint probability density function $p(y, t)$ of Y_t as a function of time t by exploiting the fundamental equation:

$$p(y, t) = \int_{-\infty}^{\infty} p(y, t|y', t')p(y', t')dy' \qquad (44)$$

where $p(y, t|y', t')$ denotes the conditional probability density function of Y_t given that $Y_{t'} = y'$. For small time increments $\Delta t = t - t'$, $p(y, t|y', t')$ will be referred to as the short-time transition probability density function. It can be demonstrated that for a numerical solution of a stochastic differential equation, the short-time transition probability density function can always be given as an analytical, closed-form expression. Hence, if an initial probability density function, $p_0(y) = p(y, t = 0)$, is given, then Eq. 44 can be invoked repeatedly to produce the time evolution of $p(y, t)$. If the stochastic differential equation has an invariant measure there exists a stationary probability density function $p_s(y)$, then, eventually, assuming that $p_0(y) \neq p_s(y)$, $p(y, t)$ will approach this stationary probability density function. The number of times Eq. 44 has to be repeatedly used to reach the stationary situation depends, of course, on the dynamic system and on the specified initial probability density function $p_0(y)$.

Monte Carlo Simulation Methods

Generality

The probabilistic structure and the statistical moments of the response of any type of nonlinear mechanical systems can be evaluated using simulation techniques, like the Monte Carlo method. This method operates in the time domain by repeating a great number of times deterministic

analyses, each one consisting in a step-by-step nonlinear analysis of the structure subjected to an earthquake record.

In order to evaluate properly the statistics of the response, each earthquake record must belong to the same "family," in the sense that all the records must be compatible with the same power spectral density (PSD) function.

The major advantage of Monte Carlo simulation is that it can deal with almost any type of nonlinearity, maybe using the same commercially available packages that are used for the deterministic analyses. Its only disadvantage is that it is time-consuming, as many analyses are requested in order to have reliable estimates of the statistical properties of the response.

The number of needed runs may be many thousands, even if 300–400 analyses can give acceptable results in practical applications.

Probably the most important and delicate part of Monte Carlo simulation is the generation of realizations of the stochastic process or field that represents the earthquake.

In principle the earthquakes may be modeled as nonstationary non-Gaussian random fields, even if the null-mean Gaussian approximation is considered acceptable by most authors.

If the analyzed system is a building, it is reasonable, due to the reduced area of ground that it occupies, to model the process as one-dimensional. If the hypothesis that the earthquake induces ground acceleration in one direction only may be accepted, the process is univariate.

In the case of structures that have a notable extension over the ground, like long span bridges, the earthquake input is not the same in different part of the system and the process must be considered as multidimensional. If only one-direction component of the ground acceleration is considered, it is also univariate. Otherwise it is multivariate.

Univariate One-Dimensional Stochastic Processes

The basic method to generate a realization of a univariate, one-dimensional (1V-1D) stationary stochastic process $f_0(t)$ with zero mean and one-sided PSD function $G_{f_0}(\omega)$ was proposed by Shinozuka (Shinozuka 1972).

The process may be simulated by the series

$$f(t) = \sqrt{2} \sum_{n=0}^{N-1} A_n \cdot \cos(\omega_n t + \Phi_n) \qquad (45)$$

where

$$A_n = \sqrt{G_{f_0}(\omega_n)\Delta\omega}, \ n = 0, 1, 2, \ldots N-1 \qquad (46)$$

$$\omega_n = n\Delta\omega, \qquad (47)$$

$$\Delta\omega = \frac{\omega_u}{N} \qquad (48)$$

ω_u is the upper cutoff frequency and the phase angle Φ_n is the realization of a random variable uniformly distributed over the interval $[0, 2\pi]$.

Shinozuka and Deodatis (1991) showed that it must be

$$A_0 = 0 \text{ and } S_{f_0}(\omega_0 = 0) = 0 \qquad (49)$$

to ensure that the realization of the process is ergodic.

Univariate Multidimensional Stochastic Processes

For the sake of simplicity, let us consider the special case of a univariate two-dimensional stochastic process (1V-2D), which may represent the random field of the one-direction ground acceleration interesting a great-extension structure frozen in time.

The process $f_0(x_1, x_2)$, having cross-PSD $G_{f_0}(k_{1n_1}, k_{2n_2})$, can be simulated by the series

$$f(x_1, x_2) = \sqrt{2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \left[A_{n_1 n_2} \cos\left(k_{1n_1} x_1 + k_{2n_2} x_2 + \Phi^{(1)}_{n_1 n_2}\right) + \tilde{A}_{n_1 n_2} \cos\left(k_{1n_1} x_1 - k_{2n_2} x_2 + \Phi^{(2)}_{n_1 n_2}\right) \right]$$

$$(50)$$

where

$$G_{f_0}(k_1, k_2), \ \tilde{A}_{n_1 n_2} = \sqrt{G_{f_0}(k_{1 n_1}, -k_{2 n_2}) \Delta k_1 \Delta k_2} \tag{51}$$

$$k_{1 n_1} = n_1 \Delta k_1, \ k_{2 n_2} = n_2 \Delta k_2, \tag{52}$$

$$\Delta k_1 = \frac{k_{1u}}{N_1}, \ \Delta k_2 = \frac{k_{2u}}{N_2} \tag{53}$$

$$n_1 = 0, 1, 2, \ldots N_1 - 1, \ n_2 = 0, 1, 2, \ldots N_2 - 1 \tag{54}$$

and

$$A_{0 n_2} = A_{n_1 0} \text{ for } n_1 = 0, 1, 2, \ldots N_1 - 1 \text{ and} \\ n_2 = 0, 1, 2, \ldots N_2 - 1 \tag{55}$$

$$\tilde{A}_{0 n_2} = \tilde{A}_{n_1 0} \text{ for } n_1 = 0, 1, 2, \ldots N_1 - 1 \text{ and} \\ n_2 = 0, 1, 2, \ldots N_2 - 1 \tag{56}$$

k_{1u} and k_{2u} are the upper cutoff wave numbers and the phase angles $\Phi_{n_1 n_2}^{(1)}$ and $\Phi_{n_1 n_2}^{(2)}$ are the realization of two statistically independent random variables uniformly distributed over the interval $[0, 2\pi]$.

The extension to the case of 1D-3V process $f_0(x_1, x_2, t)$, including also the time variability, is straightforward.

Earthquake Power Spectra

A well-known PSD of the ground motion available in literature is the so-called Kanai-Tajimi spectrum. It is based on the hypothesis that the ground acceleration during earthquakes may be considered as a filtered band-limited white-noise process expressed by the following function:

$$G(\omega) = \frac{\left[1 + 4\xi_g^2 \cdot (\omega/\omega_g)^2\right]}{\left[1 - (\omega/\omega_g)^2\right]^2 + 4\xi_g^2 \cdot (\omega/\omega_g)^2} G_0 \tag{57}$$
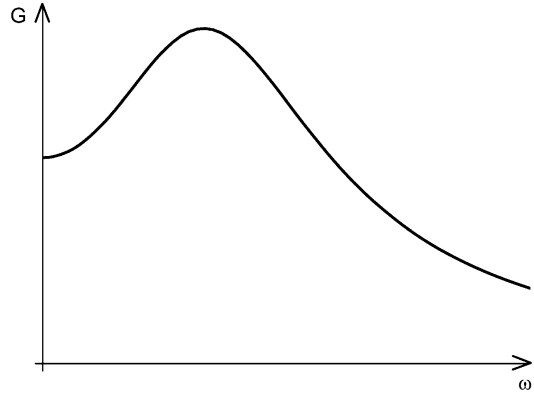
Stochastic Analysis of Nonlinear Systems, Fig. 5 Kanai-Tajimi power spectrum

where G_0 is the constant value of the spectrum of the white-noise process and ω_g and ξ_g are the predominant ground frequency and the ground damping.

The values of $\omega_g = 4\pi$ and $\xi_g = 0.60$ have been suggested as representative of earthquakes on firm ground.

The typical aspect of the Kanai-Tajimi PSD is shown in Fig. 5.

Anyway in most cases it is better to derive the PSD from the response spectrum.

For this purpose let us consider the following SDOF deterministic dynamical system:

$$\ddot{x}(t) + 2\xi\omega\dot{x}(t) + \omega x(t) = f_0(t) \tag{58}$$

acted upon by a 1V-1D stationary stochastic process $f_0(t)$ which represents the ground acceleration. If the process is Gaussian, its probabilistic structure is fully represented by its PSD G_{f_0}.

The random dynamic analysis of the system in the frequency domain allows finding the PSD of the response $x(t)$, which is the system displacement:

$$G_x(\omega) = G_{f_0}(\omega)|H(\omega)|^2 \tag{59}$$

The variance of the response is represented by the area under the PSD function:

$$\sigma_x^2 = \int\limits_0^\infty G_{f_0}(\omega)|H(\omega)|^2 d\omega \tag{60}$$

where $|H(\omega)|^2$ is the squared frequency-domain transfer function between loading and system response:

$$|H(\omega)|^2 = \frac{1}{\left(\omega_n^2 - \omega^2\right)^2 + 4\xi^2\omega_n^2\omega^2} \qquad (61)$$

with ω_n the natural frequency of the system.

The variance of the pseudo-acceleration response is

$$\sigma_{\ddot{x}}^2 = \omega_n^4\sigma_x^2 = \omega_n^4\int\limits_0^\infty G_{f_0}(\omega)|H(\omega)|^2 d\omega$$

$$\approx G_{f_0}(\omega_n)\omega_n\left(\frac{\pi}{4\xi} - 1\right) \qquad (62)$$

The ordinates $\ddot{x}_{s;p}$ of the acceleration response spectrum are the maxima of the response $\ddot{x}(t)$ of an oscillator whose frequency ω varies from 0 to some stipulated maximum value. The parameter "s" is the duration of the strong motion and "p" is the non-exceedance probability of the maxima.

The solution of the equation

$$\ddot{x}_{s;p} = r_{s;p}\sigma_{\ddot{x}} \qquad (63)$$

implies the evaluation of the peak factor $r_{s;p}$ by solving the corresponding "first-passage" problem.

First of all, it must be reminded that the moment of order i of the generic one-sided PSD $S(\omega)$ is

$$\lambda_i = \int\limits_0^\infty \omega^i S(\omega) d\omega \qquad (64)$$

while a measure where the spectral mass is concentrated along the frequency axis is

$$\Omega = \sqrt{\lambda_2/\lambda_0} \qquad (65)$$

parameter that resembles the root mean square of a random variable.

A measure of the spread of the PSD function is

$$\delta = \sqrt{1 - \lambda_1^2/\lambda_0\lambda_2} \qquad (66)$$

Given the above, an approximate expression for the peak factor $r_{s;p}$ is given by

$$r_{s;p} = \sqrt{2\log\left\{2n\left[1 - \exp\left(-\delta_{\ddot{x}}(s)\sqrt{\pi\log 2n}\right)\right]\right\}} \qquad (67)$$

where

$$n = \frac{\Omega_{\ddot{x}}(s)s_0/2\pi}{-\log p} \qquad (68)$$

$$s_0 = s \cdot \exp\left[-2\left\{\sigma_{\ddot{x}}^2(s)/\sigma_{\ddot{x}}^2(s/2) - 1\right\}\right] \qquad (69)$$

$\Omega_{\ddot{x}}(s)$ and $\delta_{\ddot{x}}(s)$ are the spectral moments of the response.

Then the PSD function is

$$G_{f_0}(\omega_n) \approx \frac{1}{\omega_n\left(\dfrac{\pi}{4\xi_s} - 1\right)} \frac{1}{\sqrt{\dfrac{\omega_n^2 S_v(\omega_n)_{s,p}^2}{r_{s,p}^2} - \int\limits_0^{\omega_n} G_{f_0}(\omega) d\omega}} \qquad (70)$$

where ξ_s is the time-dependent damping for the duration s

and S_v is the velocity response spectrum.

As the PSD $G_{f_0}(\omega)$ appears on both sides of Eq. 65, an iterative computational procedure must be used.

Summary

The present chapter is devoted to the probabilistic analysis of the random response of nonlinear structural systems exposed to the random excitation due to earthquakes. The system response is nonlinear due to the nonlinear character of the

material properties and to the effects of large structural displacements.

The attention is focused on systems having deterministic properties, including deterministic boundary conditions, taking into account those types of nonlinearity that can really occur during earthquakes.

After a brief review of the types of nonlinear behaviors that can be expected, some available methods for the analysis of the response are discussed.

They can be classified into analytical and simulation methods. Among the first ones are presented: the Fokker-Planck-Kolmogorov (FPK) equation, the equivalent linearization, the perturbation method, and the stochastic averaging. The simulation methods based on Monte Carlo techniques are then presented.

Cross-References

▶ Plastic Hinge and Plastic Zone Seismic Analysis of Frames
▶ Reinforced Concrete Structures in Earthquake-Resistant Construction
▶ Stochastic Analysis of Linear Systems

References

Atalik T, Utku S (1976) Stochastic linearization of multi-degree of freedom nonlinear systems. Earthquake Eng Struct Dyn 4(4):411–420

Booton R (1953) The analysis of nonlinear control systems with random inputs. Polytechnic Inst, Brooklyn

Bouc R (1967) Forced vibrations of mechanical systems with hysteresis. Proceedings of the Fourth Conference on Nonlinear Oscillation, Prague

Breccolotti M, Materazzi A, Venanzi I (2008) Identification of the nonlinear behavior of a cracked RC beam through the statistical analysis of the dynamic response. Struct Control Health Monitor 15:416–435

Caughey T (1953) Response of nonlinear systems to random excitation. Lecture Notes California Institute of Technology, Pasadena, Calif

Caughey T (1960) Random excitation of a system with bilinear hysteresis. J Appl Mech 27:649–652

Crandall S (1963) Perturbation techniques for random vibration of nonlinear systems. J Acoust Soc Am 35(11):1700–1705

Fokker A (1914) The median energy of rotating electrical dipoles in radiation fields. Annalen Der Physik 43:810–820

Iwan WD (1973) A generalization of the concept of equivalent linearization. Int J Nonlinear Mech 8:279–287

Iwan WD, Lutes LD (1968) Response of the bilinear hysteretic system to stationary random excitation. J Acoust Soc Am 48(3):545–552

Khasminskii R (1966) A limit theorem for the solution of differential equations with random right-hand sides. Theory Probab Appl 11:390–405

Kolmogorov A (1931) Ober analytische Methoden in der Wahrscheinlichkeitsrechnung. Math Ann 104:415–458

Landau P, Stratonovich R (1962) Theory of stochastic transitions of various systems between different states. Moscow University, Vestinik

Lin Y, Cai G (1995) Probabilistic structural dynamics advanced theory and applications. McGraw-Hill, New York

Masing G (1926) Eigenspannungen und Verfestigung beim Messing (in German). In Proceedings of the 2nd international congress of applied mechanics, Zurich, pp 332–335

Naess A, Johnsen J (1993) Response statistics of nonlinear, compliant offshore structures by the path integral solution method. Probab Eng Mech 8(2):91–106

Park R, Kent D, Sampson R (1972) Reinforced concrete members with cyclic loading. ASCE J Struct Div 98(7):1341–1360

Pirrotta A, Santoro R (2011) Probabilistic response of nonlinear systems under combined normal and Poisson white noise via path integral method. Prob Eng Mech 26(1):26–32

Piszczec K, Niziol J (1986) Random vibration of mechanical system. Ellis Horwood, Chichester

Planck M (1915) Sgr. preuss. Akad. Wiss. p 512

Popov E, Bertero V, Krawinkler H (1972) Cyclic behavior of three r.c. flexural members with high shear. In: EERC report 72-5. Earthquake Engineering Research Center, University of California, Berkeley

Roberts J, Spanos P (1986) Stochastic averaging: an approximate method of solving random vibration problems. Int J Nonlinear Mech 21(2):111–134

Shinozuka M (1972) Monte Carlo solution of structural dynamics. Comput Struct 2(5/6):855–874

Shinozuka M, Deodatis G (1991) Simulation of stochastic processes by spectral representation. Appl Mech Rev 44(4):191–204

Spanos P-TD (1976) Linearization techniques for nonlinear dynamical systems, EERL 76-04. California Institute of Technology, Pasadena, Calif

Stratonovich R (1963) Topics in the theory of random noise. Gordon & Breach, New York

Sun J, Hsu C (1990) The generalized cell mapping method in nonlinear random vibration based upon short-time Gaussian approximation. J Appl Mech 57:1018–1025

Takeda T, Sozen M, Nielson N (1970) Reinforced concrete response to simulated earthquakes. Proc ASCE J Struct Div 96(ST12):2257–2573

VanMarcke E, Yanev P, De Estrada M (1970) Response of simple hysteretic systems to random excitation, s.l., Research report R70-66, Department of Civil Engineering, MIT

Wehner M, Wolfer W (1983) Numerical evaluation of path-integral solutions to Fokker–Planck equations. Phys Rev A 27(5):2663–2670

Wen Y (1976) Method for random vibration of hysteretic systems. ASCE J Eng Mech 120:2299–2325

Stochastic Finite Elements

Carsten Proppe
Institut für Technische Mechanik, Karlsruhe
Institute of Technology, Karlsruhe, Germany

Introduction

The finite element method (FEM) has become the dominant computational method in structural engineering. In general, the input parameters in the standard FEM assume deterministic values. In earthquake engineering, at least the excitation is often random. However, considerable uncertainties might be involved not only in the excitation of a structure but also in its material and geometric properties. A rational treatment of these uncertainties needs a mathematical concept similar to that underlying the standard FEM. Thus, FEM as a numerical method for solving boundary value problems has to be extended to stochastic boundary value problems. The extension of the FEM to stochastic boundary value problems is called stochastic finite element method (SFEM).

The first developments of the SFEM can be traced back at least to Cornell (1970), who studied soil settlement problems, and to Shinozuka (Astill et al. 1972), who combined FEM with Monte Carlo simulation for reliability analysis of structures with random excitation, random material properties, or random geometric properties. He introduced random fields and discretized them based on spectral representation theory (Shinozuka 1971) of wide-sense homogeneous random fields.

The expression SFEM came in use in the early 1980s (Contreras 1980; Baecher and Ingra 1981). Der Kiureghian and Ke (1988) defined SFEM as "a finite element method which accounts for uncertainties in the geometry or material properties of a structure, as well as the applied loads" where "the uncertainties are usually spatially distributed over the region of the structure and should be modelled as random or stochastic fields." The distinguishing feature of an SFEM is that it involves the discretization of the random field and the computation of solution statistics.

Discretization methods for random fields replace the random field by a finite set of random variables. They can be broadly classified into two groups: direct discretization schemes and series expansion techniques. Direct discretization schemes either assign the values of the random field at a given set of nodes to the finite set of random variables (point discretization schemes) or compute the values from local averages of the random field over a spatial domain (local averaging schemes).

Possible sets of nodes for point discretization schemes are the nodes of the finite element mesh (Hisada and Nakagiri 1981; Liu et al. 1986), the centroids of the finite elements (Der Kiureghian and Ke 1988), or the Gauss points used by the integration rules of the finite elements (Brenner and Bucher 1995). If the set of nodes for the point discretization scheme of the random field corresponds to the set of nodes of the finite elements, the shape functions for the representation of the random field may differ from those for the displacements (Liu et al. 1986). Point discretization schemes yield a positive definite correlation matrix that is easily computed. The distribution function of the random variables is the same as for the random field. Nearly all schemes can be applied to Gaussian as well as non-Gaussian random fields. The point discretization schemes that do not use different shape functions have the disadvantage that the FE mesh depends on the correlation structure of the random field and that the shape and size of all these elements should be the same. In general, the discretization produces a

huge number of random variables and leads to inefficient numerical procedures.

In contrast, local averaging schemes may yield accurate results even on rather coarse FE meshes. They can be based directly on averages of the random parameter field (Vanmarcke and Grigoriu 1984) or on integrals computed over the domain of a finite element, e.g., for the determination of the element stiffness matrix (Deodatis 1991). However, direct averaging yields random variables whose distribution functions are difficult to obtain (except for Gaussian random fields) and the approximation of non-rectangular elements may lead to a nonpositive definite covariance matrix (Matthies et al. 1997), and averaging based on element integration introduces again a dependence of the FE mesh on the correlation structure.

Series expansion techniques represent the random field by a series involving countably infinite random variables and a complete set of deterministic functions. The discretization is obtained by truncation of the series expansion. Lawrence (Lawrence 1987) considers a series expansion with an a priori set of orthogonal deterministic base functions that are multiplied by random variables. In the Karhunen-Loève (KL) expansion, the set of orthogonal base functions is obtained as eigenfunctions of a homogeneous Fredholm integral equation of second kind that involves the covariance kernel of the random field. The KL expansion is optimal in the sense that it reduces the mean-square error resulting of its truncation. In principle, it can be applied to homogeneous as well as inhomogeneous random fields and to Gaussian as well as non-Gaussian random fields. However, in practical cases, the KL expansion is applied to Gaussian random fields almost exclusively because the random variables in the expansion are then independent and standard normally distributed. In general, the eigenfunctions have to be computed numerically and the accuracy of the numerical solution influences the accuracy of the random field representation. For strongly correlated random fields with smooth covariance kernel, few terms of the series expansion are sufficient to represent the random field.

However, according to Stefanou and Papadrakakis (2007), homogeneity and ergodicity of sample functions generated by the KL expansion are questionable. The spectral representation method approximates a homogeneous Gaussian random field by a superposition of harmonics with fixed frequencies and random amplitudes or random frequencies and amplitudes. Due to the central limit theorem, the generated samples are Gaussian, if a sufficient number of harmonics (at least 128, according to Stefanou (2009)) are superposed.

In some cases, non-Gaussian random fields can be represented by nonlinear memoryless transformations of Gaussian random fields (Yamazaki and Shinozuka 1988). The memoryless transformation introduces a compatibility condition between the marginal distribution function and the autocorrelation function of the non-Gaussian random field. If this condition is not satisfied, only an approximation of the non-Gaussian random field by a nonlinearly transformed Gaussian random field is possible. Iterative procedures have been proposed (Deodatis and Micaletti 2001; Yamazaki and Shinozuka 1988) that calibrate the power spectral density of the Gaussian random field in order to approximately match the marginal distribution function and the autocorrelation function of the non-Gaussian random field. The underlying Gaussian random field can then be discretized by any of the methods described above.

A series expansion of a non-Gaussian random field in terms of independent random variables is obtained by projection on a set of orthogonal polynomials. This method is known as polynomial chaos expansion and is discussed extensively in Ghanem and Spanos (1991), where Hermite polynomials and Gaussian random variables are considered. In Xiu and Karniadakis (2002), other families of orthogonal polynomials and corresponding distributions of the random variables are discussed. The polynomial chaos expansion is convergent in mean-square sense. Field and Grigoriu (2004) critically discuss polynomial chaos expansions and show that the accuracy of the PC approximation is not always improved as additional terms are retained and

that the polynomial chaos expansion might be computationally demanding due to the large number of expansion coefficients that have to be computed.

Once the random field involved in the stochastic boundary value problem has been discretized, a solution method has to be adopted in order to solve the boundary value problem numerically. The choice of the solution method depends on the required statistical information of the solution. If only the first two statistical moments of the solution are of interest (*second moment analysis*), the perturbation method can be applied. However, if a *full probabilistic analysis* is necessary, Galerkin schemes can be utilized or one has to resort to Monte Carlo simulations eventually in combination with a von Neumann series expansion.

The perturbation method starts with a Taylor series expansion of the solution, the external loading, and the stochastic stiffness matrix in terms of the random variables introduced by the discretization of the random parameter field. The unknown coefficients in the expansion of the solution are obtained by equating terms of equal order in the expansion. From this, approximations of the first two statistical moments can be obtained. The perturbation method is computationally more efficient than direct Monte Carlo simulation. However, higher-order approximations will increase the computational effort dramatically, and therefore accurate results are obtained for small coefficients of variation only.

In the spectral SFEM, the random parameter fields are discretized by a KL or a polynomial chaos expansion, the solution is expanded with Hermite polynomials, and a Galerkin approach is applied to solve for the unknown expansion coefficients. The theoretical foundation has been laid in Deb et al. (2001) and Babuška et al. (2005), where local and global polynomial chaos expansions for linear elliptic boundary value problems with stochastic coefficients were investigated and where a priori error estimates have been proved for a fixed number of terms of the KL expansion.

Instead of a Galerkin projection, several authors employed collocation schemes for the determination of the unknown coefficients in the approximation scheme (Acharjee and Zabaras 2007; Babuška et al. 2007; Baroth et al. 2007; Huang et al. 2007). This leads to nonintrusive algorithms that allow to combine the solution procedure with repetitive runs of a finite element (FE) solver for deterministic problems. A nonintrusive algorithm based on least squares regression has been presented recently in Berveiller et al. (2006).

In the following, discretization methods for the random parameter field are illustrated, and a mathematical theory for the approximate solution of stochastic elliptic boundary value problems involving a discretized random parameter field that is represented as a superposition of independent random variables is outlined. In the random domain, global and local polynomial chaos expansions are employed. The relation between local approximations of the solution and Monte Carlo simulation is considered, and reliability assessment is briefly discussed. Finally, an example serves to illustrate the different solution procedures.

Discretization of Random Fields

Let D be a convex bounded open set in $\mathbb{R}^n$ and $(\Omega, \mathcal{F}, P)$ be a complete probability space, where Ω is the set of outcomes, $\mathcal{F}$ the σ-field of events, and $P : \mathcal{F} \to [0 : 1]$ a probability measure. A function $\alpha : D \times \Omega \to \mathbb{R}^m$ is a *random field*, if $\alpha(x, \omega)$ is a random variable for any $x \in D$. In the following, we consider scalar valued random fields ($m = 1$).

The finite dimensional distribution of order q of α at $x_1, x_2, \ldots, xq \in D$ is the probability of the set $\cap_{i=1}^q \{\alpha(x_i, \omega) \leq \alpha_i\}$. The random field is *homogeneous* (in strict sense), if the finite dimensional distributions are invariant under a space shift and thus depend only on the space lag.

Suppose that the random field α is square integrable on $D \times \Omega$ and denote by $E[X] = \int_\Omega x \mathrm{d}P$ the expectation of the random variable $X(\omega)$. Then, the mean, correlation, and covariance are given by $E[\alpha(x, \omega)]$, $R(x, y) = E$

$[\alpha(x, \omega)\alpha(y, \omega)]$, and $C(x, y) = E[(\alpha(x, \omega) - E[\alpha(x, \omega)])(\alpha(y, \omega) - E[\alpha(y, \omega)])$, respectively. A random field is *homogeneous* (in weak sense), if the mean is constant and the correlation depends only on the space lag $y - x$.

Bochner's theorem allows to introduce the spectral distribution $S(v)$ of a weakly homogeneous random field with continuous correlation function by

$$R(x) = \int_{\mathbb{R}^n} \exp(ix \cdot v) dS(v). \qquad (1)$$

If S is absolutely continuous, the Radon-Nikodym derivative $s(v)$ is called *spectral density function*: $dS(v) = s(v)\, dv$.

Spectral Representation
If α is a weakly homogeneous mean-square continuous random field, it can be represented by

$$\alpha(x, \omega) = \int_{\mathbb{R}^n} \exp(ix \cdot v) dW(v), \qquad (2)$$

where the random field $W(v)$ has mean zero and satisfies $E[dW(v)\, dW(\mu^*)] = \delta(v - \mu)\, dS(v)$. The asterisk denotes complex conjugation and δ is the Dirac δ-distribution.

This distribution can be used to approximately represent a homogeneous Gaussian random field by a superposition of harmonics. For a real-valued random field, one obtains the representation

$$\alpha(x, \omega) = \int_{\mathbb{R}^n} (\cos(v \cdot x) dU(v) + \sin(v \cdot x) dV(v)), \qquad (3)$$

where $E[dU(v)^2] = E[dV(v)^2] = dS(v)$. Starting from a partition of the wave number domain, the increments $dU(v)$ and $dV(v)$ are approximated by

$$\begin{aligned} \Delta U(v_i) &= \sqrt{2\Delta S} \cos \phi_i; \\ \Delta V(v_i) &= \sqrt{2\Delta S} \sin \phi_i; \\ \Delta S &= S(v_{i+1}) - S(v_i) \end{aligned} \qquad (4)$$

where ϕ_i are random variables uniformly distributed on $[0, 2\pi]$. A representation involving a finite number of random variables is thus

$$\sum_{i=1}^{M} \sqrt{2\Delta S} \cos(v_i \cdot x + \phi_i). \qquad (5)$$

As a consequence of the central limit theorem, it converges for $M \to \infty$ to a Gaussian random field with the same mean value and autocorrelation structure as the target Gaussian random field.

Karhunen-Loève Expansion
Due to the properties of the covariance function, the operator $T : L^2(D) \to L^2(D)$,

$$T_u = \int_D C(x, y) u(x) dx, \qquad (6)$$

is compact and self-adjoint and thus admits a spectrum of decreasing nonnegative eigenvalues $\{\lambda_i\}_{i=1}^{\infty}$. The corresponding eigenfunctions $\{\phi_i(x)\}_{i=1}^{\infty}$ are orthonormal in $L^2(D)$. The random variables given by

$$\xi_i(\omega) = \frac{1}{\sqrt{\lambda_i}} \int_D (\alpha(x, \omega) - E[\alpha](x)) \phi_i(x) dx \quad (7)$$

are uncorrelated (but in general not independent), have zero mean and unit variance, and allow to represent the random field by the KL expansion

$$\alpha(x, \omega) = E[\alpha](x) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i(\omega) \phi_i(x) \quad (8)$$

that converges in $L^2(D \times \Omega)$ (Loève 1977). Conditions for stronger convergence properties are given in Babuška et al. (2005). The KL expansion is usually truncated by retaining only the first M terms. In order to keep the computational effort small, a fast decay of the spectrum of Eq. 6 is important. It is shown in Todor and Schwab (2006) that fast eigenvalue decay corresponds to smoothness of the covariance function. Moreover, for a decreasing correlation length, the number M of retained terms increases if the accuracy of

the approximation is kept constant. The KL expansion reduces to the spectral representation method for homogeneous random fields defined over an infinite domain (Huang et al. 2001).

For a prescribed, uniformly bounded random field $\alpha(x, \omega)$, the random variables $\xi_i(\omega)$ in Eq. 8 would be dependent non-Gaussian random variables whose joint distribution function is very difficult to identify. If, on the other hand, independent but bounded distributions are prescribed for $\xi_i(\omega)$, the random field $\alpha(x, \omega)$ is not necessarily bounded for $N \rightarrow \infty$. Thus, one is left with Gaussian distributions for $\xi_i(\omega)$ and $\alpha(x, \omega)$, with transformations of Gaussian random fields or with some situations, where nonnegative distributions for $\xi_i(\omega)$ lead to meaningful (e.g., Erlang) distributions for $\alpha(x, \omega)$.

The KL expansion is optimal in the sense that the error measured in $L^2(D \times \Omega)$ resulting from a truncation after M terms is smaller than for any other linear combination of M functions. It is a representation for homogeneous as well as nonhomogeneous random fields. However,

several authors (Grigoriu 2006; Stefanou and Papadrakakis 2007; Sudret and Der Kiureghian 2000) observed problems regarding the homogeneity of samples generated from the truncated expansion.

Kriging

Application of Kriging to the discretization of random fields has been introduced by Li and Der Kiureghian (1993). The random field $\alpha(x, \omega)$ is approximated by a linear function of M nodal values $\alpha_i(\omega) = \alpha(x_i, \omega)$

$$\tilde{\alpha}(x, \omega) = \phi_0(x) + \sum_{i=1}^{M} \alpha_i(\omega)\phi_i(x). \quad (9)$$

The functions $\phi_i(x)$, $i = 0, \ldots, M$ are determined by minimizing in each point x the variance of the error $\alpha(x, \omega) - \tilde{\alpha}(x, \omega)$ under the condition that $\tilde{\alpha}(x, \omega)$ is an unbiased estimator of $\alpha(x, \omega)$. This yields

$$E[\alpha(x, \omega)] = \phi_0(x) + \sum_{i=1}^{M} E[\alpha_i(\omega)]\phi_i(x),$$
$$\sum_{i=1}^{M} \phi_i(x)Cov[\alpha_i(\omega)\alpha_j(x)] = Cov[\alpha(x, \omega)\alpha_j(\omega)], \quad j = 1, 2, \ldots, M, \quad (10)$$

with the covariance operator

$$Cov[X, Y] = E[(X - E[X])(Y - E[Y])]. \quad (11)$$

Li and Der Kiureghian (1993) introduced a spectral decomposition of the nodal covariance matrix. They showed that the maximum error of the KL expansion is not always smaller than the error of Kriging for a given number of retained terms. The point-wise variance error estimator of the KL expansion for a given order of truncation is smaller than the error of Kriging in the interior of the discretization domain but larger at the boundaries. Note however that the

KL expansion provides the lowest mean error over the domain.

Polynomial Chaos Expansion

In contrast to the KL expansion, the polynomial chaos expansion does not need an a priori knowledge of the covariance structure of the random field. A functional representation of the random field in terms of a vector of basic random variables ξ,

$$\alpha(x, \omega) = f(x, \xi), \quad (12)$$

is necessary. Such a representation is given for the solution field of a stochastic boundary value

problem, if the random parameter field has been discretized.

A nonlinear expansion of the functional $f(x, \xi)$ is obtained by projecting it on a set of polynomials $\{\Gamma_p\}$ the basic random variables ξ. The space spanned by $\{\Gamma_p\}$ is called the pth homogeneous chaos. The polynomial Γ_p, called the polynomial chaos of order p, is a polynomial of order p that is orthogonal to all polynomials with order less than p.

Assuming symmetry of the polynomials, which is always possible (Ghanem and Spanos 1991), the random field can be approximated by

$$
\begin{aligned}
\tilde{\alpha}(x, \omega) = {} & \phi_0(x) + \sum_{i_1=1}^{N} \phi_{i_1}(x) \Gamma_1\left(\xi_{i_1}\right) \\
& + \sum_{i_1=1}^{N} \sum_{i_2=1}^{i_1} \phi_{i_1 i_2}(x) \Gamma_2\left(\xi_{i_1}, \xi_{i_2}\right) \\
& + \sum_{i_1=1}^{N} \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} \phi_{i_1 i_2 i_3}(x) \Gamma_3\left(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}\right) + \dots
\end{aligned}
\tag{13}
$$

or more briefly by

$$
\tilde{\alpha}(x, \omega) = \sum_{i=1}^{M} \Phi_i(x) \Psi_i(\xi).
\tag{14}
$$

The expansion is convergent in the mean-square sense. For Gaussian random variables, the polynomial chaos expansion is described in more detail in Ghanem and Spanos (1991). Xiu and Karniadakis (2002) extended the approach to general families of orthogonal polynomials, the Wiener-Askey chaos.

Field and Grigoriu (2004) pointed out some limitations of polynomial chaos approximations. They demonstrated that the convergence rate of polynomial chaos approximations may be slow, that the accuracy of the polynomial chaos approximation is not always improved by adding terms, that higher-order moments may be inaccurate, and that the polynomial chaos approximations of homogeneous non-Gaussian processes may not be homogeneous.

Transformation Techniques for Non-Gaussian Random Fields

Transformation techniques for non-Gaussian random fields seek to represent the non-Gaussian random field as a nonlinear transformation of a Gaussian random field:

$$
\alpha(x, \omega) = F^{-1}(\Phi(g(x, \omega))),
\tag{15}
$$

where Φ is the standard Gaussian cumulative distribution function, F is the non-Gaussian marginal cumulative distribution function of $\alpha(x, \omega)$, and $g(x, \omega)$ is the underlying Gaussian random field.

The transformation imposes a correlation structure to $\alpha(x, \omega)$, namely,

$$
R(\xi) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F^{-1}(\Phi(u)) F^{-1}(\Phi(v)) \phi\left(u, v, R_g(\xi)\right) d\,u\, d\, v,
\tag{16}
$$

where $\phi\left(u, v, R_g(\xi)\right)$ denotes the joint Gaussian probability density function. If the correlation structure of $\alpha(x, \omega)$ does not match the prescribed values, one has to resort to nonlinear transformations that match the target marginal distribution and/or the correlation structure approximately (Grigoriu 1998).

The nonlinear transformation technique can be combined with any series expansion schemes described above for the underlying Gaussian random field. It allows to calculate analytically many important quantities, such as crossing rates and extreme value distributions.

Phoon et al. (2002, 2005) used the KL expansion for the simulation together with an iterative mapping scheme to fit the target marginal distribution function of non-Gaussian random fields. The method allows to simulate homogeneous as well as nonhomogeneous random fields.

Stochastic Linear Elliptic Boundary Value Problems

Consider the following model problem with stochastic operator and deterministic input on $\overline{D} \times \Omega$: find $\overline{D} \times \Omega \to \mathbb{R}$, such that P-almost surely

$$
\begin{aligned}
-\nabla \cdot \alpha(x, \omega) \nabla u(x, \omega) &= f(x) \ \text{ on } \ D, \\
u(x, \omega) &= 0 \ \text{ on } \ \partial D.
\end{aligned}
\tag{17}
$$

It is assumed that the deterministic input function $f(x)$ is square integrable and that the random field $\alpha : D \times \Omega \to \mathbb{R}$ is bounded and coercive, i.e., there exists positive constants $a_{\min}$, $a_{\max}$, such that

$$
P(\omega \in \Omega : a_{\min} < \alpha(x, \omega) < a_{\max} \ \forall x \in D) = 1
\tag{18}
$$

and that the random field has a continuous and square-integrable covariance function.

We are interested in the probability that a functional $F(u)$ of the solution $u(x, \omega)$ exceeds a threshold F_0, i.e., we want to evaluate the integral

$$
P_F = \int_{\Omega} \chi_{(F_0, \infty)}(F(u(x, \omega))) dP(\omega),
\tag{19}
$$

where $\chi_I(\cdot)$, the indicator function, assumes the value 1 in the interval I and vanishes elsewhere.

The variational formulation of the stochastic boundary value problem necessitates the introduction of the Sobolev space $H_0^1(D)$ of functions having generalized derivatives in $L^2(D)$ and vanishing on the boundary ∂D with norm $\|u\|_{H_0^1(D)} = \left(\int_D |\nabla u|^2 dx \right)^{1/2}$, the space $L_P^2(\Omega)$ of square-integrable random variables, and the tensor product space $H_0^1(D) \otimes L_P^2(\Omega)$ of $H_0^1(D)$-valued random fields with finite second order moments, equipped with the inner product

$$
(u, v)_{H_0^1(D) \otimes L^2(\Omega)} = \int_{\Omega} \int_D \nabla_u(x, \omega) \cdot \nabla_v(x, \omega) dx dP(\omega).
\tag{20}
$$

The variational formulation of the stochastic linear elliptic boundary value problem Eq. 17 then reads: find $u \in H_0^1(D) \otimes L_P^2(\Omega)$, such that for all $v \in H_0^1(D) \otimes L_P^2(\Omega)$,

$$
\int_{\Omega} \int_D \alpha(x, \omega) \nabla u \cdot \nabla v dx dP(\omega) = \int_{\Omega} \int_D f(x) v(x, \omega) dx dP(\omega).
\tag{21}
$$

The assumptions on the random field $\alpha(x, \omega)$ guarantee the continuity and coercivity of the bilinear form in Eq. 21, and thus, the existence and uniqueness of a solution to Eq. 21 follows from the Lax-Milgram lemma.

Numerical Solution of the Stochastic Boundary Value Problem

In a first step, the random parameter field is discretized and replaced by a finite sum of random variables. Assume that a suitable approximation is given by a linear combination of continuous and independent random variables $\xi_i(\omega)$ with zero mean and unit variance,

$$
\alpha_M(x, \omega) = E[\alpha](x) + \sum_{i=1}^{M} \xi_i(\omega) \phi_i(x),
\tag{22}
$$

where $\Gamma_i = \xi_i(\Omega)$ are bounded intervals in $\mathbb{R}$.

Under these assumptions, the stochastic variational problem involving the random field $\alpha_M(x, \omega)$ has the following deterministic equivalent: find $u \in H_0^1(D) \otimes L_P^2(\Gamma)$, such that

$$
\begin{aligned}
\int_{\Gamma} \int_D \alpha_M(x, y) \nabla_x u(x, y) \cdot \nabla_x v(x, y) dx p(y) dy \\
= \int_{\Gamma} \int_D f(x) v(x, y) dx p(y) dy,
\end{aligned}
\tag{23}
$$

for all $v \in H_0^1(D) \otimes L_P^2(\Gamma)$, where $p \in L^{\infty}(\Gamma)$ is the joint probability density functions of the random variables $\xi_i(\omega)$, $i = 1, 2, \ldots, M$; see Babuška et al. (2005). Here, $\Gamma = \prod_{i=1}^{M} \Gamma_i \subset \mathbb{R}^M$. This variational formulation is now discretized on finite dimensional approximation spaces. For $H_0^1(D)$, a family of standard finite element approximation spaces $X_h \subset H_0^1(D)$ of continuous piecewise linear functions in a regular

S

triangulation T_h of D with mesh parameter h is considered. Following Deb et al. (2001), discontinuous finite elements are applied on the domain Γ, which is partitioned into elements $\gamma = \prod_{i=1}^{M}(a_i^{\gamma}, b_i^{\gamma})$, with $(a_i, b_i) \subset \Gamma_i$. Elements of $Y_k^q \subset L_p^2(\Gamma)$ are functions that are polynomials of degree at most q (i.e., $y_1^{q1} y_2^{q2} \ldots y_M^{qM} \in y_k^q$ if $q_1 + q_2 + \cdots + q_M \leq q$) when restricted to each element $\gamma \subset \Gamma$. The parameter $k = (k_1, k_2, \ldots k_M)$, with $k_i = \max_{\gamma \subset \Gamma} |b_i^{\gamma} - b_i^{\gamma}|$, represents the mesh parameter. If the partition consists of a single element only and the degree q is varied, global approximations are obtained.

Denote with $N_i(x)$, $i = 1 \ldots N$ a basis of X_h and with $\psi_k(y)$, $k = 1 \ldots P$, a basis of Y_k^q. The solution $u(x, y)$ is approximated on $X_h \otimes Y_k^q$ by

$$\sum_{i=1}^{N} \sum_{k=1}^{P} u_{ik} N_i(x) \psi_k(y). \tag{24}$$

In order to determine the unknown coefficients u_{ik}, this representation is inserted into Eq. 23 together with test functions $v(x, y) = N_j(x)\psi_l(y)$, $j = 1, 2, \ldots, N$, $l = 1, 2, \ldots, P$, yielding

$$\sum_{j=1}^{P} \int_{\Gamma} (\boldsymbol{K}^{(0)} + \sum_{s=1}^{M} \boldsymbol{K}^{(s)} y_s) \psi_j(y) \psi_l(y) p(y) dy \boldsymbol{u}_j$$
$$= \boldsymbol{f} \int_{\Gamma} \psi_l(y) p(y) dy, \tag{25}$$

where

$$K_{ij}^{(0)} = \int_D E[\alpha](x) \nabla N_i(x) \cdot \nabla N_j(x) dx,$$
$$K_{ij}^{(s)} = \int_D \phi_s(x) \nabla N_i(x) \cdot \nabla N_j(x) dx,$$
$$f_i = \int_D f(x) N_i(x) dx, \quad i, j = 1, 2, \ldots, N, \tag{26}$$

and u_j is the $N \times 1$ dimensional matrix obtained from u_{ij} for fixed value of j. In structural mechanics, the matrices $\boldsymbol{K}^{(s)}$, $s = 0, 1, \ldots, M$ can be

interpreted as finite element stiffness matrices with a spatial variation of Young's modulus.

If the basis $\{\psi_k\}_{k=1}^{P}$ consists of discontinuous finite element base functions, the system of equations Eq. 25 decouples and can be treated separately for each element γ. Moreover, due to the presence of terms in one single variable y_s, $s = 1, 2, \ldots, M$ on the left-hand side of Eq. 25, it is possible to construct the basis $\{\psi_k\}_{k=1}^{P}$ in a way that the problem decouples into problems that have the same size as the deterministic FE problem (Babuška et al. 2005). To see this, consider the polynomial basis $\{\psi_i^{\gamma}\}_{i=1}^{P_{\gamma}}$ in one element γ only. This basis can be constructed by multiplying polynomials in one single variable y_s, $s = 1, 2, \ldots, M$. Let $\boldsymbol{P}(y)$ be the monomial basis $(1, y, y^2, \ldots, y^q)^T$. We want to find a transformation matrix $\boldsymbol{S}_{\gamma}^{(s)}$, such that the transformed basis $\psi_{\gamma}^{(s)}(y) = \boldsymbol{S}_{\gamma}^{(s)} \boldsymbol{P}(y)$ satisfies

$$\int_{a_s^{\gamma}}^{b_s^{\gamma}} \psi_i^{(s)}(y) \psi_j^{(s)}(y) p_s(y) dy = \delta_{ij} \text{ and}$$
$$\int_{a_s^{\gamma}}^{b_s^{\gamma}} \psi_i^{(s)}(y) \psi_j^{(s)}(y) y_s p_s(y) dy = \lambda_i^{(s)} \delta_{ij}. \tag{27}$$

Defining the matrices $\boldsymbol{A}_{\gamma}^{(s)}$ and $\boldsymbol{B}_{\gamma}^{(s)}$ by

$$A_{\gamma ij}^{(s)} = \int_{a_s^{\gamma}}^{b_s^{\gamma}} P_i(y) P_j(y) y p_s(y) dy$$
$$B_{\gamma ij}^{(s)} = \int_{a_s^{\gamma}}^{b_s^{\gamma}} P_i(y) P_j(y) p_s(y) dy, \tag{28}$$

it is easily seen that $\lambda_i^{(s)}$ are the eigenvalues and $\boldsymbol{S}_{\gamma}^{(s)}$ is the matrix of the eigenvectors of the generalized eigenvalue problem

$$\boldsymbol{A}_{\gamma}^{(s)} \boldsymbol{S}_{\gamma}^{(s)T} = \boldsymbol{B}_{\gamma}^{(s)} \boldsymbol{S}_{\gamma}^{(s)T} \boldsymbol{\Lambda}, \text{ with } \boldsymbol{S}_{\gamma}^{(s)} \boldsymbol{B}_{\gamma}^{(s)} \boldsymbol{S}_{\gamma}^{(s)T} = \boldsymbol{I}. \tag{29}$$

Now, by introducing multiindices $j = (j_1, j_2, \ldots, j_M)$ for the basis function $\psi_{\gamma i} = \prod_{s=1}^{M} \psi_{i_s}^{(s)}(y_s)$ on each element γ, we find from Eq. 25

$$\left(\boldsymbol{K}_{\gamma}^{(0)} + \sum_{s=1}^{M} \boldsymbol{K}_{\gamma}^{(s)} \right) \boldsymbol{u}_{\boldsymbol{j}} = \boldsymbol{f}_{\gamma} \int_{\gamma} \psi_{\gamma \boldsymbol{j}}(y) p(y) dy, \quad 1 \leq j_i \leq q + 1, \quad i = 1, 2, \ldots, M \tag{30}$$

due to the orthogonality properties Eq. 27 of this basis.

As can be seen from this equation, the parallelization of the algorithm for computing the local chaos approximations is easily possible due to the fact that:

1. The approximations on each element γ are independent.
2. The linear systems for the expansion coefficients decouple due to the choice of biorthogonal polynomials instead of Hermite polynomials.

As a consequence, any degree of parallelization (from coarse grained to fine grained) is possible, depending on the number of processors at disposal. Note, however, that the introduction of biorthogonal polynomials is possible only if the random variables ξ_i are independent and that approximations with biorthogonal polynomials require in general an upper limit q for the polynomial degree in each single variable y_s, $s = 1, \ldots, M$, i.e., $y_1^{q_1} y_2^{q_2} \ldots y_M^{q_M}$ is an approximation function, if $q_s \leq q$, $s = 1, 2, \ldots$, M. For the same level q, this truncation leads to much more expansion coefficients than requiring $q_1 + q_2 + \cdots q_M \leq q$, especially if M is large.

Relationship with Monte Carlo Simulation

There is a close correspondence between Monte Carlo simulation and local polynomial approximations. The equations for Monte Carlo simulation can be obtained if instead of the Galerkin approximation, a collocation method is applied with respect to Γ, leading to

$$\left(\mathbf{K}^{(0)} + \sum_{s=1}^{M} \mathbf{K}^{(s)} y_s^j \right) u(y^j) = \mathbf{f}, j = 1, 2, \ldots, N,$$

(31)

with N sampling points y^j, $j = 1, 2, \ldots, N$, where finite element solutions are obtained.

If the vectors $u(y^j)$ are computed by means of Eq. 31, it is not obvious how to interpolate the solution in probability space for other values of y. A simple interpretation in terms of a local approximation would be a partition of Γ into N subdomains, where each subdomain contains the nearest neighbors of the sampling point y_j. In each subdomain, $u(y)$ is then approximated by a constant value, namely, the value of the sampling point associated with the subdomain. If the number of sampling points tends to infinity, the error of the approximation vanishes.

This approach can be related to Latin hypercube sampling, if the number of subdomains γ is very high and the approximations are not computed for every subdomain of the partition. For computing n samples according to the Latin hypercube sampling method, Γ is partitioned into n^N subdomains of equal size. Then, the $n \times N$ matrix Π containing random permutations of $1, \ldots, n$ and an $n \times N$ matrix $G R$ of independent and uniformly over $[0, 1]$ distributed random numbers are constructed, and the sampling plan $S = (\Pi - R)/n$ is established. Mapping the elements of the sampling plan to Γ via $\xi_{ij} = P_G^{-1}(S_{ij})$, $i = 1, \ldots, n, j = 1, \ldots$, N, where $P_G^{-1}(\cdot)$ is the inverse of the standard Gaussian distribution, one computes the solution only in those subdomains, in which the sample $\xi_i = (\xi_{i1}, \ldots, \xi_{iN})$, $i = 1, \ldots, n$ falls. In this way, a hybrid method, i.e., a combination of sampling techniques with approximation techniques, is obtained that leads to a considerable reduction of the sampling variance.

Evaluation of Response Quantities

Once the algebraic problem is solved and the approximation coefficients have been determined, an expression of the displacement field that depends on the input random variables has been obtained. This expression can be considered as a response surface. This response surface has local character (Proppe 2008) and depends on the size and location of the elements γ, if a partition of Γ is adopted.

Approximations for the moments of $u(x, \omega)$ can be computed by evaluating this expression

with respect to the input random variables. Computation of approximations for the distribution of $u(x, \omega)$ or the failure probability Eq. 19 is a more complex task, and resort to Monte Carlo simulation via the obtained expression for $u(x, \omega)$ seems to be the easiest way to accomplish it.

For solving reliability problems, it is very helpful to consider the approximation for the displacements as a local response surface, which leads to a functional relationship between the input random variables and $u(x, \omega)$. It is then possible to compute the most probable point of failure (MPP), i.e., the point ξ with $F(u(x, \xi)) = F_0$ with lowest Euclidean norm, and to refine the grid at its vicinity. In this way, it is possible to solve the reliability problem with a high degree of accuracy in an adaptive manner.

The MPP may also be useful for the evaluation of the integral in Eq. 19 by means of variance-reduced Monte Carlo simulation (importance sampling). To this end, a sampling density $\tilde{p}(y)$ is introduced by shifting the original probability density function $p(y)$ of the random variables ξ_i, $i = 1, \ldots, M$, to the previously obtained MPP, and Eq. 19 is approximated by

$$P_F \approx \sum_{j=1}^{N} \chi_{(F_0,\infty)}\big(F\big(u\big(x, y^j\big)\big)\big) \frac{p(y^j)}{\tilde{p}(y^j)} \tilde{p}\big(y^j\big),$$

(32)

where the sampling points y^j, $j = 1, 2, \ldots, M$ are generated according to $\tilde{p}(y)$ and $u(x, y^j)$ is computed from the approximation on the element that contains y^j.

The accuracy of the MPP is influenced by the FE mesh, the truncation level M of the random parameter field, the partition of Γ, and the choice of the ansatz functions both in spatial and random domain. These parameters can be gradually adapted such that the MPP is computed with a prescribed accuracy.

Example

In order to illustrate the stochastic finite element techniques, consider a clamped thin square plate under uniform in-plane tension (cf. Ghanem and
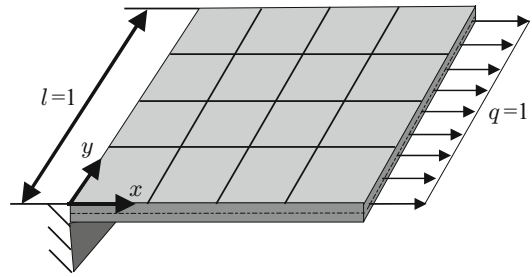
Stochastic Finite Elements, Fig. 1 Thin square plate under uniform in-plane tension

Spanos 1991 and Fig. 1). The product of Young's modulus and the thickness of the plate is assumed to be an isotropic Gaussian random field with covariance function

$$C(x_1,y_1;x_2,y_2) = \sigma \exp\left(-\frac{|x_1 - x_2|}{l_c} - \frac{|y_1 - y_2|}{l_c}\right),$$

(33)

standard deviation $\sigma = 0.2$, correlation length $l_c = 1$, and unit mean value. The Poisson's ratio is 0.3. The plate has unit length and the external excitation is deterministic and of unit magnitude. The longitudinal displacement of one of the free corners is considered in the following.

Approximation of Higher-Order Moments: Skewness and Kurtosis

The random field has been represented by two random variables ($M = 2$). Figures 2 and 3 display the approximations obtained with polynomials up to third order for the skewness and the kurtosis of the displacement of the free corner with increasing number of intervals of equal probability. Reference values have been obtained by Monte Carlo simulation with $30 \cdot 10^6$ samples. For one interval, the result corresponds to a global approximation. It can be seen that global approximations are rather inexact for the skewness and kurtosis and that it is sufficient to split the intervals into two parts in order to improve the results significantly.

Hybrid Method

Figures 4 and 5 display relative (with respect to the estimated values) 95 % confidence intervals

Stochastic Finite Elements,
Fig. 2 Approximation of the skewness
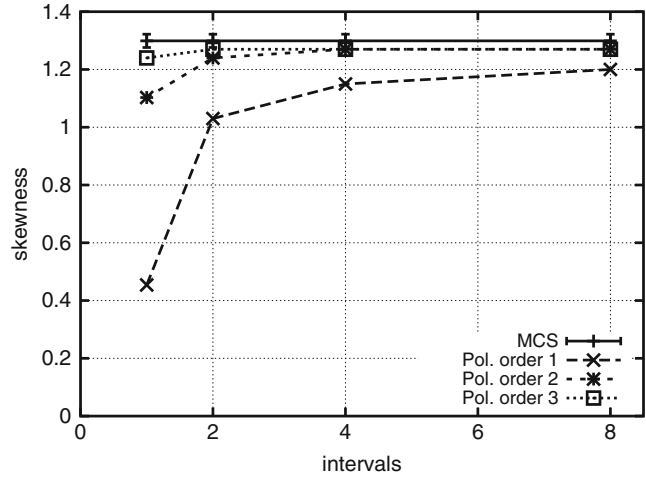
Stochastic Finite Elements,
Fig. 3 Approximation of the kurtosis
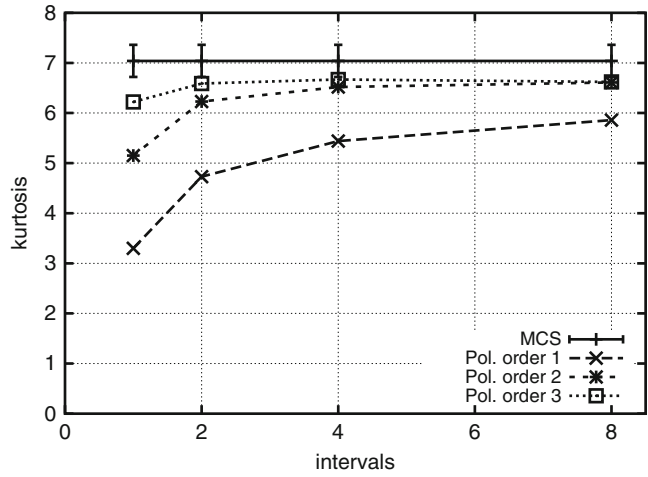
Stochastic Finite Elements,
Fig. 4 Approximation of the skewness by Latin hypercube sampling techniques

S
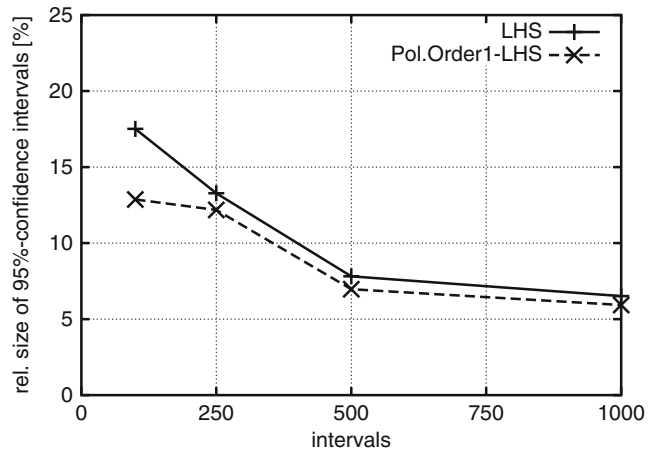
Stochastic Finite Elements, Fig. 5 Approximation of the kurtosis by Latin hypercube sampling techniques
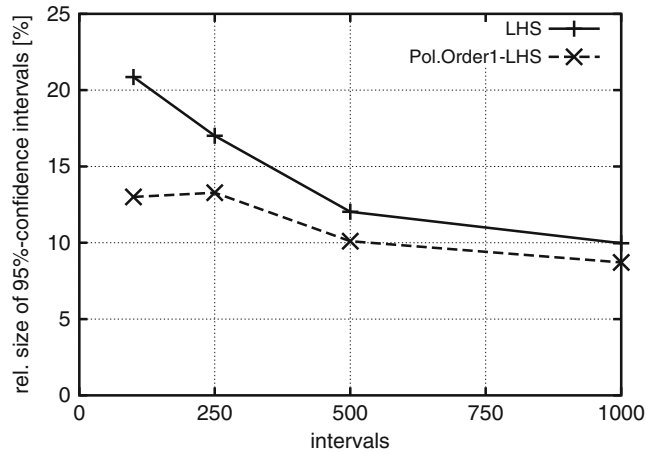
Stochastic Finite Elements, Fig. 6 Accuracy of estimates for the MPP
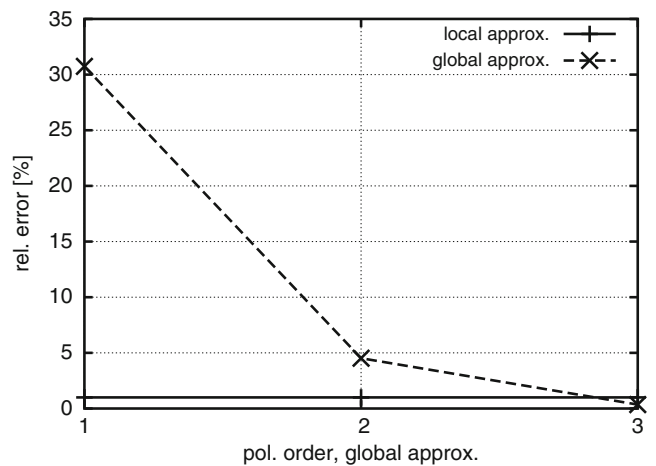
for the estimation of the skewness and kurtosis obtained from Latin hypercube sampling and Latin hypercube sampling combined with linear approximation in the sampling intervals. A three-term ($M = 3$) representation of the random field by truncated Gaussian random variables has been employed. For a low to moderate number of intervals, the confidence intervals obtained by the hybrid method are narrower than those of the standard Latin hypercube sampling.

Reliability Assessment

For most reliability approximation techniques, it is necessary to compute the MPP. In order to obtain this point, a constrained optimization problem has to be solved. This implies expensive function calls (e.g., FE solutions) for the computation of the displacements. These computations can be bypassed, if the approximations obtained from the numerical solution of the stochastic boundary value problem are employed as response surfaces. Figure 6 compares the relative error (i.e., the relative Euclidean distance) between the "true" MPP computed with FE calls and approximations obtained with local and global response surfaces of the displacement. The random field has been represented by four random variables ($M = 4$). The functional $F(u)$ is the value of the displacement field at the free corner and the threshold value $F0$ has been set to 2.0. In order to obtain local approximations, an approximation of the MPP has been computed with a linear global approximation, and an interval of length 1σ has been inserted at that point for

Stochastic Finite Elements, Fig. 7 Prediction of the failure probability
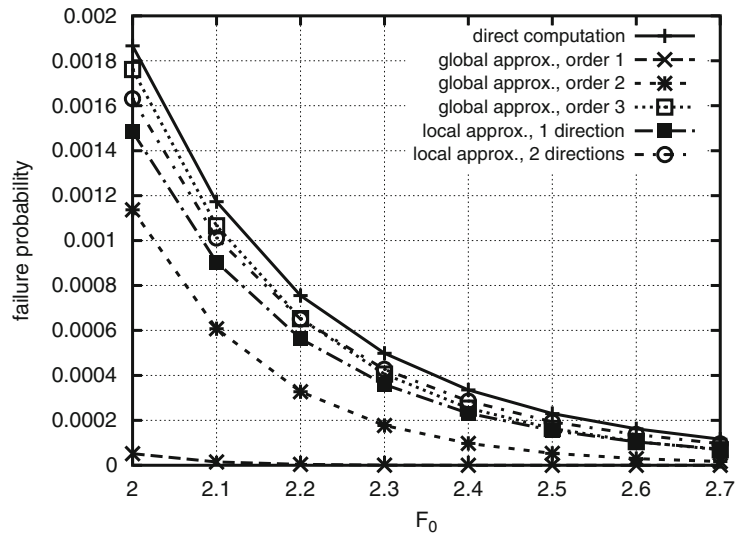
the coordinate with largest partial derivative of the longitudinal displacement for the free corner at the MPP. Linear polynomials have been applied for local approximations. From Fig. 6, one can see that this method yields very accurate approximations for the MPP, while for global approximations, polynomials of third degree are necessary in order to achieve a comparable accuracy.

From this fact, a corresponding approximation quality of the failure probability can be deduced, if variance-reduced sampling techniques which rely on the MPP together with the abovementioned response surface techniques are employed. This can be clearly seen from Fig. 7, where the probability of failure obtained with various approximation techniques has been plotted over the threshold value F_0. Importance sampling at the predicted MPP with 30 batches of 10,000 samples has been employed for each failure probability estimate. For global approximations, again polynomials of third degree are necessary in order to obtain accurate predictions. On the other hand, local approximations with linear polynomials lead already to quite accurate results, if only the principle direction is partitioned.

In Fig. 8, the relative error of the failure probability for $F_0 = 2.3$ is displayed over the corresponding number of deterministic FE runs.

This allows to compare the efficiency of local and global approximations. From Fig. 8, it is evident that local approximations have considerable advantages over global approximations. This behavior is even more pronounced for smaller values of F_0.

Finally, for $F_0 = 2.7$ (corresponding to a probability of failure of 8.2×10^{-5}) and a tolerance of 2 % for the Euclidean norm of the MPP, the procedure described in section "Relationship with Monte Carlo Simulation" yields $M = 4$ and a partition of Γ into 48 elements leading to an overall effort of 768 deterministic FE runs for the reliability estimation problem. This is considerably lower than 10^4 runs for importance sampling and 10^5 runs for a direct Monte Carlo simulation.

Concluding Remarks

For the SFE solution of stochastic boundary value problems, a mathematical theory is available that is in many aspects comparable to its deterministic counterpart, the finite element method. However, for random fields with short correlation lengths (requiring a high number M of random variables), the solution methods become inefficient, due to the series expansion of the solution. This is also true for most other SFE approximations, be they global or local, and recourse to efficient sampling

Stochastic Finite Elements,
Fig. 8 Efficiency of global and local approximation techniques for the prediction of the failure probability
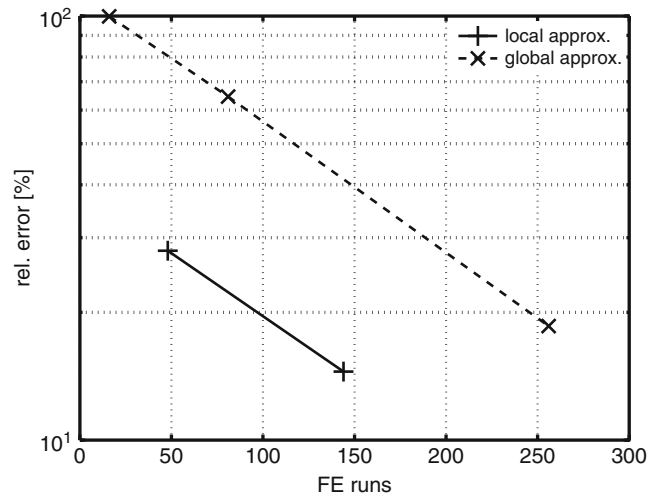
techniques combined with efficient model reduction might be the only way to circumvent this problem.

A fundamental question that still has to be addressed in detail concerns the error of the solution due to the discretization of the random parameter field. Beyond this aspect of verification, the validation of the random field model itself, either from experimental data or from information pertaining to the microscale, remains an important issue (cf. the critique of the SFE method raised, e.g., in Ostoja-Starzewski 2011).

References

Acharjee S, Zabaras N (2007) A non-intrusive stochastic Galerkin approach for modeling uncertainty propagation in deformation processes. Comput Struct 85(5–6):244–254

Astill CJ, Noseir SB, Shinozuka M (1972) Impact loading on structures with random properties. J Struct Mech 1:63–77

Babuška IM, Tempone R, Zouraris GE (2005) Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation. Comput Methods Appl Mech Eng 194(1):1251–1294

Babuška I, Nobile F, Tampone R (2007) A stochastic collocation method for elliptic partial differential equations with random input data. SIAM J Numer Anal 45:1005–1034

Baecher GB, Ingra TS (1981) Stochastic FEM in settlement predictions. J Geotech Eng Div 107:449–463

Baroth J, Bresslotette P, Chauvière C, Fogli M (2007) An efficient SFE method using Lagrange polynomials: application to nonlinear mechanical problems with uncertain parameters. Comput Methods Appl Mech Eng 196:4419–4429

Berveiller M, Sudret B, Lemaire M (2006) Stochastic finite element: a non intrusive approach by regression. Rev Eur Mècanique Numèrique 15:81–92

Brenner CE, Bucher CG (1995) A contribution to the SFE-based reliability assessment of nonlinear structures under dynamic loading. Probab Eng Mech 10:265–273

Contreras H (1980) The stochastic finite-element method. Comput Struct 12:341–348

Cornell CA (1970) First order uncertainty analysis in soils deformation and stability. In: Proceedings of first international conference of statistics and probability in soil and structural engineering, Hong Kong

Deb MK, Babuška IM, Oden JT (2001) Solution of stochastic partial differential equations using Galerkin finite element techniques. Comp Methods Appl Mech Eng 190:6359–6372

Deodatis G (1991) Weighted integral method I: stochastic stiffness matrix. J Eng Mech 117:1851–1864

Deodatis G, Micaletti RC (2001) Simulation of highly skewed non-Gaussian stochastic processes. Trans ASCE, J Eng Mech 127:1284–1295

Der Kiureghian A, Ke JB (1988) The stochastic finite element method in structural reliability. Probab Eng Mech 3:83–91

Field RV Jr, Grigoriu M (2004) On the accuracy of the polynomial chaos approximation. Probab Eng Mech 19(1–2):65–80

Ghanem RG, Spanos PD (1991) Stochastic finite elements: a spectral approach. Springer, New York

Grigoriu M (1998) Simulation of stationary non-Gaussian translation processes. J Eng Mech 124:121–126

Grigoriu M (2006) Evaluation of Karhunen-Loève, spectral and sampling representations for stochastic processes. J Eng Mech 132:179–189

Hisada T, Nakagiri S (1981) Stochastic finite element method developed for structural safety and reliability. In: Proceedings of third international conference on structural safety and reliability, Trondheim, pp 395–408

Huang SP, Quek ST, Phoon KK (2001) Convergence study of the truncated Karhunen-Loève expansion for simulation of stochastic processes. Int J Numer Methods Eng 52:1029–1043

Huang S, Mahadevan S, Rebba R (2007) Collocation-based stochastic finite element analysis for random field problems. Probab Eng Mech 22:194–205

Lawrence M (1987) Basis random variables in finite element analysis. Int J Numer Methods Eng 24:1849–1863

Li CC, Der Kiureghian A (1993) Optimal discretization of random fields. J Eng Mech 119:1136–1154

Liu WK, Belytschko T, Mani A (1986) Probabilistic finite elements for nonlinear structural dynamics. Comput Methods Appl Mech Eng 56:61–81

Loève M (1977) Probability theory. Springer, Berlin

Matthies HG, Brenner CG, Bucher CG, Guedes Soares C (1997) Uncertainties in probabilistic numerical analysis of structures and solids – stochastic finite elements. Struct Saf 19:283–336

Ostoja-Starzewski M (2011) Stochastic finite elements: where is the physics? Theor Appl Mech 38:379–396

Phoon KK, Huang HW, Quek ST (2002) Simulation of second-order processes using Karhunen-Loève expansion. Comput Struct 80:1049–1160

Phoon KK, Huang SP, Quek ST (2005) Simulation of strongly non-Gaussian processes using Karhunen-Loève expansion. Probab Eng Mech 20:188–198

Proppe C (2008) Estimation of failure probabilities by local approximation of the limit state function. Struct Saf 30(4):277–290

Shinozuka M (1971) Simulation of multivariate and multidimensional random processes. J Acoust Soc Am 49:357–367

Stefanou G (2009) The stochastic finite element method: past, present and future. Comput Methods Appl Mech Eng 198:1031–1051

Stefanou G, Papadrakakis M (2007) Assessment of spectral representation and Karhunen-Loève expansion methods for the simulation of Gaussian stochastic fields. Comput Methods Appl Mech Eng 196:2465–2477

Sudret B, Der Kiureghian A (2000) Stochastic finite element methods and reliability – state of the art. Technical report, UCB/SEMM-2000/08, Department of Civil & Environmental Engineering, University of California, Berkeley

Todor RA, Schwab C (2006) Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients. Research Report No. 2006–05, Seminar für Angwandte Mathematik, ETH Zürich

Vanmarcke E, Grigoriu M (1984) Stochastic finite element analysis of simple beams. J Eng Mech 109:1203–1214

Xiu D, Karniadakis G (2002) The Wiener-Askey polynomial chaos for stochastic differential equations. SIAM J Sci Comput 24:619–644

Yamazaki F, Shinozuka M (1988) Digital generation of non-Gaussian stochastic fields. Trans ASCE, J Eng Mech 114:1183–1197

Stochastic Ground Motion Simulation

Sanaz Rezaeian[1] and Xiaodan Sun[2]
[1]U.S. Geological Survey, Golden, CO, USA
[2]Southwest Jiaotong University, Chengdu, China

Synonyms

Earthquake ground motion simulation; Nonstationary stochastic process; Stochastic models; Strong ground motion simulation; Synthetic accelerograms

Introduction

Strong earthquake ground motion records are fundamental in engineering applications. Ground motion time series are used in ▶ response-history dynamic analysis of structural or geotechnical systems. In such analysis, the validity of predicted responses depends on the validity of the input excitations. Ground motion records are also used to develop ground motion prediction equations (GMPEs) for intensity measures such as spectral accelerations that are used in ▶ response-spectrum dynamic analysis. Despite the thousands of available strong ground motion records, there remains a shortage of records for large-magnitude earthquakes at short distances or in specific regions, as well as records that sample specific combinations of source, path, and site characteristics. The limited number of recordings has become problematic in the emerging field of

S

▶ performance-based earthquake engineering (PBEE), which considers the entire spectrum of structural response, from linear to grossly nonlinear and even collapse, and thereby requires ground motions with various levels of intensity for various earthquake design scenarios (e.g., a design scenario can be defined by the earthquake magnitude, distance, and site conditions). To obtain the desired ground motions for the purposes of PBEE, it is common engineering practice to scale or modify acceleration time series that were recorded during previous earthquakes to represent certain ground motion characteristics for the design of structural or geotechnical systems. However, scaling and modification methods can significantly alter other ground motion characteristics and result in unrealistic earthquake ground motion time series. Synthetic ground motions can be used instead to replace or supplement recorded motions when scarcity of previously recorded motions becomes a problem, provided they accurately capture characteristics of real earthquake ground motions and their natural variability.

Synthetic ground motions can be based on deterministic or stochastic simulations. A determinist model is one in which variables are uniquely determined, and the model performs the same way for a given set of initial conditions. Conversely, in a stochastic model, randomness is present and variables are not described by unique values, but rather by probability distributions. In earthquake engineering, deterministic ground motion simulation is commonly referred to as "physics-based" ground motion simulation. These simulation models synthesize the earthquake source by defining a source model (e.g., kinematic or dynamic rupture models) and describe the seismic wave travel path by defining a material model (e.g., seismic velocity model); then, they utilize numerical methods (e.g., finite element or finite difference methods) to estimate the solution to the wave propagation equation. In other words, these models explicitly incorporate the physics of the propagation of seismic waves. They produce realistic synthetic ground motions at low frequencies (typically <1Hz). But they are computationally intensive due to the fine

discretization of the medium and require a thorough knowledge of the environment (i.e., earthquake source, crustal structure, material properties, basin and local site effects); this information is not available for many regions or is not accessible to engineers for well-studied regions. On the other hand, stochastic simulation models are simple and powerful tools that incorporate what is known about ground motion, source, path, and site characteristics into simple functional forms, and as a result are widely used to simulate earthquake ground motions. There is a specific entry dedicated to ▶ physics-based ground motion simulation in this encyclopedia by Taborda and Roten that provides more details on deterministic ground motion simulation models. Here, we focus on stochastic ground motion simulation.

A word of caution may be in order here about the terminology "physics-based." The use of this term for deterministic models is mostly out of convenience and should not imply that stochastic models are not based on physics. In fact, some experts disagree with this terminology because many stochastic models are valid representations of the physics of earthquakes, and although they do not necessarily solve the mathematical problem of wave propagation, they integrate the physical characteristics of the earthquake source, path effects, and site effects implicitly through their formulations and parameters.

Both deterministic and stochastic simulations can be used for response-history dynamic analysis, but only stochastic simulations can be utilized for stochastic dynamic (i.e., random vibration) analysis, because the latter analysis method requires a random process model of the earthquake ground motion. Synthetic ground motions are particularly useful for nonlinear dynamic analysis due to the scarcity of recorded motions for large-magnitude earthquakes that are capable of causing nonlinear responses. Two approaches are available for nonlinear dynamic analysis of structures subjected to earthquakes: (1) nonlinear response-history analysis by the use of a selected set of ground motion time series and (2) nonlinear stochastic dynamic analysis by the use of a stochastic representation of the ground motion.

Well-developed methods and tools are available for nonlinear response-history analysis; a set of recorded or synthetic ground motions or a collection of both can be used for this type of analysis. Nonlinear stochastic methods are not as developed, but research in that direction is continuing (e.g., Au and Beck 2003; Der Kiureghian and Fujimura 2009). Several stochastic ground motion simulation models exist that can satisfy the requirements of these nonlinear stochastic methods. These methods typically require the input excitation to be represented in a discrete form and in terms of a finite number of random variables. For example, in Der Kiureghian and Fujimura (2009), it is necessary to represent the input excitation in the following form $\mu(t) + \sum_{i=1}^{n} s_i(t)u_i$, where $\mu(t)$ is the deterministic mean of the random process, n is the discretization of the time series and provides a measure of the resolution, and u_i and $s_i(t)$ are, respectively, representatives of standard normal random variables and deterministic basis functions of time.

Regardless of the method of simulation, the resulting synthetic time series should properly represent the main characteristics of real earthquake ground motions. For most engineering applications, the intensity, duration, and frequency content of the input ground motion are the characteristics that control the response of structural and geotechnical systems. In addition to the duration of motion, which results from variation of the ground motion intensity over time, variation of the frequency content of the motion is also an important characteristic that influences the response, particularly in nonlinear analysis. Consequently, two important characteristics of real earthquake ground motions that separate them from a simple stochastic process (e.g., a ▶ white-noise signal) are temporal and spectral nonstationarities. Temporal nonstationarity refers to the nonstationarity of ground motion in the time domain (i.e., variation of intensity with time). Spectral nonstationarity refers to the nonstationarity of ground motion in the frequency domain (i.e., variation of frequency

content with time). Whereas most simulation methods account for temporal nonstationarity, which controls most common intensity measures and duration of motion, many ignore the spectral nonstationarity, which can be very important in nonlinear dynamic analysis.

Many synthetic ground motion models exist, but not all represent the natural variability of real earthquake ground motions. Some models underestimate the natural variability of ground motions because they underestimate the variability in their model parameters. On the other hand, if correlations between the model parameters are not taken into account or are underestimated, the variability of the resulting ground motions can be overestimated compared to real earthquake ground motions. In probabilistic analysis, it is important to capture the natural variability of ground motions in the simulations to accurately quantify the variation of response due to the variation in input excitation.

In summary, stochastic ground motion simulations have many applications. They can be used in linear and nonlinear response-history dynamic analysis as well as in stochastic dynamic analysis, or they can be used to develop GMPEs for various intensity measures. Stochastic models are generally simpler, more widely used, and more accurate at high frequencies than deterministic (physics-based) models. A good stochastic ground motion model must represent the temporal and spectral nonstationary characteristics of real earthquake ground motions and must properly represent their natural variability. More details on stochastic ground motion models, their classification, formulation, parameters, and practicality are given below. Discussions on adjustment of models for special situations including multicomponent, near-fault, and multi-station simulations are also provided.

Classification of Ground Motion Simulation Models

There are two main categories of models for generating synthetic ground motions, source-based models and site-based models.

Source-based models explicitly describe the fault rupture process at the source, the propagation of the resulting seismic waves through the ground medium, and the effects of local site conditions on the seismic waves to generate a time series at a specific site. Site-based models describe the ground motion time series as it is observed at a specific site by fitting a stochastic process to previously recorded motions with known earthquake and site characteristics; in this way, they implicitly account for the source, path, and site effects. Due to their need for empirical recordings of ground motions, site-based models are sometimes referred to as empirical models, but note that this terminology is not strictly correct because many source-based models also use data to empirically identify their parameters.

Source-Based Models

Source-based models can be deterministic or stochastic. An early review of source-based models is presented by Zerva (1988). Deterministic (physics-based) simulations were briefly described in the Introduction and have their own entry in this encyclopedia. They can produce realistic accelerograms at low frequencies (typically <1 Hz) but often need to be adjusted for high frequencies by combining with a stochastic component, resulting in hybrid models. For many years, source-based simulations have been utilized to develop GMPEs for locations where earthquakes are rare and the collection of recorded ground motions is sparse, for example, in the stable continental regions of the Central and Eastern United States (e.g., Somerville et al. 2001); research is continuing to improve these simulations and the resulting GMPEs.

One family of widely used synthetic ground motions are the source-based stochastic simulations based on the work of David Boore and a number of other researchers in the past several decades (e.g., Boore 2003; Beresnev and Atkinson 1998; Motazedian and Atkinson 2005; Boore 2009). This simulation method roots in the work of McGuire and Hanks (1980), which identifies the Fourier amplitude spectrum (FAS) of a ground motion considering the source, path, and site effects and then combines it with a random phase spectrum. This method assumes ground motion to be a band-limited white Gaussian noise with finite duration. The white-noise is then adjusted to have the following FAS:

$$Y(M_0, f, R) = S(M_0, f) \times P(R, f) \times G(f) \times I(f)$$

where $S(M_0, f)$ is the source spectrum with f denoting the frequency and M_0 representing the seismic moment; R is the distance from the earthquake source to the site and $P(R, f)$ accounts for the wave propagation effects, including the geometrical attenuation with distance, which is usually described by a three-segmented function, and the energy dissipation (combining intrinsic and scattering attenuation); $G(f)$ accounts for the effects of local site conditions, including near-surface amplification and reduction of high frequencies that result from the path-independent loss of energy; and finally, $I(f)$ is the indicator of the instrument or the type of motion (i.e., acceleration, velocity, or displacement). More details are provided in Boore (2003).

This method was first developed to model far-field ground motions where the earthquake source can be considered as a point (Boore 1983), resulting in "point-source stochastic models." For simulation of ground motions closer to the earthquake source, the method was improved to consider the rupture progress on a finite fault (Beresnev and Atkinson 1998), resulting in "finite-fault stochastic models." In finite-fault modeling, the fault is discretized into many subfaults, and each subfault is treated as a point-source. The ground motion from each subfault is modeled using the point-source stochastic model with the amplitude spectrum described in the above equation. The total ground motion at a site is the superposition of the contributions of all subfaults with a proper time lag considering the difference from the triggering time of the subfault and from the travel time between the subfault and the site. Using finite-fault modeling, the synthetic motion is able to capture some near-fault characteristics of ground motion, such as rupture directivity and hanging-wall effect.

Most improvements in source-based stochastic modeling focus on enhancing the source spectrum, $S(M_0, f)$. The originally used source spectrum was according to the ω-square model with single corner frequency (Aki 1967)

$$S(M_0, f) = \frac{CM_0}{1 + \left(\frac{f}{f_c}\right)^2}$$

where C is a constant related to the radiation pattern and f_c is the corner frequency inversely controlled by the dimension of the source. By applying the above equation in finite-fault modeling, however, the sum of acceleration amplitude spectrum at high-frequency range $(f > f_c)$ strongly depends on the discretization scheme of the fault. The proper subfault size is required to be within 5–15 km in order to obtain a reliable synthesis (Beresnev and Atkinson 1998). To solve this problem, Motazedian and Atkinson (2005) made a conceptual improvement by introducing a "dynamic corner frequency" in the source spectrum

$$S_{ij}(f) = \frac{CM_{0ij}H_{ij}}{1 + \left(\frac{f}{f_{cij}}\right)^2}$$

where f_{cij} is the dynamic corner frequency for the ij-th subfault and H_{ij} is the scaling factor to compensate the high-frequency spectral loss from using a dynamic corner frequency.

The dynamic corner frequency of the ij-th subfault depends on the cumulative number of subfaults ruptured by the time the subfault is triggered. Therefore, each subfault will have a different corner frequency in the source spectrum. Generally, a subfault that is triggered in the early part of the rupture will have a higher corner frequency than a subfault triggered in the late part. This describes the difference in the frequency contents of the generated motions from different subfaults. The most significant advantage of using a dynamic corner frequency is that the high-frequency spectral amplitude is no longer dependent on the size of the subfault; thus a subfault can be as small as 1 km, which allows the finite-fault modeling to be used for

small- or intermediate-magnitude events. One of the latest improvements to the source-based stochastic simulation is the work of Boore (2009) that allows selection of different functions for calculating the scaling factor H_{ij}.

Source-based stochastic models are simple and practical compared to deterministic (physics-based) models. Availability of source codes for these models has increased their use in recent years. For example, EXSIM_dmb is a simulation software that is developed by Boore and can be downloaded from his website (http://www.daveboore.com/software_online.html), and EXSIM_beta is a simulation software that is developed by Motazedian and can be found on his website (http://http-server.carleton.ca/~dariush/); both EXSIM_dmb and EXSIM_beta are based on stochastic finite-fault modeling with dynamic corner frequency. One of the disadvantages of source-based stochastic models is that they assume stationarity of the frequency content with time, which is usually not the case in real earthquake ground motions as will be discussed in more detail in the next section. Also, care should be taken in using these models for probabilistic analyses because they usually underestimate the variability in ground motions by fixing their parameters instead of assigning a probability distribution to them.

In general, source-based models tend to heavily employ seismological principles to describe the source mechanism and wave travel path and depend on physical parameters that vary significantly from region to region. This limits their use in regions where seismological data are lacking. This is more critical for deterministic (physics-based) models than it is for stochastic models that use simpler parameters and models for the source, path, and site effects. In the current practice, most engineers prefer using methods of scaling and spectrum matching of recorded motions instead of incorporating source-based models. This is partly due to the lack of understanding the seismological principles underlying these models and the fact that these models require a thorough knowledge of the source, wave path, and site characteristics, which typically are not available to a design engineer.

S

In this regard, source-based stochastic models have an advantage compared to deterministic (physics-based) models; but the user is still required to make certain assumptions depending on the region in order to define the models for the source $S(M_0, f)$, path $P(R, f)$, and site $G(f)$. More recently developed site-based stochastic models can be more useful in engineering applications because their relatively simple formulations facilitate time-efficient simulations (similar to source-based stochastic models) and their parameters are expressed directly in terms of earthquake and site characteristics such as magnitude and distance that describe a design scenario and are readily available to a design engineer.

Site-Based Models

Site-based models are stochastic and represent the earthquake ground motion time series by a stochastic process. A large number of site-based stochastic ground motion models have been developed. Formal reviews are presented by several authors including Liu (1970), Shinozuka and Deodatis (1988), and Conte and Peng (1997). The papers by Rezaeian and Der Kiureghian (2008, 2010, and 2012) provide reviews and are representative of the more recent work. As previously mentioned, a good synthetic ground motion model must represent both the temporal and the spectral nonstationary characteristics of earthquake ground motions. Whereas temporal nonstationarity can be easily modeled in site-based stochastic simulations by time-modulating a stationary ▶ stochastic process, spectral nonstationarity is not as easy to model. Nevertheless, this spectral nonstationarity is of particular importance in nonlinear dynamic analysis because of the moving resonance effect of nonlinear structures (i.e., when the natural period of a structure varies during the ground shaking period due to inelastic behavior and gets closer to the varying frequency content of the ground motion) and should not be ignored. In general, for a stochastic model to be of practical use in earthquake engineering, it should be parsimonious, i.e., it must have as few parameters as possible. Preferably, the model parameters should provide physical insight into the characteristics of the motion. Furthermore, the model should refrain from complicated analysis, involving extensive processing of recorded motions for parameter identification.

Site-based stochastic ground motion simulation models can be classified into four categories:

Filtered white-noise processes	Processes obtained by passing a white-noise signal through a filter. These models are appealing due to the efficient digital simulation of sample functions for a filtered white-noise process (i.e., generating realizations). Many models of this type use subsequent modulation in time to achieve temporal nonstationarity, but these processes have essentially time-invariant frequency content. This type of simulation can be improved by varying the filter parameters in time to achieve nonstationarity of the frequency content (e.g., Rezaeian and Der Kiureghian 2008).
Filtered Poisson processes	Processes obtained by passing a train of Poisson pulses through a linear filter. Through modulation in time, these processes can possess both temporal and spectral nonstationarities (e.g., Lin 1986). However, matching with recorded ground motions is difficult.

Auto-regressive-moving-average (ARMA) models	By allowing the model parameters to vary with time, these models (e.g., Conte et al. 1992) can have both temporal and spectral nonstationarities. However, it is difficult to relate the model parameters to any physical aspects of the ground motion.
Time-varying spectral representation	Instead of working in the time domain, as is done in the previous three categories, these models work in the frequency domain and use various forms of spectral representation (e.g., Conte and Peng 1997). The focus in these models is in developing a time-varying spectral representation by matching to a target recorded ground motion. These models require extensive processing of the target motion. Virtually all these models assume the ground motion to be a zero-mean Gaussian process.

Simulating a Single Ground Motion Time Series Using a Site-Based Stochastic Model

Model Formulation

Regardless of the specific formulation, every site-based model is essentially a stochastic process with a certain number of parameters. As an example, the general formulation for a site-based stochastic model based on the filtered white-noise process is given below (Rezaeian and Der Kiureghian 2008)

$$a(t) = q(t, \boldsymbol{\alpha}) \times f(t, \boldsymbol{\lambda})$$

where $a(t)$ is the acceleration time series with t denoting time; $q(t, \boldsymbol{\alpha})$ is a deterministic nonnegative time-modulating function with $\boldsymbol{\alpha}$ representing a vector of parameters; and $f(t, \boldsymbol{\lambda})$ is a filtered white-noise process with $\boldsymbol{\lambda}$ representing a vector of parameters. Several models have been proposed for the $q(t, \boldsymbol{\alpha})$ modulating function, e.g.,

- The piecewise function by Housner and Jennings (1964)
- The double exponential function proposed by Shinozuka and Sato (1967)
- The gamma function proposed by Saragoni and Hart (1974)

These models or their modified versions have been used by many other modelers. In general, a time-modulating function gradually increases from zero to achieve a nearly constant intensity that represents the "strong shaking" phase of an earthquake ground motion and then gradually decays back to zero (see Fig. 1). The parameters $\boldsymbol{\alpha}$ control the intensity and shape of the function. Although it is not a common practice, two or more modulating functions can be combined to simulate ground motions that have more than a single strong shaking phase, a feature possibly caused by multiple sub-events.

The conventional "filtered white-noise" process $f(t, \boldsymbol{\lambda})$ is the stationary response of a linear time-invariant filter subjected to a white-noise process. White-noise $w(t)$ is a stationary random process in time that has a zero mean and a constant spectral density for all frequencies. The response of a linear filter to a white-noise process may be calculated by using the Duhamel convolution integral, and hence the general formulation of a filtered white-noise process can be written in the following form:

$$f(t, \boldsymbol{\lambda}) = \int_{-\infty}^{t} h(t - \tau, \boldsymbol{\lambda}) w(\tau) d\tau$$

where $h(t, \boldsymbol{\lambda})$ denotes the impulse response function or IRF of a linear filter and the parameters $\boldsymbol{\lambda}$ can be the natural frequency and damping of the

S

$$a(t) = q(t, \boldsymbol{\alpha}) \left\{ \frac{1}{\sigma_f(t)} \int_{-\infty}^{t} h[t - \tau, \boldsymbol{\lambda}(\tau)] w(\tau) d\tau \right\}$$

Time modulating function	Unit-variance process
Controls temporal nonstationarity	Controls spectral nonstationarity

×

Fully-nonstationary process

Stochastic Ground Motion Simulation, Fig. 1 Development of a fully nonstationary stochastic process according to Rezaeian and Der Kiureghian (2008)

filter or other parameters that shape the filter response. Different filters are used and various tricks are incorporated by modelers to make the resulting process "fully nonstationary" and representative of real earthquake ground motions. One approach taken by Rezaeian and Der Kiureghian (2008) is illustrated in Fig. 1, where the filtered white-noise process is normalized by its standard deviation $\sigma_f(t)$ to separate the temporal from the spectral characteristics; and the filter parameters are functions of time, $\boldsymbol{\lambda}(t)$, to achieve spectral nonstationarity. Like all other models that use a time-modulating function, temporal nonstationarity is achieved by $q(t, \boldsymbol{\alpha})$. In this figure, the nonstationary characteristics of ground motions are demonstrated using the trace of an example stochastic process.

Model Parameters

Given a target accelerogram (i.e., a recorded ground motion), the model parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\lambda}(t)$ are identified by fitting certain characteristics of the stochastic process to those of the target accelerogram. The approaches to parameterization and model fitting are entirely different from model to model. Currently, there are no standards on the exact characteristics that modelers should match to identify their parameters for a target motion. However, all these different characteristics are designed to capture the intensity, duration, and frequency content of the ground motion at certain points in time or in frequency. In some models, this is done over the entire duration of the recorded ground motion and its entire frequency spectrum; in this case, the parameters are identified using the following generalized formulation:

$$\hat{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmin}} \int_0^{t_n} \left[f_{obj}(\text{recorded}) - f_{obj}(\text{simulation}) \right] dt$$

where $\mathbf{p}$ is a vector of model parameters and t_n denotes the entire duration of the recorded ground

motion. Linear or nonlinear optimization methods are used to minimize the difference between an objective function for the recorded motion, f_{obj}(recorded), and the same function for the simulation, f_{obj}(simulation), resulting in the vector of identified parameters $\hat{\mathbf{p}}$ for that particular recorded motion. The misfit of the objective functions can be characterized by their difference, as is shown in the above equation, or by their absolute difference, or by their square root difference. Many objective functions are used in the literature by various modelers. Objective functions such as the expected cumulative energy of the process or the variance of the process can be used to identify the parameters of the time-modulating function; objective functions such as the zero-level up-crossing rate can be used to identify the frequency parameters (e.g., Yeh and Wen 1990; Rezaeian and Der Kiureghian 2008).

Once a set of model parameters has been identified, $\hat{\mathbf{p}}$, the model formulation is used to simulate realizations of the ground motion. These realizations are all different due to the stochasticity of the model (e.g., the white-noise process in the above example), but they all have the same model parameters and expected characteristics similar to those of the target motion. In this way, the target accelerogram may be regarded as a single realization of the ground motion process for a specified set of model parameters, while the simulated motions may be regarded as other random samples of the process for the same set of model parameters.

Post-Processing for Long Spectral Periods (Low Frequency)

In general, site-based stochastic ground motion models fail to match the response spectrum associated with the target accelerogram in the long spectral period range and tend to overestimate the structural response (typically beyond 2–4 s depending on the model). This is because a stochastic model developed for an acceleration process cannot guarantee zero velocity and displacement residuals (final values at the end of the record) upon integration of a sample realization. This shortcoming of site-based models has been recognized by some modelers who "post-process" the simulations by extending

baseline correction methods or high-pass filtering methods used for recorded accelerograms to simulated motions in order to overcome this disadvantage (e.g., Papadimitriou 1990; Liao and Zerva 2006). For example, Rezaeian and Der Kiureghian (2008) use a high-pass filter representing a critically damped oscillator to post-process the stochastically simulated ground motion to obtain zero velocity and displacement residuals and appropriate response spectral values for periods as long as 5–10 s. An example of a simulated ground motion before and after post-processing is shown in Fig. 2. Even though the difference between the acceleration processes is insignificant, the integration over time results in unacceptably high nonzero velocity and displacement residuals as well as high spectral intensities at long periods before post-processing. After the post-processing, the residuals are zero and the spectral content is lower at long periods.

Simulating a Suite of Ground Motions for a Specified Design Scenario

Here, a design scenario refers to a set of earthquake and site characteristics such as earthquake magnitude, source-to-site distance, and soil conditions of the site. In the broader context of PBEE, an ensemble of ground motions that represents all possible ground shakings for a specified design scenario is of interest, not just ground motion simulations with characteristics similar to those of a previously observed motion. A good stochastic ground motion model must properly represent the natural variability of real earthquake ground motions. It also must be practical for use by a practicing engineer.

Variability

The variability among an ensemble of simulated ground motions for a specified design scenario comes from two sources: (1) the randomness of ground motions for a specified set of model parameters (the stochasticity mentioned in the discussion of the model parameters) and (2) the randomness of the model parameters for a specified design scenario. The former is due to the stochastic nature

S
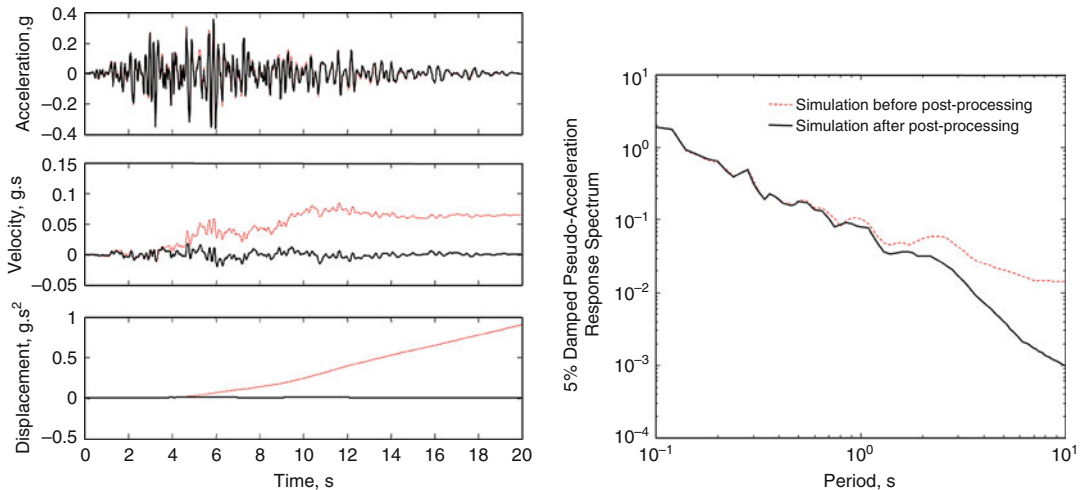
Stochastic Ground Motion Simulation, Fig. 2 An example stochastic simulation before and after post-processing using a high-pass filter with a cutoff frequency of 0.2 Hz

of the model and it is achieved when fitting to and simulating a target accelerogram. The latter is referred to as "parametric uncertainty." Both should be modeled properly to reproduce in the synthetic motions the variability present in real earthquake ground motions.

The underestimation of parametric uncertainty results in the underestimation of the natural variability of ground motions. This has been a major problem in the current practice of seismic hazard analysis and the generation of synthetic ground motions. A few exceptions are some recent site-based models (e.g., Rezaeian and Der Kiureghian 2010), where model parameters are fitted to a database of ground motions and are randomized at the time of simulation to achieve variability similar to the sample database. Most source-based models account for the variability in ground motions by manually varying the source parameters; whether the realizations follow the "true" underlying probability distribution for each parameter is the question in these models.

To properly account for parametric uncertainty, a good model needs to randomly generate parameters according to their underlying probability distributions and also to account for the correlations between the parameters. Probability distributions for parameters and their correlations can be estimated empirically or based on the physical meaning of a parameter. If correlations between the parameters are ignored, the variability in the resulting ensemble of ground motions may be overestimated because the parameter realizations are not well constrained.

Practicality

There is a need for generating synthetic motions that correspond to a specified design scenario, but most site-based stochastic models limit their scope to generating synthetic motions similar to a target recorded ground motion. In such models, no attempt is made to select a different but appropriate set of model parameters for the specified design scenario. Moreover, for a model to be practical and to be used in the engineering community, it should be expressed in terms of variables that are readily available to a design engineer.

In practice, GMPEs are widely used because they predict measures of ground motion intensity in terms of common earthquake and site characteristics that define a design scenario to a practicing engineer. Typical intensity measures modeled by GMPEs are peak ground motion values (i.e., peak ground acceleration, velocity, and displacement) and elastic response spectra at specified oscillator periods and damping ratios. A few GMPEs for inelastic response spectra have also been developed. These GMPEs are useful for

linear response-spectrum analysis or crude nonlinear analysis, but not for response-history analysis, as they do not predict ground motion time series. Such simplified analysis methods have proven to be adequate for simple code-based design purposes, but not for PBEE. Stochastic ground motion models can complement GMPEs by generating ground motion time series in terms of similar earthquake and site characteristics used by GMPEs. This is done by developing predictive relations for the model parameters in terms of commonly used earthquake and site characteristics.

To adjust a stochastic model for practical use, the model parameters are identified for a database of recorded ground motions. Then, through regression analyses and empirical modeling methods, predictive relations for each parameter in terms of earthquake and site characteristics are developed to relieve the design engineer from the task of predicting the parameter values for a design scenario (e.g., Sabetta and Pugliese 1996; Stafford et al. 2009; Rezaeian and Der Kiureghian 2010). A generic form of predictive relations for a model parameter p is given below

$$\Phi^{-1}[F_P(p)] = \mu(\text{Earthquake, Site, } \boldsymbol{\beta}) + e$$

where Φ^{-1} denotes the inverse of the standard normal cumulative distribution function, F_P denotes the marginal cumulative distribution function estimated for parameter p, μ is the conditional mean of $\Phi^{-1}[F_P(p)]$ given the earthquake and site characteristics of interest, and $\boldsymbol{\beta}$ is the vector of regression coefficients. The regression error e is a zero-mean normally distributed random variable with standard deviation σ representing the uncertainty of the parameter. The earthquake and site characteristics are typically the earthquake magnitude, source-to-site distance, faulting mechanism, and various proxies to describe the site soil conditions such as V_{S30} (i.e., the time-averaged shear wave velocity in the top 30 m of soil). These regression variables change depending on the region and the parameter under consideration.

Given the predictive relations as described above, one can calculate the mean of each model parameter given a set of earthquake and site characteristics. The error can then be randomized to generate the realizations of the model parameter. To generate the realizations of multiple parameters, correlations between the parameters should be taken into account in the simulation.

Extensions of Site-Based Stochastic Models

Multiple-Component Simulation

Earthquake ground motions are multidimensional. Neglecting the rotational components, three translational components are usually expressed in two orthogonal horizontal directions and one vertical direction. For proper dynamic analysis of complex structural and geotechnical systems, for example, asymmetric structures that are vulnerable to torsion or structural components that are sensitive to vertical ground shaking, it is important to simulate multicomponent ground motion time series. The vast majority of stochastic models are restricted to simulating a single horizontal ground motion component, but the research to simulate multiple components is continuing and under development (e.g., bidirectional ground motion simulation by Rezaeian and Der Kiureghian (2012).

When modeling and simulating multiple ground motion components, differences as well as similarities and dependencies between the components must be taken into account. Because the ground motion components are generated from the same earthquake source and from seismic waves that travel through the same ground medium, high correlations between model parameters of the components are expected. Correlations and coherencies between the components can be estimated empirically by analyzing a large number of recorded ground motion components.

For stochastic ground motions, in addition to cross-correlations between the model parameters, dependencies between the stochastic components (e.g., the seed white-noise processes) should also be taken into account. This is a difficult task, but Penzien and Watabe (1975) showed that there is a unique set of orthogonal axes, referred to as "principal axes," along which the three translational

S

components of ground motion may be considered uncorrelated. These are the major, intermediate, and minor principal axes along which the ground motion components have intensities in decreasing order and the stochastic processes (e.g., the seed white-noise processes) are statistically independent. Although the stochastic processes can be assumed statistically independent, there can still be high correlations between the model parameters of the two ground motion components (see Rezaeian and Der Kiureghian 2012). Based on examination of few real accelerograms, Penzien and Watabe (1975) suggested that the major principal axis usually points in the general direction of the epicenter, the intermediate principal axis is horizontal and perpendicular to the major principal axis, and the minor principal axis is almost vertical. The concept of principal axes has been the basis for many subsequent studies on multicomponent stochastic modeling of ground motions, but more studies are required to confirm the direction of the principal components. If simulation is done along the horizontal principal axes, an orthogonal pair of acceleration time series $a_1(t)$ and $a_2(t)$ can easily be rotated to the horizontal direction of interest using a rotation matrix

$$\begin{bmatrix} a_{1,\theta}(t) \\ a_{2,\theta}(t) \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} a_1(t) \\ a_2(t) \end{bmatrix}$$

where $a_{1,\theta}(t)$ and $a_{2,\theta}(t)$ are the counterclockwise rotations by angle θ.

Near-Fault Simulation

Ground motions close to fault ruptures (within about 30 km), i.e., ▶ near-fault ground motions, can have distinct characteristics such as the directivity effect in the fault-normal direction and the fling step in the fault-parallel direction. These characteristics can have a strong influence on the structural response and are usually modeled separately and then added to a site-based stochastic ground motion simulation.

A site near a fault may experience forward directivity, which occurs when the fault rupture propagates toward the site with a velocity almost equal to the shear wave velocity. The resulting ground motion time series is characterized by the presence of a two-sided, long-period, large-amplitude velocity pulse that typically shows as a long-period pulse in the response spectrum and may impose extreme demands on structures. Because near-fault recordings that contain directivity pulses are relatively scarce and because scaling and modification methods cannot create a pulse in a record that does not initially contain a pulse, simulated near-fault ground motions are needed. But only a few stochastic models currently exist that simulate the forward directivity effect (e.g., Mavroeidis and Papageorgiou 2003; Dabaghi et al. 2011). This is usually done by modeling a long-period velocity pulse that is a function of the earthquake design scenario, including the relative location of the site to the fault rupture. The simulated velocity pulse is added to the derivative of the simulated acceleration time series resulting in a simulated pulse-like near-fault ground motion.

Not all near-fault ground motions contain a pulse. For example, backward directivity occurs when the fault rupture propagates away from the site. The resulting ground motion tends to be of low intensity and long duration. A valid near-fault ground motion model should take into account both the pulse-like and non-pulse-like cases.

The fling step is caused by the permanent displacement of the fault and is usually characterized by a one-sided velocity pulse in the fault-parallel direction. Modeling of the fling step is not as well developed in stochastic ground motion modeling, but its effect on structural response is not as extreme as the forward directivity effect. The fling step is difficult to model because information is usually lost in the typical processing of recorded ground motions.

Spatial Simulation (Multiple Stations)

Most stochastic ground motion models are developed for single-station simulations. To simulate ground motions from a single earthquake event at multiple stations that are located close to each other, the spatial coherency between records should be taken into account. This coherency usually depends on the frequency content of the ground motions and separation distance between the stations and can be estimated from the power spectral density functions of recorded motions.

Several models exist in the literature that can be used to adjust a ground motion time series for spatial variation (e.g., Liao and Zerva 2006; Konakli and Der Kiureghian 2012, Konakli et al. 2014). These models can be applied to a stochastic simulation of ground motion at a single station to generate consistent simulations at nearby stations.

Validation of Ground Motion Simulation Models

Although many stochastic ground motion simulation models exist, no standard validation techniques are available. There are almost as many validation approaches undertaken as there are simulation models. Comparisons against empirical data and trusted models in engineering practice aid in identifying the limitations of the proposed simulation models and may encourage implementation of simulations in practice. The next step in the field of ground motion simulation that has been gaining attention in recent years is development of standard techniques for ground motion simulation validation for use in various engineering applications (see, e.g., the efforts of the Southern California Earthquake Center ground motion simulation validation technical activity group http://collaborate.scec.org/gmsv/Main_Page).

Summary

Stochastic ground motion simulations are used to replace or supplement recorded earthquake ground motions when recorded motions are sparse. Stochastic ground motion models are categorized into source-based and site-based models. Definitions and examples for each category are given. Source-based simulations explicitly model the source spectrum, path, and site effects. Site-based stochastic models are classified into four groups depending on their formulations. Regardless of the formulation, the model must exhibit the characteristics of real earthquake ground motions including temporal and spectral nonstationarities. Site-based models simulate a recorded accelerogram by identifying the model parameters through optimization and fitting into certain target characteristics of real earthquake ground motions. The parameters are then used in the formulation of the stochastic process. Post-processing is usually required for site-based stochastic models to ensure zero residuals and proper modeling of long-period spectral values. Simulating a suite of ground motions for a specified design scenario is discussed, where the natural variability of ground motions must be modeled properly by accounting for parametric uncertainty and correlations between the model parameters. Practicality of a model is also discussed, where the model parameters are expressed in terms of earthquake and site characteristics that define a design scenario and are readily available to a design engineer. Finally, adjusting site-based stochastic ground motion simulations for special cases including multicomponent, near-fault, and multi-station simulations is discussed. Currently, many research efforts are directed toward developing ground motion simulation validation procedures for various engineering applications.

Cross-References

▶ Physics-Based Ground-Motion Simulation

References

Aki K (1967) Scaling law of seismic spectrum. J Geophys Res 72:1217–1231

Au SK, Beck JL (2003) Subset simulation and its applications to seismic risk based on dynamic analysis. J Eng Mech ASCE 129:901–917

Beresnev IA, Atkinson G (1998) Stochastic finite-fault modeling of ground motions from the 1994 northridge, California, earthquake (I): validation on rock sites. Bull Seismol Soc Am 88(6):1392–1401

Boore DM (2003) Simulation of ground motion using the stochastic method. Pure Appl Geophys 160:635–676

S

Boore DM (1983) Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. Bull Seismol Soc Am. 73:1865–1894

Boore DM (2009) Comparing stochastic point-source and finite-source ground-motion simulations: SMSIM and EXSIM. Bull Seismol Soc Am 99:3202–3216

Conte JP, Peng BF (1997) Fully nonstationary analytical earthquake ground-motion model. J Eng Mech ASCE 12:15–24

Conte JP, Pister KS, Mahin SA (1992) Nonstationary ARMA modeling of seismic motions. Soil Dyn Earthq Eng 11:411–426

Dabaghi M, Rezaeian S, Der Kiureghian A (2011) Stochastic simulation of near-fault ground motions for specified earthquake and site characteristics. In: Proceedings of the 11th international conference on applications of statistics and probability in civil engineering (ICASP11), Zurich, pp 759–760

Der Kiureghian A, Fujimura K (2009) Nonlinear stochastic dynamic analysis for performance-based earthquake engineering. Earthq Eng Struct 38:719–738

Housner GW, Jennings PC (1964) Generation of artificial earthquakes. J Eng Mech Div ASCE 90:113–150

Konakli K, Der Kiureghian A (2012) Simulation of spatially varying ground motions including incoherence, wave-passage and differential site-response effects. Earthq Eng Struct 41:495–513

Konakli K, Der Kiureghian A, Dreger D (2014) Coherency analysis of accelerograms recorded by the UPSAR array during the 2004 Parkfield earthquake. Earthq Eng Struct 43:641–659

Liao S, Zerva A (2006) Physically compliant, conditionally simulated spatially variable seismic ground motions for performance-based design. Earthq Eng Struct 35:891–919

Lin YK (1986) On random pulse train and its evolutionary spectral representation. Probabilist Eng Mech 1:219–223

Liu SC (1970) Synthesis of stochastic representations of ground motions. Bell Sys Tech J 49:521–541

Mavroeidis GP, Papageorgiou AS (2003) A mathematical representation of near-fault ground motions. Bull Seismol Soc Am 93(3):1099–1131

McGuire RK, Hanks TC (1980) RMS accelerations and spectral amplitude of strong ground motion during the San Fernando, California, earthquake. Bull Seismol Soc Am 70(5):1907–1919

Motazedian D, Atkinson GM (2005) Stochastic finite-fault modeling based on a dynamic corner frequency. Bull Seismol Soc Am 95(3):995–1010

Papadimitriou K (1990) Stochastic characterization of strong ground motion and application to structural response. Earthquake Engineering Research Laboratory, California Institute of Technology, Pasadena, Report No. EERL 90–03

Penzien J, Watabe M (1975) Characteristics of 3-dimensional earthquake ground motions. Earthq Eng Struct 3:365–373

Rezaeian S, Der Kiureghian A (2008) A stochastic ground motion model with separable temporal and spectral nonstationarities. Earthq Eng Struct 37:1565–1584

Rezaeian S, Der Kiureghian A (2010) Simulation of synthetic ground motions for specified earthquake and site characteristics. Earthq Eng Struct 39:1155–1180

Rezaeian S, Der Kiureghian A (2012) Simulation of orthogonal horizontal ground motion components for specified earthquake and site characteristics. Earthq Eng Struct 41:335–353

Sabetta F, Pugliese A (1996) Estimation of response spectra and simulation of nonstationary earthquake ground motions. Bull Seismol Soc Am 86:337–352

Saragoni GR, Hart GC (1974) Simulation of artificial earthquakes. Earthq Eng Struct 2:249–267

Shinozuka M, Deodatis G (1988) Stochastic process models for earthquake ground motion. Probabilist Eng Mech 3:114–123

Shinozuka M, Sato Y (1967) Simulation of nonstationary random process. J Eng Mech ASCE 93:11–40

Somerville P, Collins N, Abrahamson N, Graves R, Saikia C (2001) Ground motion attenuation relations for the Central and Eastern United States – Final report, 30 June 2001: Report to U.S. Geological Survey for award 99HQGR0098, 38 p

Stafford PJ, Sgobba S, Marano GC (2009) An energy-based envelope function for the stochastic simulation of earthquake accelerograms. Soil Dyn Earthq Eng 29:1123–1133

Yeh CH, Wen YK (1990) Modeling of nonstationary ground motion and analysis of inelastic structural response. Struct Safety 8:281–298

Zerva A (1988) Seismic source mechanisms and ground motion models, review paper. Probabilist Eng Mech 3:64–74

Stochastic Structural Identification from Vibrational and Environmental Data

Minas D. Spiridonakos and Eleni N. Chatzi
Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, Institute of Structural Engineering, Zurich, CH, Switzerland

Synonyms

ARMA model; Environmental conditions; Output-only identification; Polynomial chaos expansion; Structural identification; Vibration response

Introduction

In recent years, Structural Health Monitoring (SHM) of engineering structures has become an important area of research. This is a broad term which by large also deals with the online detection and identification of structural damage and therefore holds a critical significance for the case of large-scale civil structures where a failure incidence is connected not only to high cost losses but potentially to human loss in the worst-case scenario.

Vibration-based methods are today the fastest growing research area in the SHM field (Fassois and Kopsaftopoulos 2013). However, while such methods are already used for industrial mechanical structures, these are still far from being successfully implemented in large-scale civil structures such as high-rise buildings and bridges. The main reason for this discrepancy is the fact that these structures do not operate in the controlled environment of an industry or a laboratory facility but rather they are affected by a number of uncontrollable environmental conditions, such as temperature, temperature spatial gradients due to solar irradiation, temporal temperature gradient due to thermal capacity, humidity, and others (Sohn 2007). The shortcomings of vibration-based SHM methods have been recognized already in one of the first review papers on the subject (Doebling et al. 1998) exposing the advantages and capabilities of several methodologies in this area.

For the aforementioned reason, one is faced with the problem of identifying a dynamic model that is able to represent a given structure for an extended spectrum of operational conditions – if not its entire life cycle. Toward this end, the combined vibrational response and environmental data, which are normally available in current SHM systems, have to be utilized along with appropriate models that can account for the variability of structural dynamics due to environmental conditions. The approach detailed herein treats this issue through a functional representation that introduces this dependence into the model's parameters.

Structures Operating Under Varying Conditions and Their Responses

A linear viscously damped structure may be described by the following lumped parameter, time-varying ordinary differential equation:

$$\boldsymbol{M}(t)\ddot{\boldsymbol{x}}(t) + \boldsymbol{C}(t)\dot{\boldsymbol{x}}(t) + \boldsymbol{K}(t)\boldsymbol{x}(t) = \boldsymbol{u}(t), \qquad \text{for } t \in \left[t_0, t_f\right] \quad (1)$$

where t designates analog time with t_0 denoting the start and t_f the end of the structure's lifetime or inspection interval, $\boldsymbol{x}(t)$ the structural displacements response vector, $\boldsymbol{u}(t)$ the applied excitation forces vector, and $\boldsymbol{M}(t)$, $\boldsymbol{C}(t)$, $\boldsymbol{K}(t)$ the mass, damping, and stiffness matrices of the structure, respectively.

The variability of structural response may be usually attributed to a small number of parameters related with either inherent properties of the structure or exogenous random variables, e.g., environmental and loading conditions, geometry, mass distribution, and so on. Gathering these L parameters in a single vector $\boldsymbol{\xi}(t) = [\xi_1(t), \xi_2(t), \ldots, \xi_L(t)]^{\mathrm{T}}$ the previous model may be rewritten as

$$\boldsymbol{M}(\boldsymbol{\xi}(t))\ddot{\boldsymbol{x}}(t) + \boldsymbol{C}(\boldsymbol{\xi}(t))\dot{\boldsymbol{x}}(t) + \boldsymbol{K}(\boldsymbol{\xi}(t))\boldsymbol{x}(t) = \boldsymbol{u}(t) \quad (2)$$

Focusing on environmental conditions and assuming that they are characterized by slow (compared to the structural dynamics) variation, the previous model may be rewritten as

$$\boldsymbol{M}(\boldsymbol{\xi})\ddot{\boldsymbol{x}}(t) + \boldsymbol{C}(\boldsymbol{\xi})\dot{\boldsymbol{x}}(t) + \boldsymbol{K}(\boldsymbol{\xi})\boldsymbol{x}(t) = \boldsymbol{u}(t), \quad \text{for } t \in [t_a, t_b] \subset [t_0, t_f] \quad (3)$$

The latter may be considered as a local time-invariant model which represents the structure when monitored during a small time interval. Within this interval the input parameter vector,

S

$\boldsymbol{\xi}$ and as a consequence the mechanical properties of the structure, may be considered to be "fixed" to constant values. Then, the input parameter vector $\boldsymbol{\xi}$ may be considered as a realization of the random vector Ξ with joint probability density function (pdf) $f_\Xi(\boldsymbol{\xi})$. In this way the initial linear time-varying (LTV) model of Eq. 1 may be written as a linear time-invariant (LTI) model:

$$
\begin{aligned}
& \boldsymbol{M}(\Xi)\ddot{\boldsymbol{x}}(t) + \boldsymbol{C}(\Xi)\dot{\boldsymbol{x}}(t) \\
& + \boldsymbol{K}(\Xi)\boldsymbol{x}(t) = \boldsymbol{u}(t), \quad \text{with } \Xi(\omega) \in \Omega
\end{aligned} \tag{4}
$$

which is however characterized by parameters that are depending on the input variable vector Ξ. In the equation above, Ω designates the event space of the random vector Ξ, and ω an elementary event of this space.

The above continuous-time model obviously has a discrete-time counterpart, which for the simpler single-input single-output case may be approximated by a scalar difference equation of the following form (Andersen 1997):

$$
\begin{aligned}
& x(t) + a_1(\Xi)x[t-1] + \cdots + a_{n_a}(\Xi)x[t-n_a] \\
& = b_0(\Xi)u[t] + b_1(\Xi)u[t-1] + \cdots \\
& \quad + b_{n_b}(\Xi)u[t-n_b], \quad \text{for } t = 1, \ldots, N
\end{aligned} \tag{5}
$$

in which t designates normalized discrete time (absolute time normalized by the sampling period T_s), N the signals' length in samples, $x[t]$, $u[t]$ the discretized versions of the observed vibration response and excitation force, respectively, $a_i(\Xi)$, $b_i(\Xi)$ the model parameters depending on Ξ, and finally n_a, n_b the equation orders.

Structural identification is an inverse problem which concerns the estimation of an appropriate mathematical representation of the structural dynamics based on vibrational measurements. In the case of ambient loading, the excitation may not be measured with accuracy, and thus, the input sequence $u[t]$ in the theoretical model of Eq. 5 has to be replaced by an unknown excitation

sequence $e[t]$. The latter is typically considered to be a normally independently distributed (NID) sequence, that is, a white noise process, with zero mean and constant variance σ_e^2. Therefore, the analytical model of Eq. 5 is approximated by a model of the following form:

$$
\begin{aligned}
& \underbrace{x(t) + a_1(\Xi)x[t-1] + \cdots + a_{n_a}(\Xi)x[t-n_a]}_{\text{AR part}} \\
& = \underbrace{e[t] + c_1(\Xi)e[t-1] + \cdots + c_{n_c}(\Xi)e[t-n_c],}_{\text{MA part}} \\
& \mathrm{e}[t] \sim \text{NID}\big(0, \sigma_e^2\big), \quad \text{for } t = 1, \ldots, N
\end{aligned} \tag{6}
$$

which has the form of a conventional ARMA model (Ljung 1999), with the important difference that it is characterized by stochastic parameters which depend on the input variable vector Ξ. These parameters may actually be represented by a deterministic mapping describing their relation with the input random variables. Specifically, assuming that the ARMA model parameters $a_i(\Xi)$ and $c_i(\Xi)$ have finite variance, they admit the following representation (Soize and Ghanem 2004):

$$
\begin{aligned}
a_i(\Xi) &= \sum_{j=1}^{\infty} a_{i,j}\varphi_{\mathbf{d}(j)}(\Xi) \, (i = 1, \ldots, n_a), \\
c_i(\Xi) &= \sum_{j=1}^{\infty} c_{i,j}\varphi_{\mathbf{d}(j)}(\Xi) \, (i = 1, \ldots, n_c)
\end{aligned} \tag{7}
$$

with $a_{i,j}$ and $c_{i,j}$ designating unknown deterministic coefficients of projection onto the probability space of Ξ and $\varphi_{\mathbf{d}(j)}$ are multivariate basis functions that are orthonormal with respect to the joint pdf of Ξ, that is,

$$
E\big[\varphi_{\boldsymbol{\alpha}}(\Xi)\varphi_{\boldsymbol{\beta}}(\Xi)\big] = \begin{cases} 1, & \text{for } \boldsymbol{\alpha}=\boldsymbol{\beta} \\ 0, & \text{otherwise} \end{cases} \tag{8}
$$

and $\boldsymbol{d}(j)$ is the multi-indices of the multivariate basis. Each probability density function may be

associated with a well-known family of orthogonal polynomials. For instance, a normal distribution is associated with Hermite polynomials while a uniform distribution with Legendre. A list of the most common probability density functions along with the corresponding orthogonal polynomials and the relations for their construction may be found in Xiu and Karniadakis (2002). For the case of independent random variables Ξ_j the expansion of Eq. 7 is called polynomial chaos (PC) expansion (Blatman and Sudret 2010), and the resulting models may be referred to as PC-ARMA models.

Of course, for purposes of practicality, the infinite series of PC expansion of Eq. 7 must be truncated by selecting an appropriate functional subspace consisting of a finite number of terms. A usual approach lies in the selection of a functional subspace consisting of polynomial basis functions with total maximum degree P, that is, $|\boldsymbol{d}(j)| = \sum_{\ell=1}^{L} d(j, \ell) \leq P$ for all j. In this case the dimensionality of the functional subspace is equal to

$$p = \frac{(L + P)!}{L! P!} \qquad (9)$$

where L is the number of random variables and P the maximum basis total degree. For instance, for a random input vector of dimension equal to three ($L = 3$) and maximum degree equal to two ($P = 2$), the multi-indices vectors are:

	l		
	1	2	3
$\boldsymbol{d}(1)$	0	0	0
$\boldsymbol{d}(2)$	1	0	0
$\boldsymbol{d}(3)$	0	1	0
$\boldsymbol{d}(4)$	0	0	1
$\boldsymbol{d}(5)$	2	0	0
$\boldsymbol{d}(6)$	0	2	0
$\boldsymbol{d}(7)$	0	0	2
$\boldsymbol{d}(8)$	1	1	0
$\boldsymbol{d}(9)$	1	0	1
$\boldsymbol{d}(10)$	0	1	1

In this way, the resulting PC-ARMA model is fully parametrized in terms of a finite number of deterministic coefficients of projection $a_{i,j}$ and $c_{i,j}$ while a specific PC-ARMA model structure is fully defined by the model orders n_a, n_c, and the basis total degree P.

At this point it should be noted that similar models have been used for the structural identification and damage detection issues in a number of recent studies (Functionally Pooled ARX models; see Kopsaftopoulos and Fassois 2013 and the references therein) and automatic control within the Linear Parameter Varying (LPV) models framework (Tóth 2010, Zhao et al. 2012).

PC-ARMA Model Identification

The complete PC-ARMA structural identification problem may be stated as follows: Given a set of dynamic vibration response signals obtained from the structure under study for various measured environmental conditions, select the PC-ARMA model structure and estimate the corresponding model parameters that best fit the available records. The issues of PC-ARMA model parameter estimation and model structure selection are discussed in the following sections.

Model Parameter Estimation
As already mentioned, the estimation of a PC-ARMA model refers to the determination of the coefficients of projection assembled in the parameter vector

$$\boldsymbol{\theta} = \begin{bmatrix} a_{1,1} a_{1,2} \cdots a_{n_a,p} \, c_{1,1} \, c_{1,2} \cdots c_{n_c,p} \end{bmatrix}^T$$

or a subset of $\boldsymbol{\theta}$ selected by the procedure described in the next section. Its estimation is based on the availability of structural vibration response data and measurements of the environmental conditions of interest. In order to identify a stochastic model that will be able to describe the structure's dynamics for the wide range of its operational spectrum, data from a sufficient number of operating conditions should be available.

Considering K such datasets have been acquired during an equal number of experiments conducted under different input random vector realizations

S

$\boldsymbol{\xi}_k = \left[\boldsymbol{\xi}_{1_k} \boldsymbol{\xi}_{2_k} \ldots \boldsymbol{\xi}_{L_k} \right]^T$, for $k = 1, 2 \ldots, K$. Based on the validity of the assumptions described in the previous section, it is considered that the recorded vibration response signals $\boldsymbol{x}_k^N = \{ x_k[1], x_k[2], \ldots, x_k[N] \}^T$ follow the PC-ARMA model of Eqs. 6 and 7:

$$
\begin{aligned}
x_k[t] &+ \sum_{j=1}^{p} a_{1,j} \varphi_{\mathbf{d}(j)}(\boldsymbol{\xi}_k) x_k[t-1] + \cdots + \sum_{j=1}^{p} a_{n_a,j} \varphi_{\mathbf{d}(j)}(\boldsymbol{\xi}_k) x_k[t-n_a] \\
&= e_k[t] + \sum_{j=1}^{p} c_{1,j} \varphi_{\mathbf{d}(j)}(\boldsymbol{\xi}_k) e_k[t-1] + \cdots + \sum_{j=1}^{p} c_{n_c,j} \varphi_{\mathbf{d}(j)}(\boldsymbol{\xi}_k) e_k[t-n_c], \\
&\text{for } t = 1, \ldots, N, \text{ and } k = 1, \ldots, K
\end{aligned} \tag{9a}
$$

$$
e_k \sim \text{NID}\left(0, \sigma_{e_k}^2 \right) \tag{9b}
$$

$$
E\{ e_i[t] e_j[t-\tau] \} = \begin{cases} \sigma_{e_i}^2 \delta_{t,\tau} & \text{for } i=j \\ 0 & \text{for } i \neq j \end{cases} \tag{9c}
$$

with $\delta_{i,j}$ designating the Kronecker delta function. In the relationships above, the uncorrelatedness of the residual series between different vibration response records has been added to the model assumptions.

The estimation of the PC-ARMA model parameter vector $\boldsymbol{\theta}$ may be based on the minimization of the Prediction Error (PE) criterion. The PE criterion consists of the sum of squares of the model's one-step-ahead prediction errors for the complete set of datasets. As it may be easily shown, the one-step-ahead prediction errors

coincide with the model's residual sequence $e_k[t]$ (Ljung 1999), and thus, $\boldsymbol{\theta}$ has to be estimated through the following minimization problem:

$$
\begin{aligned}
\boldsymbol{\theta} &= \arg \min_{\theta} \left\{ \sum_{k=1}^{K} \sum_{t=1}^{N} \frac{(x_k[t] - \hat{x}_k[t|t-1])^2}{\gamma_k^2} \right\} \\
&= \arg \min_{\theta} \left\{ \sum_{k=1}^{K} \sum_{t=1}^{N} \frac{\hat{e}_k[t]^2}{\gamma_k^2} \right\}
\end{aligned} \tag{10}
$$

with $\hat{x}_k[t|t-1]$ denoting the model's one-step-ahead prediction, γ_k user defined weights, and argmin minimizing argument with respect to the indicated quantity. Toward this end, Eq. 9 may be rewritten in a regression type form as

$$
x_k[t] = -\sum_{i=1}^{n_a} \sum_{j=1}^{p} a_{i,j} \varphi_{\mathbf{d}(j)}(\boldsymbol{\xi}_k) x_k[t-i] + \sum_{i=1}^{n_c} \sum_{j=1}^{p} c_{i,j} \varphi_{\mathbf{d}(j)}(\boldsymbol{\xi}_k)[t-i] + e_k[t]
$$

$$
\underbrace{\begin{bmatrix} -\varphi_{\mathbf{d}(1)}(\boldsymbol{\xi}_k) x_k[t-1] \\ -\varphi_{\mathbf{d}(2)}(\boldsymbol{\xi}_k) x_k[t-1] \\ \vdots \\ -\varphi_{\mathbf{d}(p-1)}(\boldsymbol{\xi}_k) x_k[t-n_a] \\ -\varphi_{\mathbf{d}(p)}(\boldsymbol{\xi}_k) x_k[t-n_a] \\ \varphi_{\mathbf{d}(1)}(\boldsymbol{\xi}_k) e_k[t-1] \\ \varphi_{\mathbf{d}(2)}(\boldsymbol{\xi}_k) e_k[t-1] \\ \vdots \\ \varphi_{\mathbf{d}(p-1)}(\boldsymbol{\xi}_k) e_k[t-n_c] \\ \varphi_{\mathbf{d}(p)}(\boldsymbol{\xi}_k) e_k[t-n_n] \end{bmatrix}}_{\varphi[\boldsymbol{\xi}_k, t]^T} \cdot \underbrace{\begin{bmatrix} a_{1,1} \\ a_{1,2} \\ \vdots \\ a_{n_a,p-1} \\ a_{n_a,p} \\ c_{1,1} \\ c_{1,2} \\ \vdots \\ c_{n_c,p-1} \\ c_{n_c,p} \end{bmatrix}}_{\boldsymbol{\theta}} + \hat{\boldsymbol{e}}_k[t] \tag{11}
$$

or by stacking all time instants in a single vector

$$
\underbrace{\begin{bmatrix} x_k[1] \\ x_k[2] \\ \vdots \\ x_k[N] \end{bmatrix}}_{x_k^N} = \underbrace{\begin{bmatrix} \boldsymbol{\varphi}[\boldsymbol{\xi}_k,1]^T \\ \boldsymbol{\varphi}[\boldsymbol{\xi}_k,2]^T \\ \vdots \\ \boldsymbol{\varphi}[\boldsymbol{\xi}_k,N]^T \end{bmatrix}}_{\boldsymbol{\Phi}(\xi_k)} \cdot \underbrace{\begin{bmatrix} a_{1,1} \\ a_{1,2} \\ \vdots \\ a_{n_a,p-1} \\ a_{n_a,p-1} \\ c_{1,1} \\ c_{1,2} \\ \vdots \\ c_{n_c,p-1} \\ c_{n_c,p} \end{bmatrix}}_{\boldsymbol{\theta}} + \underbrace{\begin{bmatrix} e_k[1] \\ e_k[2] \\ \vdots \\ e_k[N] \end{bmatrix}}_{e_k^N}
$$

$$(12)$$

Finally, by pooling all the available datasets, the following regression model is obtained:

$$
\underbrace{\begin{bmatrix} \boldsymbol{x}_1^N \\ \boldsymbol{x}_2^N \\ \vdots \\ \boldsymbol{x}_K^N \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} \boldsymbol{\Phi}(\boldsymbol{\xi}_1) \\ \boldsymbol{\Phi}(\boldsymbol{\xi}_2) \\ \vdots \\ \boldsymbol{\Phi}(\boldsymbol{\xi}_K) \end{bmatrix}}_{\boldsymbol{\Phi}(\xi)} \cdot \boldsymbol{\theta} + \underbrace{\begin{bmatrix} \boldsymbol{e}_1^N \\ \boldsymbol{e}_2^N \\ \vdots \\ \boldsymbol{e}_K^N \end{bmatrix}}_{E}
$$

$$(13)$$

where $\boldsymbol{\Phi}(\xi)$ is the regression matrix. Thus, the minimization problem of Eq. 7 may be rewritten as follows:

$$
\boldsymbol{\theta} = \arg \min_{\boldsymbol{\theta}} \left\{ \boldsymbol{E}^T \cdot \boldsymbol{\Gamma} \cdot \boldsymbol{E} \right\} \tag{14}
$$

where $\boldsymbol{\Gamma} = \boldsymbol{I}_N \bigotimes \boldsymbol{\gamma}$ with $\boldsymbol{\gamma} = [\gamma_1 \cdots \gamma_k]^T$ denoting the vector of weights, $\boldsymbol{I}_N$ the $N \times N$ identity matrix, and $\bigotimes$ the Kronecker product.

Due to the nonlinear dependence of the a priori, unknown residual sequence $\boldsymbol{E}$ on the parameter vector $\boldsymbol{\theta}$, the last equation leads to a nonlinear-weighted least squares problem, which has to be tackled by nonlinear optimization methods. However, nonlinear least squares techniques are sensitive to the initial parameter values and if no accurate estimates are available, the minimization procedure is very likely to converge to a local minimum. In order to avoid potential inaccurate convergence problems associated with arbitrary initial estimates, initial values for the coefficients of projection may be obtained by identifying conventional ARMA models for each of the K data

sets and expanding their parameters onto the PC basis (Zhao et al. 2012).

However, in the simpler PC-AR case, the regression model of Eq. 13 is linear, and thus, $\boldsymbol{\theta}$ may be estimated by the linear weighted least squares (WLS) estimator:

$$
\hat{\boldsymbol{\theta}} = \left(\boldsymbol{\Phi}(\boldsymbol{\xi})^T \cdot \boldsymbol{\Gamma} \cdot \boldsymbol{\Phi}(\boldsymbol{\xi}) \right)^{-1} \left(\boldsymbol{\Phi}(\boldsymbol{\xi})^T \cdot \boldsymbol{\Gamma} \cdot \boldsymbol{Y} \right) \tag{15}
$$

The simplest and obvious choice for the weights γ_k is to set them equal to unity, with this selection leading to the ordinary least squares (OLS) estimator. However, it has been shown for the similar FP-ARX models (Kopsaftopoulos and Fassois 2006) that the weights that lead to consistent and asymptotically efficient estimates are those equal to the standard deviation of the residual sequence, that is,

$$
\gamma_k = \sigma_{e_k}, \quad \text{for } k = 1, 2, \ldots, K \tag{16}
$$

Due to the fact that the true residual standard deviation is typically unknown, it has to be replaced by an appropriate estimate, typically obtained via the OLS estimation which is used as a first stage. The WLS estimator may then be realized in an iterative manner till convergence of the PE criterion is achieved.

Model Structure Selection

Model structure selection is the optimization procedure during which models corresponding to various candidate "structures" are estimated and weighted according to an appropriately selected fitness criterion. Model "fitness" may be considered in terms of various criteria, such as the residual sum of squares (RSS)

$$
\text{RSS} = \sum_{k=1}^{K} \sum_{t=1}^{N} e_k[t]^2 \tag{17a}
$$

or the equivalent mean squared error (MSE)

$$
\text{MSE} = \frac{\text{RSS}}{NK} \tag{17b}
$$

However, the limitation of using such criteria is that they may monotonically decrease with

increasing model structure and as a result may lead to overfitting. For this reason, caution has to be exercised and a compromise has to be made between model accuracy and model structure complexity.

The model structure selection procedure for a PC-ARMA model concerns the determination of the set of 3 integers (n_a, n_c, P) for obtaining the best fitting model, and a combinatorial approach, that is, estimating PC-ARMA models for various integer sets, may be used. However, a simpler approach that may lead to sparse PC basis by selecting only a subset of the initial PC functional subspace may be described as follows (Poulimenos and Fassois 2006):

Phase I. AR/MA orders selection. In order to "decouple" the selection of the model orders (n_a, n_c) from that of functional subspaces, their interaction has to be minimized. For this reason a high-degree P and the complete set of PC basis functions may be initially adopted. When employing these, AR/MA model order selection may be achieved through trial-and-error techniques based on the values of the fitness function.

Phase II. PC basis subspace selection. The aim of this phase is the optimization of the PC functional subspace, in the sense of reducing model complexity without significantly reducing model accuracy. This may be accomplished by dropping, one at a time, the PC basis functions whose removal does not cause significant loss of the criterion employed.

A flowchart of the complete PC-ARMA identification procedure is given in Fig. 1.

Validation

Once a PC-ARMA model has been obtained, it must be validated with respect to the assumptions behind the identification method. A standard validation procedure in the simple ARMA case involves the testing of Gaussianity and, in particular, the whiteness of the identified model's one-step-ahead prediction error sequence. In the PC-ARMA case this assumption has to be validated for each of the recorded datasets, while the
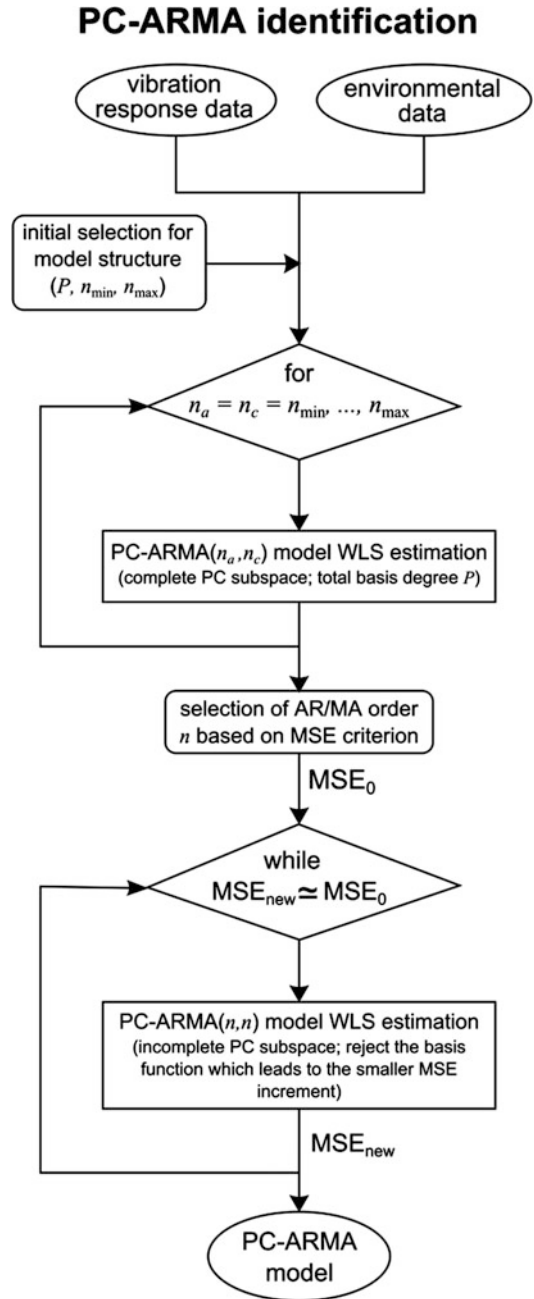
Stochastic Structural Identification from Vibrational and Environmental Data, Fig. 1 Flowchart of the PC-ARMA model identification procedure

additional assumption of cross-uncorrelatedness between the datasets has to be validated.

The validation procedure of the estimated PC-ARMA model may be based on the typical
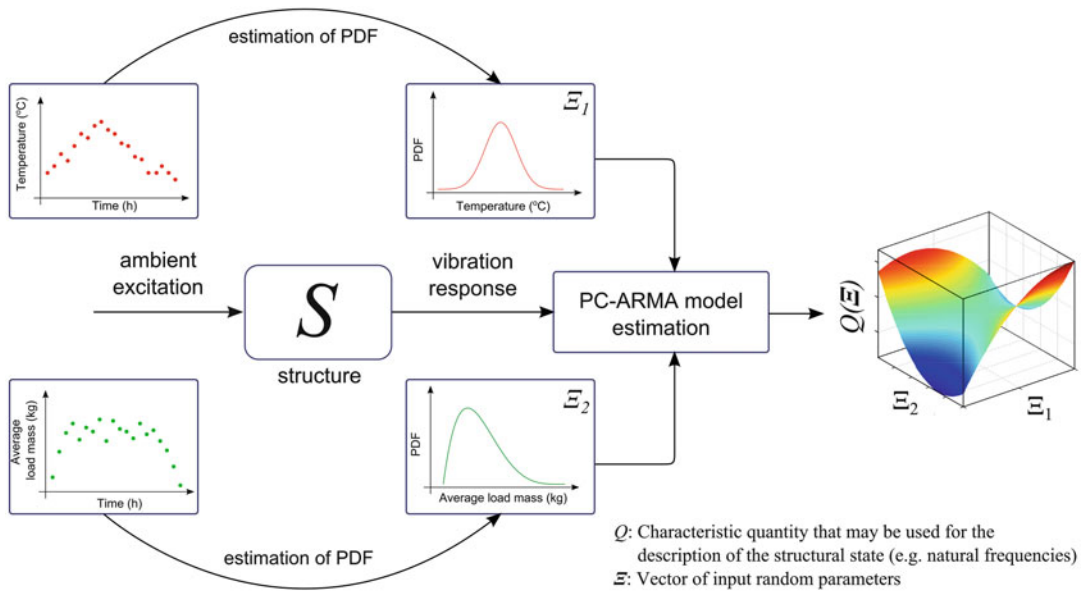
Stochastic Structural Identification from Vibrational and Environmental Data, Fig. 2 The complete stochastic structural identification framework based on PC-ARMA models

cross-validation principle which presumes the separation of the processed signal into an estimation set and a smaller validation set of data (Ljung 1999).

Summary of the Method

Step 1. Collect K datasets of vibration response and environmental variables measurements from the structure under study for a range of operating conditions.

Step 2. Fit a probability density function to the environmental variables based on the available data and select the corresponding family of PC basis functions.

Step 3. Select a high total degree for the PC basis P; estimate PC-ARMA models of various orders (e.g., considering common AR/MA order, from $n_a = n_c = n_{\min}$ up to $n_a = n_c = n_{\max}$); and select the final orders based on the "fitness" criterion (e.g., MSE).

Step 4. Estimate PC-ARMA (n_a, n_c) models by rejecting each time a number of PC basis function from the complete PC subspace. Reject the functions whose removal does not imply significant reduction of the "fitness" criterion value.

The complete stochastic structural identification framework based on PC-ARMA models is illustrated in the abstract schematic representation of Fig. 2. It is noted that the estimated PC-ARMA model parameters may be used for the calculation of the structure's modal characteristics, such as the modal frequencies and damping ratios, as a function of the input variable vector Ξ for the entire event space Ω.

Numerical Case Study

The modeling of a 3-DOF system (Fig. 3) is presently considered for demonstrating the workings of the PC-ARMA identification framework. The mechanical properties of the system are:

$$m_1 = 10^2, \quad m_2 = 10^{1.5}, \quad m_3 = 10 \ (\text{kg})$$

$$k_1 = 10^6, \quad k_2 = 10^{5.5}, \quad k_3 = 10^5 \ (\text{N/m})$$

$$c_1 = 100, \quad c_2 = 10, \quad c_2 = 10, \ (\text{N/ms}^{-1})$$

The system is considered to operate under varying conditions of temperature and humidity.
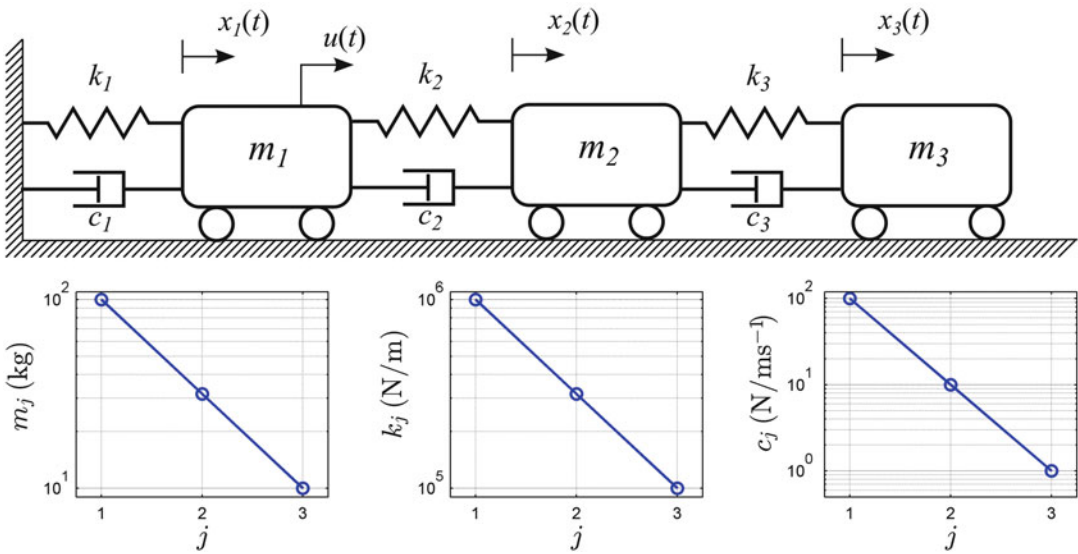
Stochastic Structural Identification from Vibrational and Environmental Data, Fig. 3 Three degree of freedom system and its properties

The system's stiffness coefficients are considered to be linearly depended on the temperature through the following relationship:

$$k_i' = k_i[1 - (T - T_{\mathrm{ref}})k_T]$$

with $k_T = 5 \times 10^{-3}$ denoting the thermal coefficient and $T_{\mathrm{ref}} = 20\ {}^\circ\mathrm{C}$ the reference temperature, while humidity is modeled as mass increment due to trapped humidity. For the sake of simplicity, in this test case study, it is assumed that the temperature is uniform with no spatial gradients or temporal variations within each experiment, while the mass load is evenly distributed over the system.

In order to identify an inverse model for this system under various operating conditions, the system's response (acceleration of mass m_3) is simulated for various operating condition scenarios. Specifically, 50 samples of the added mass and temperature variables are drawn, by means of the Latin hypercube sampling (LHS) method, from their theoretical pdfs which are considered to correspond to a log-normal distribution with mean value equal to 1 and variance equal to 0.25 for the mass load, that is, $\Xi_1 \sim \ln N(1, 0.25)$, and a normal distribution with mean value equal to

15 and variance equal to 25 for the temperature, that is, $\Xi_2 \sim N(15, 25)$. The 50 samples of the input variables are shown in Fig. 4.

In a real-world structural identification application, where no information is available regarding the true pdfs of the input random vector, someone could use maximum likelihood estimation fitting of the environmental condition data values to a parametric distribution. The results of such a fitting of the data onto pdfs are shown in Fig. 5. Based on this fitting and after transforming the pdf of the mass load into a normal distribution by using the natural logarithm, the Hermite polynomials may be selected for the construction of the multivariate PC basis functions.

For each set of these "environmental" conditions, a white noise realization of the excitation force $u(t) \sim N(0, 10^4)$ is used for the excitation of the system. System's response is obtained by using the Runge-Kutta algorithm for 1,800 points with sampling frequency equal to 64 Hz. After discarding the first 1,000 samples in order to avoid transient effects, the rest of the data are separated into an estimation set consisting of 500 samples and a validation set of 300 samples.
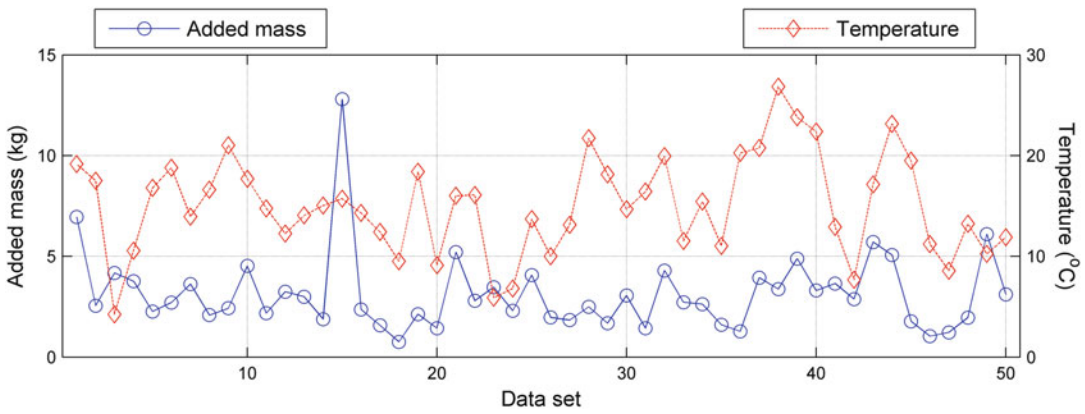
Stochastic Structural Identification from Vibrational and Environmental Data, Fig. 4 Input variable values for the 50 simulation experiments conducted
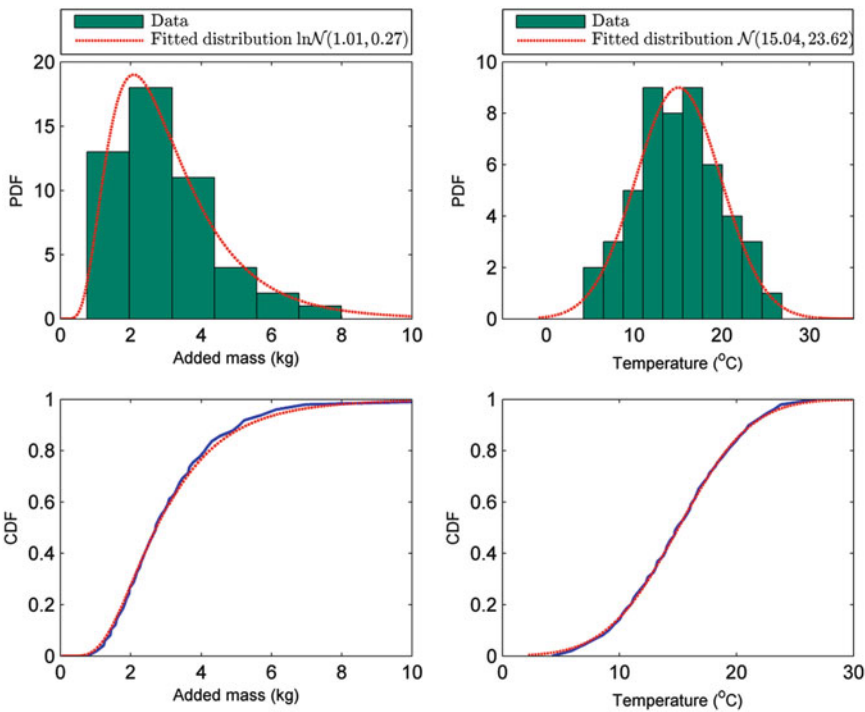
Stochastic Structural Identification from Vibrational and Environmental Data, Fig. 5 The empirical density and distribution functions calculated based on the 50 realizations of the input variables, compared to the corresponding probability density and cumulative distribution functions estimated by the maximum likelihood method

The estimation dataset (50 sets $\times$ 500 samples) is used for the identification of the system through simple PC-AR models in what follows. PC-AR model structure selection is based on the MSE criterion and the procedure described in section 2.2. The maximum total degree of the PC multivariate basis is initially selected equal to 5, which using Eq. 9 with $P = 5$, and
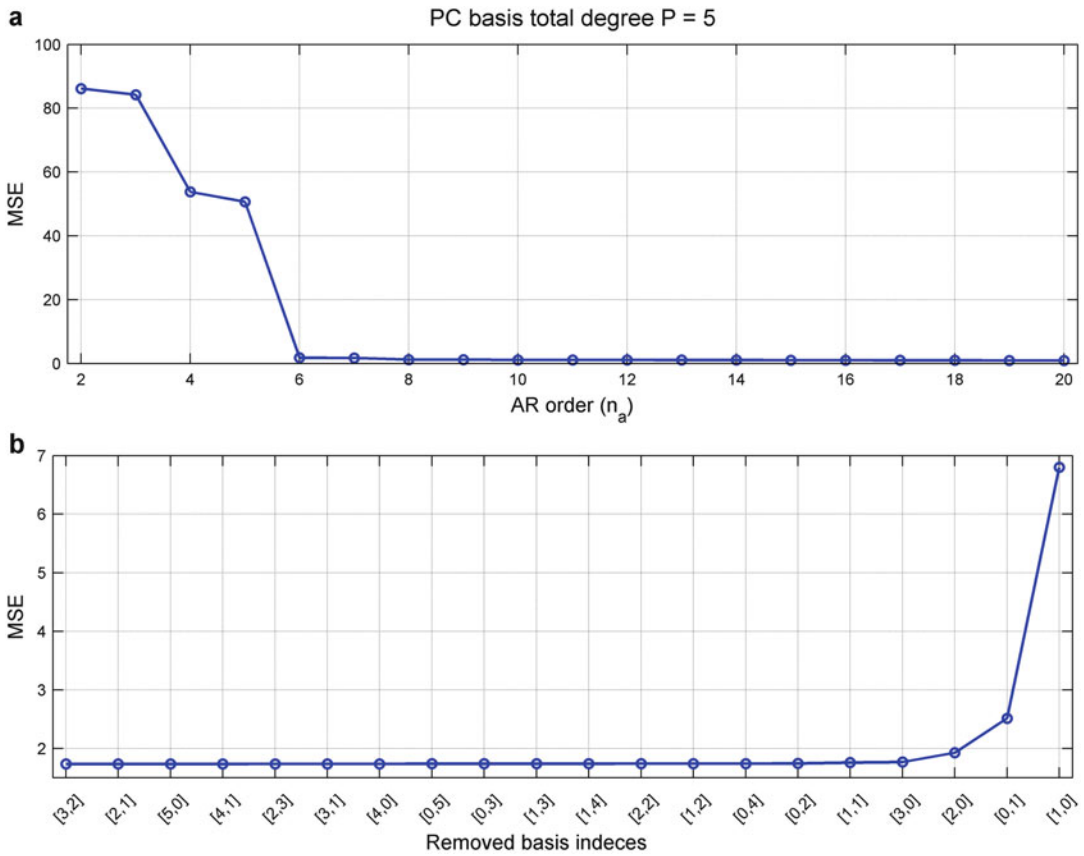
Stochastic Structural Identification from Vibrational and Environmental Data, Fig. 6 PC-AR model structure selection results: (**a**) MSE criterion values for WLS-based estimated PC-AR models with $n_a = 2$, 3, ..., 20 and $P = 5$, and (**b**) MSE criterion values obtained from a PC-AR(6) model by dropping, one at a time, the PC basis functions of the initial subspace

$L = 2$ gives a total number of 21 candidate basis functions. These functions are firstly used for the AR order selection by estimating PC-AR models of various orders ($n_a = 2, 3, \ldots, 20$) by means of the WLS estimator. The results illustrated in Fig. 6a show an AR order equal to six as appropriate for capturing the dynamics of the simulated system. At the second stage of the structure selection procedure, a sparse PC basis is sought. As shown in Fig. 6b, this comes down to a basis of only 4 basis functions which are given in Table 1. It may be observed that the basis indicates a linear relationship of the PC-AR model parameters with respect to temperature and quadratic with respect to humidity (mass load).

Stochastic Structural Identification from Vibrational and Environmental Data, Table 1 Multivariate indices of the finally selected PC basis functions of the PC-AR (6) model

	l	
	1	2
$d(1)$	0	0
$d(2)$	1	0
$d(3)$	0	1
$d(4)$	2	0

The estimated PC-AR(6) model is successfully validated by calculating the one-step-ahead prediction errors for the validation dataset (50 sets × 300 samples). Indicative results for the autocorrelation function of the resulting

Stochastic Structural Identification from Vibrational and Environmental Data, Fig. 7 PC-AR(6) model validation results. (**a**) The autocorrelation function for the prediction error sequence of the validation set (simulation experiment 2; time lags 100; 95 % confidence limits) and (**b**) the covariance matrix of the prediction error matrix E
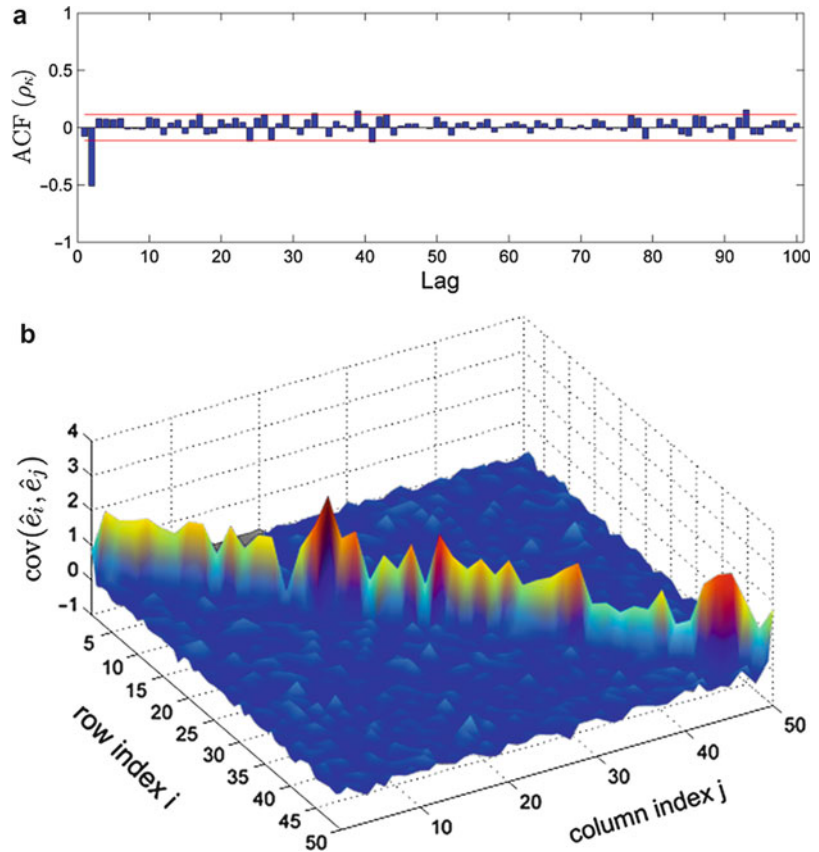
prediction error sequence corresponding to the second simulation experiment are shown in Fig. 7a, while also the assumption of uncorrelatedness between the various prediction error sequences may be qualitatively validated by visually inspecting the residual covariance matrix which is shown in Fig. 7b.

Finally, the estimated coefficients of projection may be used for calculating the PC-AR(6)-based estimates of the system natural frequencies and thus for constructing a surface that shows the dependence of the natural frequencies of the 3-DOF system on the two input random variables. The results are plotted against the theoretical curves in Fig. 8. It may be observed that there is a good agreement between the theoretical and the estimated natural frequencies with the highest errors, which are below 1 Hz, occurring at extreme values of the input variables which were not included in the estimation set.

Summary

Mechanical properties, and thus the dynamic behavior of an engineering structure, are significantly affected by environmental and operational conditions. A complete identification procedure has to lead to a consistent representation of the structural dynamics for a great percentage, if not the whole, range of the structure's operational spectrum. This may be achieved either by a multi-model approach, that is, estimating local models for every new set of environmental conditions data or using functional models that incorporate into their parameters the structure's dependency on a set of random input variables which correspond to the measured environmental conditions. In this way, after the estimation (training) of the functional model on a given set of measurements, its parameters may be extrapolated for any new set of environmental conditions. Presently, the functional models used are of the
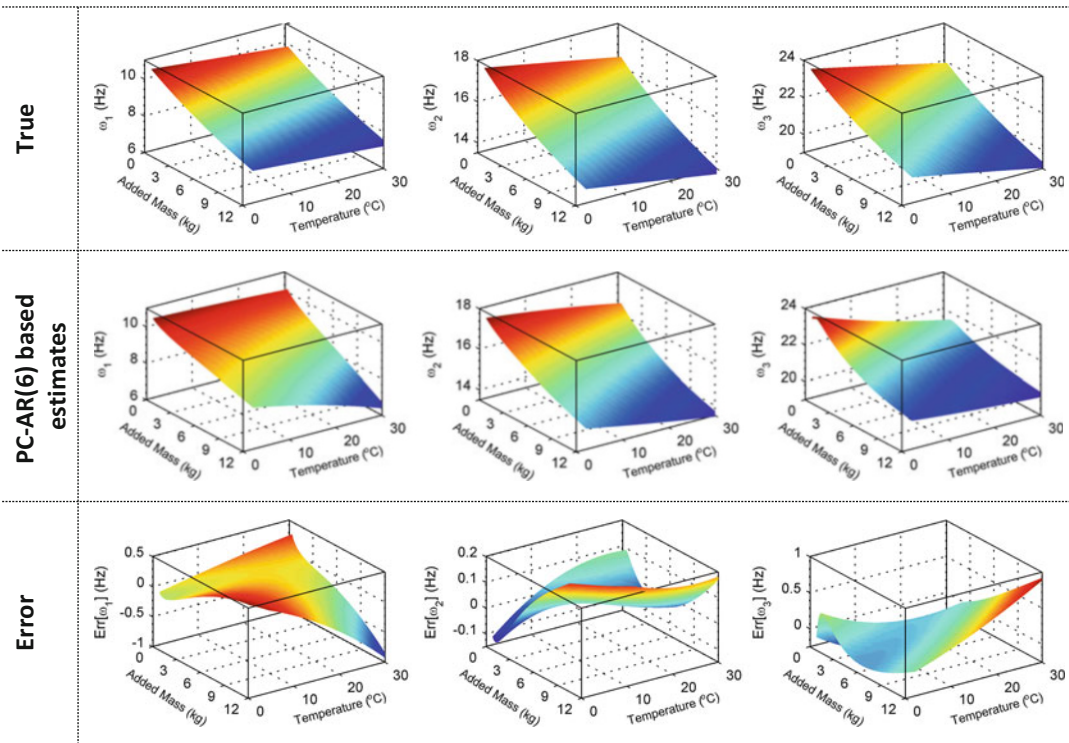
Stochastic Structural Identification from Vibrational and Environmental Data, Fig. 8 PC-AR(6)-based natural frequency estimates compared to the theoretical curves. The natural frequencies are plotted versus the input variables, that is added mass and temperature

ARMA form with parameters that are expanded on polynomial chaos basis. The latter has the important feature of being orthogonal to the probability density function of the input variable vector, providing in this way the best convergence rate with increasing functional subspace.

Cross-References

▶ Operational Modal Analysis in Civil Engineering: An Overview
▶ Response Variability and Reliability of Structures
▶ Stochastic Analysis of Linear Systems
▶ System and Damage Identification of Civil Structures
▶ Uncertainty Quantification in Structural Health Monitoring
▶ Vibration-Based Damage Identification: The Z24 Bridge Benchmark

References

Andersen P (1997) Identification of civil engineering structures using vector ARMA models. Aalborg University, Department of Mechanical Engineering

Blatman G, Sudret B (2010) An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. Probabilist Eng Mech 25:183–197

Doebling S, Farrar C, Prime M (1998) A summary review of vibration-based damage identification methods. Shock Vib Dig 30:91–105

Fassois SD, Kopsaftopoulos FP (2013) Statistical time series methods for vibration based structural health monitoring. In: Ostachowicz W, Guemes A (eds) New trends in structural health monitoring. Springer, Wien/London, pp 209–264

Kopsaftopoulos FP, Fassois SD (2006) Identification of stochastic systems under multiple operating conditions: the vector dependent FP–ARX parameterization. In: Proceedings of the 14th Mediterranean conference on control and automation. Ancona

Kopsaftopoulos FP, Fassois SD (2013) A functional model based statistical time series method for vibration based damage detection, localization, and magnitude estimation. Mech Syst Signal Process 39:143–161

Ljung R (1999) System identification: theory for the user, 2nd edn. Prentice Hall, Upper Saddle River

Poulimenos AG, Fassois SD (2006) Parametric time-domain methods for non-stationary random vibration modelling and analysis: a critical survey and comparison. Mech Syst Signal Process 20:763–816

Sohn H (2007) Effects of environmental and operational variability on structural health monitoring. Philos Trans R Soc A 365:539–560

Soize C, Ghanem R (2004) Physical systems with random uncertainties: chaos representations with arbitrary probability measure. SIAM J Sci Comput 26:395–410

Tóth R (2010) Identification and modeling of linear parameter-varying systems. Lecture notes in control and information sciences, vol 403. Springer, Berlin

Xiu D, Karniadakis GE (2002) The Wiener–Askey polynomial chaos for stochastic differential equations. SIAM J Sci Comput 24:619–644

Zhao Y, Huang B, Su H, Chu J (2012) Prediction error method for identification of LPV models. J Process Control 22:180–193

Strengthened Structural Members and Structures: Analytical Assessment

Georgia E. Thermou
Department of Civil Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece

Synonyms

Assessment; Buildings; Interface; Jacketing; Repair; Retrofitting; Shear transfer; Strengthening; Structural interventions; Structures

Introduction

The catastrophic earthquakes that struck several countries over the last 20 years brought to light the susceptibility of the existing building stock (Fig. 1). Old-type reinforced concrete (R/C) structures are characterized by insufficient reinforcement detailing (lack of stirrups for ensuring a certain ductility level, indirect supports, insufficient anchorages of bars), nonuniform distribution of stiffness and/or mass along the height of the building, insufficient foundation system, poor quality of materials, and various other weaknesses such as increased loading due to change of use and corrosion of reinforcement.

The majority of multistory R/C buildings in southern Europe were built in the first half of the twentieth century. Structures were designed for gravity loads only by implementing the allowable stress design philosophy which did not allow any control of the mode of failure and the corresponding deformation capacity of the individual members. Taking Greece as an example, the first seismic code was introduced in 1959, and R/C walls were introduced in construction in the 1960s (often without extending up to the foundation of the building). The modern seismic codes were introduced more than 20 years later, in the mid-1980s. The multistory R/C buildings of the 1950s represent the cutting edge of the construction technology for gravity load-designed frame buildings. Information regarding the material, detailing, and geometrical characteristics of representative low- to medium-rise (up to 8 stories) R/C buildings real structures found in the urban areas built between the 1920s and 1960s is presented in Table 1 (Thermou and Palaioxorinou 2013). The lack of a continuous vertical load path along the height of the buildings is a common feature of the buildings of that era. There is no typical floor since the dimensions of the columns and beams change from story to story. Often in-plan column layout does not follow a grid pattern, hence leading to indirect supports. Representative typical floor plan layouts are presented in Fig. 2.

The result of these systematic deficiencies of existing buildings is a decreased level of seismic protection, increased seismic vulnerability, and hence extensive damage expected in future seismic excitations. This is a rather alarming issue considering the socioeconomic impact of severely damaged buildings or collapses in future strong ground motions. The recommended solution for this category of buildings, which comprises the vast majority of the existing stock, is retrofitting (strengthening) with a view to upgrading their seismic capacity and meeting the current standards for seismically designed structures. Adopting global intervention (GI)

S

Strengthened Structural Members and Structures: Analytical Assessment, Fig. 1 Damage patterns in old-type R/C members and structures; typical multistory R/C residential buildings with pilotis in Greece (Source of figures: author's personal files)

Strengthened Structural Members and Structures: Analytical Assessment, Table 1 Information of the typical construction practice followed in the 1920s–1960s in Greece (Thermou and Palaioxorinou 2013)

Construction practice characteristics	
Concrete grade	B120 ÷ B160[a] ($f_{ck} \cong$ 8–10 MPa)
Steel grade (longitudinal and transverse reinforcement)	Smooth StI ($f_{yk} =$ 220 MPa)[b]
Typical column cross section dimensions	250 ÷ 600 mm
Typical diameters of column longitudinal reinforcement bars	Ø14 ÷ Ø20
Typical column transverse reinforcement	Ø6/250 ÷ 300 mm[c]
Typical column longitudinal reinforcement ratio	7‰ ÷ 9‰ < 10‰
Typical beam cross section dimensions	150 × 300 ÷ 300 × 600 mm
Typical diameters of beam longitudinal reinforcement bars	Ø10 ÷ Ø18
Typical beam transverse reinforcement	Ø6/200 ÷ 250 mm[c]
Typical thickness of masonry infill walls	100 mm for internal, 200 mm for external
Typical wall cross section dimensions	Length: 1.5 ~ 3.6 m, thickness: 150 ~ 250 mm, boundary elements: length 200 mm, 4Ø12, web: #Ø8/250
Anchorage/lap splice construction practice	Longitudinal reinforcement with hooks with arbitrary lengths
	Stirrups anchored with 90° hooks
	Unconfined lap splices

[a]As per Greek Royal Decree of 18/2/1954. Later concrete grade B225 ($f_{ck} \cong$ 14 MPa) was introduced in construction
[b]As per DIN 1045 (1936). Later longitudinal reinforcement steel grade was changed to StIII ($f_{yk} =$ 400÷420 MPa)
[c]Diameter Ø8 was rarely applied

methods (e.g., R/C jacketing, R/C infill walls) leads to the modification of strength and stiffness (Fig. 3). They are used in systems with a high flexibility to sway, in torsionally unbalanced systems, or where strengthening of the existing building is required (e.g., increase of base shear strength in buildings with an open first story, Fig. 1). GI methods cannot be used separately from local intervention (LI) methods since they are used to accommodate premature failure modes (Thermou et al. 2012a). LI methods as, for example, the intervention methods with composite materials are considered not to alter the stiffness or flexural strength of the retrofitted
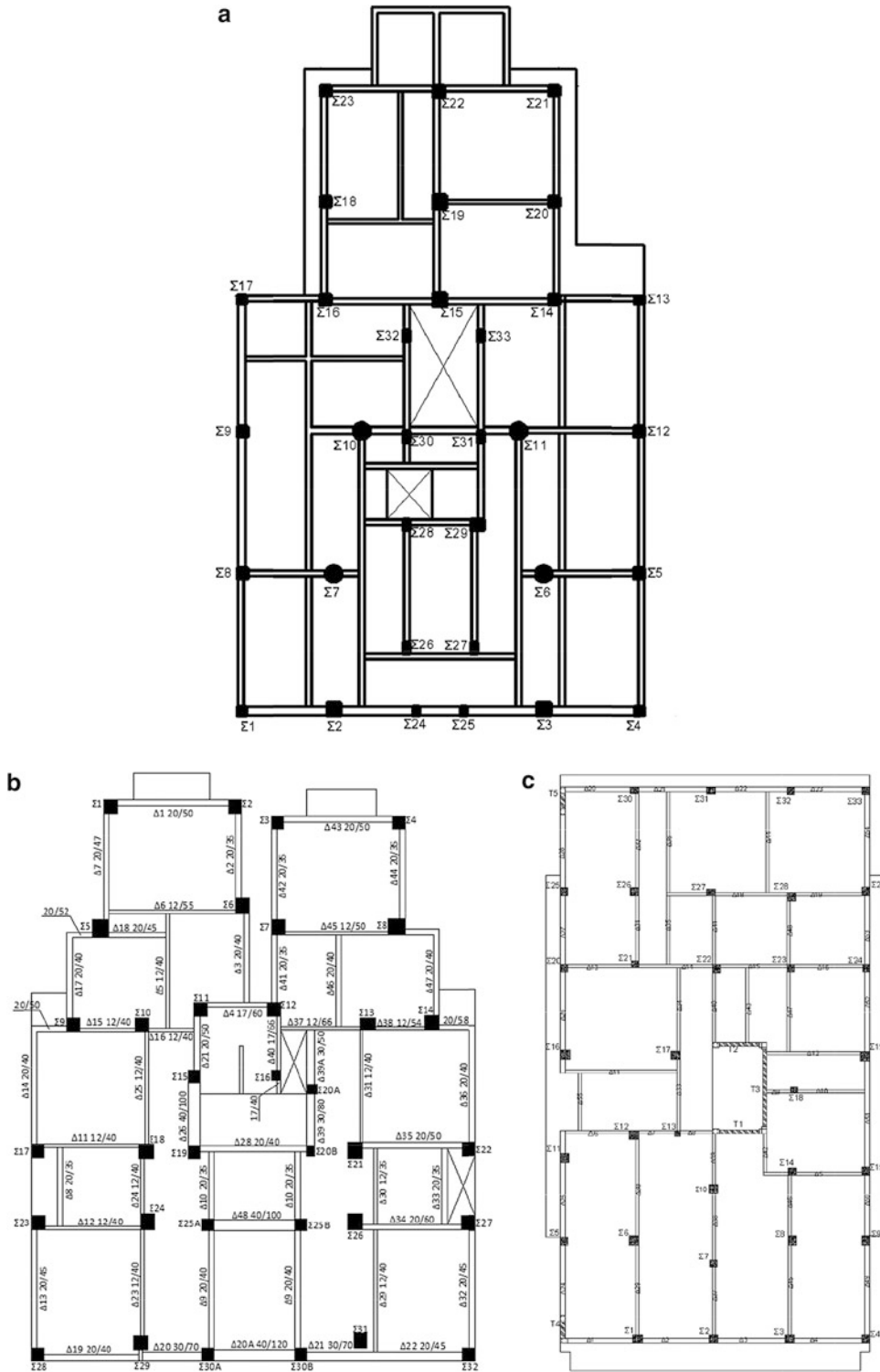
Strengthened Structural Members and Structures: Analytical Assessment, Fig. 2 Representative ground floor plan layouts of the R/C multistory buildings in Greece of the (**a**) 1920s, (**b**) 1950s, and (**c**) 1960s
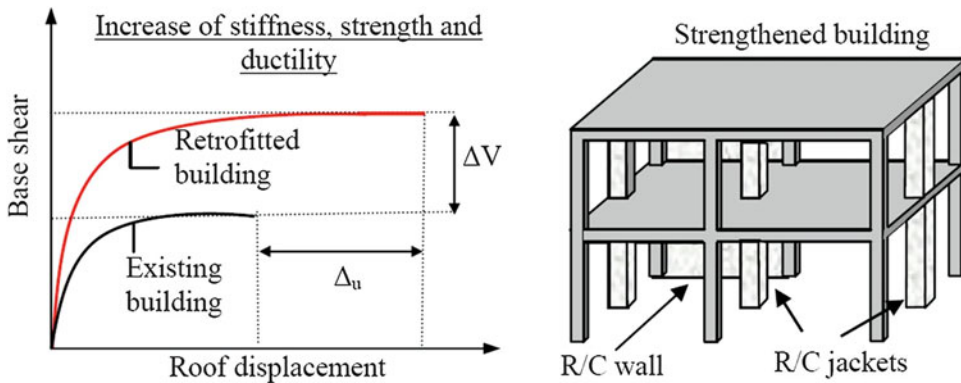
Strengthened Structural Members and Structures: Analytical Assessment, Fig. 3 Retrofitting strategy: stiffness, strength, and ductility increase (Thermou and Elnashai 2006)

members, but only affect their post-yielding deformation capacity through confinement and by suppressing premature modes of failure (e.g., brittle shear or lap splice failure).

The decision on the extent of the interventions (i.e., which floors, which elements of the floor, and which design parameters are going to be modified) and its impact on the modification of selected global response indices may be assessed through analytical modeling of the strengthened structure. This implies that there should sufficient information relative to the effectiveness of the intervention methods adopted at member level. The addition of a new concrete layer (e.g., flexural strengthening of beams) or a new R/C member (e.g., R/C jacketing, R/C infill walls) in various retrofit techniques entails issues related to the connection between existing and newly cast concrete. The response of the composite member, and subsequently of the whole structure, depends largely on the response characteristics of the interface, since slip takes place and shear transfer mechanisms are mobilized. Estimating response indices such as stiffness and strength in a reliable manner requires the development of analytical models that will consider the phenomena that take place along interfaces due to slip and their interaction.

This contribution aims to present the framework for the analytical assessment of strengthened structures when the intervention methods adopted result in composite members with multiple monolithic phases. The slippage introduced due to the discontinuity in the normal strain along the interfaces is rather determinant for the response at member level. One of the objectives is to highlight the role of the interface characteristics and present the parameters that influence the interconnection between the existing and new member. The shear transfer mechanisms mobilized due to relative slippage of the contact interfaces are presented in the following, and their interaction is discussed. Information is also provided for the influence of cyclic loading on the shear resistance mechanisms. The fundamentals for the development of analytical models which take into account slip at the interfaces are discussed. The interrelation between the effects that the changes at local level may have on global level is identified as a key issue in the assessment of strengthened structures. Moreover, the alternative simplified design approach suggested by codes, where the monolithicity factors are utilized as to indirectly account for the adverse effect of slippage on response, is presented. The last section of this entry focuses on the alternative analytical design approaches that could be followed for one of the most popular global intervention techniques, the R/C jacketing technique.

Implications of the Addition of New Concrete Layers or New Elements

The response of a composite beam, as the one shown in Fig. 4, depends on the degree of

Strengthened Structural Members and Structures: Analytical Assessment,
Fig. 4 Strain distribution profiles for (a) no connection; (b) perfect and (c) partial connection between beam A and beam B. (d) Shear flow along the interface
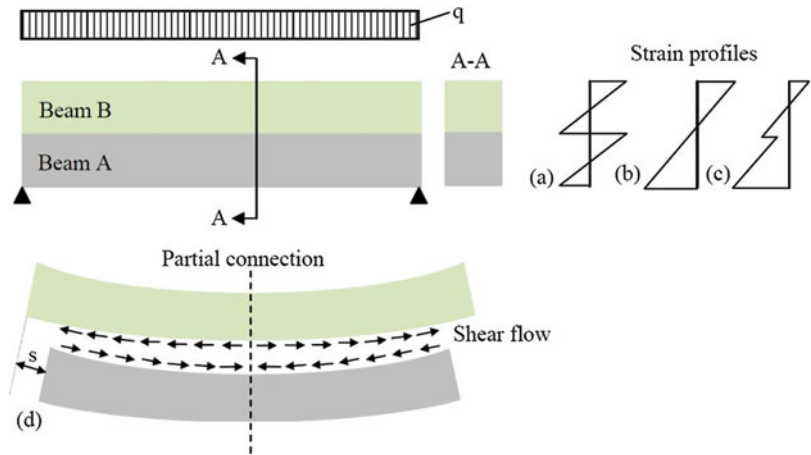
interconnection between beam A and beam B. In one extreme case, the case of zero friction (no connection), beam B could slide relative to beam A without mobilizing any kind of shear resistance (Fig. 4a). In an ideal situation, the composite members (beam A and beam B) would behave as monolithic (perfect connection, Fig. 4b). In real conditions, partial connection between the two bodies is expected, which implies that relative slip, s, between the two bodies is allowed, thus mobilizing the shear transfer mechanisms at the interface (Fig. 4c). The discontinuity in the strain profile corresponds to the slip, s, at the horizontal interfaces (Fig. 4c). Curvature, which is the slope of the strain profile on the cross section, is the same between the various slipping parts. This is a common assumption in the analysis of composite (layered) beams (e.g., glulam timber beams, composite steel beams, etc.). The shear flow, T, at the interface is estimated by

$$T = \frac{V \cdot S}{I} \qquad (1)$$

where S is the first moment of area, I is the moment of inertia of the composite cross section, and V is the shear force on the member.

Flexural enhancement of beams with the addition of a new reinforced concrete layer in the tension zone follows the principles presented for the composite beam of Fig. 4. The same approach can also be followed in the case of composite cross sections where an outer shell is placed around the core of the cross section, as in the case of R/C jacketing of an existing column, or when an infill wall is added to an existing bay by incorporating the existing columns.

The preparation of the interface and the interconnection measures taken between existing and new member play a key role in the interfacial force transfer and relative slip occurring along the contact surface. The bond between the monolithic phases of the composite member can be enhanced by roughening of the substrate surface, addition of dowels placed through holes driven on the substrate surface, and when applicable (e.g., R/C jacketing technique) welding of the new longitudinal bars to the existing ones through deformed reinforcing bars bent into a U-shape. Moreover, experimental evidence has shown that differential shrinkage and stiffness of new-to-old concrete interfaces, as well as the compressive strength of the added concrete to an existing concrete substrate, influence the response of composite system where concrete-to-concrete interfaces exist (e.g., Choi et al. 1999). For example, in the experimental study conducted by Júlio et al. (2006), it was found that the increase of the concrete compressive strength of the new layer compared to that of the existing one leads to an enhancement of the compressive strength of the interface, i.e., taking into account the compressive strength of the weakest concrete seems to be conservative. Recently, Santos and Júlio (2014) proposed

Strengthened Structural Members and Structures: Analytical Assessment, Fig. 5 Slip at a concrete interface crossed by reinforcement
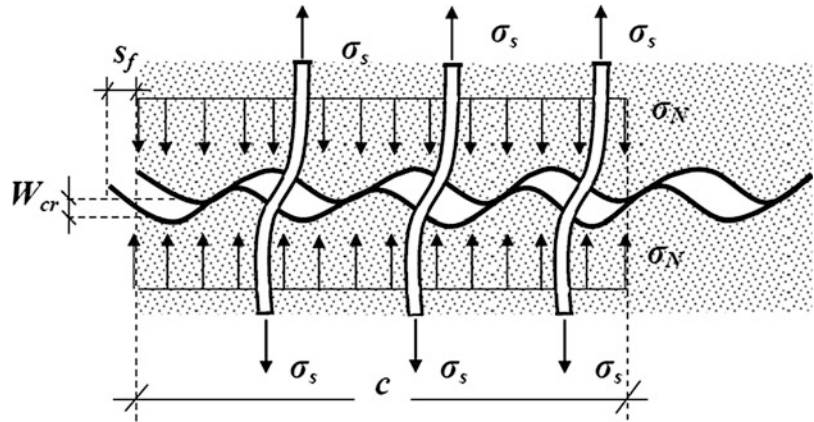
design expressions where the effects of differential shrinkage and differential stiffness at the concrete-to-concrete interface are considered.

Shear Transfer Along Concrete-to-Concrete Interfaces

Shear transfer along interfaces has been a subject of continuous research since the 1970s. A variety of analytical models are available for modeling the main shear transfer mechanisms (friction and dowel resistance) by considering that they act separately or jointly. There are models where all forces are transferred through reinforcement (e.g., Birkeland and Birkeland 1966; Walraven 1981), whereas others that, apart from reinforcement contribution, include a cohesion term (e.g., Mattock and Hawkins 1972; Vecchio and Collins 1986; Tassios 1986). In the modified compression field theory (MCFT) (Vecchio and Collins 1986), the role of the cohesion term is played by the aggregate interlock mechanism, whereas in the model of Tassios (1986), by friction owing to the clamping action of reinforcement normal to the interface. In the rest of the models, the cohesion term corresponds to the friction resistance developed along the interface.

Shear transfer is affected by the roughness of the sliding planes, by the characteristics of the reinforcement, by the compliance of concrete, and by the state of stress in the interface zone.

Mechanisms that resist sliding (slip) are (i) aggregate interlock between contact surfaces, including any initial adhesion of the jacket concrete on the substrate; (ii) friction owing to the clamping action of reinforcement normal to the interface; and (iii) dowel action of any properly anchored reinforcement crossing the sliding plane. Dowel action develops by three alternative mechanisms, namely, by direct shear, by kinking, and by flexure of the bars crossing the contact plane. The relationship that describes the contribution of the individual shear transfer mechanisms is

$$
\begin{aligned}
\tau_{\mathrm{tot}} &= \tau_{\mathrm{agr}} + \tau_f + \tau_D = \tau_{\mathrm{agr}} + \mu\sigma_N + \tau_D \Leftrightarrow \\
\tau_{\mathrm{tot}} &= \tau_{\mathrm{agr}} + \mu(\sigma_c + \rho\sigma_s) + \tau_D \\
&= \tau_{\mathrm{agr}} + \mu(vf_c + \rho\sigma_s) + \tau_D
\end{aligned}
$$

(2)

where τ_{agr} represents the shear resistance of the aggregate interlock mechanism, μ is the interface shear friction coefficient, σ_N is the normal clamping stress acting on the interface, and τ_D is the shear stress resisted by dowel action, F_D, in cracked reinforced concrete. The clamping stress represents any normal pressure, p, externally applied on the interface, but also the clamping action of reinforcement crossing the contact plane, σ_s is the axial stress of the bars crossing the interface, ρ is the corresponding reinforcement area ratio, $v = \mathrm{N}/(\sigma_c f_c) = \sigma_c/f_c$ is the normalized axial load at the interface of

area A_c, and f_c is the concrete compressive strength (Fig. 5). The first two terms in Eq. 2 collectively represent the *contribution of concrete* as they depend on the frictional resistance of the interface planes.

As it is mentioned in Model Code (2010) and the Greek Code for Interventions (KANEPE 2013), in real structures the various mechanisms interact, thereby affecting each other as a function of the shear slip. The aggregate interlock including any initial adhesive decreases at low slip values, whereas the maximum contribution of frictional and dowel resistance occurs at different slips. Therefore, Model Code (2010) as well as the Greek Code for Interventions (KANEPE 2013) introduce interaction factors which consider that the dowels crossing the interface are subject to bending and axial forced simultaneously and that the maximum values of the different shear resistance mechanisms occur at different slips. The interaction factors depend on various parameters such as the magnitude of the expected slip at the interface, the diameter of the reinforcing bars crossing the interface, the concrete compressive strength, cyclic loading, etc. Equation 2 is thus modified according to Model Code (2010) as follows:

$$\tau_{\text{tot}} = \tau_{\text{agr}} + \mu(\sigma_c + \kappa \cdot \rho\sigma_s) + \alpha \cdot \rho \cdot \sqrt{f_y \cdot f_c} \leq \beta \cdot v \cdot f_c \quad (3)$$

where the coefficients κ and α refer to the interaction factors for interface roughness and flexural resistance, respectively, whereas β is a coefficient related to the compressive struts. For the case where dowels are driven into the interface, $\kappa = 0.5$ for roughened interfaces and $\kappa = 0$ for smooth interfaces, whereas in the latter case $\alpha = 1.5$ denoting that the main resistance mechanism is provided by the dowel action. Model Code (2010, Chap. 7) provides detailed information regarding the design of the interface in shear and how the interaction coefficients are modified for different surface roughness.

The Greek Code for Interventions (KANEPE 2013) considers also the interaction between the

shear transfer mechanisms by estimating the shear resistance of the interface as

$$\tau_{\text{tot}} = \beta_F \tau_f + \beta_D \tau_D \quad (4)$$

where the β_F and β_D are the interaction factors of the friction and dowel resistance, which are equal to $\beta_F = 0.4$ and $\beta_D = 0.7$ for slip values $s \leq 1.00$ mm. These coefficients can be further decreased to 0.5 when either the expected slip value cannot be predicted or when the external compressive force at the interface is negligible.

The friction resistance, τ_f (Eq. 2), is calculated by considering the coefficient of friction, μ, which depends on the classification of the roughness of the interface. The majority of the experimental studies provide a qualitative description of the roughness of the interface after application of the selected type of treatment (e.g., sandblasting, use of pneumatic chipping device). It is obvious that the characterization by visual inspection only may lead to inaccuracies. Therefore, the roughness of the interface should be quantified by specific indicators. Although several methods have been developed that can successfully quantify the roughness of concrete interfaces (Santos and Júlio 2013), design codes in their majority still provide qualitative descriptions of the interface (e.g., smooth, rough, or very rough). An exception is Model Code (2010) which suggests two methods: the sand area method and the one where the average roughness is estimated (i.e., measuring the average deviation of the profile from a mean line).

From the above, it is seen that describing in detail the mechanisms mobilized along interfaces due to slip and their interaction is a complex mechanical issue which becomes even more complicated considering the response of the interface under cyclic loading conditions where degradation should also be accounted for. The research conducted on the derivation of constitutive models that describe the combined shear force resistance mobilized along interfaces due to sliding both under cyclic imposed deformations is limited. Palieraki et al. (2012) based on the previous studies of Tassios and Vintzileou (1987) and Vintzileou and Tassios (1986, 1987)

S

Strengthened Structural Members and Structures: Analytical Assessment,
Fig. 6 Response to symmetric cyclic loading: (**a**) friction and (**b**) dowel resistance
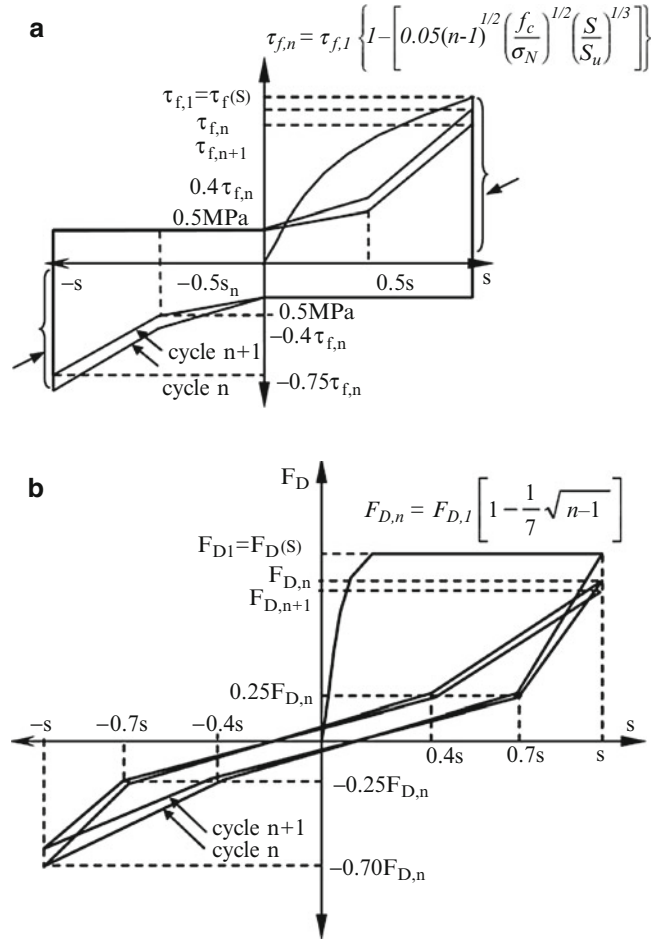
and after carrying out a new experimental program proposed the degradation rules for frictional and shear resistance, which were also adopted by the Greek Code for Interventions (KANEPE 2013). The frictional and dowel resistance is reduced at each cycle, n, according to (see Fig. 6)

$$\tau_{f,n} = \tau_{f,1} \cdot \tau_{\mathrm{deg}}$$
$$= \tau_{f,1} \left\{ 1 - \left[0.05(n-1)^{1/2} \left(\frac{f_c}{\sigma_N} \right)^{1/2} \left(\frac{s}{s_u} \right)^{1/3} \right] \right\} \tag{5}$$

$$F_{D,n} = F_{D,1} \cdot D_{\mathrm{deg}} = F_{D,1} \left[1 - \frac{1}{7} \sqrt{n-1} \right] \tag{6}$$

where $\tau_{f,1}$ and $F_{D,1}$ are the peak frictional and dowel resistance value, respectively, attained in

the first cycle, n is the number of cycles, s is the slip value, s_u is the ultimate slip value, σ_N is the normal clamping stress acting on the interface, and f_c is the concrete compressive strength.

Analytical Assessment of Strengthened Members and Structures

Analytical Model: Slip at the Interface

The success of repair/strengthening, as quantified by the improvement of strength and stiffness of the retrofitted members, depends on its entirety on the degree of collaboration between the various monolithic phases of the composite member. The response of the retrofitted member is modified proportionally to the slip that takes place along the interface. The calculation of slip

and its effects on the mechanical characteristics of the retrofitted members presuppose the use of analytical models where the interface characteristics need to be explicitly modeled. This comprises a quite challenging subject of continuous research.

In case of R/C jacketing of columns, Thermou et al. (2007, 2012b) have developed an analytical model for predicting the response under monotonic and reversed cyclic loading of structural members with *old-type* detailing, strengthened with R/C jacketing. The analytical model introduces one additional degree of freedom between the existing member and its outer R/C shell, thus allowing slip to take place at the interface between the existing member and the jacket. Shear resistance mechanisms, such as aggregate interlock, friction, and dowel action, are mobilized to resist slip. The magnitude of shear flow sustained along the contact surface is calculated by considering the states of stress of the composite member at a cracked cross section and at a point between successive cracks, in order to introduce in the flexural behavior the effect of the moment gradient (shear force magnitude). A detailed description of the analytical model for R/C jacketed members is presented in the last section of this article. The same approach has also been implemented after the necessary modifications for predicting the response of thin reinforced self-compacting three-jacketed beams (Chalioris et al. 2014). Moreover, Tsioulou and Dritsos (2011) developed an analytical model for predicting the flexural capacity of beams strengthened by the addition of a concrete layer in the tension or compression zone. The algorithm developed aimed at the evaluation of slip and shear stress distribution along the interface.

The analytical models, as the ones previously presented, may provide information relative to the effectiveness of the selected intervention method applied at member level. Stiffness and strength of the retrofitted (composite) members are estimated and assessed. The interrelation between the effects that the changes at local level may have on global level is a key issue in the assessment of strengthened structures. These

may be evaluated rapidly through standard Rayleigh-type or Stodola-type iteration (Clough and Penzien 1993), using secant-to-yield stiffness values for the individual members (Thermou et al. 2012c). The estimated fundamental translational response shape and period may guide definition of retrofit objectives. Additional lateral strengthening of the building (owing to the implicit relationship between stiffness and strength) by means of controlled stiffness addition along the building height may be required. Proportioning the stiffness of the individual floors is determined so as to even out large discrepancies in relative drift between successive floors detected in the fundamental response shape pattern.

Simplified Design Approach: The Use of Monolithicity Factors

A pragmatic design approach commonly adopted by codes of practice considers the monolithic approach for the analysis of composite members making use of properly defined *monolithicity factors* for obtaining the mechanical properties of the strengthened member. The redesign procedure is thus simplified substantially and is used extensively in practice for various intervention methods for different types of structural members (slabs, beams, columns, walls, or footings). The values given so far to these modification factors are empirical or semiempirical. The use of reliable analytical models (e.g., Thermou et al. 2007, 2012b) can be utilized for the derivation of monolithicity factors after extensive study of their sensitivity to the characteristics of the intervention method (e.g., Kappos et al. 2012).

The use of monolithicity factors simplifies substantially design calculations and is applicable to various types of structural members (slabs, beams, columns, walls, foundation elements) and intervention methods (e.g., R/C jacketing, R/C infill walls). These reduction factors are used to obtain the strength and deformation indices of the jacketed members and are applied to the respective properties of monolithic members with identical geometry. The monolithicity factors are defined as follows:

$$K = \frac{\text{Response index of the composite member}}{\text{Response index of the monolithic member with identical geometry}} \quad (7)$$

The various monolithicity factors usually used in design are:

– The monolithicity factors that refer to the deformation capacity indices such as the chord rotations at yield and ultimate, which are defined as

$$\text{Chord rotation at yield} : K_{\theta y} = \frac{\theta_{y,J}}{\theta_{y,M}};$$

$$\text{Chord rotation at ultimate} : K_{\theta y} = \frac{\theta_{u,J}}{\theta_{u,M}} \quad (8)$$

– The strength related monolithicity factors

$$\text{Shear strength} : K_v = \frac{V_{J,\max}}{V_{M,\max}};$$

$$\text{Moment at yield} : K_{My} = \frac{M_{y,J}}{M_{y,M}} \quad (9)$$

– The stiffness monolithicity factor

$$\text{Stiffness at yield} : K_k = \frac{K_{y,J}}{K_{y,M}} \quad (10)$$

where the subscripts J and M correspond to the composite (jacketed) cross section and to the identical monolithic cross section, θ_y, θ_u are the chord rotations at yield and ultimate, $V_{\max}$ is the maximum strength of the cross section, M_y is the yield moment, and K_y is the secant flexural stiffness at yield, defined as the ratio of the yield moment (M_y) to the yield curvature (φ_y).

Eurocode 8, Part 3 (§A.4.2, 2005), and the Greek Code for Interventions (KANEPE 2013) suggest the monolithicity factors presented in Table 2 to be used in the design of R/C members (slabs, columns, beams, and foundation system) when strengthened with additional reinforced concrete layers. In the case of Eurocode 8, Part3 (§A.4.2, 2005), the values of the monolithicity factors for columns of Table 2 should be used under the assumptions of: (i) full composite

action between old and new concrete, (ii) application of full axial load to the jacketed member, and (iii) application of the concrete properties of the jacket over the full section of the element. It is noted that the monolithicity factors for the design approach as proposed by the Greek Code for Interventions (KANEPE 2013) are subject to certain limitations, i.e., that the target strength increase of the jacketed member should not exceed twice that of the original. On the other hand, considering the code minima regarding the percentage of longitudinal reinforcement of the jacket, but also of the entire composite cross section (equal to 1 %), and the pertinent detailing rules (minimum thickness of the jacket is 70 mm), it is seen that the strength of the strengthened member far exceeds twice its original strength. In the case of EC8-Part 3 (2005), there is no restriction related to the increase of resistance of the R/C member due to jacketing.

In case of R/C column jacketing, the sensitivity of the monolithicity factors to the construction materials of the existing cross section (core) and the jacket, as well as to the percentage of longitudinal reinforcement of the jacket for increasing axial load, has been highlighted by Thermou et al. (2007). The analytical model for R/C jacketed columns, which is presented in the last section of this article, was utilized by Thermou et al. (2007, 2014) for performing an extensive parametric study for the derivation of monolithicity factors. For the range of parameters considered in their study, the lower and upper limits for the monolithicity factors were found to be $K_{\theta y} = 1.17 \sim 4.85$, $K_{\theta u} = 0.45 \sim 4.13$, $K_{My} = 0.32 \sim 0.99$, $K_r = 0.35 \sim 1.02$ and $K_k = 0.19 \sim 0.95$. Lampropoulos et al. (2012) utilized a computational (finite element) model of experimental specimens, and for the specific range of parameters examined, the monolithicity factors received the following values: $K_{\theta y} = 1.05 \sim 3.00$, $K_{\theta u} = 0.95 \sim 2.85$, $K_r = 0.70 \sim 0.90$ and

Strengthened Structural Members and Structures: Analytical Assessment, Table 2 Suggested values of the monolithicity factors by EC8-Part 3 (2005) and KANEPE (2013)

Element type	K_i	K_V	K_{My}	K_k	$K_{\theta y}$	$K_{\theta u}$
Slab	KANEPE (2013)	0.95		0.85	1.15	0.85
Column	EC8-Part 3 (2005)	0.90	1.00	–	1.05[a]	1.00
					1.20[b]	
	KANEPE (2013)	0.90[c]	–	0.80[c]	1.25[c]	0.75[c]
Beam	KANEPE (2013)	0.85	–	0.80	1.25	0.75
Foundation	KANEPE (2013)	0.90		0.70	1.30	0.80

[a]Roughening
[b]Rest of measures or no measures
[c]Roughening and connection measures

$K_k = 0.25 \sim 0.75$. The comparison of the code suggested monolithicity factors to the experimental values for monolithicity factors revealed that there is large dispersion in the case of monolithicity factors $K_{\theta y}$, $K_{\theta u}$, and K_k (Thermou et al. 2014). This observation has to be further assessed by considering the limited range of parameters of the experimental database (Thermou et al. 2011) as well as the fact that deformation and stiffness values are difficult to measure experimentally and "ultimate" conditions are not defined in a uniform way in all tests (Thermou et al. 2014). The suggested values for the monolithicity factors to be used in EC8-Part 3 (2005) and KANEPE (2013) apply for the specific properties of the construction materials and level of applied axial load. These limits are defined by the experimental data as: (i) percentage of the longitudinal reinforcement of the existing cross section 0.81 ~ 2.01 %; (ii) percentage of the longitudinal reinforcement of the jacket 0.75–1.64 %; (iii) concrete compressive strength of the existing cross section 23–56 MPa; (iv) concrete compressive strength of the jacket 18–69 MPa; and (v) dimensionless axial load (with the assumption that it is applied to the jacketed cross section) 0–0.23.

Analytical Assessment of R/C Jacketed Columns

Reinforced concrete (R/C) jacketing is one of the most commonly applied methods for the rehabilitation of concrete members. Jacketing is considered to be a global intervention method if longitudinal reinforcement placed in the jacket passes through holes drilled in the slab and new concrete is placed in the beam-column joint (Fig. 7). The main advantage of the R/C jacketing technique is the uniformly distributed lateral load capacity throughout the structure, thereby avoiding concentrations of lateral load resistance, which occur when only a few shear walls are added. A disadvantage of the method is the presence of beams which may require most of the new longitudinal bars in the jacket to be bundled into the corners of the jacket. Because of the presence of the existing column, it is difficult to provide crossties for the new longitudinal bars, which are not at the corners of the jacket.

The analytical model presented in section "Analytical Model for R/C Jacketed Columns" explicitly accounts for the slip at the interface between the existing member (core) and the jacket. The transfer of normal and shear stresses in the interface between the external layers of the jacket and the existing member determines the composite action developed by the strengthened members (Thermou et al. 2007). An alternative to this approach is presented in section "Simplified Analytical Expressions for R/C Jacketed Columns" where simplified analytical expressions in combination with the monolithicity factors may be used for designing the retrofit scheme.

Analytical Model for R/C Jacketed Columns
The proposed analytical model for predicting the flexural response of existing R/C members strengthened with concrete jacketing under

S

Strengthened Structural Members and Structures: Analytical Assessment, Fig. 7 Column R/C jacketing (Source of figures: author's personal files)
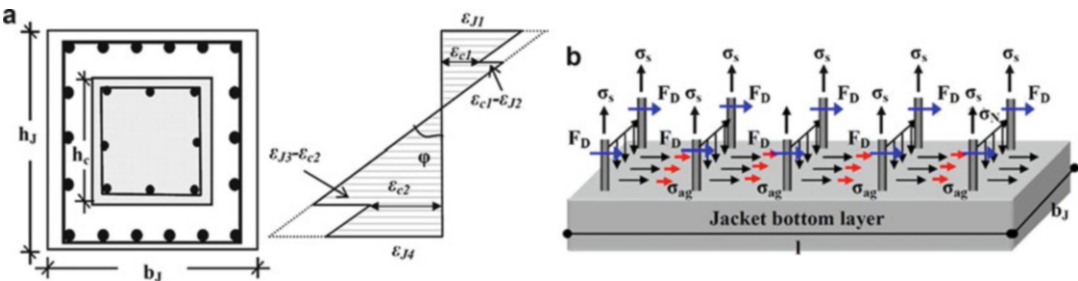
Strengthened Structural Members and Structures: Analytical Assessment, Fig. 8 (**a**) Strain profile of the jacketed cross section; (**b**) shear transfer mechanisms at the interface between the jacket and the core

monotonic and cyclic loading conditions introduces a degree of freedom allowing the relative slip at the interface between the existing member and the jacket (Thermou et al. 2007). Slip along the member's length is attributed to the difference in normal strains at the contact interfaces (Fig. 8a). For flexural analysis, the cross section is divided into three layers which bend with the same curvature, φ (Fig. 8a). The two external layers represent the contribution of the jacket, whereas the internal one represents both the core (existing cross section) and the web of the jacket shell. Slip at the interface mobilizes the shear transfer mechanisms such as aggregate interlock, friction due to clamping action, and dowel action provided by the stirrup legs of the jacket and by the dowels placed at the interface between the core and the jacket in case that such a connection measure is taken (Fig. 8b).

According to the analytical model of Thermou et al. (2007) for R/C jacketed members, shear transfer at the interface between the existing member and the jacket takes place between half crack intervals along the length of the jacketed member, as commonly considered in bond analysis. At the initial stages of loading, cracks form only at the external layers (jacket) increasing in number with increasing load, up to crack stabilization. This occurs when the jacket steel stress at the crack, $\sigma_{s,cr}$, exceeds the limit (fib 2010)

$$\sigma_{s,cr} > f_{\text{ctm},J} \frac{1 + \eta \rho_{s,\text{eff}}}{\rho_{s,\text{eff}}} \quad (11)$$

where $f_{ctm,J}$ is the tensile strength of concrete, $\eta(=E_s/E_c)$ is the modular ratio, and $\rho_{s,\text{eff}}$ is the effective reinforcement ratio defined as the total steel area divided by the area of mobilized
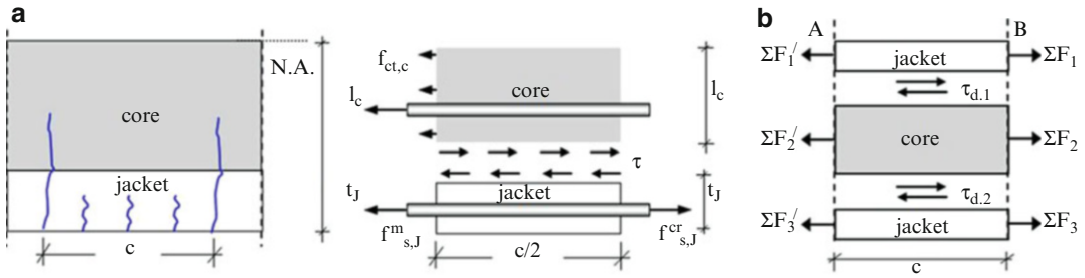
Strengthened Structural Members and Structures: Analytical Assessment, Fig. 9 (**a**) Free body equilibrium in the tension zone of the core of the composite section; (**b**) section equilibrium between adjacent cracks; (**c**) crack spacing

concrete in tension, usually taken as a circular domain with a radius of $2.5D_b$ around the bar (Model Code 2010). Using the same considerations in the combined section it may be shown that a number of the external cracks penetrate the second layer (core) of the jacketed member (Fig. 8a). The distance between those cracks, taken as c, is a key element of the proposed methodology (Fig. 9a).

After crack stabilization and assuming that the neutral axis depth is about constant in adjacent cross sections, from the free body equilibrium in the tension zone of the core of the composite section (Fig. 9a), the crack spacing is defined as follows (Thermou et al. 2007):

$$c = \frac{0.64 \cdot b_J \, l_c \, f_{ct,c}}{n_c \, D_{b,c} \, f_{b,c} + n_J \, D_{b,J} \, f_{b,J}} \quad (12)$$

where b_J is the width of the jacketed cross section, l_c is the height of the tension zone in the core of the composite cross section, $f_{ct,c}$ is the tensile strength of concrete core, n_c, n_J are the number of bars in the tension steel layer of the core and the jacket, respectively, $D_{b,c}$, $D_{b,J}$ are the bar diameter of the core and jacket longitudinal reinforcement, respectively, and $f_{b,c}$ and $f_{b,J}$ are the average bond stress of the core and the jacket reinforcement layer, respectively.

Shear stress demand at the interfaces, $\tau_{d,i}$, is determined by examining the cross section along the height and along a member length equal to the distance between successive cracks (Fig. 9b). The layer force resultant ΣF_i (sum of concrete and steel forces at each layer), for the externally

applied axial load, N_{ext} (considered to be applied to the jacketed section), is used to calculate the vertical shear stress demand in the member, $\tau_{d,i}$ (Fig. 9b). With the assumption that the shear flow, q, reversal takes place at length equal to $c/2$ (where c is the crack spacing), the average stress demand $\tau_{d,i}$ is equal to

$$\tau_{d,i} = \frac{\Sigma F_i}{0.5 \, c \, b_J} \quad (13)$$

where ΣF_i is the layer force resultant, b_J is the width of the jacketed cross section, and c is the crack spacing length.

Shear Transfer Mechanisms Under Monotonic and Cyclic Loading

The monotonic and cyclic behavior of interfaces is described by the constitutive model developed by Vassilopoulou and Tassios (2003) where the models of friction and dowel resistance of Tassios and Vintzileou (1987), Vintzileou and Tassios (1986, 1987) are adopted. The interface model accounts for the combined shear force resistance mobilized along interfaces due to sliding both under monotonic and cyclic imposed deformations. This model was further modified by Palieraki et al. (2012) and adopted by the current Greek Code for Interventions (2013) (see Fig. 6). In case of cyclic loading, additional modifications and extensions were applied (Thermou et al. 2012b).

Solution Algorithm

The objective of the calculation algorithm at each loading step was twofold: simultaneous

Fig. 10 Comparison of moment–curvature response histories with the corresponding experimental ones (SS1 and SS3 specimens from Rodriguez and Park (1994), R and E from Vandoros and Dritsos (2008))
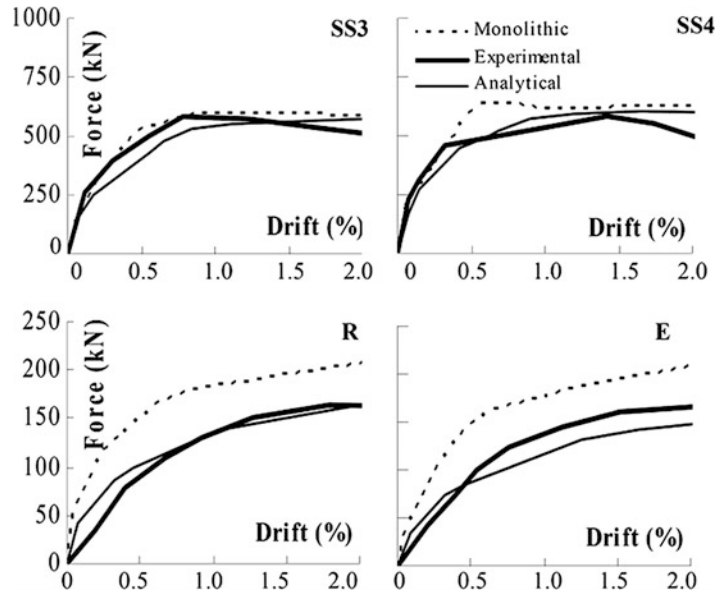
establishment of equilibrium between the shear stress capacity and demand at the interfaces for relative slip, s, and force equilibrium at the cross section. An iterative procedure was followed, and equilibrium is established until convergence is achieved. Due to the complexity of the proposed solution algorithm, a program was necessary to be developed where fiber analysis was considered. The analytical model was set up to predict moment–curvature response curves. More details can be found in Thermou et al. (2012b).

Comparison with Experimental Data
The validity of the proposed analytical model for predicting the flexural response of R/C jacketed members was examined by comparing: (i) for monotonic loading, the analytical lateral load versus lateral displacement curves along with the envelope experimental curves of the recorded lateral load versus lateral displacement hysteretic loops from various experimental studies (Thermou et al. 2007) (the monotonic moment–curvature response curves were derived according to the analytical model and then converted to force–displacement response curves). In Fig. 10, the experimental lateral load versus drift curves of a representative number of test specimens are compared to the corresponding analytical curves; (ii) For cyclic loading, the

moment–curvature histories derived by the analytical model with those for the specimens studied by Bousias et al. (2007). The decision to select this experimental study was based not only on its scope, but also on the fact that it provides detailed test results in terms of moment–curvature curves. The comparison between the experimental and the analytical moment–curvature histories for some of the specimens of Bousias et al. (2007) is presented in Fig. 11. It is observed that the analytical curves reasonably match the strength and stiffness level of the experimental curves at each loading step. When slip is taken into account, pinching is more pronounced in the analytical model, indicating less energy dissipation at each loading cycle, which is an indication of conservatism.

Simplified Analytical Expressions for R/C Jacketed Columns
The composite cross section is considered as monolithic, assuming that there is full connection in the interface between old and new concrete. The monolithicity factors may be utilized to modify the response indices in order to indirectly account for the adverse effect of slip on the interface (see section "Simplified Design Approach: The Use of Monolithicity Factors"). A simplified model of a typical jacketed cross section of a
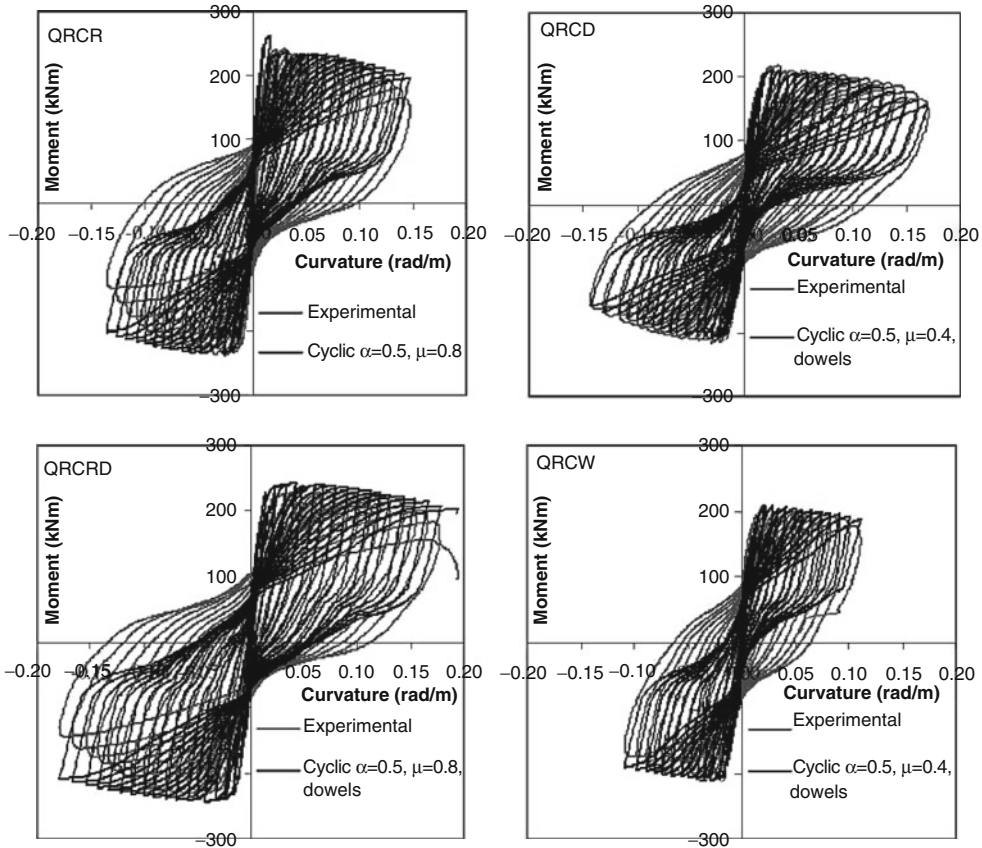
Strengthened Structural Members and Structures: Analytical Assessment, Fig. 11 Comparison of moment–curvature response histories with the corresponding experimental ones for specimens QR/CR, QR/CD, QR/CRD, and QR/CW from Bousias et al. (2007)

column member is depicted in Fig. 12 (Thermou et al. 2012c). Dimensions of the initial section are b_c and h_c, whereas those of the section after jacketing are b_J, h_J. Compression zone depth c is expressed as a fraction of the depth of the jacketed section, $c = \xi_y^J \cdot d_J$. For calculations of yield moment and flexural stiffness, all reinforcement is considered to act at the location of the added (jacket) reinforcement. Note that according to Steiner's theorem, existing tension longitudinal reinforcement (equal to compression reinforcement, Fig. 12), given by the area ratio ρ_c, contributes to flexural stiffness of the jacketed section through the term $\rho_c b_c d_c (0.5h_c - d_c)^2$; thus, in order to maintain the same contribution in this calculation, the equivalent amount that is transferred to the location of jacket reinforcement

is $\rho_c b_c d_c (0.5h_c - d_c)^2/(0.5h_J - d_J)^2$. Furthermore, any other existing web longitudinal reinforcement is neglected, as it is considered to have a small influence on post-jacketed flexural strength. Thus, the equivalent longitudinal tensile reinforcement ratio ρ_e, of the jacketed cross section (total reinforcement area divided by the total area of the jacketed member) is given by

$$\rho_e = \rho_J + \rho_c \frac{(0.5h_c - d_c)^2}{(0.5h_J - d_J)^2} \cdot \frac{b_c h_c}{b_J h_J} \qquad (14)$$

where $\rho_c (= A_c/b_c h_c)$ and $\rho_J (= A_J/b_J h_J)$ are the tension longitudinal reinforcement ratios of the existing cross section (equal to compression reinforcement) and the jacket, respectively, h_c and h_J are the heights of the existing and jacketed cross
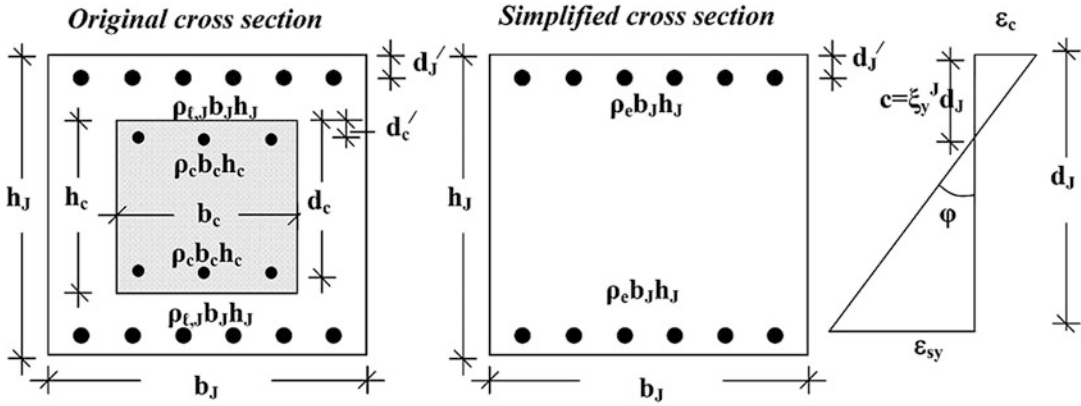
Strengthened Structural Members and Structures: Analytical Assessment, Fig. 12 Simplified model for R/C jacketed members (existing web longitudinal reinforcement is neglected as it is considered to have a small influence on post-jacketed flexural strength)

sections, respectively, b_c and b_J are the widths of the existing and jacketed cross sections, respectively, and d_c and d_J are the depths of the existing and jacketed cross sections, respectively.

The jth member translational stiffness (secant to yield) is defined as

$$K_{y,j}^J = \frac{12 E_c I_{y,j}^J}{h_{st}^3} = \frac{12 M_{y,j}^J}{\varphi_{y,j}^J h_{st}^3} \tag{15}$$

The moment at yield, $M_{y,j}{}^J$, at the center of gravity of the simplified jacketed cross section is estimated equal to

$$M_{y,j}^j = \varphi_{y,j}^J b_j h_J^3 E_c \big[0.40\rho \big[\xi_y(-1 - 0.25 N_c) \\ + 1.15 n_c + 0.10 \big] + 0.25\xi_y^2 \big(1 - 0.66\xi_y\big) \big] \tag{16}$$

Hence, the jth member translational stiffness is equal to

$$K_{y,j}^J = \frac{b_{J,j} h_{J,j}^3 E_c}{h_{st}^3} \cdot \Big\{ 4.8\rho_{e,j} \Big[1.15 \cdot \eta_c - \xi_{y,j}^J \\ \cdot (1 + 0.25\eta_c) + 0.1 \Big] + 3\big(\xi_{y,j}^J\big)^2 \big(1 - 0.66\xi_{y,j}^J\big) \Big\} \tag{17}$$

where ρ_e is the equivalent longitudinal tensile reinforcement ratio, $\eta_c(=E_s/E_c)$ is the modular ratio of steel and concrete, and $\xi_{y,j}{}^J$ is the normalized depth of compression zone of the R/C jacketed member.

Depending on the magnitude of the axial load ratio, $v_j = N_j/b_J h_J f_c'$, yielding may occur when either the tension steel reinforcement reaches yielding or the compressive concrete strain exceeds the limit of linear response in the compressive stress–strain envelope, estimated in the range of $\varepsilon_c = 1.8 \cdot f_c'/E_c$. From the basic cross-sectional equilibrium, the normalized depth of compression zone, $\xi_{y,j}{}^J$, associated with these two alternative definitions of phenomenological yielding is obtained from:

(a) Upon yielding of tension steel

$$\xi_{y,j}^J = - \left[(2\eta_c - 1)\rho_{e,j} + \frac{v_j f_c'}{E_c \varepsilon_{sy}} \right] \\ + \left\{ \left[(2\eta_c - 1)\rho_{e,j} + \frac{v_j f_c'}{E_c \varepsilon_{sy}} \right]^2 \\ + 2 \left[(1.10\eta_c - 0.10)\rho_{e,j} + \frac{v_j f_c'}{E_c \varepsilon_{sy}} \right] \right\}^{0.5} \tag{18a}$$

where f_c' is the concrete compressive strength and ε_{sy} is the steel strain at yielding.

(b) At the onset of concrete strain nonlinearity, $\varepsilon_c = 1.8 \cdot f_c'/E_c$

$$\xi_{y,j}^J = - \left[(2\eta_c - 1)\rho_{e,j} + 0.55v_j \right]$$
$$+ \left\{ \left[(2\eta_c - 1)\rho_{e,j} + 0.55v_j \right]^2 + 2(1.10\eta_c)\rho_{e,j} \right\}^{0.5}$$
$$(18b)$$

From the above, the secant to yield stiffness and the moment at yield may be calculated directly by utilizing Eqs. 15 and 16, respectively, after being multiplied by the adequate monolithicity factors according to Table 2.

References

Birkeland PW, Birkeland HW (1966) Connections in precast concrete construction. J ACI 63(3):345–368

Bousias S, Biskinis D, Fardis MN, Spathis A-L (2007) Strength, stiffness, and cyclic deformation capacity of the concrete jacketed members. ACI Struct J 104(5):521–531

Chalioris KE, Thermou GE, Pantazopoulou SJ (2014) Behaviour of rehabilitated R/C beams with self-compacting concrete jacketing – analytical model and test results. Constr Build Mat J 55(31):257–273

Choi D-U, Fowler DW, Jirsa JO (1999) Interface shear strength of concrete at early ages. ACI Struct J 96(3):343–348

Clough RW, Penzien J (1993) Dynamics of Structures, 2nd edn. McGraw-Hill International Editions, New York, p 738

Deutsches Institut für Normung [German Standards Institute] (DIN) (1936) Beton und Stahlbetonbau: Bemessung und Ausführung. DIN 1045, Berlin [Concrete and reinforced concrete: design and construction]

Eurocode 8 (2005) Design of structures for earthquake resistance, Part 3: assessment and retrofitting of buildings. EN1998-3, European Committee for Standardization (CEN), Brussels

fib (2013) Model Code for Concrete Structures 2010. Ernst & Sohn, Hoboken, New Jersey

Greek Code for Interventions (KANEPE) (2013) Earthquake Planning and Protection Organization, Athens

Júlio ENBS, Branco FAB, Silva VD, Lourenco JF (2006) Influence of added concrete compressive strength on adhesion to an existing concrete substrate. Build Environ 41:1934–1939

Kappos AJ, Thermou GE, Papanikolaou VK (2012) Investigation of the behaviour of old type detailing R/C columns strengthened with R/C jackets, Report. Earthquake Planning and Protection Organization 179 (in Greek)

Lampropoulos AP, Tsioulou OT, Dritsos SE (2012) Monolithic coefficient values for design when seismically strengthening RC columns with jackets. J Earthq Eng 16:1023–1042

Mattock HA, Hawkins MN (1972) Shear transfer in reinforced concrete- recent research. PCI J 17(2):55–75

Palieraki V, Vintzileou E, Zeris C (2012). Behaviour of interfaces in repaired/strengthened R/C elements subjected to cyclic actions: experiments and modelling. In: Third international symposium on life-cycle and sustainability of civil infrastructure systems (IALCCE'12), Vienna

Rodriguez M, Park R (1994) Seismic load tests on reinforced concrete columns strengthened by jacketing. ACI Struct J 91(2):150–159

Royal Decree (1954) Design regulation of reinforced concrete building works. Royal Decree (18.2/26.07.1954), Ministry of Public Works, Greece (in Greek)

Santos PMD, Júlio ENBS (2013) A state-of-the-art review on roughness quantification methods for concrete surfaces. Constr Build Mat J 38:912–923

Santos PMD, Júlio ENBS (2014) Interface shear transfer on composite concrete members. ACI Struct J 111(1):113–121

Tassios TS (1986) Fundamental mechanisms of force-transfer across reinforced concrete critical interfaces. In: CEB workshop, West Germany, pp 381–397

Tassios T, Vintzileou E (1987) Concrete-to-concrete friction. ASCE J Struct Eng 113(4):832–849

Thermou GE, Elnashai AS (2006) Seismic retrofit schemes for R/C structures and local–global consequences. J Progr Struct Eng MatWiley InterSci 8(1):1–151

Thermou GE, Palaioxorinou T (2013) Rapid evaluation of non-ductile multistory RC buildings. In: Proceedings of ECCOMAS thematic conference "computational methods in structural dynamics and earthquake engineering (COMPDYN 2013)", Kos, paper no 1459

Thermou GE, Pantazopoulou SJ, Elnashai AS (2007) Flexural behavior of brittle R/C members rehabilitated with concrete jacketing. J Struct Eng 133(10):1373–1384

Thermou, GE, Papanikolaou VK, Kappos AJ (2011) Analytical model for predicting the response of old-type columns rehabilitated with concrete jacketing under reversed cyclic loading. In: Proceedings of ECCOMAS thematic conference "computational methods in structural dynamics and earthquake engineering (COMPDYN 2011)", Corfu, paper no 137

Thermou GE, Pantazopoulou SJ, Elnashai AS (2012a) Global interventions for seismic upgrading of substandard R/C buildings. J Struct Eng 138(3):387–401

Thermou GE, Papanikolaou VK, Kappos AJ (2012b) Cyclic response of R/C jacketed columns including modelling of the interface behavior. In: Proceedings of 15th world conference on earthquake engineering, Lisbon, 24–28 Sep 2012, paper no 2855

Thermou GE, Elnashai AS, Pantazopoulou SJ (2012c) Retrofit Yield Spectra–a practical device in seismic rehabilitation. Earthq Struct 3(2):141–168

Thermou GE, Papanikolaou VK, Kappos AJ (2014) Monolithicity factors for the design of R/Ccolumns strengthened with R/C jackets. In: Proceedings of

S

2nd European conference on earthquake engineering and seismology, Istanbul, paper no 766

Tsioulou OT, Dritsos SE (2011) A theoretical model to predict interface slip due to bending. Mat Struct 44:825–843

Vandoros KG, Dritsos SE (2008) Concrete jacket construction detail effectiveness when strengthening R/C columns. Constr Build Mater 22:264–276

Vassilopoulou I, Tassios P (2003) Shear transfer capacity along a R/C crack under cyclic sliding. In: Proceedings fib symposium, Technical Chamber of Greece, Athens, paper no 271

Vecchio FJ, Collins MP (1986) The modified compression field theory for reinforced concrete elements subjected to shear. J ACI 83(2):219–231

Vintzileou E, Tassios TS (1986) Mathematical models for dowel action under monotonic and cyclic conditions. Mag Concr Res 38(134):13–22

Vintzileou E, Tassios TS (1987) Behavior of dowels under cyclic deformations. ACI Struct J 84(1):18–30

Walraven JC (1981) Fundamental analysis of aggregate interlock. Struct Div ASCE 107(ST11):2245–2270

Strengthening Techniques: Bridges

Chris P. Pantelides
Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, UT, USA

Synonyms

Bridge; Damage; Design; Earthquakes; Rehabilitation; Repair; Retrofitting; Seismic; Strengthening

Introduction

The goal of seismic design is to prevent life-threatening damage and to allow moderate damage that can be repaired after an earthquake. From a life safety perspective, the most serious bridge damage occurs when the deck collapses, because of inadequate seat details, column failure, or cap beam-column joint failure. The collapse of the link span of the San Francisco-Oakland Bay Bridge during the 1989 Loma Prieta earthquake (Fig. 1a) was caused by failure of two bolted connections under the span's supporting truss which led to its unseating (Buckle et al. 1990). The bridge failure of the Hanshin elevated expressway during the 1995 Great Hanshin Kobe earthquake (Fig. 1b) is attributed to column failure caused mainly by poor anchorage of the circular hoop ends which were lap spliced in the cover concrete (Park 1996) and inadequate amount and details of the longitudinal and transverse reinforcement (Anderson et al. 1996). This entry presents a summary of observed damage in past earthquakes and seismic strengthening techniques of existing bridges. The stages of evolution of bridge design in the USA from the 1970s including vulnerabilities and typical design details are outlined.

Seismic Bridge Design

The seismic design of bridges is separated into three periods, based on the State of California provisions (Caltrans 2006).

Pre-1971 Design

In this period, it was recognized that earthquakes generate forces proportional to the structure's dead weight. Until 1965, the maximum lateral seismic design force was 6 % of the structural dead weight. Vulnerabilities of bridges built during this period include: (a) column shear failure, (b) column longitudinal reinforcement pullout, and (c) unseating of expansion hinges. Typical design details of the period are 13 mm diameter column ties at 300 mm spacing (regardless of column or bar size), very short seat widths at expansion joints (150–200 mm), inadequate lap splices of column bars from the footing (20 bar diameters), inadequate development of column bars into the footing (20 bar diameters without standard hooks), and lap-spliced column ties in the cover concrete (absence of seismic hooks).

1971–1994 Design

After the 1971 San Fernando earthquake, the importance of ductility and detailing was recognized, and the concept of capacity design was

Strengthening Techniques: Bridges, Fig. 1 Bridge failure: (**a**) San Francisco-Oakland Bay Bridge (© California Department of Transportation (2003) (CALTRANS)); (**b**) Hanshin elevated expressway (https://www.fhwa.dot.gov/publications/publicroads/96fall/imgs/p96au18.jpg (FHWA))

introduced in the design code. Vulnerabilities of bridges built during this period include: (a) column shear failure of plastic hinge regions, (b) shear failure of flared columns, and (c) unseating of expansion joint hinges. In retrofitted bridges, vulnerabilities were found at expansion joint hinge restrainers, particularly for skewed bridges. Typical details for this period include: closer spacing and improved column shear detailing (spacing 100–150 mm, but no confinement/anti-buckling requirement for plastic hinges), top reinforcement in footings and pile caps (but no shear reinforcement), column longitudinal splices were prohibited at maximum moment locations, short seat widths at expansion joints (300 mm), no cap beam-column joint reinforcement, and poor column flare details.

Post-1994 Design

After the 1994 Northridge earthquake, a capacity design philosophy was adopted that would ensure ductile flexural failure of the columns while all other bridge components remained elastic. Typical design details include: tight confinement reinforcement in plastic hinge regions (100 mm spacing), long seat widths at expansion joints (600 mm), improved flare column details (gap between top of flare and superstructure), no lap splices in plastic hinge zones, shear reinforcement in footings, and cap/column and footing/column joint reinforcement.

Damage to Bridges in Past Earthquakes

After an earthquake, lessons are learned which influence seismic strengthening of existing bridges as well as the development of new design codes. This section describes damage and its location as observed in bridges from past earthquakes (Caltrans 2006). Deck damage includes complete collapse of a span, as shown in Figs. 1a and 2a, b, or movement of the deck, as shown in Fig. 2c. Cap beam, cap beam-column joint, and girder shear damage are shown in Fig. 3. Damage to columns can be flexural as shown in Fig. 4 or shear damage as shown in Fig. 5; shear damage is brittle and frequently leads to collapse. Soil settlements can occur around columns, and gaps between columns and soil may become large as shown in Fig. 6. Abutment failures are also common, as shown in Fig. 7.

Bridge Seismic Strengthening

There is historic evidence which supports the assertion that seismic strengthening of bridges is effective. Damage in past earthquakes, such as the 1994 Northridge earthquake, has been reported from cable restrainers breaking through the hinge diaphragm; however, no bridge retrofitted after the 1989 Loma Prieta earthquake experienced any serious damage (Yashinsky 1998). Similarly, most of the observed structural

S

Strengthening Techniques: Bridges, Fig. 2 Deck failure: (**a**), (**b**) span collapse (© California Department of Transportation (2006)); (**c**) span displacement
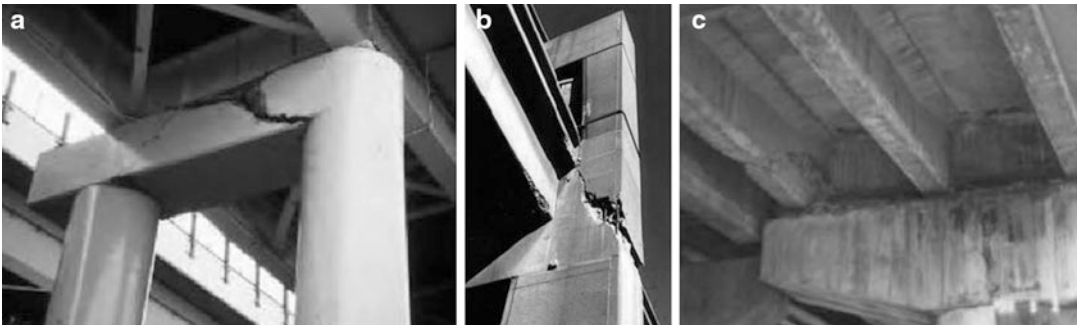
Strengthening Techniques: Bridges, Fig. 3 Cap beam, cap beam-column joint, and girder shear damage (Cap beam, girder shear damage: © California Department of Transportation (2006); beam-column joint: © California Department of Transportation (2003))

Strengthening Techniques: Bridges, Fig. 4 Column flexural damage (© California Department of Transportation (2006))

Strengthening Techniques: Bridges, Fig. 5 Column shear damage (© California Department of Transportation (2006))

Strengthening Techniques: Bridges, Fig. 6 Soil settlement and gap around columns (© California Department of Transportation (2006))
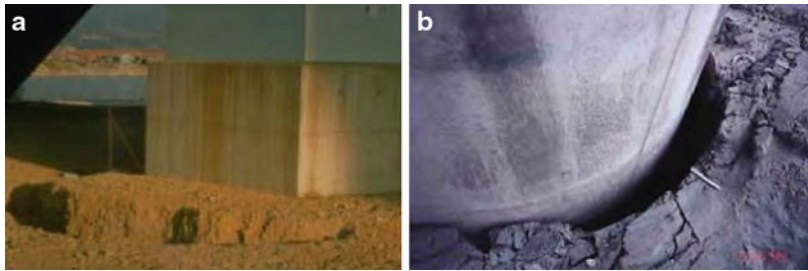
Strengthening Techniques: Bridges, Fig. 7 Abutment failures (© California Department of Transportation (2006))
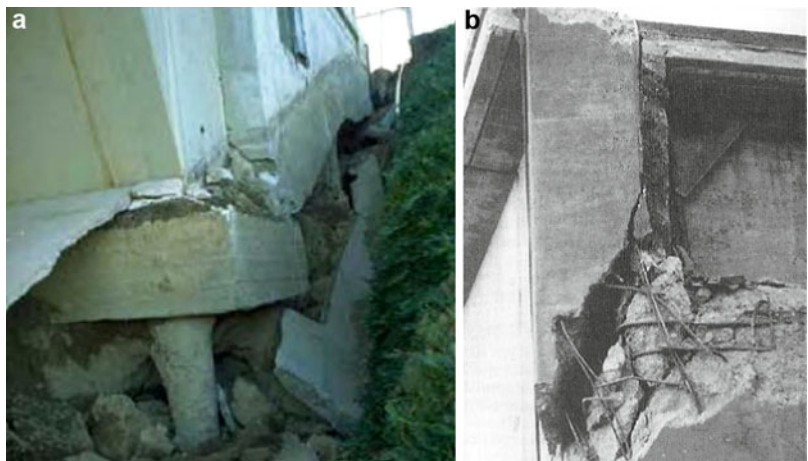
damage in the 2011 Great East Japan earthquake has reportedly occurred in older bridges that had not yet been retrofitted after the 1995 Great Hanshin Kobe earthquake or only partially so (Marsh et al. 2011).

Several measures are used to strengthen bridges: restrainers, steel girder plate connections, seat extenders, column strengthening, infill wall between columns, footing and abutment strengthening, and cap beam-column joint strengthening. In addition, retrofit measures such as seismic isolation reduce the demand imposed on a bridge system, and in this sense they improve its seismic performance significantly.
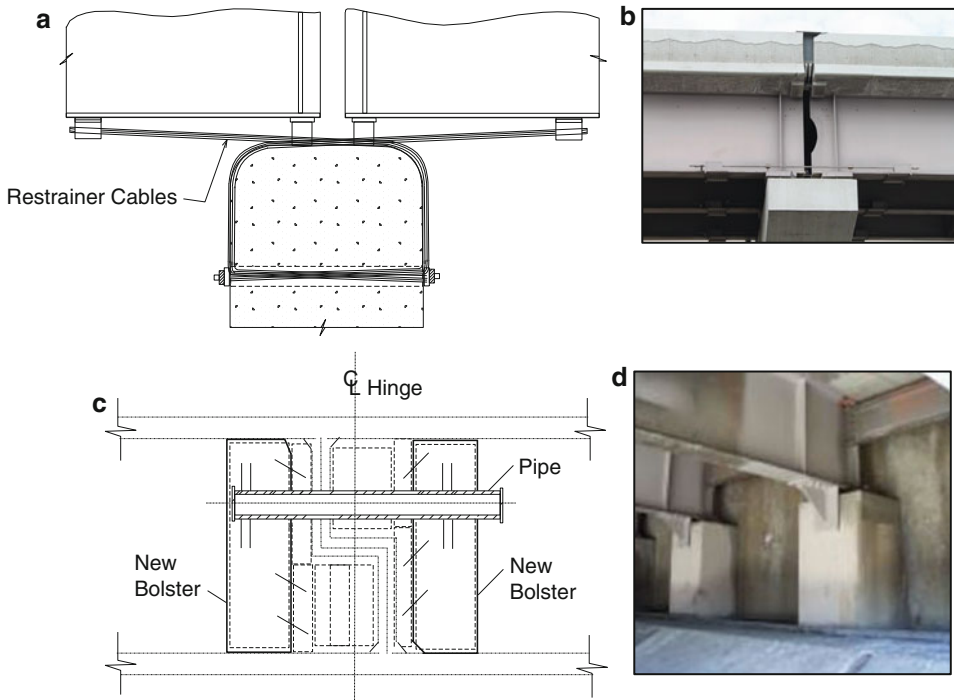
Strengthening Techniques: Bridges, Fig. 8 (**a**) Cable restrainer, (**b**) steel bar restrainer, (**c**) pipe seat extender, (**d**) bumper blocks

Restrainers/Expansion Joint Seat Width Extenders/Steel Plates/Bumper Blocks

Cable restrainers (Fig. 8a) and longitudinal steel bar restrainers (Fig. 8b) are used to prevent girders from falling off their supports (Yashinsky 1998). The typical cable restrainer is a galvanized 19 mm diameter steel cable with galvanized cold-swaged fittings and 25 mm diameter threaded studs at each end. Loss of support of the superstructure can be achieved using beam seat extenders, pipe seat extenders shown in Fig. 8c, steel plates used to connect two steel girders at expansion joints, and bumper blocks as shown in Fig. 8d (Pantelides et al. 2004).

Failure of restrainer cables has been observed in the 1989 Loma Prieta, 1994 Northridge, and 1995 Kobe earthquakes (FHWA 2006). Restrainers ruptured, anchorage plates pulled through concrete diaphragms, fittings pulled away from cables, and anchorage nuts worked loose from the ends of cable units. Restrainers alone may not be the best solution in many cases

because of the need to limit restrainer forces to protect other structural elements, while increasing restrainer forces to accomplish their intended purpose. Alternatives that should be considered when loss of support is a possibility include seat extensions, bearing strengthening, and replacement with conventional or isolation bearings.

Column Jacketing: RC/Steel/Fiber Reinforced Polymer (FRP)

Strengthening of concrete columns can be achieved by reinforced concrete (RC) jackets. This is an intrusive and time-consuming procedure, but in some cases it is necessary. However, the effect of strengthening a column using RC jackets should be examined carefully as this method will increase column flexural strength and stiffness more than a steel jacket, FRP composite jacket, or wire wrap, with potentially undesirable effects on bridge performance.

Steel jacketing of circular columns by circular steel jackets (Fig. 9a), and of rectangular columns by elliptical steel jackets, was found to be
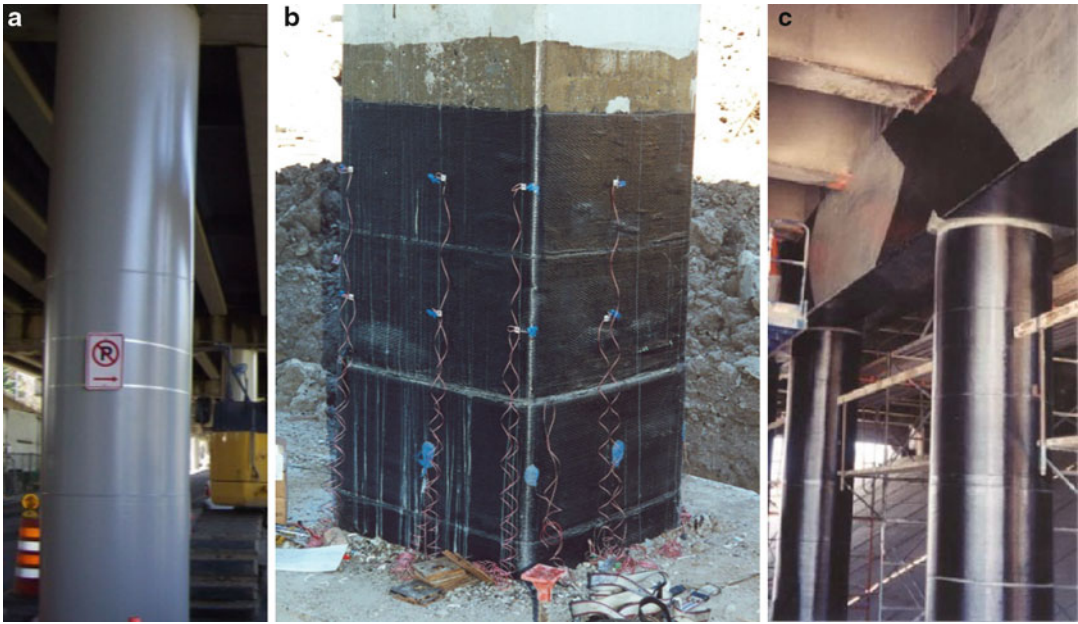
Strengthening Techniques: Bridges, Fig. 9 Column jackets: (**a**) steel (Photo credit: Washington State Department of Transportation. http://sdotblog.seattle.gov/2012/ 06/21/fauntleroy-expressway-wearing-new-jackets/#sthash. nlxozUnm.dpbs), (**b**) FRP square, (**c**) FRP circular

effective in enhancing shear strength and flexural ductility (Priestley et al. 1994). Encasing circular columns with a bonded steel jacket was found to inhibit bond failures in lap splices of longitudinal reinforcement in plastic hinge regions (Chai et al. 1991). Rectangular steel jackets are effective in enhancing the performance of shear critical columns. These jackets can improve column ductility by eliminating the brittle shear mode of failure, but it is important to recognize that the failure mode may then shift to a flexural one for which the rectangular jacket can provide only limited assistance.

Seible et al. (1997) used continuous FRP jackets for flexural hinge confinement of circular and rectangular columns, shear strengthening, and lap splice clamping. Full-scale tests (Fig. 9b) have shown that significant increases can be gained in the ductility of bridge column bents strengthened with FRP composites (Pantelides et al. 1999, 2007). FRP composites have been used widely for seismic strengthening of existing bridges (Priestley et al. 1996; Ogata and Osada 2000), as shown in Fig. 9c for the State

Street Bridge on I-80 in Salt Lake City, Utah (Pantelides et al. 2004).

The advantage of using FRP composites stems from their superior resistance to corrosion, high stiffness-to-weight ratio, high strength-to-weight ratio, and ability to control the material's behavior by selecting the proper fiber orientation. FRP composites must be protected against ultraviolet radiation; even though carbon fibers themselves are not affected by moisture, the quality of the epoxy matrix is critical in the performance of the composite. Carbon fiber systems have shown loss of tensile strength at high temperatures due to moisture absorption. Another consideration when using FRP systems is their high initial cost; however, they have a superior life-cycle cost as compared to steel because of their resistance to corrosion. In shear and flexural critical applications, bond of FRP composites to concrete is a critical parameter which must be considered in the design (ACI 440 2002); mechanical anchorage of FRP laminates to concrete has been used successfully to postpone debonding.

Another effective strengthening technique is to wrap prestressing wire under tension around a column (Hawkins et al. 1999). Rectangular steel columns have been strengthened with steel angles and circular steel columns with circular steel pipes (Nishikawa et al. 1998).

Infill Wall Strengthening

Infill shear walls are reinforced concrete walls cast in place between the columns of a multicolumn bent, as shown in Fig. 10. The reinforced concrete columns can have a circular or rectangular shape. They have been used successfully to increase transverse shear capacity. These walls prevent the formation of plastic hinges in the columns during transverse loading and help overcome deficiencies in the flexural or shear strength of the cap beam. To be effective, infill shear walls should be designed to act compositely with existing members. This is done by providing drilled and bonded dowels in the columns and the bottom of the cap beam, so that shear is transferred at the interfaces through a shear friction mechanism; this may require the existing concrete surfaces to be roughened. A footing should be provided under the infill wall which is tied into the existing column footings. There needs to be sufficient reinforcement to transfer all seismic forces, and the potential for differential settlement between the existing and new footings should be considered.

Foundation Strengthening

Footings that support columns may be structurally unable to resist the forces transmitted from those columns. This usually occurs when there is a lack of reinforcement in the top of the footing. A concrete overlay can be designed to provide increased negative moment capacity in the footing. Existing positive moment reinforcement is utilized along with the additional footing depth provided by the overlay to increase positive moment capacity. For isolated pile caps a grade or link beam is utilized to improve shear capacity, as shown in Fig. 11. In addition, longitudinal reinforcement in reinforced concrete piles may be terminated too close to the top of the pile, resulting in insufficient tensile capacity; in this
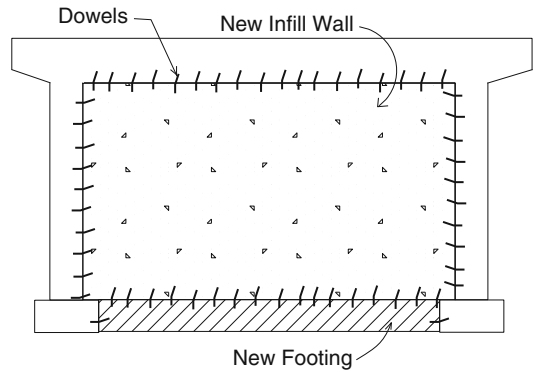
Strengthening Techniques: Bridges, Fig. 10 Pier strengthening using infill wall

case, it is desirable to add hold-down anchors or supplemental tension piles. An alternative is to drill holes into the piles from the top of the pile cap and epoxy high strength steel bars into the pile cap and piles (Pantelides et al. 2001); these vertical bars are shown in Fig. 11 (Pantelides et al. 2007).

Abutment Strengthening

Strengthening of seat-type abutments in the longitudinal and transverse directions can reduce forces on the piers. Anchor slabs are used to increase the ability of abutments to carry both longitudinal and transverse loads, and resist displacement in either direction, as shown in Fig. 12. Transverse abutment shear keys and anchors can reduce forces on the piers (FHWA 2006).

Cap Beam Joints: RC Bolsters/Posttensioning/ Steel Jackets/FRP Jackets

Cap beam-column joints are subjected to large stresses and are vulnerable to damage in earthquakes transverse to the bridge. Existing cap beams and joints were not usually designed to behave in a ductile manner, so any strengthening measure must ensure that these elements are either capable of accommodating the ductility demands placed on them or are capable of elastically resisting the forces that will result from plastic hinging in the columns.

RC bolsters are used to improve flexural capacity, with or without prestressing. Another
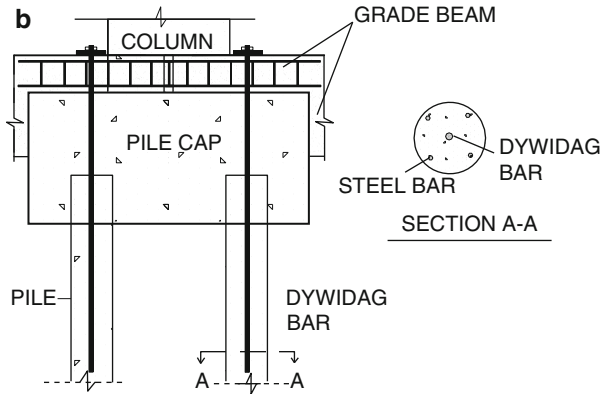
Strengthening Techniques: Bridges, Fig. 11 Footing strengthening using grade beam and pile cap to pile steel bars
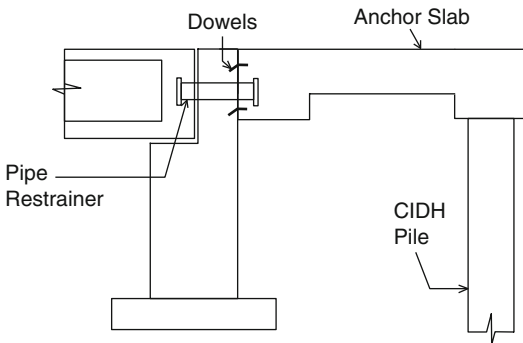
Strengthening Techniques: Bridges, Fig. 12 Abutment strengthening using anchor slab

strengthening method is to use external posttensioning as shown in Fig. 13a. Steel jackets bonded to the sides of the joint and anchored to the bent cap are also effective (Thewalt and Stojadinovic 1995). Finally, FRP jackets in the form of an ankle wrap have been tested (Gergely et al. 2000) and used to strengthen column to cap beam-column joints as shown in Fig. 13b (Pantelides et al. 2004). The nominal principal tensile stress developed in the beam-column joint is used to design the carbon FRP composite layers. When strengthening of these joints is required to improve the longitudinal response, transversely prestressed bolsters may be used.

Seismic Isolation: Elastomeric Bearings/ Friction Pendulum Bearings

Seismic isolation bearings can shift the natural frequency of the bridge away from the region of dominant earthquake energy; in addition, they increase damping and reduce and redistribute the earthquake-induced lateral forces to levels within the elastic capacities of the substructure and foundation (Mayes et al. 1994). An elastomeric bearing (Fig. 14a) consists of alternating layers of rubber and steel plates and often has a lead core to dissipate seismic energy. A friction pendulum bearing (Fig. 14b) consists of a stainless steel concave dish and articulated slider surfaced with a composite liner. During an earthquake, the slider moves back and forth on the concave dish; the spherical surfaces of the slider and the dish define a motion similar to that of a pendulum. A friction pendulum bearing isolates a bridge from an earthquake through pendulum motion and absorbs energy through friction.

Summary

The entry presents a number of methods for seismic strengthening of existing bridges. A brief review of seismic bridge design evolution and
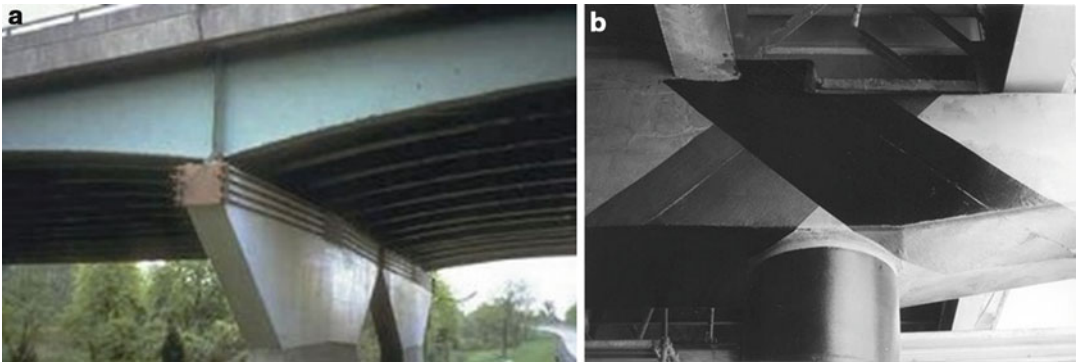
Strengthening Techniques: Bridges, Fig. 13 Cap beam-column joint strengthening: (**a**) posttensioning (Photo Credit: VSL International Ltd. http://vsl-sg.com/sps-images/4-004.jpg), (**b**) FRP jacket
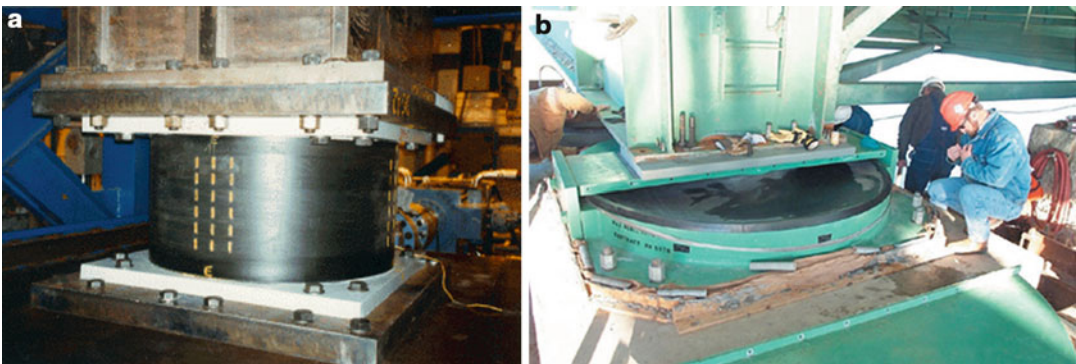
Strengthening Techniques: Bridges, Fig. 14 Bridge bearings: (**a**) elastomeric (Photo Credit Federal Highway Administration. https://www.fhwa.dot.gov/publications/ publicroads/99marapr/images/seismic3.gif), (**b**) friction pendulum (Photo credit: Federal Highway Administration)

an overview of damage to structural elements of bridges have been presented. Strengthening techniques covered include restrainers, expansion joint seat width extenders, column jackets, footings, abutments, and cap beam-column joints. Retrofit measures such as seismic isolation reduce the demand imposed on a bridge system thus improving the seismic performance. Hinges and abutments are the two areas where retrofitted bridges continue to experience damage. The 1994 Northridge and 2011 Great East Japan earthquakes have shown that retrofitted bridges did not experience any serious damage. This clearly demonstrates that seismic strengthening of bridges is valuable and should be continued. A unique feature of earthquakes is that they often expose the weaknesses of a structural system. Despite the

progress made in designing new and strengthening existing bridges, it is likely that future earthquakes will expose additional vulnerabilities of existing and retrofitted bridges.

Cross-References

▶ Bridge Foundations
▶ Resilience to Earthquake Disasters
▶ Retrofitting and Strengthening of Structures: Basic Principles of Structural Interventions
▶ Seismic Resilience
▶ Seismic Strengthening Strategies for Existing (Code-Deficient) Ordinary Structures
▶ Strengthened Structural Members and Structures: Analytical Assessment

References

American Concrete Institute, ACI Committee 440 (2002) Guide for the design and construction of externally bonded FRP systems for strengthening concrete structures. ACI 440.2R-02, Farmington Hills

Anderson DL, Mitchell D, Tinawi RG (1996) Performance of concrete bridges during the Hyogo-Ken Nanbu (Kobe) earthquake on January 1995. Canadian J Civil Eng 23:714–726

Buckle IG, Goodson M, Cassano R, Douglas B, Hegemier G, Imbsen R, Jones D, Liu D, Mayes R, North P, Priestley N, Rojahn C, Seible F, Selna L, Viest I (1990) Bridge structures. Earthq Spectra 6(S1):151–187

California Department of Transportation (2003) The race to seismic safety: protecting California's transportation system. Department of Transportation Advisory Board Report, Sacramento, 186 pp

California Department of Transportation (2006) Visual catalog of reinforced concrete bridge damage. Structure Maintenance and Investigations Division, Sacramento, 350 pp

Chai YH, Priestley MJN, Seible F (1991) Seismic retrofit of circular bridge columns for enhanced flexural performance. ACI Struct J 88(5):572–584

Federal Highway Administration (2006) Seismic retrofitting manual for highway structures: part 1-bridges. U.S. Department of Transportation, McLean, 656 pp

Gergely I, Pantelides CP, Reaveley LD (2000) Shear strengthening of R/C T-joints using CFRP composites. J Compos Constr 4(2):56–64

Hawkins NM, Kaspar II, Karshenas M (1999) Seismic retrofit of Poplar Street interchange in East St. Louis. Structures Congress. ASCE, Reston

Marsh ML, Buckle IG, Yen WP, Dasti S, Frost D, Monzon E (2011) Bridge performance during the Great East Japan Earthquake of March 11, 2011. http://www.pwri.go.jp/eng/ujnr/tc/g/pdf/27/27-1-3_Marsh.pdf

Mayes RL, Choudhury D, Crooks RS, Jones DM, Knight RP (1994) Seismic isolation retrofit of existing bridges. In: Proceedings of the materials engineering conference. ASCE, Reston, pp 1179–1186

Nishikawa K, Yamamoto S, Natori T, Terao K, Yasunami H, Terada M (1998) Retrofitting for seismic upgrading of steel bridge columns. Eng Struct 20(4–6):540–551

Ogata T, Osada K (2000) Seismic retrofitting of expressway bridges in Japan. Cem Concr Compos 22:17–27

Pantelides CP, Gergely I, Reaveley LD, Volnyy VA (1999) Retrofit of R/C bridge pier with CFRP advanced composites. J Struct Eng 125(10):1094–1099

Pantelides CP, Gergely J, Reaveley LD (2001) In-situ verification of rehabilitation and repair of reinforced concrete bridge bents under simulated seismic loads. Earthq Spectra 17(3):507–530

Pantelides CP, Alameddine F, Sardo T, Imbsen R (2004) Seismic retrofit of State Street bridge on Interstate 80. J Br Eng 9(4):333–342

Pantelides CP, Duffin JB, Reaveley LD (2007) Seismic strengthening of reinforced-concrete multicolumn bridge piers. Earthq Spectra 23(3):635–664

Park R (1996) An analysis of the failure of the columns of a 600 m length of the Hanshin elevated expressway during the Great Hanshin Earthquake of 17 January 1995. Bull N Z Natl Soc Earthq Eng 29(2):73–82

Priestley MJN, Seible F, Xiao Y, Verma R (1994) Steel jacket retrofitting of reinforced concrete bridge columns for enhanced shear strength-part 2: test results and comparison with theory. ACI Struct J 91(5):537–551

Priestley MJN, Seible F, Calvi GM (1996) Seismic design and retrofit of bridges. Wiley, New York, 704 pp

Seible F, Priestley MJN, Hegemier GH, Innamorato D (1997) Seismic retrofit of RC columns with continuous carbon fiber jackets. J Compos Constr 1(2):52–62

Thewalt CR, Stojadinovic B (1995) Behavior of bridge outrigger knee joint systems. Earthq Spectra 11(3):477–509

Yashinsky M (1998) Performance of bridge seismic retrofits during Northridge Earthquake. J Br Eng 3(1):1–14

Strengthening Techniques: Code-Deficient R/C Buildings

Andreas P. Lampropoulos
School of Environment and Technology,
University of Brighton, Brighton, UK

Synonyms

Epoxy grouting; Interfaces; Monolithicity coefficients; RC jacketing; Repair; Shrinkage; Strengthening

Introduction

Nowadays, many buildings need to be upgraded, especially in earthquake-prone areas, because of increased strength demands introduced by modern design codes or damage due to strong earthquakes. As a result, many techniques have been developed to improve the performance of the existing buildings. There is a wide range of applications of fiber-reinforced polymers (FRPs) in the form of rebars, plates, and sheets for the

strengthening of the existing buildings. These materials can be used for flexural and/or shear strengthening of existing structures, and a comprehensive description of this technique can be found in Essay 382,109 "Retrofitting and Strengthening of Contemporary Structures: Materials Used."

In this entry, special emphasis is given on the technique of reinforced concrete (RC) jacketing. RC jacketing is a widespread technique used for the strengthening of existing columns, beams, and walls. The use of epoxy resins for the crack repair of damaged structural elements is also presented.

In earthquake-prone areas, the majority of the old buildings could be characterized as "code deficient," and these buildings need to be upgraded.

"Code-deficient" buildings are the buildings that do not satisfy the requirements proposed by the new codes (Eurocode 8 2005; Code of Structural Interventions 2012; fib 2010). There are different factors responsible for structural deficiencies in existing buildings. The structural system of many old buildings was designed with configuration problems. Lack of regularity in geometry, in strength or stiffness, in plan, or in elevation are common characteristics of several existing buildings. Also, in the past, approximations and simplifications were adopted in carrying out structural analysis. Computers were not available, 3D analysis was not feasible, even 2D analysis was rarely used, and (continuous) beams and columns were considered as independent elements. Some critical requirements concerning the behavior of structures under earthquakes, such as ductility, capacity design, and detailing provisions (minimum amount of stirrups, lower limits for compressive strength, upper limit for tensile reinforcement), were also ignored. Also, in the past, the seismic actions used for the design of structures were much lower than the actions currently used for the design of new structures (Dritsos 2012).

As a consequence, the majority of existing buildings are deficient and in need of strengthening. Also, some of the old buildings are already damaged after previous strong earthquakes. The type of intervention that will be used for

strengthening deficient buildings depends on the level of damage and the requirements. In case of lightly damaged concrete elements, where there are no requirements for further upgrading of the structure, the technique of epoxy grouting is widely used. Epoxy grouting can be used to improve the performance and the durability of cracked reinforced concrete elements.

In case of heavily damaged structural elements or structures that need to be upgraded to conform with modern codes, seismic strengthening is required. There are a number of available techniques for the earthquake strengthening of existing buildings, such as jacketing of load-bearing elements, addition of infill walls or bracing systems, and installation of energy dissipation systems. For the jacketing of structural elements, steel, fiber-reinforced polymers (FRPs), and reinforced concrete can be used. The decision for the selection of the most suitable technique is not always easy and depends on a number of parameters such as cost, required time, and disruption of the occupants. Detailed descriptions of a number of different repair and strengthening techniques have been published by Tassios (2009) and Dritsos (2005).

In this entry, the technique of repair of damaged buildings with epoxy grouting and the strengthening with RC jackets are described in detail, and practical recommendations are presented. Special focus is given on the modeling and design of the jacketed elements. Critical parameters for the effectiveness of this technique are the bond between the old and new concrete elements and shrinkage of the new concrete. To take into account these two parameters, finite element method can be used. However, for the practical design of strengthened elements the use of monolithicity coefficients is proposed, which is a simplified procedure. Monolithicity coefficients are used to correlate the strength, stiffness, and deformation at yield and failure of strengthened columns to those of respective monolithic columns, the behavior of which can be determined by following conventional concrete design procedures. Monolithicity coefficients for jackets with different characteristics are presented at the end of this entry.

Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 1 (**a**) Preparation of cracks before the use of epoxy resin and (**b**) repaired columns. http://www.episkeves.civil.upatras.gr

Repairing Existing RC Structures Using Epoxy Resins and Bonded Mortar

Epoxy grouting is one of the most commonly used techniques for the repair of cracked structural elements. For the successful application of this technique, the surface of the damaged or cracked section needs to be properly prepared (Fig. 1a). Depending on the type of damage, different epoxy resin types and installation procedures can be used (Fig. 1b).

In this section, the technique of epoxy resin injection for crack sealing and the use of epoxy-bonded concrete replacement for the repair of damaged concrete are described.

Epoxy Resin Injection for Crack Sealing

The resin characteristics and the application methods used for the epoxy resin injection depend on the environmental and concrete temperature, the crack width, the thickness of the concrete member, and the extent of cracking.

Substrate Preparation

The first critical step for the successful application of this method is the preparation of the substrate. Any existing contaminants should be removed to improve substrate conditions. All the existing coatings along the crack, at intervals of not less than the thickness of the concrete member or 10 cm (ETEP 2012), should be removed.

Port Installation

Entry ports (tube-like devices that provide for the successful transfer of the epoxy resin under pressure into the crack) together with "cap" crack sealing should be installed on the face of the crack to ensure containment of the epoxy as it will be injected under pressure into the crack (ACI 2003; Fig. 2). In case of crack penetration throughout a section, the use of cap seals on both sides of the crack elements is suggested (ACI 2003).

The distance of the ports depends on the opening at the face of the crack. The recommended port distance values proposed by ETEP (2012) are presented in Table 1.

Before the cap seal installation, the location of the widest portion of the crack should be marked. The material used for the seal cap should be carefully prepared with accurate batching of components and with a consistent application of the material over the crack. The use of 25 mm wide and 5 mm thick cap seals is proposed by ACI (2003).

S

Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 2 Installation of ports for crack sealing with epoxy resins
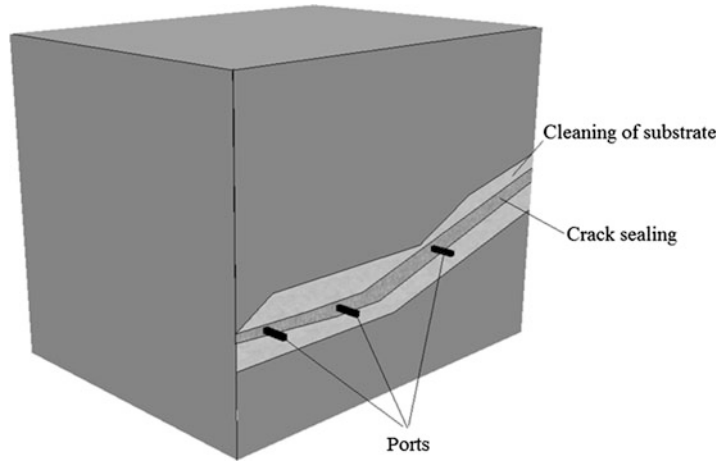
Strengthening Techniques: Code-Deficient R/C Buildings, Table 1 Recommended port spacing for different crack widths (ETEP 2012)

Crack width (mm)	Port spacing (mm)
0.3–0.5	100
0.5–1	100–135
1–2	135–170
2–3	170–200
3–10	200

Epoxy Resin Injection

Epoxy resins cure to form solids with high strength and relatively high modulus of elasticity. Epoxy resin used for crack injection should be a 100 % solid resin that meets the requirements of specification ASTM C-881 (2010) for type I or IV, grade 1 (low viscosity), class B or C (depending on the ambient temperature). If the purpose of injection is to restore concrete to its original load-bearing capacity, a type IV epoxy should be used. If full restoration of load-bearing capacity is not envisaged, a type I epoxy is sufficient. No solvents or nonreactive diluents are permitted in the resin (USBR 1997). Class B epoxy is used between 4 °C and 16 °C, while class C is used above 16 °C (ASTM C-881 2010).

The injection should start from the bottom in the case of vertical cracks, while for horizontal cracks the starting point should be the widest crack section. The injection should continue until refusal or bleeding of the adjacent port (ACI 2003; ETEP 2012).

Strengthening Techniques: Code-Deficient R/C Buildings, Table 2 Recommended resin viscosity and injection pressure values for different crack widths (ETEP 2012)

Crack width (mm)	Dynamic viscosity of the resin (cps)	Required injection pressure (MPa)
0.3–0.5	1,000	–
	500	0.8
	250	0.4
	130	0.2
0.5–1	1,000	0.8
	500	0.4
	250	0.2
1–2	1,000	0.4
	500	0.2
	250	0.1
2–3	1,000	0.3
	500	0.1
	250	0.005

The injection pressure and the dynamic viscosity of the resin depend on the crack width. In Table 2, the recommended values proposed by ETEP (2012) are presented.

Upon completion of the resin injection, the ports and the cap seals will be removed by heat chipping or grinding (ACI 2003).

Quality Control

The first check should be done by hand 48 h after the injection, to check if the coagulation

(polymerization) of the epoxy resin has been completed.

Then, an additional check is required to ensure that the crack is filled with the resin. Cores should be taken with a diameter of 25–50 mm and depth equal to the thickness of the cracked element and at least equal to 15 cm (ETEP 2012). At least one core is required for 30 m crack length, and the core will be visually inspected to check if the crack is filled to the appropriate level with resin. The cores can also be tested for compressive and tensile loading following ASTM C42 (1999) and (ACI 2003). The removed core should be patched with a non-shrink or expansive cementitious or epoxy grout (ETEP 2012; ACI 2003).

The use of an endoscope is also recommended for quality control. In this method, holes with a diameter 2 mm larger than the diameter of the tube of the endoscope should be drilled, to a depth equal to the thickness of the cracked element and at least 15 cm. At least two points (in two different crack segments) for every 30 m of crack injected (total length of the injected crack segments) should be inspected. There are also non-destructive methods suggested for the inspection of the repaired elements. Use of the ultrasonic pulse velocity method is also an option (ETEP 2012; ACI 2003). At least three measurements should be taken (in three different crack segments) for every 30 m of crack injected (total length of the injected crack segments) (ETEP 2012).

Epoxy-Bonded Mortar or Concrete for Concrete Repair

Epoxy-bonded mortar or concrete is used for repairs involving shallow replacement of concrete. Epoxy-bonded mortar is used for depth of replacement less than 40 mm, while epoxy-bonded concrete is used in depths between 40 and 140 mm (USBR 1997). Shallow concrete replacements are subject to poor curing conditions as a result of moisture loss to evaporation and to capillary absorption by the existing concrete element with subsequent poor bond to the existing structure. The use of epoxy-bonded mortar or concrete can be used to ensure adequate bond between the old and the new material

(USBR 1997). The surface preparation of the old concrete and the materials used for the repair are critical parameters for the effectiveness of the technique.

Old Concrete Preparation

The first step for the old concrete preparation is to saw cut the perimeter of the damaged area to a depth of 25–40 mm in order to provide a retaining boundary for the compaction and consolidation of the repair material. Then, the deteriorated concrete needs to be removed to create a sound substrate for the bonding of the new repair material (USBR 1997). There are several techniques for the damaged concrete removal such as the use of high-pressure hydroblasting, jackhammer, shotblasting, and dry or wet sandblasting (USBR 1997).

Materials Used for Epoxy-Bonded Repair

The epoxy resin used for epoxy-bonded mortar repairs (less than 40 mm depth) should be two component, 100 % solid type, meeting the requirements of specification ASTM C-881 (2010) for type III, grade 2 (medium viscosity), class B or C (depending on the ambient temperature).

The material used in this technique is conventional concrete and epoxy resin bonding agent. The concrete should be the same as the old concrete but with consistency (slump value) within 40 mm. The resin should be of two components, 100 % solid type, meeting the requirements of specification ASTM C-881 (2010) for type II or type V, class B or C (depending on the ambient temperature). Type II can be used for use in non-load-bearing applications to bond fresh to hardened concrete, while for load-bearing applications, type V should be used (ASTM C-881 2010).

Strengthening Using Additional Layers/ Jackets

This technique involves casting of an additional reinforced concrete element in connection to the existing load-bearing elements and is extensively used to upgrade existing buildings (Fig. 3).
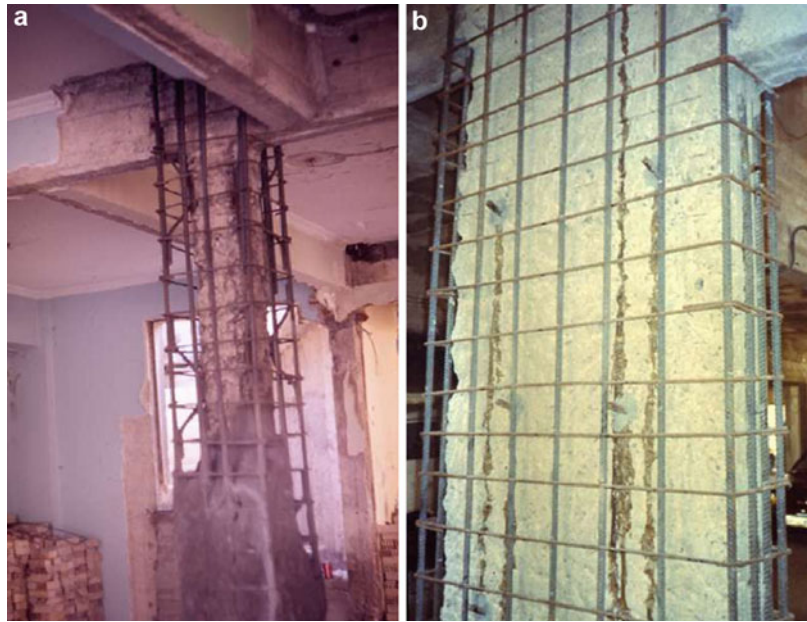
Two crucial issues for the effectiveness of this technique are the interface between the old and the new concrete and shrinkage of the new concrete layer or jacket (Lampropoulos and Dritsos 2011). Concrete shrinkage induces an initial slip at the interface which in some cases can be high enough and even lead to debonding (Beushausen and Alexander 2006, 2007). Also, tensile stresses are developed at the new concrete, which may lead to a reduced concrete compressive strength due to biaxial stress state (Lampropoulos and Dritsos 2010, 2011; Lampropoulos et al. 2012a). Practical recommendations for the strengthening of the additional elements using additional layers or jackets are presented in section "Interface Preparation", while in section "Analysis and Design of Strengthened Elements", analysis and design methods for strengthened/composite specimens are presented.

Practical Recommendations for Jacketing Procedure

Interface Preparation

In order to guarantee a sufficient connection between contact surfaces, the shear force acting at the interface should be lower than, or equal to, the shear resistance.

Four mechanisms contribute to the shear resistance at the interface: concrete-to-concrete adhesion, concrete-to-concrete friction (Fig. 4a), the connecting action from either steel bars placed across the interface between the old and the new concrete (Fig. 4b), and bent-down bars welded between the bars of the old and the new concrete (Fig. 4c; Dritsos 2007). These four mechanisms can be subdivided into the two groups of unreinforced and reinforced interfaces, depending on whether or not additional steel is placed across the interface or welded between the bars of the old and the new concrete (Dritsos 2007).

The roughening of the interfaces is crucial for the performance of the strengthened specimens especially in case of unreinforced interfaces. The roughening of the surface of the existing elements can be achieved using jet with water and sand mixture or light air equipment or electric needle (Code of Structural Interventions 2012). In Fig. 5, the use of air chipping hammer for roughening the surface of a beam is presented.

There are two main techniques currently used to characterize the interface roughness. The one of the two is a qualitative approach, wherein the surface of the existing elements is compared to standard surface samples with different roughness

Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 4 Friction, steel bars across the interface, and bent-down bars used in concrete-to-concrete interfaces

Concrete-to-concrete friction

Steel bars across the interface

Bent down bars
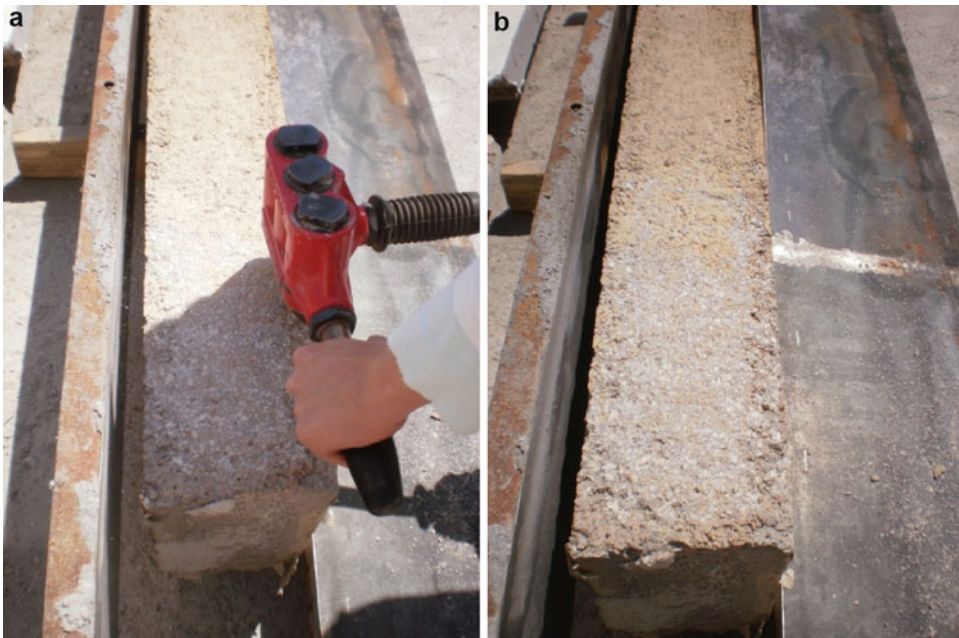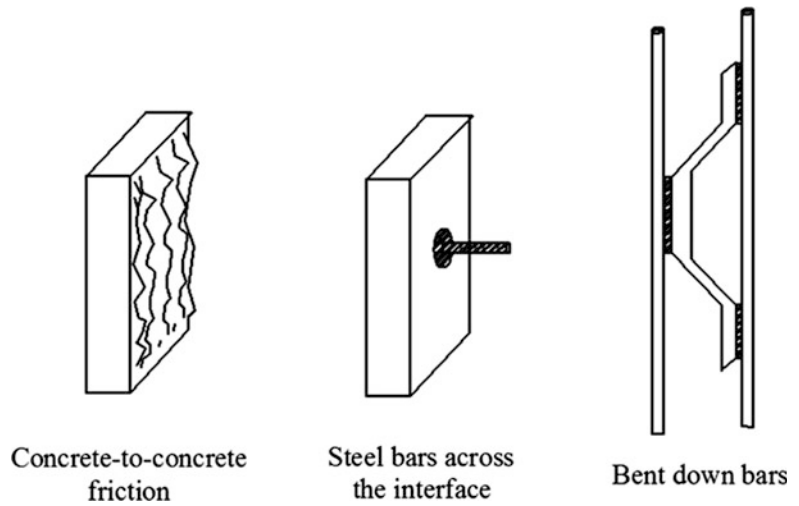
Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 5 (**a**) Roughening procedure and (**b**) roughened surface

grades. The second one, which is a quantitative approach, is the sand patch test. This method consists in uniformly spreading a standard volume of sand (V) and then measuring the diameter (D) of the covered area from which the roughness is calculated ($=4 \text{ V}/\pi\text{D}^2$). This method presents the advantage of being inexpensive and easy to perform, both in place and in laboratory, and as all interfaces are horizontal top surfaces, it is easy to be applied (ASTM E965 2003; Santos and Júlio 2010, 2013).

Based on the roughness value, the surface can be classified as very smooth when the roughness is not measurable, smooth when the roughness is

Fig. 6 Construction detail of a typical jacketed column (Tassios 2009)

Anchorage with double bolts

Epoxy resin injection

At least one bent down bar in the joint

New longitudinal bars

New stirrups

Bent down bars

Concrete jacket

Existing longitudinal reinforcement
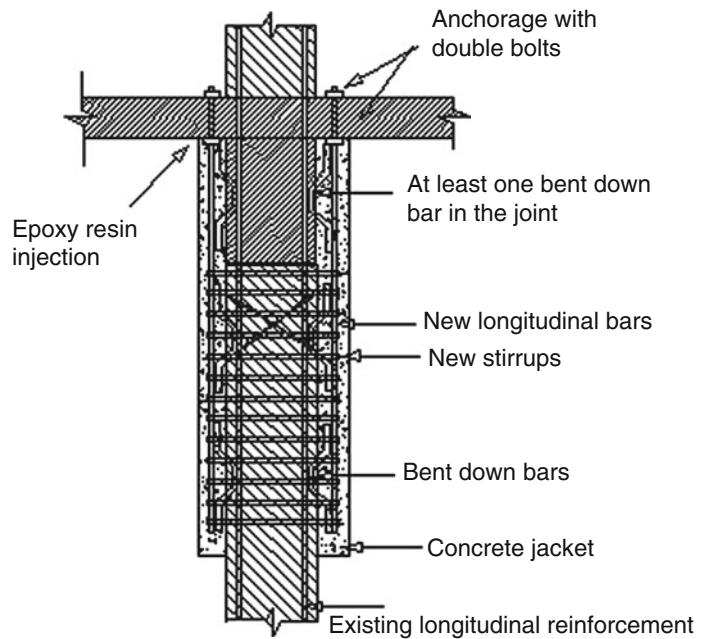
lower than 1.5 mm, rough when the roughness is higher than 1.5 mm, and very rough when the roughness is higher than 3 mm (Santos and Júlio 2013; fib 2010).

Apart from the two techniques presented above, there are other "relatively new" techniques used to determine the roughness of the interface (i.e., outflow meter, mechanical stylus, circular texture meter, digital surface roughness meter, microscopy, ultrasonic, slit-island method, roughness gradient method, photogrammetric method, shadow profilometry, air leakage method, processing of the digital image (PDI) method, two-dimensional laser roughness analyzer (2D-LRA) method, and three-dimensional laser scanning method). Most of these methods are based on optical techniques, and their main benefits are the improved accuracy and the fact that they are not destructive methods and can be applied both in the laboratory and in situ (Santos and Júlio 2010).

Casting of New Concrete and Reinforcement Details

For the casting of new concrete, small aggregate size of about 2 mm is normally used because of the lack of space in the jacket or the additional

layer, while use of non-shrinkage concrete is also recommended (Júlio et al. 2003).

The anchorage of the added longitudinal reinforcement to the existing element is essential to prevent debonding. In case of jacketed columns, continuity between floors is recommended (Júlio et al. 2003). However, it is not always feasible to continue a jacket to the other floors. In such cases, or in order to minimize occupant disturbance, the following construction details are required (Fig. 6) (Tassios 2009):

(i) Filling of the gap between the jacket and the slab with epoxy resin injection to ensure effective compressive load transfer. The epoxy resin should be injected at least one month after the casting of the jacket.

(ii) The longitudinal bars of the jacket should be anchored to the slab with double bolts to ensure tensile load transfer.

(iii) Welding of at least one bent-down bar with each reinforcing bar of the initial column.

In case of "open" jackets, where the jacket cannot be cast on all four sides of the column, special provisions are required to prevent debonding. In three-sided jackets, epoxy grout

Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 7 Typical shear stress-slip model for roughened interface and for interface with dowels
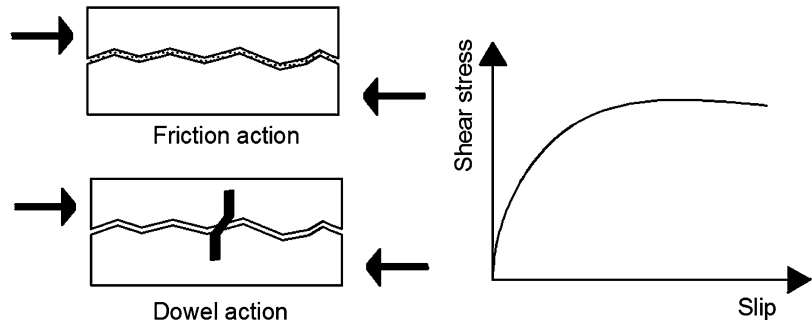
should be used to integrate the stirrups of the jacket into the initial column. In case of strengthening with additional layers only on one or two sides, the connection of the new with the old reinforcement using bent-down bars is essential (Dritsos 2005).

Analysis and Design of Strengthened Elements

To take into account the effect of the interface and concrete shrinkage, finite element analyses can be conducted (Lampropoulos and Dritsos 2011). However, in practice, it is not feasible to perform sophisticated finite element analyses and take into account all these parameters, so the use of monolithicity coefficients is proposed by the codes (Eurocode 8 2005; Code of Structural Interventions 2012). Monolithicity coefficients correlate the characteristics of the strengthened members to those of the respective monolithic ones. This procedure is quite old; however, most of the proposed monolithicity coefficients are empirical, and single values are proposed without accounting for the different values of the normalized axial load, the thickness of the jacket, and the shrinkage of the new concrete parameters that considerably affect the response of the strengthened specimens (Lampropoulos et al. 2012b). In the literature, there are studies (Dritsos 2007; Thermou et al. 2007; Thermou et al. 2014) where monolithicity coefficients values are proposed, and the effect of jacket concrete shrinkage has been ignored. It must be stated that the restrained shrinkage of the jacket concrete significantly influences the response of strengthened columns and cannot be overlooked (Lampropoulos and Dritsos 2010; Lampropoulos

et al. 2012a). Detailed investigations have recently been published where monolithicity coefficient values have been calculated for different values of the normalized axial load and jacket thicknesses (Lampropoulos et al. 2012b; Thermou et al. 2014). The effect of jacket's restraint shrinkage has also been studied (Lampropoulos et al. 2012b), and it has been shown that in some of the examined cases, the "code"-proposed values are not conservative. In the following section (section "Concrete-to-Concrete Interface"), different models for the shear stress calculation at the interface are presented. The proposed method for modeling concrete shrinkage effects is presented in section "Concrete Shrinkage Effect".

Concrete-to-Concrete Interface

There are a number of available models for the calculation of the interface strength between old and new concrete. In most of these models, the shear stress at the interface between the old and new concrete is related to the slip at the interface (Tassios 1986; Tassios and Vintzileou 1987; CEB-FIP 1993; Code of Structural Interventions 2012). In these studies, concrete-to-concrete interfaces with different roughening grades were examined. A typical shear stress-slip model for roughened interfaces with and without steel dowels is presented in Fig. 7.

For the shear resistance due to cohesion, the following model is proposed by CEB-FIP (1993) and Code of Structural Interventions (2012). The distribution of the shear stress with the slip at the interface is considered linear until the limit value for the shear resistance (τ_{fud}) which is the shear stress value for slip equal to 0.01-0.02 mm.

After that point, shear resistance is considered to be constant and equal to the limit value (τ_{fud}). The τ_{fud} value depends on the interface type Eq. 1.

$$\tau_{fud} = \begin{cases} 0.25 \cdot f_{ct}, & \text{smooth interface} \\ 0.75 \cdot f_{ct}, & \text{rough interface} \\ f_{ct}, & \text{shotcreting} \end{cases} \quad (1)$$

where:

τ_{fud} is the shear resistance at the interface,
f_{ct} is the concrete tensile strength.

The shear stress-slip distribution is presented in Fig. 8.

In case of smooth interfaces, there are also similar models for the shear stress distribution due to friction (Tassios and Vintzileou 1987; CEB-FIP 1993; Code of Structural Interventions 2012; Fig. 9).

According to this model, the shear stress distribution is linearly increasing up to the point of its limit value, while after that point the stress remains constant. The maximum shear stress point is calculated using Eq. 2.

$$\tau_{fud} = 0.4 \cdot \sigma_c, \quad s_{fu} = 0.15 \cdot \sqrt{\sigma_c} \quad (2)$$

where
s_{fu} is the slip at maximum shear strength (τ_{fud}) and σ_c is the design compressive strength at the interface.

Various models have also been proposed for roughened interfaces. The first model for roughened interfaces is presented in Fig. 10 (Tassios 1986; CEB-FIP 1993).

According to this model, the shear strength of roughened interfaces is linearly increasing up to a slip equal to 0.1 mm. At this point, the shear stress is equal to half of its maximum value. For slip values higher than 0.1 mm and lower than 2.0 mm, the shear resistance can be derived using Eq. 3.

$$\left(\frac{\tau_f}{\tau_{fud}}\right)^4 - 0.5 * \left(\frac{\tau_f}{\tau_{fud}}\right)^3 = 0.3 * s - 0.03$$
$$\text{where} \quad \tau_{fud} = 0.4 * \left(f_c^2 * \sigma_c\right)^{1/3}$$
$$(3)$$
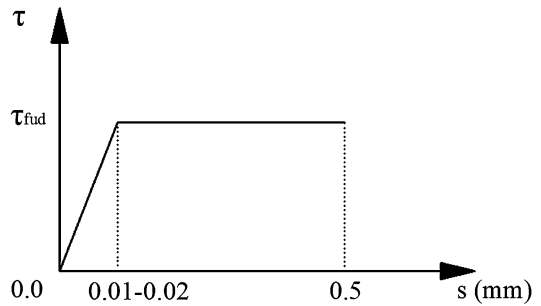
Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 8 Shear stress-slip model for concrete-to-concrete interfaces due to cohesion (CEB-FIP 1993; Code of Structural Interventions 2012)

A similar model is presented in the Code of Structural Interventions (2012), where the shear stress distribution with the slip is defined using two different equations, before and after half of the maximum slip value Eq. 4 (Fig. 11):

$$\frac{s}{s_{fu}} \leq 0.5 \rightarrow \frac{\tau_f}{\tau_{fud}}$$
$$= 1.14 \cdot \sqrt[3]{\frac{s}{s_{fu}}}, \frac{s}{s_{fu}} > 0.5 \rightarrow \frac{\tau_f}{\tau_{fud}}$$
$$= 0.81 + 0.19\frac{s}{s_{fu}} \quad \text{where} \quad \tau_{fud} = 0.4 * \left(f_c^2 * \sigma_c\right)^{1/3}$$
$$(4)$$

The recommended maximum shear stress at the interface (τ_{fud}) proposed by Model Code 2010 (fib 2010) is in the range of 1.5–2.5 MPa for strengthened interfaces with sandblasting, while for a well-roughened interface with high-pressure hydroblasting, the maximum shear stress is in the range of 2.5–3.5 MPa. These values are valid for concrete with strength lower than C50/60.

Values for the coefficients of friction and cohesion at the interface of strengthened elements have also been proposed. The suggested values according to CEB (1983), together with the coefficient of friction values proposed by Eurocode 2 (1996) and Model Code 2010 (fib 2010) are presented in Table 3.

From the values of Table 3, it can be observed that the design values for the coefficient of friction proposed by Eurocode 2 (1996) are lower for

Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 9 Shear stress-slip models due to friction for smooth interfaces (Tassios and Vintzileou 1987; CEB-FIP 1993; Code of Structural Interventions 2012)
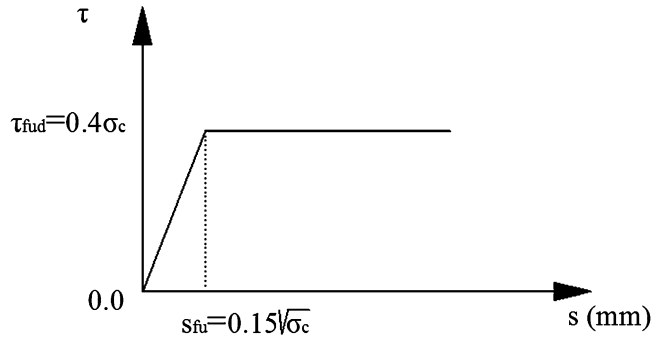
Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 10 Shear stress-slip distribution due to friction for roughened interfaces (Tassios 1986; CEB-FIP 1993)

$$\left(\frac{\tau_f}{\tau_{fud}}\right)^4 - 0.5 \cdot \left(\frac{\tau_f}{\tau_{fud}}\right)^3 = 0.3 \cdot s - 0.03$$

$$\tau_{fud} = 0.4 \cdot \left(f_c^2 \cdot \sigma_c\right)^{1/3}$$
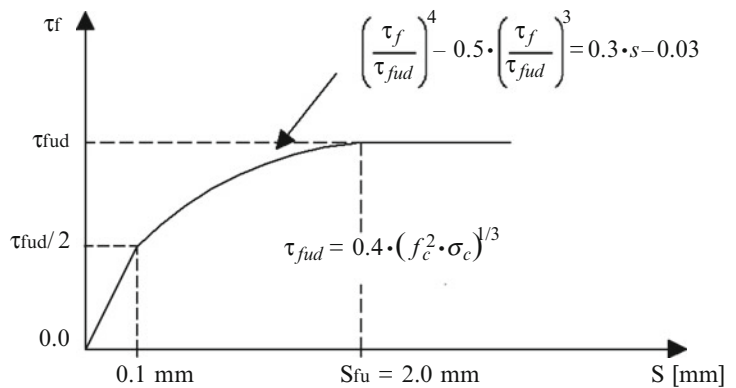
Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 11 Shear stress-slip distribution due to friction in roughened interfaces (Code of Structural Interventions 2012)
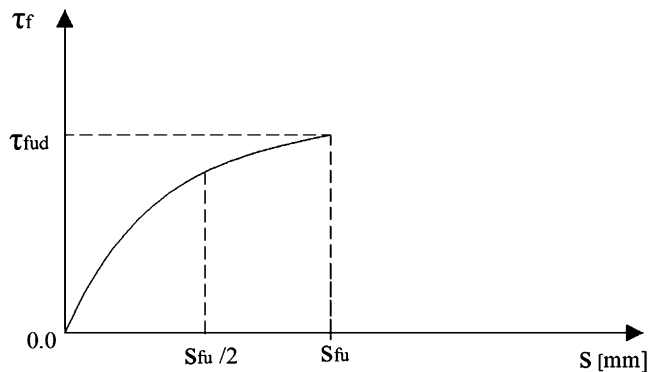
well-roughened interfaces compared to the values proposed by CEB Bulletin No. 162 (CEB 1983) and Model Code 2010 (fib 2010). The coefficient of friction values proposed by Model Code 2010 (fib 2010) are close to the values proposed by CEB (1983).

In the study of Tassios and Vintzileou (1987) and in the Code of Structural Interventions (2012), an analytical relationship for the calculation of coefficients of friction for different values of the normal to the interface stresses is proposed (Eq. 5).

$$\mu = 0.44 * \left(\frac{\sigma_c}{f_c}\right)^{-2/3} \qquad (5)$$

Another investigation has been conducted for dowel-reinforced interfaces, and models for the distribution of the shear stress with the friction

Strengthening Techniques: Code-Deficient R/C Buildings, Table 3 Coefficient of friction and cohesion values for strengthened elements with various roughening levels (CEB 1983; Eurocode 2 1996; fib 2010)

	CEB Bulletin No. 162 (1983)		Eurocode 2 (1996)	Model code 2010 (fib 2010)
	Coefficient of friction	Cohesion (MPa)	Coefficient of friction	Coefficient of friction
Well-roughened interface	1.5	1.9	0.9	1–1.4
Roughened interface	0.9	1.7	0.7	0.7–1
Smooth interface	0.7	1.0	0.6	0.5–0.7
Very smooth interface	–	–	0.5	–

Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 12 Shear stress-slip relationship for dowel action (Vintzileou and Tassios 1986)
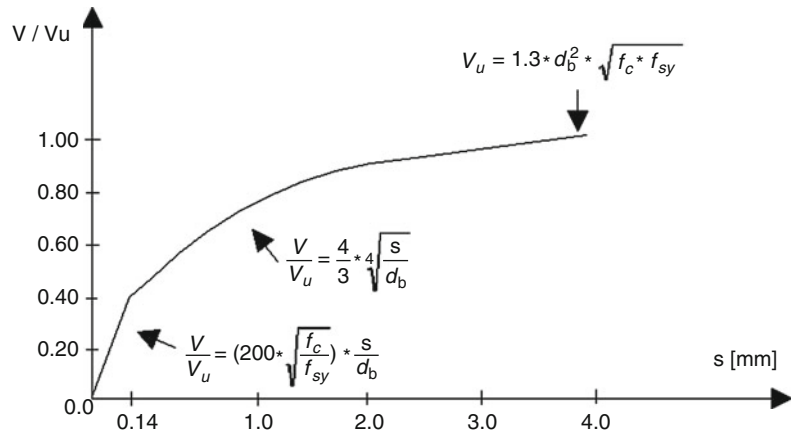
have been proposed (Vintzileou and Tassios 1986; CEB-FIP 1993; Code of Structural Interventions 2012). In the model proposed by Vintzileou and Tassios (1986), the variation of the shear stress with the slip is defined by two relationships, depending on the slip values (Eq. 6):

$$\frac{V}{V_u} = \left(200 * \sqrt{\frac{f_c}{f_{sy}}}\right) * \frac{s}{d_b} \quad \text{for} \quad s \le 0.14\,\text{mm, and} \quad \frac{V}{V_u} = \frac{4}{3} * \sqrt[4]{\frac{s}{d_b}} \quad \text{for} \quad s > 0.14\,\text{mm} \tag{6}$$
$$\text{where} \quad V_u = 1.3 * d_b^2 * \sqrt{f_c * f_{sy}}$$

where:

V	is the shear stress (N),
V_u	is the shear strength (N),
f_{sy}	is the yield stress of steel (MPa), and
d_b	is the diameter of the dowels (mm).

This model is presented in Fig. 12.

A similar bilinear model has been proposed by CEB (1983). The equation for the initial branch ($s \le 0.14$ mm) is similar to Eq. 6 with the only difference that a coefficient equal to 250 is proposed instead of 200. After the initial elastic branch, there is a second linear one up to the maximum shear resistance.

A different model has been presented in CEB-FIP (1993). According to this model, the shear resistance due to dowel action linear increases up to half of the maximum shear stress.

At this point, the slip equals to 0.05 of the maximum slip and after this point Eq. 7 is used. The shear stress-slip distribution is presented in Fig. 13.

$$\frac{s}{s_u} = 1.7 \cdot \left(\left[\frac{V}{V_u} \right]^4 - 0.5 \times \left[\frac{V}{V_u} \right]^3 \right) + 0.05 \quad \text{for} \quad 0.05 \leq \frac{s}{s_u} \leq 1 \quad \text{where} \quad s_u = 2.0 \text{ mm}$$

$$V_u = \frac{1.30}{\gamma_{Rd}} \cdot d_b^2 \cdot \left[\sqrt{1 + (1.3 \cdot \varepsilon)^2} - 1.3 \cdot \varepsilon \right] \cdot \sqrt{f_{cd} \cdot f_{yd} \cdot (1 - \zeta^2)} < A_s \cdot f_{yd} / \sqrt{3} \tag{7}$$

where

γ_{Rd}	is considered equal to 1.30,
A_s	is the dowel cross section, and
e	is the load eccentricity.

In the model proposed by the Code of Structural Interventions (2012), the maximum shear strength value is calculated using the same equation proposed by CEB-FIP (1993). The variation of the shear stress with the slip is considered linear for slip equal or lower than 0.005 dB. For any value between 0.005 and 0.05 dB, the shear stress value is calculated using Eq. 8. This model is presented in Fig. 14.

$$s = 0.1 \cdot d_u + 1.80 \cdot d_u \left(\left[\frac{V}{V_u} \right]^4 - 0.5 \cdot \left[\frac{V}{V_u} \right]^3 \right)$$

$$d_u = 0.05 \cdot d_b \tag{8}$$

The model for the dowel action proposed by Model Code 2010 (fib 2010) is given by Eq. 9

$$V_u = k * A_s * \sqrt{f_{cd} * f_{yd}} \left(\frac{s}{s_u} \right)^{0.5},$$

$$s_u = 0.1 - 0.2 * d_b \tag{9}$$

where k ~ 1.6 for circular reinforcement cross section and concrete strength equal to or lower than C50/60.

There are also models for the reduction of shear stress due to cyclic loading in roughened interfaces and for dowel action (Vintzileou and Tassios 1986, 1987; Tassios and Vintzileou 1987; Code of Structural Interventions 2012). According to these models, there is a shear stress reduction after each loading cycle. Characteristic diagrams of the shear stress distribution with the slip under cyclic loading are presented in Fig. 15 (Vintzileou and Tassios 1986, 1987; Tassios and Vintzileou 1987).

The first model for the calculation of the reduced shear stress due to cyclic loading is the one proposed by Tassios and Vintzileou (1987) (Eq. 10)

$$\tau_{f,n} = \tau_{f,1} * \left(1 - [0.002 * (n-1) * (s/s_{fu}) / (\sigma_{cn} / f_c)]^{1/3} \right) \tag{10}$$

where

$\tau_{f,n}$	is the shear stress at the interface in the nth loading cycle,
$\tau_{f,1}$	is the shear stress at the interface in the first loading cycle,
N	is the number of cycles, and
σ_{cn}	is the normal to the interface stress in the nth loading cycle.

In the Code of Structural Interventions (2012), two models have been presented for the shear stress reduction in smooth (Eq. 11) and roughened interfaces (Eq. 12).

Smooth interfaces:

$$\tau_{f,n} = \tau_{f,1} * \left(1 - 0.15 * \sqrt{(n-1)} \right) \tag{11}$$

Roughened interfaces:

$$\frac{\Delta \tau_{f,n}}{\tau_{f,1}} = 0.05 \cdot \left(\frac{f_c}{\sigma_{cn}} \right)^{1/2} \cdot (n-1)^{1/2} \cdot \left(\frac{s}{s_{fu}} \right)^{1/3} \tag{12}$$

S
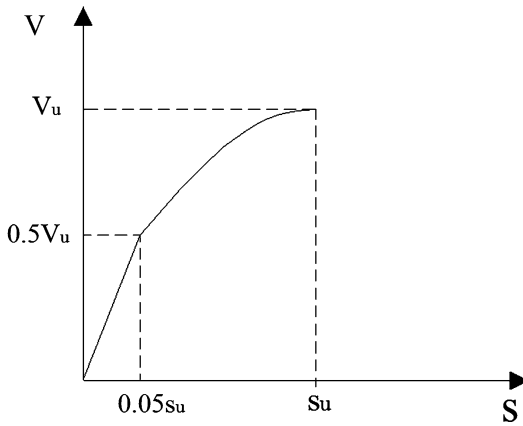
Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 13 Shear stress-slip relationship for dowel action (CEB-FIP 1993)
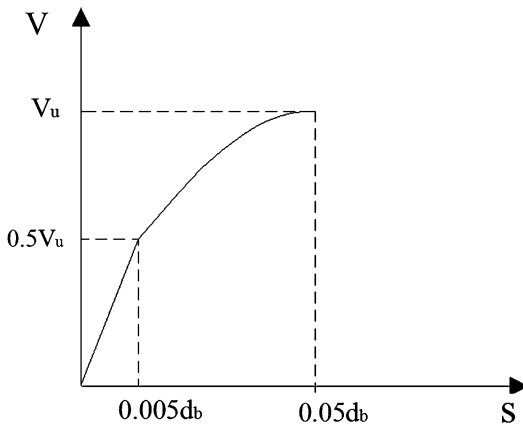
Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 14 Shear stress-slip relationship for dowel action (Code of Structural Interventions 2012)

The following model has been proposed by Vintzileou and Tassios (1986) and Vintzileou and Tassios (1987) for dowel action under cyclic loading (Eq. 13)

$$D_n = D_1 * \left(1 - \frac{1}{7} * \sqrt{(n-1)} \right) \qquad (13)$$

where

| D_n | is the maximum shear stress due to dowel action in the nth loading cycle, |
| D_1 | is the maximum shear stress due to dowel action in the 1st loading cycle. |

Concrete Shrinkage Effect

Concrete shrinkage can considerably affect the performance of strengthened elements with additional layers or jackets. Tensile stresses are induced to the new concrete of the layer or jacket while there is also an induced slip at the interface between the old and the new concrete. Denarie and Silfwerbrand (2004) presented a model where the tensile stresses of the new concrete due to restrained concrete shrinkage are proportional to coefficient "R." Coefficient "R" is characteristic of the connection between the old and the new concrete and the degree of restraint (Eq. 14).

$$\sigma = E * \varepsilon_{fsh} * R \qquad (14)$$

where

E:	modulus of elasticity of concrete,
ε_{fsh}:	free shrinkage strain, and
R:	degree of restraint.

Values for the coefficient R have been proposed by Abbasnia et al. (2005) based on experimental investigations. The risk of debonding and cracking of the new layer of strengthened beams was investigated by Beushausen and Alexander (2006, 2007) where the tensile stress relaxation due to concrete creep was also taken into consideration.

The biaxial stress state in strengthened elements due to the development of restrained shrinkage tensile stresses was numerically investigated, and the importance of the restrained concrete shrinkage in case of jacketed columns was highlighted (Lampropoulos and Dritsos 2010, 2011). An extensive experimental study was also conducted to validate the reliability of the numerical results (Lampropoulos et al. 2012a). In this study, specimens with and without steel plates (with expanded polystyrene) were cast, to simulate the concrete of the jacket under restrained and free shrinkage, respectively. Specimens were stored until testing in a room with a constant temperature and relative humidity, and shrinkage strain at the end of the specimens was measured using micrometer dial
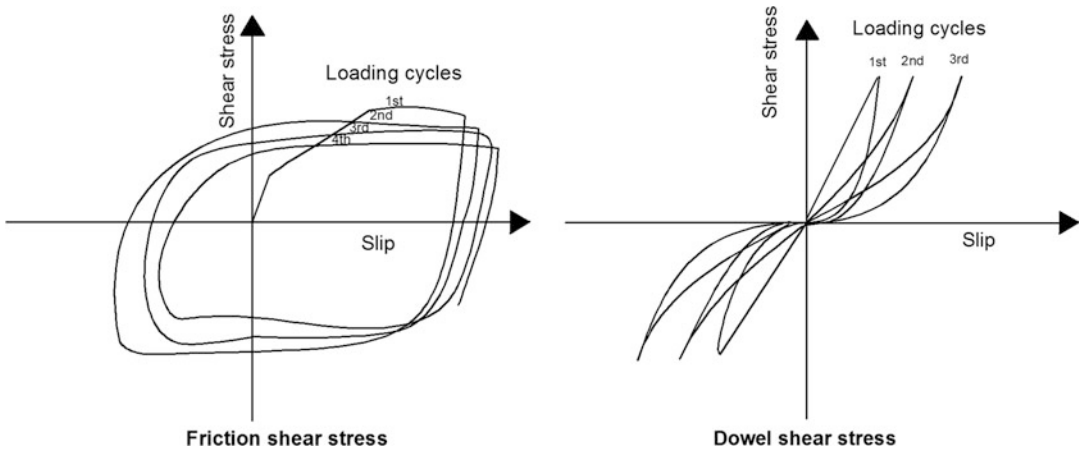
Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 15 Characteristic shear stress-slip diagrams for friction and dowel shear stress due to cyclic loading

Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 16 Specimens with and without steel plates and strain measuring setup

gauges. The specimens together with the setup for the shrinkage strain measurements are presented in Fig. 16.

Each pair of specimens (with steel plates or expanded polystyrene cast from the same batch of concrete) was tested in compression at a different age, to obtain results for different values of the shrinkage strain. The specimens with restrained and free concrete shrinkage were tested under compression (Fig. 17a). The ratio of the tensile stress to the tensile strength together with the respective reduction of the concrete strength was calculated. It was found that there is a significant reduction of concrete strength due to restrained concrete shrinkage and the subsequent biaxial stress state, and this is in agreement with the model proposed by Kupfer et al. (1969) (Fig. 17b).

In the tested square prism specimens, the reduction of the concrete strength was found to range from 20 % to 30 %. This reduction was more significant for larger values of tensile stresses caused by the restrained concrete shrinkage (Lampropoulos et al. 2012a).

Modeling of Strengthened Elements and Monolithicity Coefficients for the Design of Strengthened Elements

Finite Element Modeling

For modeling the strengthened elements, the finite element method has been used. Different assumptions have been examined, and the accuracy of the numerical models was validated using experimental data (Lampropoulos and Dritsos 2011).
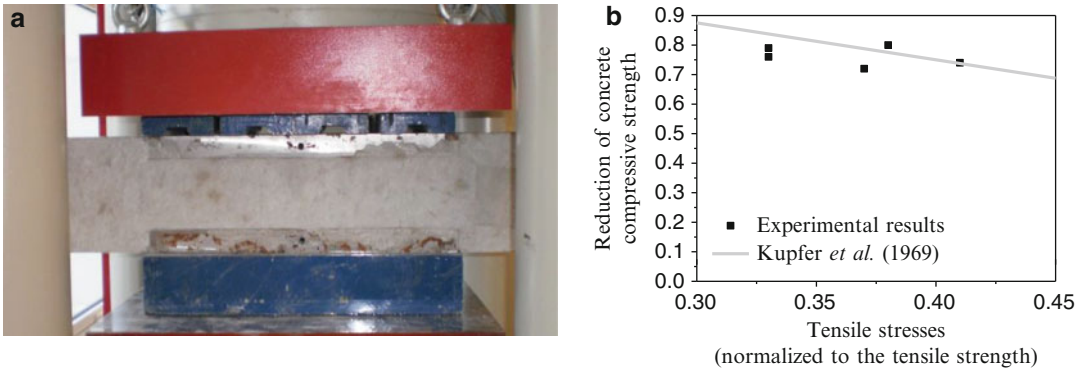
Strengthening Techniques: Code-Deficient R/C Buildings, Fig. 17 (**a**) Compressive test setup and (**b**) Biaxial stress state due to restrained concrete shrinkage (Lampropoulos et al. 2012a)

Interface Between the Old and the New Concrete The interface between the old and the new concrete can be simulated using spring elements or using special contact elements. To determine the characteristics of the interface, the models presented in section "Concrete-to-Concrete Interface" can be used. Strength deterioration due to cyclic loading should also be taken into account (Lampropoulos and Dritsos 2011).

Concrete Shrinkage Concrete shrinkage effect can be simulated by applying volumetric strain to the elements of the jacket. The effect of concrete creep should also be included to take into account the tensile stress relaxation. In a previous study (Lampropoulos and Dritsos 2010), in order to include creep effects, application of a volumetric strain equal to half of the free concrete shrinkage strain was proposed.

Monolithicity Coefficients
The accuracy of the numerical models was validated using a number of different experimental data. The same assumptions were used in a parametric study where columns with different characteristics were examined (Lampropoulos et al. 2012b). Analyses were performed for the strengthened and the respective monolithic specimens (Fig. 18), and monolithicity coefficients were calculated using Eq. 15. Monolithic specimens are specimens with exactly the same characteristics as the examined strengthened

columns, with the only difference that there is full connection (no slip) between the original column and the jacket, and concrete shrinkage is neglected.

The load and the deflection at the points of yield, maximum load, and failure of the specimens were estimated, and Eq. 15 was used to calculate monolithicity coefficient values.

$$
K_F = \frac{[F_{\max}]_{STR}}{[F_{\max}]_{MON}}, \quad K_k = \frac{\left[\frac{F_y}{\delta_Y}\right]_{STR}}{\left[\frac{F_y}{\delta_Y}\right]_{MON}},
$$
$$
K_{\delta,\theta_y} = \frac{[\delta_y]_{STR}}{[\delta_y]_{MO}}, \quad K_{\delta,\theta_u} = \frac{[\delta_u]_{STR}}{[\delta_u]_{MO}} \tag{15}
$$

where K_F, K_k, and $K_{\delta,\theta_{y,u}}$ are monolithicity coefficients for the strength, stiffness, and deflection or rotation angle at yield and failure, respectively, F_y and $F_{\max}$ are the yield and maximum load, and δ_y, δ_u are the deflections at yield and failure. Subscripts *STR* and *MO*N indicate strengthened and monolithic specimens, respectively.

The proposed design values are based on the assumption of concrete with free shrinkage strain of 800 microstrains. A roughened interface was assumed without any other connection technique (bent-down bars or dowels) which is a conservative assumption. Different values have been proposed depending on the normalized axial load (ν) and the ratio of the concrete cross-sectional area of the jacket (A_{cj}) to the cross-sectional area of

Strengthening Techniques: Code-Deficient R/C Buildings,
Fig. 18 Typical load deflection relationship for the strengthened and the respective monolithic specimens (Lampropoulos and Dritsos 2011, 2012b)
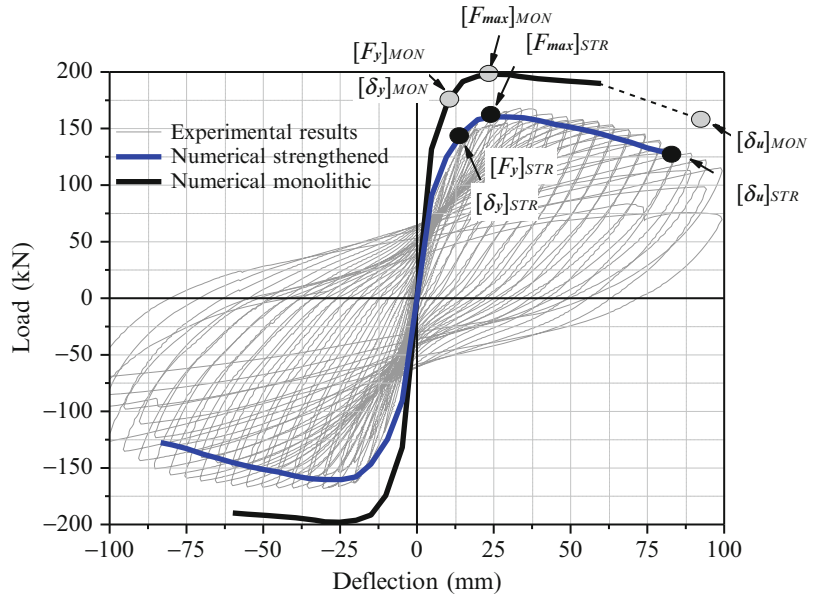
the original column (A_{co}) (A_{cj}/A_{co}). The normalized axial load (v) can be calculated using Eq. 16.

$$v = \frac{N}{A_{co} \cdot f_{co} + A_{cj} \cdot f_{cj}} \tag{16}$$

where N is the axial load, A_{co} is the cross-sectional area of the original column, f_{co} is the concrete strength of the original column, A_{cj} is the cross-sectional area of the concrete jacket, and f_{cj} is the strength of the jacket concrete.

Four- and three-sided jackets have been examined. In case of three-sided jacket, the stirrups of the jacket were merged into the initial column in order to simulate anchorage, which is essential and in practice is ensured by using epoxy grout (Lampropoulos et al. 2012b).

The proposed values for a four-sided jacket with A_{cj}/A_{co} ratio of 1.5 and normalized axial load values of 0.1, 0.2, and 0.4 are presented in Table 4 (Lampropoulos et al. 2012b together with the values proposed by existing codes (Eurocode 8 2005; Code of Structural Interventions 2012).

Table 5 presents monolithicity coefficient values for a normalized axial load of 0.2 and different A_{cj}/A_{co} ratio values for a four-sided jacket (Lampropoulos et al. 2012b).

Strengthening Techniques: Code-Deficient R/C Buildings, Table 4 Monolithicity coefficient values for a four-sided jacket with different normalized axial load values and an A_{cj}/A_{co} ratio of 1.5 (Lampropoulos and Dritsos 2012b)

		Monolithicity coefficients			
	v	K_F	K_k	K_{δ,θ_y}	K_{δ,θ_u}
Proposed design values	0.1, 0.2	0.85	0.55	1.50	1.00
	0.4	0.70	0.50	1.30	1.10
Eurocode 8, Part 3 (2005)		0.90	0.95	1.20	1.00
Code of Structural Interventions (2012)		0.90	0.80	1.25	0.80

The proposed values for the monolithicity coefficients for a three-sided jacket with an A_{cj}/A_{co} ratio of 1.0, and normalized axial load values of 0.1, 0.2, and 0.4 are presented in Table 6 (Lampropoulos et al. 2012b).

Table 7 presents proposed monolithicity coefficients values for a three-sided jacket and a normalized axial load value of 0.2 with different A_{cj}/A_{co} ratio values (Lampropoulos et al. 2012b).

Monolithicity coefficients for any case with a different A_{cj}/A_{co} ratio or normalized axial load

Strengthening Techniques: Code-Deficient R/C Buildings, Table 5 Monolithicity coefficient values for a four-sided jacket with different A_{cj}/A_{co} ratio values and a normalized axial load of 0.2 (Lampropoulos and Dritsos 2012b)

| | A_{cj}/A_{co} | Monolithicity coefficients | | | |
		K_F	K_k	K_{δ,θ_y}	K_{δ,θ_u}
Proposed design values	0.5	0.90	0.70	1.30	0.95
	1.5	0.85	0.55	1.50	1.00
	4.0	0.75	0.75	1.05	2.85
Eurocode 8, Part 3 (2005)		0.90	0.95	1.20	1.00
Code of Structural Interventions (2012)		0.90	0.80	1.25	0.80

Strengthening Techniques: Code-Deficient R/C Buildings, Table 7 Monolithicity coefficient values for a three-sided jacket with different A_{cj}/A_{co} ratio values and a normalized axial load of 0.2 (Lampropoulos and Dritsos 2012b)

| | A_{cj}/A_{co} | Monolithicity coefficients | | | |
		K_F	K_k	K_{δ,θ_y}	K_{δ,θ_u}
Proposed design values	0.5	0.85	0.55	1.70	1.15
	1.5	0.80	0.55	1.40	1.10
	4.0	0.75	0.25	3.00	1.00
Eurocode 8, Part 3 (Eurocode 8 2005)		0.90	0.95	1.20	1.00
Code of Structural Interventions (2012)		0.90	0.80	1.25	0.80

Strengthening Techniques: Code-Deficient R/C Buildings, Table 6 Monolithicity coefficient values for a three-sided jacket with different of normalized axial load values and an A_{cj}/A_{co} ratio of 1.0 (Lampropoulos and Dritsos 2012b)

| | ν | Monolithicity coefficients | | | |
		K_F	K_k	K_{δ,θ_y}	K_{δ,θ_u}
Proposed design values	0.1, 0.2	0.80	0.55	1.40	1.10
	0.4	0.70	0.35	1.90	1.70
Eurocode 8, Part 3 (2005)		0.90	0.95	1.20	1.00
Code of Structural Interventions (2012)		0.90	0.80	1.25	0.80

value from those presented in Tables 4, 5, 6, and 7 can be estimated by linear interpolation (Lampropoulos et al. 2012b).

The monolithicity coefficient values for the design of strengthened RC columns presented herein depend on the type of the jacket (four or three sided), the normalized axial load value, and the thickness of the jacket. The proposed values are more realistic than using one value to cover all situations, as is the case with present code recommendations (Eurocode 8 2005; Code of Structural Interventions 2012). It was also found that in many cases, monolithicity coefficient values proposed by codes are not conservative. The proposed values of the present work can also be used when considering other techniques to connect between the original column and the new jacket (such as dowels crossing the interface or bent-down bars welded between the two sets of longitudinal bars), as these values are conservative (only roughening at the interface was considered) and are suitable for design purposes (Lampropoulos et al. 2012b).

Summary

In this entry, the technique of concrete repair with epoxy grouting and the strengthening of existing structures with additional reinforced concrete layers and jackets were described in detail. Practical recommendations were presented for both techniques. For the strengthening of the existing structures using additional layers and jackets, recommendations for the analysis and monolithicity coefficients for the design of the strengthened elements were also presented.

The following general recommendations can be drawn:

• Epoxy resins can be used for the crack sealing and for the replacement of damaged concrete. Due consideration should be given to

successfully apply this method. In this chapter, a step-by-step procedure was presented, with detailed instructions about the epoxy resin injection technique, including details about the substrate preparation, the nozzle placement, and the quality control. The use of epoxy resins for shallow concrete replacements was also described, and information about the suitable material for these applications was presented.

- The technique of strengthening using additional layers and jackets is one of the most commonly used techniques. Two crucial parameters for the effectiveness of this technique are the connection between the old and the new elements and shrinkage of the new concrete.
- Models for the shear stress at the interface for reinforced and unreinforced interfaces and models for the strength deterioration of the interfaces due to cyclic loading were presented.
- For the modeling of the strengthened specimens, it is essential to simulate the interface between the old and the new concrete and the shrinkage effect of the new concrete. The interface can be simulated using contact elements, while concrete shrinkage effect can be simulated by applying a volumetric strain to the elements of the jacket.
- For the design of strengthened elements, monolithicity coefficient values can be used. Various values have been proposed for the design of strengthened RC columns, based on the type of the jacket (four or three sided), the normalized axial load value, and the thickness of the jacket.

Cross-References

References

Abbasnia R, Godossi P, Ahmadi J (2005) Prediction of restrained shrinkage based on restraint factors in patching repair mortar. Cem Concr Res 35:1909–1913

ACI (2003) Structural Crack repair by epoxy injection, field guide to concrete repair application procedures. ACI Committee E706, RAP Bulletin 1

ASTM C42/C42M–99 (1999) Standard test method for obtaining and testing drilled cores and sawed beams of concrete. West Conshohocken, Pennsylvania

ASTM C881/C881M (2010) Standard specification for Epoxy-resin-base bonding systems for concrete. ASTM International, West Conshohocken

ASTM E965 (2003) Standard test method for measuring pavement macrotexture depth using a volumetric technique. ASTM International, West Conshohocken

Beushausen H, Alexander MG (2006) Failure mechanisms and tensile relaxation of bonded concrete overlays subjected to differential shrinkage. Cem Concr Res 36:1908–1914

Beushausen H, Alexander MG (2007) Localised strain and stress in bonded concrete overlays subject to differential shrinkage. Mater Struct 40:189–199

CEB Bulletin No. 162 (1983) Assessment of concrete structures and design procedures for upgrading (Redesign), CEB, Paris

CEB-FIP (1993) Model code 1990. Thomas Telford, London

Code of Structural Interventions (2012) Team for development of code of interventions on reinforced concrete buildings harmonization team of code of interventions to Eurocodes. Earthquake Planning and Protection of Greece (E.P.P.O.), Athens

Denarie E, Silfwerbrand J (2004) Structural behaviour of bonded concrete overlays. In: Proceedings international RILEM workshop on bonded concrete overlays, Stockholm, pp 37–45

Dritsos SE (2005) Repair and strengthening of reinforced concrete structures. Dritsos, Patras (in Greek)

Dritsos SE (2007) Seismic strengthening of columns by adding new concrete. Bull N Z Soc Earthq Eng 40(2):49–68

Dritsos SE (2012) Seismic assessment and retrofit of existing buildings under the eurocodes framework. Assessment and Retrofitting Framework, Boku

S

ETEP (2012) Recommended technical specifications for retrofitting. Technical Chamber of Greece, Athens (in Greek)

Eurocode 2 (1996) Design of concrete structures Part 1.3: general rules – precast concrete elements and structures. British Standards Institution; ENV 1992-1-3, London

Eurocode 8 (2005) European standard EN. Design of structures for earthquake resistance, Part 3, assessment and retrofitting of buildings. CEN Technical Committee CEN/TC250, Brussels

fib Bulletin No 55 (2010) Model code 2010. Thomas, Telford

Júlio ES, Branco F, Silva VD (2003) Structural rehabilitation of columns with reinforced concrete jacketing. Prog Struct Eng Mater J 5:29–37

Kupfer H, Hilsdorf H, Rusch H (1969) Behaviour of concrete under biaxial stresses. Am Concr Inst J 66(8):656–666

Lampropoulos AP, Dritsos SE (2010) Concrete shrinkage effect on columns strengthened with concrete jackets. Struct Eng Int 20(3):234–239

Lampropoulos AP, Dritsos SE (2011) Modelling of RC columns strengthened with RC jackets. J Earthq Eng Struct Dyn 40(15):1689–1705

Lampropoulos AP, Tsioulou OT, Dritsos SE (2012a) Biaxial stress due to shrinkage in concrete jackets of strengthened columns. ACI Mater J 109(3): 331–340

Lampropoulos AP, Tsioulou OT, Dritsos SE (2012b) Monolithic coefficient values for design when seismically strengthening RC columns with jackets. J Earthq Eng 16(7):1023–1042

Santos P, Júlio E (2010) Comparison of methods for texture assessment of concrete surfaces. ACI Mater J 107(5):433–440

Santos P, Júlio E (2013) A state-of-the-art review on roughness quantification methods for concrete surfaces. Construct Build Mater 38:912–923

Tassios T (1986) Fundamental mechanisms of force-transfer across reinforced concrete critical interfaces. CEB workshops of Commissions II and IV, Karlsruhe

Tassios T (2009) Design theory for repair and strengthening. Symmetria Publications, Athens (in Greek)

Tassios T, Vintzileou E (1987) Concrete to concrete friction. Jf Struct Eng ASCE 113(4):832–849

Thermou GE, Pantazopoulou SJ, Elnashai AS (2007) Flexural behavior of brittle RC members rehabilitated with concrete jacketing. ASCE J Struct Eng 133(10): 1373–1384

Thermou GE, Papanikolaou VK, Kappos AJ (2014) Monolithicity factors for the design of R/C columns strengthened with R/C jackets. In: The 2nd European conference on earthquake engineering and seismology, Istanbul

USBR (1997) United States Department of the Interior Bureau of Reclamation Technical Service Centre, Guide to concrete repair. http://www.usbr.gov/pmts/materials_lab/concrete/

Vintzeleou E, Tassios T (1986) Mathematical model for dowel action under monotonic and cyclic conditions. Magazine of Concrete Research 38(134):13–22

Vintzileou E, Tassios T (1987) Behavior of dowels under cyclic deformations. Am Concr Inst Struct J 84(1):18–30. http://www.episkeves.civil.upatras.gr (Repair and Strengthening of Structures, University of Patras webpage)

Strengthening Techniques: Code-Deficient Steel Buildings

Konstantinos Daniel Tsavdaridis
School of Civil Engineering, University of Leeds, Leeds, UK

Synonyms

Braced frames; Composite materials; Connections; Cyclic behavior; Deformation capacity; Energy dissipation; Fuses; Repair; Seismic retrofitting; Steel buildings; Strengthening techniques

Introduction

The design of steel buildings is often governed by lateral wind loads and not seismic loads. Also, statistics indicate that the number of fatalities during earthquakes due to failure of all types of steel buildings is significantly less compared to other types of buildings. Consequently, much effort has been invested to seismically retrofit buildings having unreinforced masonry walls and reinforced concrete frames. However, recently steel buildings have received significant attention, while this interest mainly stems from the realization, following the 1994 Northridge earthquake, that the welded beam-to-column connections in moment-resisting frames were likely to fail in a brittle manner, prior the development of significant inelastic response, therefore negating the design intent and possibility causing safety hazards.

Recent research has expanded the variety and versatility of the tools available in the structural

engineer's toolbox to meet the seismic performance objectives. This entry provides an overview of how this research is expanding the available options for the seismic strengthening of steel buildings, by reporting on some selected research projects.

Structural strengthening and proving seismic resistance for steel building, but also masonry and reinforced concrete, may be done by first considering the direction of the weak links in the structures. For instance, for a heavy building with large dead load, this would be the major factor that contributes to the increase of lateral seismic load. Therefore, it is reasonable to first consider reducing the overall existing dead load and then provide the necessary strengthening technique for the lateral load-resisting system of the structure.

The use of structural steel in buildings' retrofitting can be often considered economical and efficient because:

- Steel buildings are particularly effective under performance-based design.
- Steel members exhibit ductile behavior beyond elastic limit, hence dissipate considerable amount of energy before damages occur.
- Steel members have higher strength-to-weight and stiffness-to-weight ratios; hence, the buildings attract less base shear under an earthquake.
- A better quality control is practiced in the production of the material as well as the fabrication and erection of them, while ensuring results close to the theoretical predictions.
- Steel can be generally used to retrofit all types of structures without increasing the dead weight dramatically, making the works less intrusive and time consuming.

Code-Deficient Buildings

All buildings can carry their own weight. They can usually carry a bit of snow and a few other floor loads vertically, so even badly built buildings and structures can resist some up-and-down loads. However, buildings and structures are not necessarily resistant to lateral loads, unless this has been taken into account carefully during the structural engineering design and construction phase with some earthquake-proof measures taken into consideration. It is the side-to-side load which causes the worst damage. Poorly designed buildings often collapse on the first shake. The side-to-side load can be even worse if the shocks come in waves, as taller buildings can vibrate like a huge tuning fork, while each new sway is bigger than the last one, until failure. Usually, significant weight is added in time to such code-deficient steel buildings (i.e., walls, partitions to make more and smaller rooms, etc.) or even due to extreme reinforcing techniques. The more weight there is, and the higher this weight is located in the building, the stronger the building and its foundations must be to withstand the earthquake actions. Many buildings have not been strengthened when such extra weight was added. These buildings are then more vulnerable to even a weak aftershock, perhaps from a different direction or at a different frequency, and can cause collapse. Moreover, in a lot of multistory steel buildings, the ground floor has increased headroom with taller slender columns as well as with more large openings and fewer walls. So, these columns, which carry the largest loads from both the self-weight and the cumulative sideways actions from the seismic event, are vulnerable and they are often the first to fail. It only takes one to fail for the worst disaster; therefore, it is deemed necessary to cautiously strengthen steel buildings with the most appropriate method.

The potential deficiencies are different for different types of steel buildings (i.e., steel moment frames, steel braced frames, steel frames with concrete shear walls, steel frames with infill masonry shear walls). The indicators such as the global strength and stiffness, the configuration, the load path, the component detailing, the diaphragm, and the foundation design demonstrate the performance under seismic actions and the margins for improvement in specific ways; hence, they should be studied carefully before any decision is taken.

Retrofitting of existing code-deficient steel buildings accounts for a major portion of the total cost of hazard mitigation. Therefore, it is

S

important to identify correctly the structures that need and can accept strengthening or weakening, while the overall cost should be also monitored. If appropriate, seismic retrofitting should be performed through several methods such as increasing the load, deformation, and energy dissipation capacity of the structure (FEMA 356 2000).

Code-Efficient Buildings Resistant to Earthquake

To be earthquake proof, the buildings and their foundations need to be built to be resistant to sideways loads. The lighter the building is, the less the loads are. In steel, especially in high-rise buildings, the sideways resistance mainly comes from the diagonal bracing which must be placed equally in both directions. Where possible, the diagonal bracing should be strong enough to accept tension as well as compression loads; the bolted or welded connections should resist more tension than the ultimate tension value of the brace or much more than the design load. If the sideways load is to be resisted with moment-resisting framing, then great care has to be taken to ensure that the joints are stronger than the beams and that the beams will fail before the columns. Also in such a case, special care should be provided to the foundation-to-first-floor level, avoiding soft-story effects, while the columns should be much stronger than at higher levels. The foundations could be enhanced by having a grillage of steel beams at the foundation level able to resist the high column moments and keep the foundations in place. The main beams should be fixed to the outer columns with full capacity joints, which almost mean hunched connections, and care should be taken to consider the shear within the column at these connections.

When the steel beams are able to yield and bend at their highest stressed points, without losing resistance, while the connections and the columns remain full strength, then the resonant frequency of the whole frame changes, while the energy is absorbed and evenly dissipated across the framing.

The vibration occurred from the shock waves is tend to be damped out. This phenomenon is called "plastic hinging" and is easily demonstrated in steel beams. In extreme earthquake sway, the beams should always be able to form hinges somewhere, while the columns should behave elastically. In this way, the frame can deflect and the plastic hinges can absorb energy, while the resonant frequency of the structure is altered without major loss of strength and inevitable collapse. All floors should be connected to the framing in a robust and resilient way and should be as light as possible. They should possibly span around each column and be fixed to every supporting beam using enough shear connectors (e.g., studs). An effective way of reducing the vulnerability of large buildings is to isolate them from the floors using bearings or dampers; however, this is an expensive process, and it is not applied to low to medium rise buildings which have not been classified as important, due to the content they carry and the occupancy usage.

Nothing can be though guaranteed to behave as such, even in code-based designs; hence, most of the steel buildings and especially those under-designed with older seismic codes can be considered as code-deficient buildings in certain circumstances.

Design Concept for EC8

Eurocode 8 (EC8) follows three general design concepts based on the ductility requirements and capacity design considerations of steel buildings: the concept of the low-dissipative structural behavior of DCL structures, the concept of dissipative structural behavior of DCM and DCH structures satisfying the ductility and capacity design requirement, and the concept of dissipative structural behavior with steel dissipative-controlled zones. In the latter case, when composite action may be considered from Eurocode 4 (EC4) in the presence of the steel and concrete (slab) interaction, specific measures have been stipulated to prevent the contribution of concrete under seismic conditions, hence apply general rules for steel frames.

Introduction to Strengthening Techniques

Preliminary Investigation

It is becoming preferable, both environmentally and economically, to upgrade building structures rather than to demolish them and rebuild them. Engineers assessing structures for increased or special loadings are finding that new methods of analysis using, for instance, computer models are revealing shortcomings under service and ultimate conditions. Under such circumstances, a method has to be found to bring the structure up to the required standard. There is a range of techniques which can be used on structures, but, one must take into consideration that disruption to normal stage must be minimal while work is in progress.

Evaluation and subsequent strengthening of existing structures require a realistic and pragmatic design approach. However, some of the solutions proposed by researchers do not lie within this category and they will be eliminated in the current study. Also, effective communication between the owner, structural engineer, architect, risk analyst, insurance provider, and other stakeholders is paramount to a successful finished solution. In general, in the case where additional load-carrying capacity is required of an existing building, engineers have the option of either *reinforcing* the existing framing or *adding* new framing to replace or supplement the existing.

Where a decision is made to strengthen some parts of an existing facility or a specific structural system or element, the design approach is influenced by a series of factors such as:

- Information about relevant existing conditions which is often limited
- Part of the structure to be strengthened is commonly hidden or obstructed by existing architectural or building services systems that are difficult or costly to remove
- Structural renovation work is typically constrained by the need for continuity of building operations
- Level of ductility of the existing construction may limit its strength

- The susceptibility to local buckling of outstanding flanges as well as the lack of connection ductility

Often, the nonstructural costs will likely exceed the structural costs; therefore, the true costs of a retrofit project are primarily dependent on the number of locations of work than the amount of work done in each location, and thus, this influences the structural design and analysis decisions.

The general approach to strengthen existing structures includes the following aspects:

- Risk assessment and structural vulnerability assessment
- Preliminary analysis
- Consideration of alternatives (structural and nonstructural)
- Detailed design and impact of connections

The common goal is to:

- Protect specific structural elements.
- Provide redundancy to structural systems.
- Strengthen a specific part of the structure.

Assessing Existing Conditions and Strengthening Methods

Introduction

A site visit should also be performed to inspect the building, especially for structures more than 30 years old. Some key things to look for when assessing the existing condition of a steel building are: damage to framing, noticeable corrosion, signs that modifications to the structure that may have been performed without engineering review, unusual deflections in floor framing, cracks in supported slabs, signs of foundation settlement, signs for new rooftop equipment, heavy hung piping loads, folding partitions, rigging, or other suspended loads that may have been added without proper structural engineering review (Schwinger 2014). A valuable resource available to structural engineers working with existing building structures is the *AISC Steel Design Guide 15 – AISC Rehabilitation and*

S

Retrofit Guide (Brockenbrough 2002). Other publications for further reference are ASCE 41-06 2006, FEMA 274 1997, and FEMA 547 2006.

The strengthening methods can be categorized as follows:

1. Passive against active methods.
2. Strengthening techniques:
   (a) Reinforcing beams by welding (enlarge section with plates).
   (b) Reinforcing (or weakening) connections:
       • Framing.
       • Seated angles.
       • Partial-depth end plate.
       • Replace with high-strength fasteners.
       • Add welds at the perimeter of the connection and/or properly clean existing welds.
       • Converting single- to double-shear connections by adding angles or plates.
       • Add web stiffener plate.
       • Add steel cover plates.
       • Enhance column splices.
       • Enhance braced frame connection.
   (c) Shortening span (provided that there are no fitting issues):
       • Add beams.
       • Add columns and girders.
       • Add diagonal braces.
       • Add walls with openings.
       • Add steel braced frame.
       • Add concrete, masonry, or steel plate walls.
       • Enhance strength and ductility of braced frames.
   (d) Introducing composite action:
       • Steel (partially or fully) encased with concrete
       • Shear connectors
   (e) Post-tensioning (or external pre-stressing) of beams and connections (considering eccentricities of brackets on member capacity – some need protection from corrosion, fire, and vandalism).
   (f) Openings in existing beams (using thermal cutting – plasma arc cutting is faster than oxy-fuel – while avoiding cuts at areas subjected to high shear):
       • Place reinforcement (e.g., stiffeners) before cutting holes.
   (g) Replacement of members (may be economical).
   (h) Strengthening columns.
   (i) Convert gravity frame to moment-resisting frame.

Determining Load Capacity of Existing Buildings

Knowing the yielding strength of the steel used in the framing is essential for computing the load capacity; therefore, testing should be performed to ascertain and verify the actual yield strength. One technique is to test the steel to determine its actual yield strength, in hope of finding it to be a higher value than the one that was used in the original design. Another technique applied in existing structures is to analyze the framing using the load and resistance factor design (LRFD) method (AISC 1999); in LRFD usable strength is approximately 1.5 times greater than the older allowable stress design (ASD) service level strength.

Increasing Capacity of Connections

The technique by which existing shear and moment connections can be strengthened is limited only by the imagination of the engineer. Various techniques are going to be presented thereafter, based on well-established but also recent research outcomes obtained from numerous computational analyses and experimental campaigns. It is worth to be aforementioned that the capacities of existing connections must be determined when existing framing is modified or additional capacity is sought.

Increasing Flexural Strength Floor Framing Members

There are two options for reinforcing existing flooring systems to support additional loads: (a) **add** new framing to supplement the existing framing and (b) **reinforce** the existing beams, girders, and connections. The easiest solution is usually that of reinforcing the existing structural

Strengthening Techniques: Code-Deficient Steel Buildings,
Fig. 1 Examples of strengthened beams
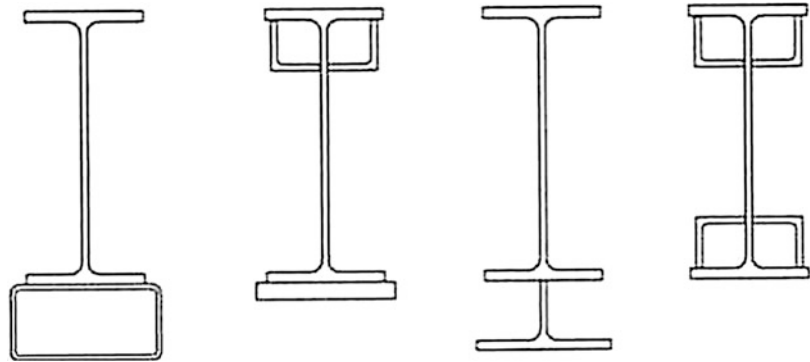
Strengthening Techniques: Code-Deficient Steel Buildings,
Fig. 2 Examples of strengthened columns
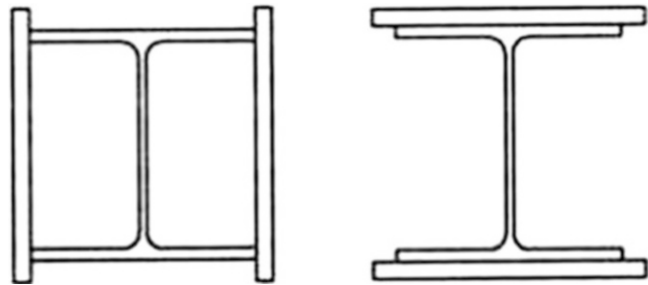
elements, provided that the floor slab has sufficient capacity to carry the loads. The most efficient way is to weld rectangular high-strength steels (HSS) to the flanges as shown in Fig. 1.

Increasing Axial Load Capacity of Columns
The buckling limit state and its variable slenderness should be determined in order to evaluate the axial load capacity of columns. Column strengthening serves both to reduce slenderness by increasing the radius of gyration of the section and to reduce stress. Column buckling is a mid-height phenomenon; therefore, increasing column stiffness between the supports, not at the supports, is required to increase column capacity. Both methods shown in Fig. 2 are effective; however, the one on the left better increases the weak axis stiffness of an H-shaped section.

Dealing with Weldability Issues
Weldability is verified by mechanical and chemical testing. The former measures ductility and the latter determines the "carbon equivalent" value.

Connecting New Frame to Existing Frame
Similarly to the connections, there are numerous ways that new framing can be connected to existing framing. Welding the new steel members to the existing members is a straightforward approach which requires less precision as compared to the bolting process, while drilling new holes through existing steel and bolting in the field. Various details for connecting new framing to an existing one can be found by Schwinger (2014).

Detailed Description of Retrofitting and Strengthening Techniques

Introduction
The performance of steel frames can be synopsized in three very different behaviors:

1. Formation of plastic hinges
2. Local and global instabilities
3. Fracture and structural discontinuity
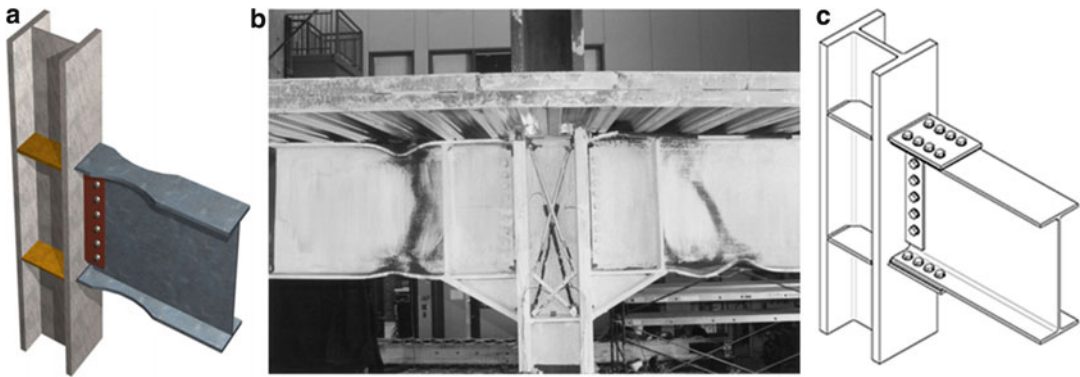
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 3 (**a**) Reduced beam section (RBS) connection (Crawford 2002), (**b**) welded haunch connection (Uang et al 2000), (**c**) bolted brackets

These three behaviors and/or combinations of them are likely to occur and govern the capacity of a connection or member with result on the structural continuity and integrity of the system. Overall, it is known from seismic studies that frame capacity is related to two different aspects of frame behavior:

1. The **member** response, as controlled by plastic rotational strength and deformation characteristics (including local and global buckling)
2. The **connection** response, as controlled by bolt fracture, premature brittle weld failure, and panel zone failure

Determining the capacity of columns is difficult as in many situations code-deficient buildings are not designed for large lateral loading.

Steel Connections: Fuses

Beam-to-Column Connections: Developing Ductile Behavior (Fuse Concept)

In parallel with the FEMA/SAC steel research program, the National Institute of Standards and Technology (NIST) and the AISC initiated a research project to upgrade existing special moment frames (SMFs) and investigate the effectiveness of two rehabilitation schemes. Modifications to pre-Northridge moment connections to achieve improved seismic performance focused on reducing or eliminating some of the

contribution factors to the brittle fractures. Brittle fractures originated in the beam flange groove welds and often propagated to rupture beam flanges or columns. A cooperative effort by NISC, AISC, the University of California at San Diego, the University of Texas at Austin, and Lehigh University examined three techniques for the retrofit of existing code-deficient steel moment connections trying to force plastic hinging of the beam away from the column face, namely, (a) the reduced beam section (RBS) concept to weaken a portion of the beam near the column so that plastic hinging would occur at the designated location, (b) the addition of a welded haunch to strengthen the steel beam near the welded connection, and (c) the use of bolted brackets to reinforce the connection (Fig. 3). More RBS patterns have been developed by Plumier in 1997 and appeared in different configurations, as it is shown in Fig. 4.

Further analytical research on the same connection complement this work, as design model and guidelines have been recommended. A target plastic rotation capacity of 0.02 rad was selected. In 2004, Engelhardt recommended (Bruneau 2004) the following as potential positive solutions:

- The use of a bottom flange RBS combined with the replacement of top and bottom beam flange groove welds with high toughness weld metal provided plastic rotations on the order of 0.02 to 0.025 radian. The presence of a
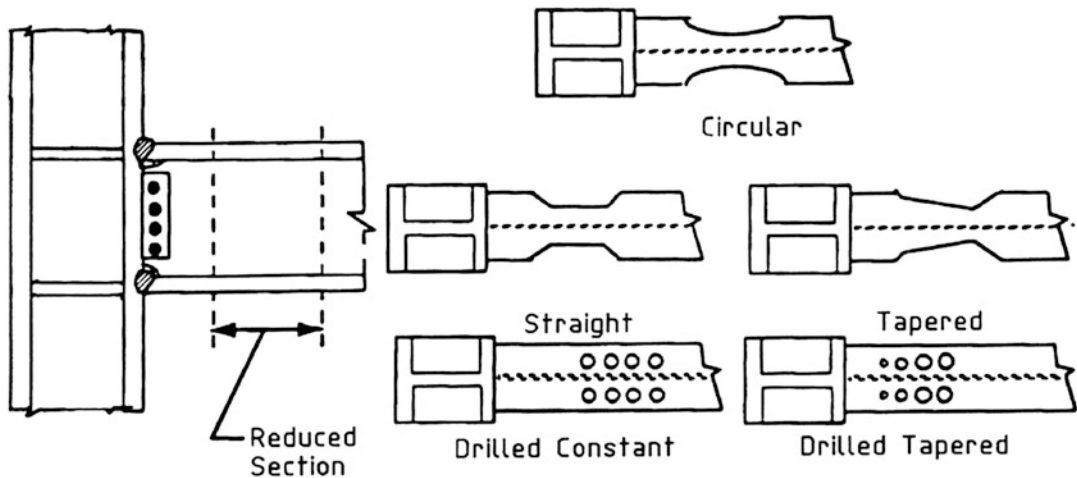
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 4   Various RBS patterns

composite slab had little effect on the performance of this retrofit technique.

- The addition of a welded bottom haunch, with the existing low toughness beam flange groove welds left in-place, resulted in significantly improved connection performance, which was dramatically influenced by the absence or presence of a composite concrete floor slab. In the former case, the specimens developed plastic rotations of 0.015 to 0.025 radian, whereas in the latter case developed plastic hinges in excess of 0.03 radian.
- The use of bolted brackets at the top and bottom flanges provided plastic rotations in excess of 0.03 radian.

Design recommendations have been provided for fully restrained, radius cut RBS connections. A step-by-step procedure is presented, with commentary for various design considerations. A similar procedure is included in FEMA 351 (2000) and AISC 358 (2005), which also provides design guidelines for other prequalified post-Northridge connections such as the bolted flange plate (BFP) moment connections, the bolted unstiffened (BUEEP) and stiffened extended end plate (BSEEP) moment connections, and the so-called ConXtech ConXL and Kaiser bolted bracket (KBB) moment connections.

It is worth to note that conventional beam theory cannot provide a reliable prediction for neither of the above structural systems. Uang et al. (2000) and Yu et al. (2000) proposed a simplified model that considers the interaction of forces and deformation compatibility between the beam and the haunches.

Exhaustive research works have been conducted on RBS connections varying the geometric characteristics of both the beam and the connection assembly itself. More recently, RBS moment-resisting connections have been also investigated by researchers in Europe using European HEA-profile sections (Pachoumis et al. 2008), since so far they have been only investigated by the US design construction practices. The result is the readjustment of the geometric characteristics of the RBS in order to apply to the European profiles. Limitation in using RBS is the shear connection between the top steel flange and the metal decking of the steel-concrete composite (SCC) slab due to the significant width reduction.

More recently, the same concept has been applied to steel frames as a strengthening-weakening technique, while introducing a circular opening (Fig. 5) in the beam's web instead, at a certain distance from the beam-to-column connection, as an effective method to improve the

Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 5 Failure mode of connection with circular web opening (Qingshan et al. 2009)

aseismic behavior of MRFs (Qingshan et al. 2009). The accurate position and size of the circular opening has been investigated through numerical modeling as well as experimental works, while the plastic hinge positions is effectively controlled. Similar studies have demonstrated the effect of various nonstandard web opening shapes (Fig. 6), in enhancing the ductility but also the strength of the connections (Tsavdaridis et al. 2014). Step-by-step procedures have been proposed to determine the most suitable geometries to achieve adequate connection strength, ductility, stiffness, and rotational capacities. Such techniques prove suitable in cases where large plastic rotations are required (i.e., larger than 0.03 rad). Tsavdaridis et al. (2014) have further proposed the use of previously patented novel elliptically based web opening shapes, which can also be used for perforated beams (e.g., cellular and castellated beams) adding numerous advantages from the manufacturing process to their life-span, while they can develop rotational capacities up to 0.05 rad with insignificant strength degradation (Fig. 7).

The so-called reduced web section (RWS) connections have been studied, yet not extensively, assessing local connection models mainly computationally with cyclic (quasi-static), and pseudo-dynamic (PSD) analyses. The results

show that the ultimate displacement of the modified buildings increases a lot due to the web openings, and thus, the building ductility is improved greatly. Moreover, brittle weld fractures can be avoided and the maximum plastic zone moved to the weakened areas. RWS connections can easily be applied to the beams of new as well as existing code-deficient buildings as the web cutting does is away from the composite slab that sits on top of the compression steel flange. It is worth to mention that different geometric characteristics and limitations of beams and columns should be used for different RWS connection types. There is a need to bring the attention and propose more experiment physical testings to validate and establish RWS connections in the current European and American practices.

Similarly, welded haunches and bolted brackets are used to move the plastic hinges away from the column, but also strengthen the existing connection and seek to maintain the original flexural capacity of the beam.

Ductile Behavior: Fuses in Bracing Members

The concept of adding ductile fuses in bracing members of steel concentrically braced frames (CBFs) resisting seismic loads is well linked to the RBS technique. Current code provisions require that steel CBFs are designed to exhibit ductile energy dissipation. Limits on brace overall slenderness ratio must be satisfied to achieve ductile inelastic behavior. It is apparent that implementation of this design approach may result in significant increases in design loads for brace connections, beams, and columns (Egloff et al. 2012).

In order to reduce seismic design loads, ductile fuses in bracing members have been recently proposed, as they control their axial resistances. Such behavior can be achieved by locally reducing the brace cross-sectional area or by introducing ductile components that yield in both tension and compression. In the former case, the reduced section might need to be confined to prevent local buckling, or it can be resized (reshaped - restructured) to yield in tension while remaining elastic in compression, which means that the
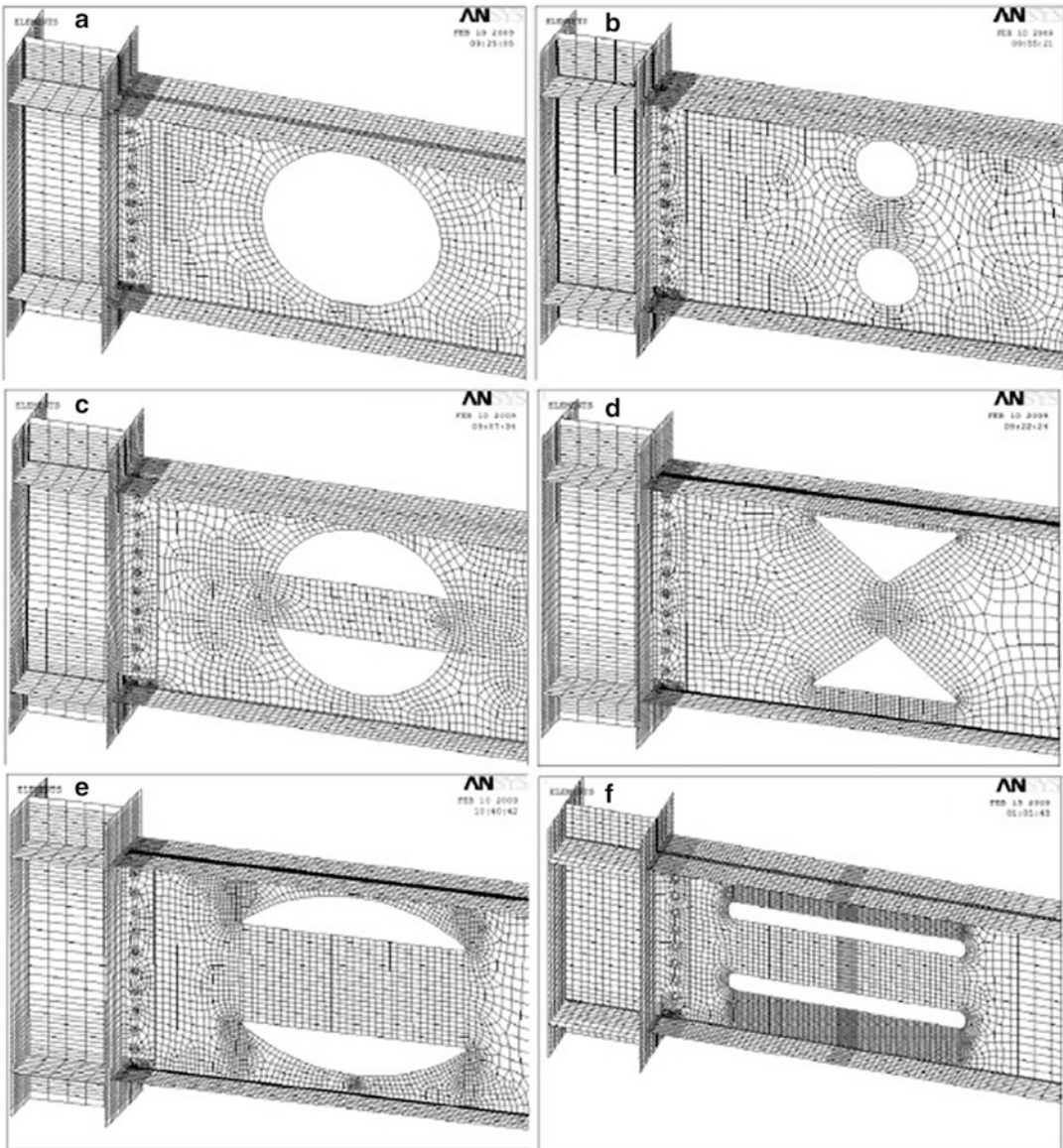
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 6 Types of perforated beam webs (Hedayat and Celikag 2009)

effect of fatigue loading will be minimized. In the latter case, the overall buckling is eliminated, and hence, the strength degradation is limited due to symmetrical hysteretic behavior. This type of fuse technique can be applied in open and closed profile sections of the bracing members, while it has been noticed that the former ones perform better exhibiting higher ductility. A special fuse (Fig. 8) for controlling the tension resistance of open bracing members has been proposed by Vincent in 2008. A part of the flange to web intersection is removed to limit the impact on the brace flexural stiffness and buckling resistance. In 2012, Egloff et al. introduced a new local buckling restraining system (LBRS) which includes two cold-formed channels that support the web and the flanges. Moreover, external cover plates can be bolted to the channels in order to
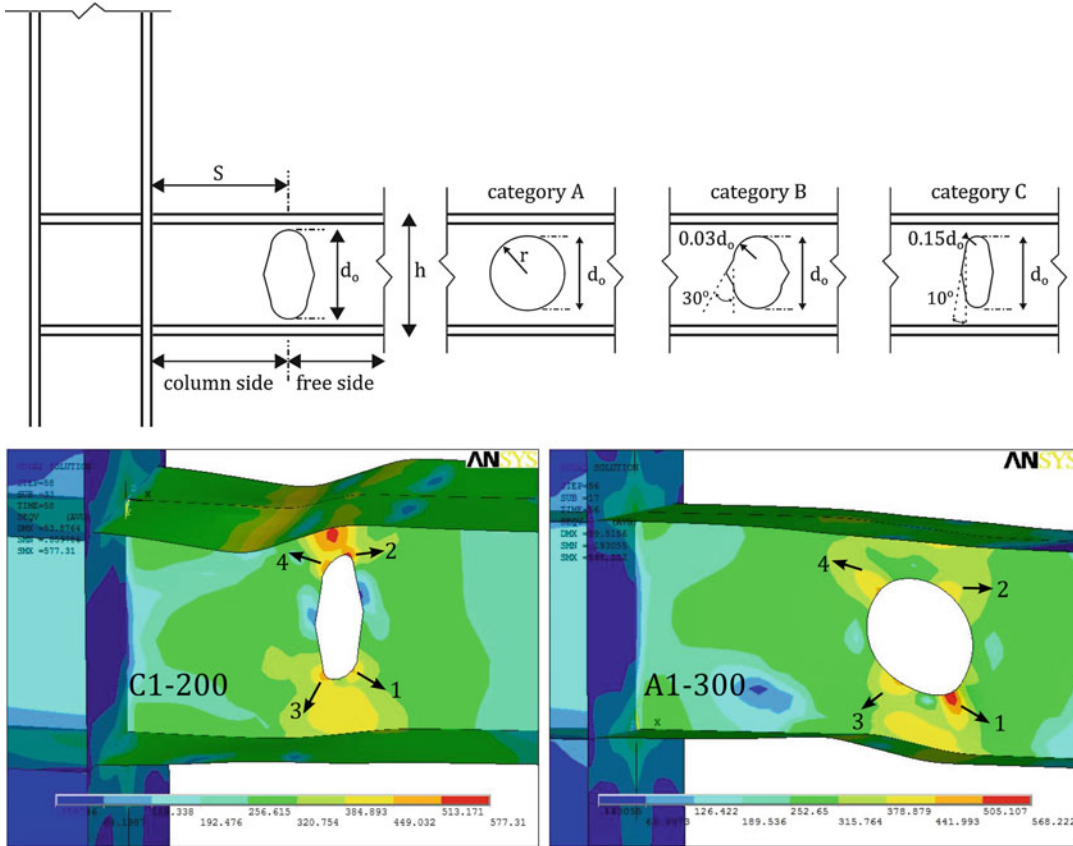
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 7 Geometric parameters of various novel perforated beam-to-column moment connections and von Mises plastic stresses (Tsavdaridis et al. 2014)
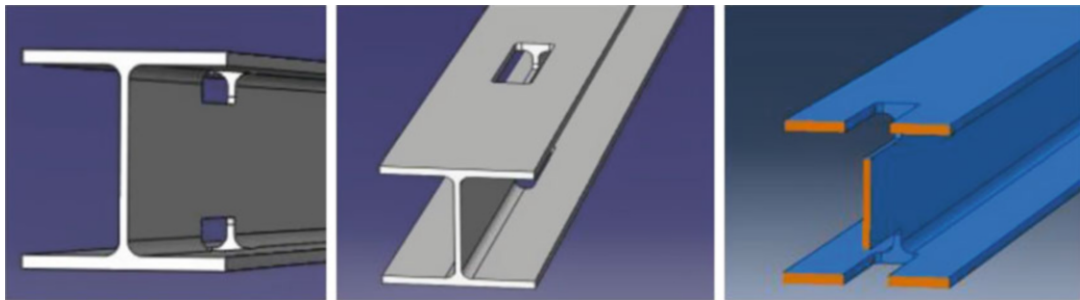
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 8 Ductile fuse for H-shaped bracing members (Vincent 2008)

prevent local buckling of the brace flanges. Splice plates can also be used to provide lateral support to the flanges in the fuse (Fig. 9). This LBRS can slip longitudinally with respect to the brace so that it does not attract any axial forces. Further

improvements and the design procedure for the fuse and the fuse LBRS are available in the literature.

Cast Connex, a high-strength steel connector for round hollow structural section brace
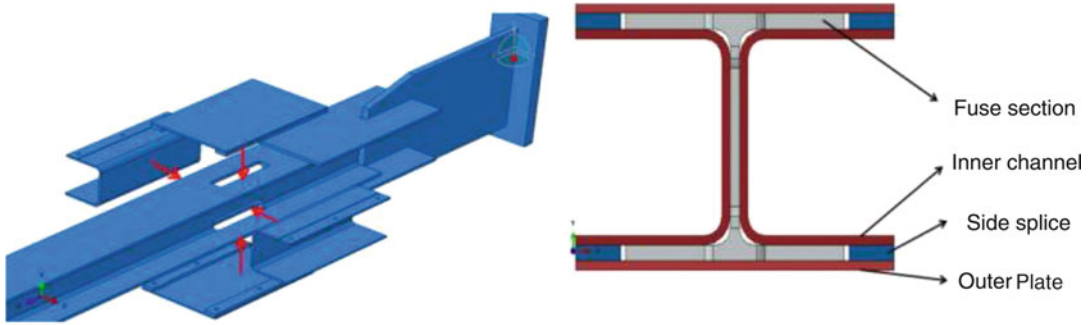
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 9 Proposed fuse local buckling restraining mechanism (Egloff et al. 2012)
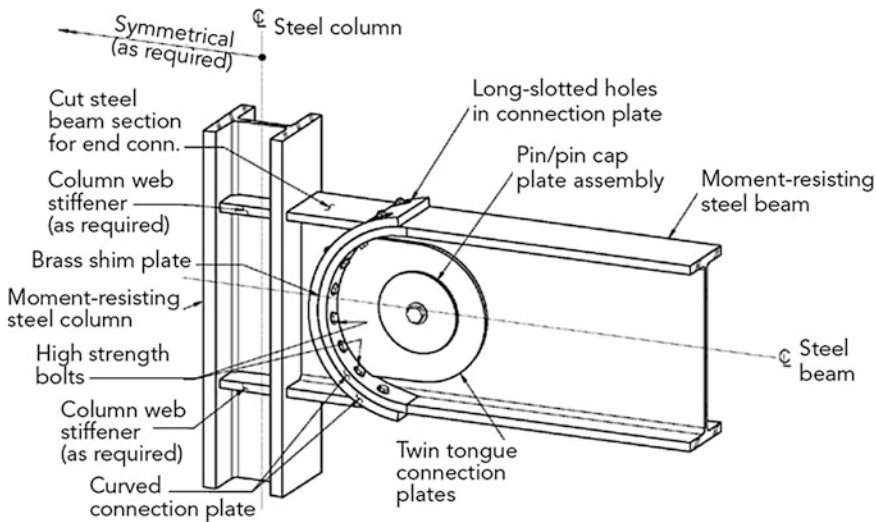
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 10 Structural details of pin-fuse connection (Coedova and Hamburger 2011)

members in CBF, has also developed a yielding fuse connector for CBF, called the Scorpion Yielding Brace System (SCBF), that relies on the flexural yielding of fingerlike plates that are specially designed to dissipate energy locally and can be used in both architecturally exposed and nonexposed braced bays.

Pin-Fuse Joints

A new type of connection, which begun its prequalification process in 2011, is the pin-fuse connection (Fig. 10) which incorporates a curved plated end connection using slip-critical bolts and a steel pin adjacent to the beam-to-column joint. The bolts are designed to slip within slotted holes allowing the pin joint to rotate, dissipating the energy through frictional resistance. This joint acts as the fuse for the system, while the rest of the steel frame can be designed to remain elastic. Following an earthquake, and avoiding damages, the frame can be adjusted to its initial position, reducing the potential for permanent residual displacements while both the connection and the frame maintain their structural integrity and reduce the need for costly structural repairs. The simplicity of the pin-fuse connection offers the ability to accommodate braces and dampers.

S
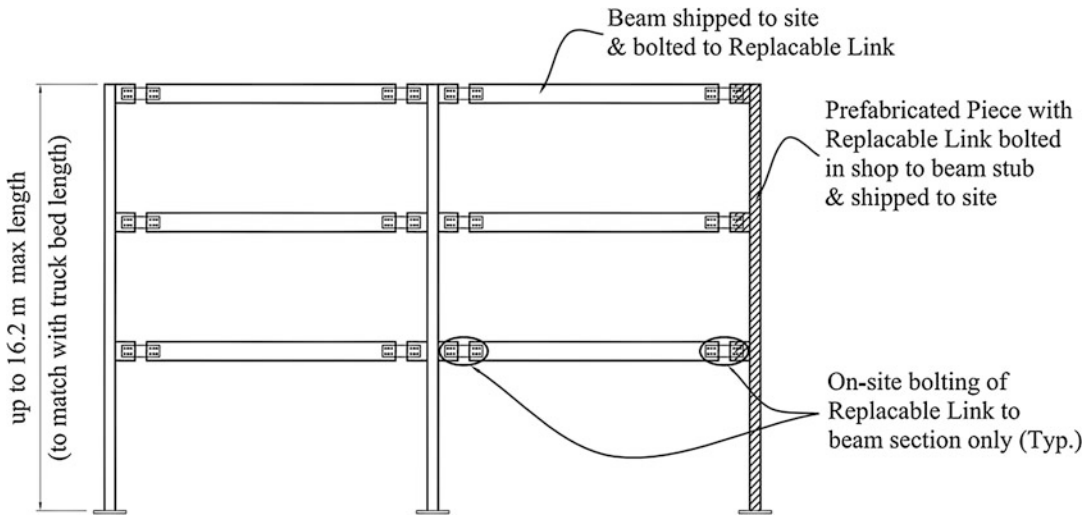
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 11 Proposed connections with replaceable link (Shen 2009)

Replaceable Links

Using RBS- and RWS-type connections has been proved very efficient for certain applications; however, there are some drawbacks. As the yielding fuse is a part of the beam, strength design and drift design of the structure are interlinked. For instance, due to increased drift requirements, the capacity of the yielding fuse may be also increased, which then leads to higher demands on the other parts of the structure including columns, floor slabs, connections, and foundations and often resulting in overdesigned buildings with increased overall cost. Further, significant damage can result in the beam from repeated inelastic deformation and localized buckling (hence crack propagation) during a design-level earthquake. As this cumulative inelastic action of the building cannot be precisely anticipated, it is not trivial to assess the extent of damage on site and the residual capability of the structure to adequately provide the required level of safety for any subsequent loading. In such a case, repair of the beam is not a straightforward procedure and it can be disruptive and costly.

The replaceable link concept (Fig. 11) effectively eliminates the aforementioned concerns instead of reducing the beam section size; dismountable dissipative elements (Fig. 12)

which can be removed and replaced with smaller flexural capacities are used at the locations of the expected inelastic actions. Consequently, the other structural elements in the frame will remain elastic during an earthquake. This efficient method of repair for MRFs allows for quick inspection and replacement of damaged links while it minimizes the disruption time. Further, the welding of critical elements of beam-to-column connections can be done in the shop while improving construction quality and significantly reducing the initial erection time (Shen 2009).

In particular, two types of replaceable links have been proposed by Shen (2009): (i) H-section with end plates and (ii) back-to-back channels eccentrically bolted to the beam web. The former one is prepared in the shop using complete joint penetration welds. The end plate (flush or extended) is then bolted to the column flange using prestressed high-strength bolts. The latter type of these double channel-built sections intended to act as truss girders in special truss moment frames, and they have been connected using welded reinforcing gusset plates. In certain circumstances, lateral bracing is deemed necessary in the region adjacent to the plastic hinge to achieve large (i.e., 0.06 rad) plastic rotation of the hinge. However, sometimes
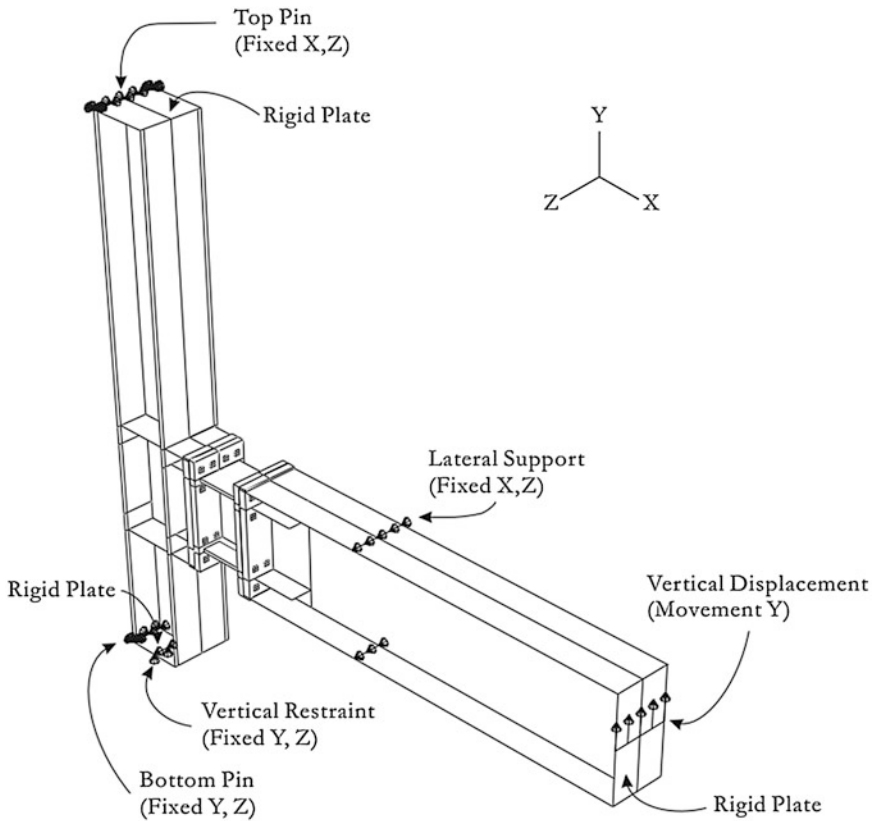
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 12 End plate model boundary conditions (Shen 2009)

large overstrength has been observed in these buildup channel sections.

The end plate links can exhibit 0.04 rad contributing to 90 % of the total story drift and demonstrate higher energy-dissipating capacity than the double channel links, while some strength degradation occurs due to ductile local flange and web buckling. On the other hand, higher story drift can be reached by the double channel links before experiencing strength degradation at 0.06 rad. The degradation has been also caused due to ductile tearing of the flanges and the webs. Overall, double channel links type has been considered to be preferable as it provides a more gradual transmission of forces at the connections via friction. For more stability, further modifications can take place enhancing the connection of the channel webs and the connection of the beam segments.

Steel Connections: Stiffeners

Introduction

Forcing the plastic deformation to the beam end away from the connection is a common practice in seismic moment-resisting frames, but in contrast to the fuse concept, this can be also achieved by increasing the relative stiffness of the column and the connection with respect to the beam. Eventually, this is an alternative in effectively controlling the position and intensity of the plastic hinge in the connection zone when such modifications are allowed.

SidePlate™ Connections

A well-promising retrofitting method for upgrading an existing traditional moment-resisting connection is shown in Fig. 13. This concept uses the so-called SidePlate™ retrofit system where the
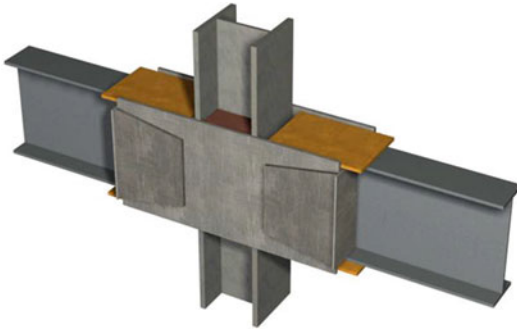
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 13 SidePlate^{TM} retrofit connection (Crawford 2002)

physical separation between the face of the column flange and the end of the beam, for mitigating the stress concentration, is achieved using parallel full-depth side plates which act as discrete continuity elements to sandwich the connecting beam(s) and column (Crawford 2002). Similar design concept can be used for steel and concrete-filled hollow section columns (Fig. 14).

Whole steel frame is eventually stiffened, and the zone panel deformation is eliminated using this type of connections due to the increased stiffness of the side plates that ultimately provide the three panel zones. This connection system uses all fillet-welded fabrication which predominately carries all shear actions as well as moments through the combination of vertical shear plates and fillet welds. The side plates should be designed with sufficient strength and stiffness to force all significant plastic behavior of the connection system into the beam.

The same system can be used to upgrade construction of deficient buildings. The difference is that an initial hole is required in each side of the plates to permit welding access, while the holes are closed with the same cutoff plate following the completion of the welding process. All new welds are again fillet welds loaded in shear along their length, whereas if there are any existing Complete Joint Penetration (CJP) welds, they are removed by air arcing to eliminate the reliance on through thickness properties and triaxial stress concentrations. More information can be found from FEMA 351 (2000).
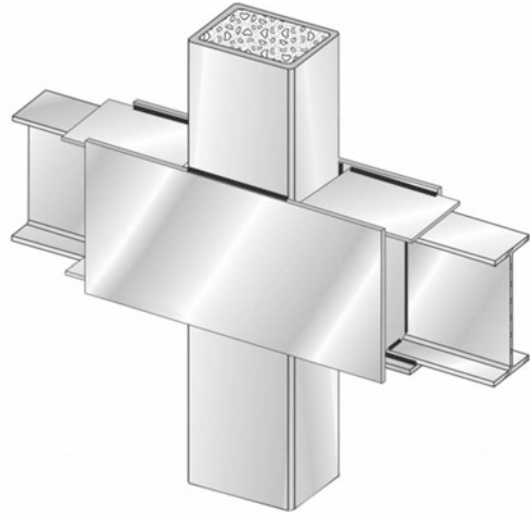
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 14 Strengthening retrofit concept: concrete-filled hollow section of column (Crawford 2002)

Stiffeners at Connections

In addition to the prequalified connections for SMFs and intermediate moment frames (IMF) presented in AISC (1999) and the SidePlate^{TM} system, research has been focused on the effect of stiffeners on the strain patterns of the welded connection zone.

For example, the effect of both internal and external stiffeners on the behavior of I-beam to hollow-column section connections (Fig. 15) has been initially thoroughly investigated by Chen and Lin (1990). It has been observed that the connections with triangular stiffeners have the lowest rigidity, in contrast to those with side stiffeners which present significantly higher moment rotation capacities, stiffness, and ductility. Moreover, the performance of the retrofitted connections with side stiffeners has been investigated (Ghobadi et al. 2009) and design guidelines proposed. The benefits of using side stiffeners have been also introduced on concrete-filled tabular (CFT) columns connected to I-beams, while stable hysteresis and adequate ductility are provided. Overall, it has been concluded that connections with both column stiffeners and top-flange stiffeners have the highest value of energy dissipation, while the beam top-flange
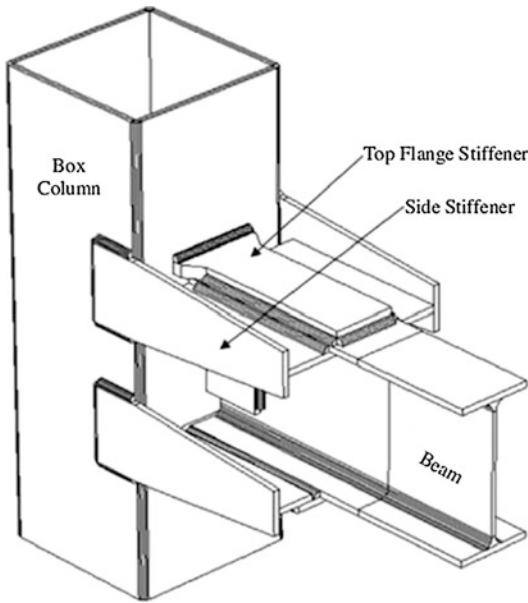
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 15 Typical I-beam to box-column connection (Kiamanesh et al. 2010)

stiffener is the most effective one, especially when it is incorporated with the column stiffeners of a hollow section.

Steel Frames: Modifications

Frame Modification at Beam's Mid-Span (Fuse Concept)

A retrofitting method which can be used for new construction as well as a strengthening technique for existing moment-resisting frames has been developed by Leelataviwat et al. (1998). This technique replaces certain beams and is introducing a ductile fuse element in shear at their mid-span instead of modifying the beam-to-column connections (Fig. 16). A braced rectangular opening is created in the web of each girder at the mid-span, to move the plastic deformation away from the critical connection regions, while ensuring the development of a ductile mechanism.

Cabling: Self-Centering Systems

The use of the cables is another promising technique, which can be applied to both slabs and connections. Placement of cables with connections to girders or cables with connections to beams is very important especially for high-rise buildings. Self-centering braces have been designed and built using prestressed aramid fiber strands in conjunction with friction pads or memory alloys. Energy dissipation is implemented using yielding seat angles, friction dampers, or energy-dissipating bars confined in tubes. Researchers have investigated self-centering column bases that use post-tensioned (PT) bars or spring-loaded wedges. Tendons can span over multiple floors, while elastomeric spring dampers and fuse bars can be used to provide energy dissipation.

Furthermore, self-centering structural systems (Fig. 17) have been proposed, for the seismic retrofit of special moment-resisting frames, by Christopoulos et al. (2002). This is a post-tensioned energy-dissipating (PTED) steel frame design, where high-strength bars or tendons provide the post-tension at each floor. Confining steel sleeves have been often used to prevent the energy-dissipating bars from buckling during cyclic inelastic loading. It has been concluded that these economical innovative systems:

- Incorporate the nonlinear characteristics of yielding structures and, thereby, limit the induced seismic forces and provide additional damping characteristics.
- Encompass self-centering properties allowing the structural system to return to its original position after an earthquake.
- Reduce or eliminate cumulative damage to the main structural elements.

Later, Garlock et al. (2004) has proposed a similar structural system with high-strength steel PT strands, after bolted replaceable top-and-seat angles have been installed (Fig. 18). Here, the vertical shear is supported by both the angles and the friction between the beam and the column, and it is expected to continue to perform if failure of one or more strands occurs. It is proved that this system can achieve greater strength and ductility.

Recently, researchers have designed and experimentally evaluated a new self-centering PT connection using yielding web hourglass shape pins (WHPs) as seismic energy dissipaters. WHPs do
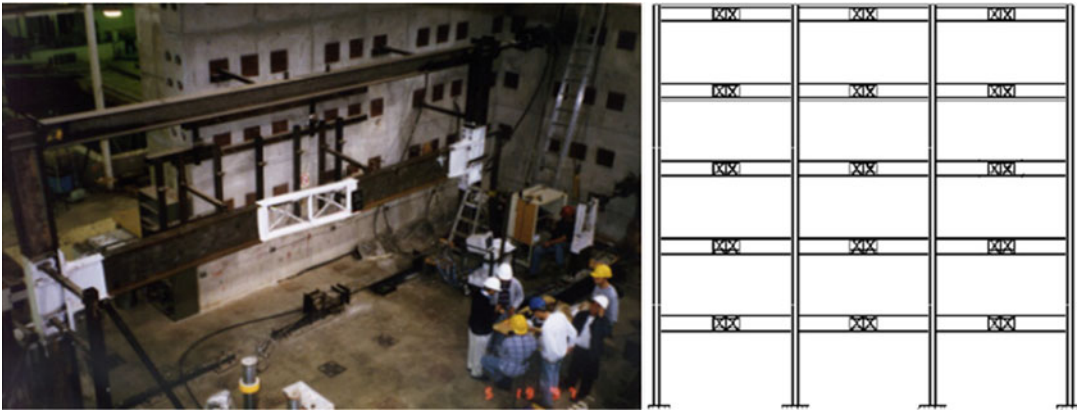
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 16 Frame modified with mid-span truss opening (Leelataviwat et al. 1998)
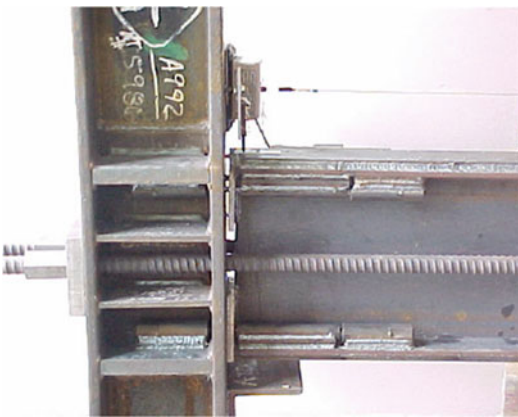
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 17 PTED system (Christopoulos et al. 2002)

not interfere with the composite slab and can be very easily replaced without the need for welding and bolting and, therefore, can significantly decrease downtime in the aftermath of a strong earthquake. Repeated experiments are described in detail and proved the reparability of the PT connection with WHPs (Vasdravellis et al. 2013).

Structural System: Adding Structural Elements (Walls, Blocks, Bracings)

Introduction

Conventional retrofitting methods include the addition of new structural elements to the system,
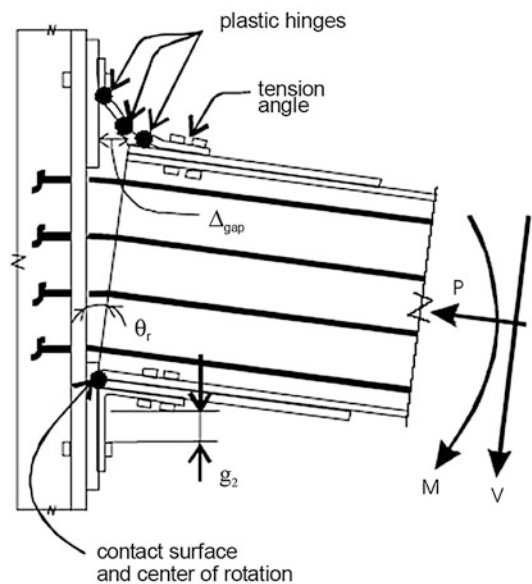
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 18 Post-tensioned moment connection with top and bottom seat angles (Garlock et al. 2004)

enlarging the existing members. Bracings, masonry (Fig. 19), and reinforced concrete (pre- and post-cast infill) shear walls are the most popular and efficient strengthening techniques as they provide lower overall cost and they are easy to use. Braces are more effective due to their much higher ductility, but the shear walls are indeed the most commonly applied method, as they also reduce the demand on the other

Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 19 Ordinary hybrid wall (Adopted by http://masonryedge.com/site/mim-archives-thestorypole-vol-38-no-4/143-filling-in-the-frame-)

structural members resisting large lateral loads, hence increasing their safety in a simple manner. The actual capacity of bearing walls has been often underestimated or even ignored. However, it can be a major contributor, and it can provide the required capacity, without the need of more complex strengthening techniques (Sarhosis et al. 2014).

Bearing Walls

Walls must go equally in both directions, and they must be strong enough to add stiffness to the steel framing system while they are tied in to any framing in order to take load in their weakest direction. Also, they must not fall apart and must remain in place after the worst shock waves, so as to retain strength for the aftershocks.

In particular, three approaches have been identified (Crawford 2002) for enhancing the resilience of a building's bearing walls under the progressive collapse and seismic scenarios. These are the following: (i) backup wall, build a second wall or gravity-carrying frame to support the existing wall; (ii) strong wall, employ fabric retrofit to control the breach area; and (iii) ductile

wall, polyurethane spray to prevent punching shear failure. Moreover, openings can be accommodated in such walls, especially with multistory buildings, while further enhancements are needed.

Steel Plate Shear Walls

In addition to concrete and reinforced concrete walls with SCC beams, research has been initially conducted by Thorburn et al. (1983) later on steel plate shear wall (SPSW) design (Fig. 20a) and retrofitting methods. SPSWs can be used as the primary lateral force resisting system in steel buildings allowing the occurrence of shear buckling. Following buckling, diagonal tension field is developed to transfer the lateral load in the panel, while the forces in beams and columns are reduced. Furthermore, the use of low-yield-strength steel panels and RBS connections as well as light-gauge cold-formed steel plates has been examined as potential applications by Berman and Bruneau (2004). The former one demonstrates an earlier onset of energy dissipation by the panel, while perforated panel specimens (Fig. 20b) can be used to control the
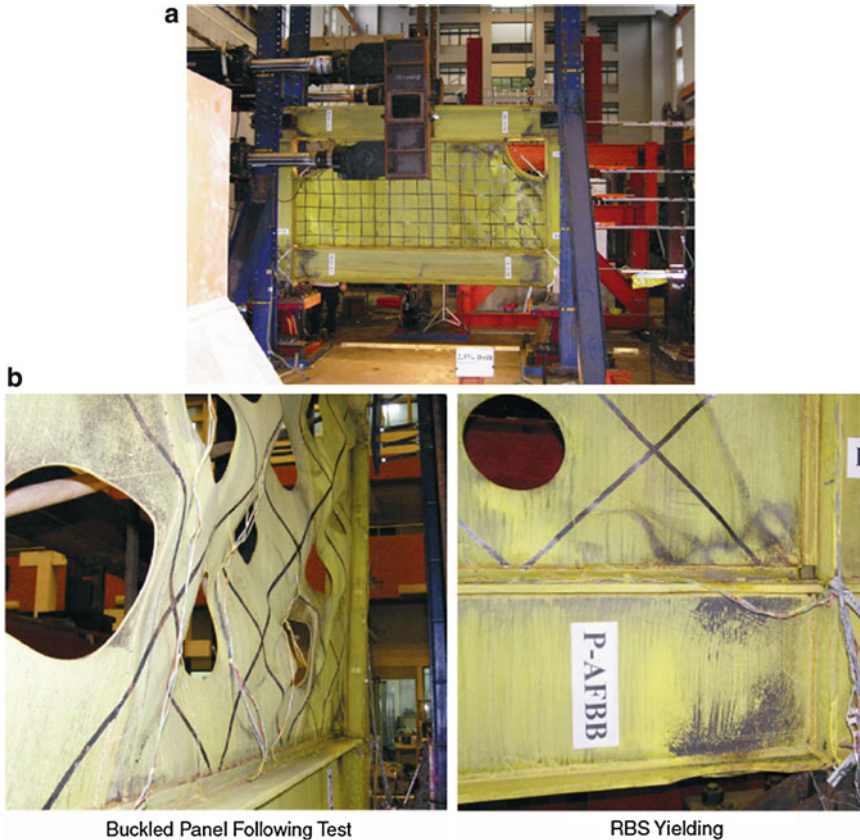
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 20 (**a**) SPSW specimen with cutout corners (*right*), (**b**) buckled panel and RBS yielding of SRW specimen (Vian and Bruneau 2004)

stiffness and overstrength issues using hot-rolled plates. This option is also useful in a retrofit situation, providing access for utilities to penetrate the predesigned system.

Braced Frames

Braced frames can be constructed of single diagonal, x and k braces, chevron and split braces, and lattice or knee braces, and they can be used in interior cores – so connections could be easily made with wall panels, as well as in the exterior. Composite braced frames are also becoming popular where concrete bracings are supporting steel frames.

A simplified design procedure for suspended zipper frames, initially proposed by Khatib et al. (1988), has been introduced by Leon and Yang (2003) and consists of inverted V-braces

adding zipper columns which connect the intersection point of the braces above the first floor. The zipper columns tie all brace-to-beam intersection points together and force compression braces in a braced bay to buckle simultaneously, hence better distribute the dissipated energy over the height of the building (Bruneau 2004). A suspension system has been proposed later, ensuring that the top story braces are designed to remain elastic, whereas all other compression braces are designed to buckle, while the suspended zipper struts are designed to yield in tension. Therefore, adequate ductility is provided, with superior seismic performance.

Engaging the fuse, the replaceable links, the self-centering, and the braced frame retrofitting concept, Eatherton et al. (2008) proposed a controlled rocking system which virtually eliminates
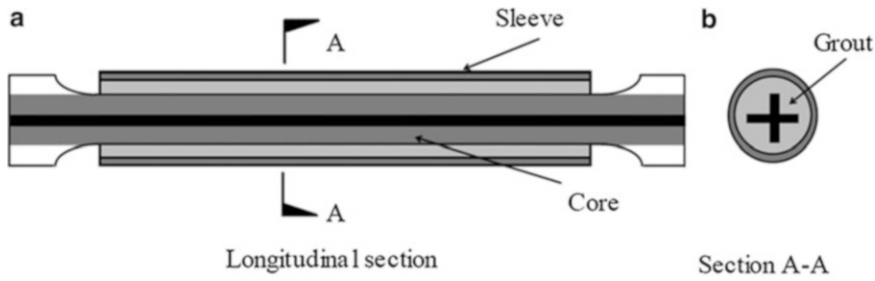
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 21 A non-buckling brace

residual drifts and concentrates the majority of structural damage in replaceable fuse elements. The system is consisted of three components: (a) a stiff steel braced frame which remains elastic, but it is not tied to the foundation and hence allowed to rock, (b) a vertical post-tensioning that strands the top of the frame down to the foundation and brings the frame back to the center, and (c) the replaceable fuses that absorb the energy as the frame rocks.

Non-Buckling Braces

Conventional braces tend to buckle under the compression cycle of the seismic load, hence dissipating little energy under compression. This causes pinching of the hysteresis loop and failure of the braces within a few cycles, due to the formation of plastic hinge close to mid-length of the member. The use of non-buckling braces (also known as buckling restrained braces or un-bonded braces) bypasses this problem. In this type of bracing system, the requirements of adequate strength to resist compression, as well as rigidity to avoid buckling, have been addressed separately to a core and a sleeve (Fig. 21).

In the last decade, buckling restrained braced (BRB) frames have received much attention in the United States, as they demonstrate stable hysteretic behavior and excellent low-cycle fatigue life characteristics. However, buckling and cracking of gusset plates are expected in certain cases, similarly to all types of braced frames.

Strengthening Members

Strengthening members by adding plates or encasing/filling them with concrete provides an effective technique to add strength, and it can be applied for a particular group of members, such as on the ground floor. Thus, concrete encasement of columns (Figs. 14), and floor beams (Fig. 22) as proposed (Tsavdaridis et al. 2013), constitute forms of strengthening techniques for new and existing steel buildings.

Materials

Innovative ways have been explored for the strengthening and rehabilitation of deficient steel buildings, due to the demand to access the specified load capacity and the deterioration as a result of corrosion. In particular, externally bonded fiber-reinforced polymer (FRP) composites can be applied to various structural members such as columns, beams, slabs, and walls in order to improve their structural performance in terms of stiffness, load-carrying and deformation capacity, and ductility, while simultaneously providing environmental durability. Generally, FRPs have been widely used mostly in applications that allow complete wrapping of the member, while attention deemed necessary to avoid brittle shear and de-bonding failures, especially prone when used on steel. In such cases, the actual member can entirely waste the strengthening application, or the composite material might harm the member itself by decreasing its ductility (Buyukozturk et al. 2003). It has been proposed that for buildings with large seismic deficiencies, a combination of conventional and FRP strengthening techniques may prove to be an effective retrofitting solution.

S
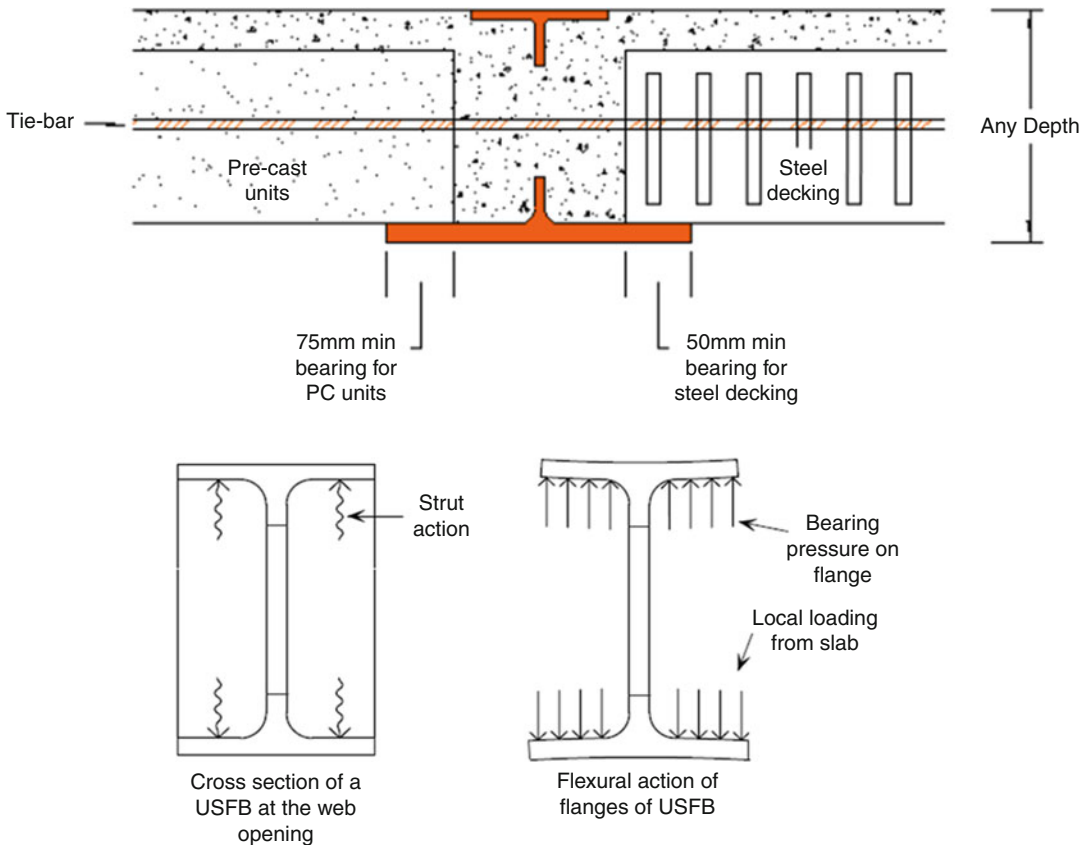
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 22 Schematic representation of the USFB system and the internal actions (Tsavdaridis et al. 2013)

The preformed high-strength carbon fiber plates currently being used for concrete structures are typically 4 mm thick. To strengthen steel beams, they would need to be at least 20 mm thick, in order to achieve a significant increase in bending moment for the steel or SCC beam. Consequently, new high-modulus carbon fiber-reinforced polymer (CFRP) materials are likely to provide solutions for steel structures' deterioration issue (Schnerch et al. 2005). Time should be allowed for the surface preparation, application of the adhesive, and curing time (usually between 4 and 8 h). Specific surface preparation and detailing are critical to ensure adequate bond interaction between steel and FRP materials, both in the short and long term, and capable of sustaining the high interfacial

stresses necessary to appreciate the full strength of these materials (Lenwari et al. 2006).

Energy Dissipation and Active/Passive Structural Control Systems

A quite different retrofitting method, which can be quite cost-efficient, is the installation of complementary energy dissipation devices in structures as a means of passive, semi-active, or active structural control systems. These are not described thoroughly here, as it is beyond the scope of the present entry. The main objective of structural control is to minimize structural vibrations improving safety and serviceability limits mainly under wind and earthquake actions. Up to date, the majority of passive energy dissipation devices have been found very effective in
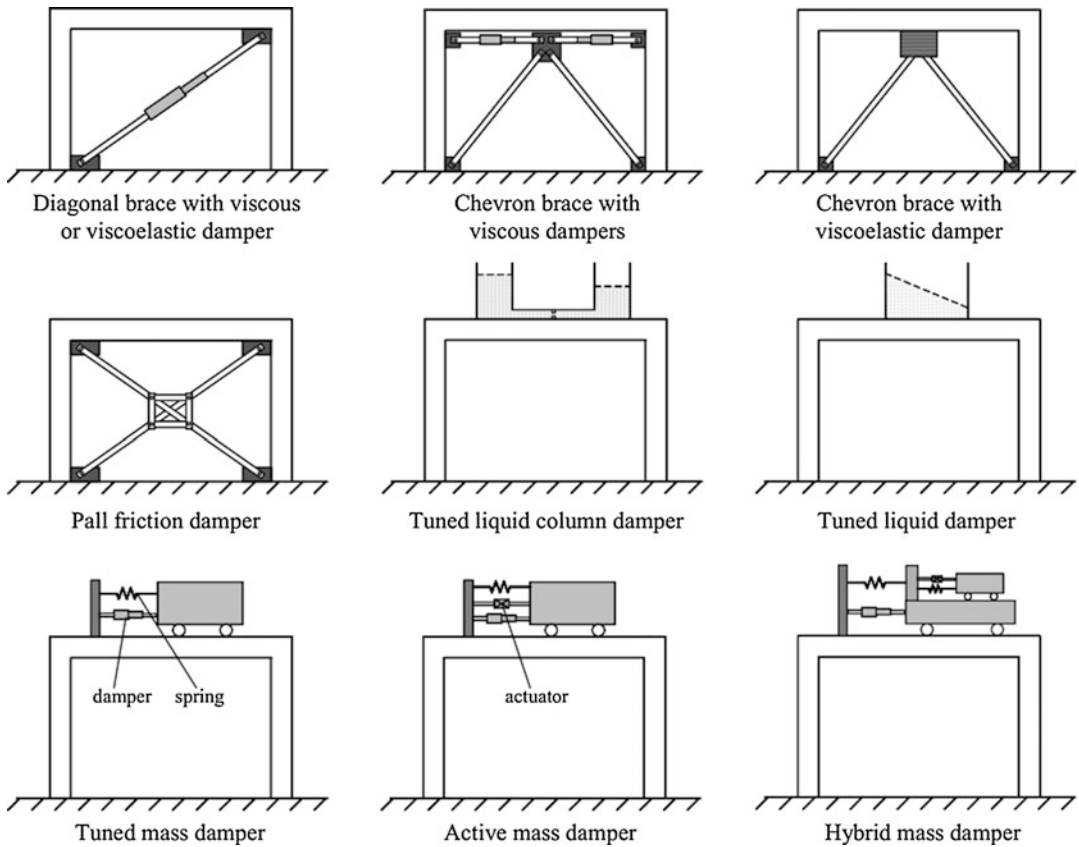
Strengthening Techniques: Code-Deficient Steel Buildings, Fig. 23 Supplemental energy dissipation devices (Adopted from MIT.edu website)

controlling the seismic response of steel frames. Further advanced techniques seem very promising such as the introduction of an "inerter" (Smith 2002; Marian and Giaralis 2014), and its combined use with the already well-operated tuned mass dampers, with scope to reduce the size of the mass required to control and dissipate the energy of high-rise buildings. Figure 23 shows the basic principles of various control systems commonly used on building structures.

There is a vast research conducted on energy dissipation devices during the past 20 years, while their use becomes more direct and apparent with the upsurge of technology. However, a diverse background of researchers is required, integrating a number of disciplines, some of which are not within the domain of traditional civil engineering. In particular, the control theory is elaborated with computer science, data processing, sensing technology and materials science using the knowledge and principles of earthquake (and wind) engineering, structural dynamics, as well as stochastic processes. It is essential to mention though, that the effectiveness of such dissipation devices is predominantly dependent of the deformation capacity of the structure. Consequently, the application of such devices to code-deficient buildings with inadequate seismic detailing or post-earthquake damages should be carefully considered, and perhaps it should be combined with an appropriate strengthening technique with deformation enhancement measures as proposed above.

Ongoing research on special strengthening techniques involving the use of simple and robust active control systems, while emphasizing on the performance of various controllers (Demetriou et al. 2014) as well as new conceptual methods introducing the ability of structures to adapt under dynamic loads (Slotboom et al. 2014), are posing great expectations.

Summary

It is worth to mention that there are factors which inhibit retrofit design when it comes to strengthening existing buildings. Many times important issues are misaddressed due to the complexity of strengthening design concepts, lack of technology understanding or even ignoring it, while uncertainties about the design of the building involved. Therefore, existing buildings should be best approached with a risk-based retrofit scheme in order to concentrate the works where they are actually needed most. In this way, more safe, effective, and cost-efficient steel buildings will stay operated in the future. This inherently requires more skilled designers and engineers, better design tools, and more truly innovative "smart" retrofitting techniques to be developed for the seismic strengthening of code-deficient steel buildings.

Cross-References

References

AISC 1999 LRFD (1999) Specification for structural steel buildings. American Institute of Steel Construction, Chicago

AISC 358 (2005) Specification for structural steel buildings. American Institute of Steel Construction, Chicago

ASCE 41–06 (2006) Seismic rehabilitation of existing buildings. American Society of Civil Engineers – ASCE, Reston

Brockenbrough RL (2002) AISC Rehabilitation Bernam on and retrofit guide – a reference for historic shapes and specifications, Steel design guide series 15. American Institute for Steel Construction, Chicago

Bruneau M (2004) Seismic retrofit of steel structures. In: Sociedad Mexicana de Ingeniera Sismica, VIII SNIS, Tlaxcala

Buyukozturk O, Gunes O, Karaca E (2003) Progress on understanding de-bonding problems in reinforced concrete and steel members strengthened using FRP composites. Construct Build Mater 18(1):9–19

Chen SJ, Lin HY (1990) Experimental study of steel I-beam to box-column moment connection. In: 4th international conference on steel structures and space frames, Singapore, pp 41–47

Christopoulos C, Filiatrault A, Uang CM (2002) Self-centering post-tensioned energy dissipating (PTED) steel frames for seismic regions. In: Structural systems research project report no. SSRP-2002/06, Department of Structural Engineering, University of California, San Diego, p 292

Coedova PP, Hamburger RO (2011) Steel connections: proprietary or public domain. Modern Steel Construct 51(10):24–31

Crawford JE (2002) Retrofit methods to mitigate progressive collapse, 2002. Online: available in May 2014

Demetriou D, Nikitas N, Tsavdaridis KD (2014) Performance of proportional-integral-derivative controllers on structures with variable damping tuned mass dampers. In: 6th World conference on structural control and monitoring, Barcelona

Eatherton MR, Hajjar JF, Deierlein GG, Krawinkler H, Billington S, Ma X (2008) Controlled rocking of steel-framed buildings with replaceable energy-dissipating fuses. In: The 14th World conference on earthquake engineering, Oct 12–17, Beijing

Egloff O, Tremblay R, Vincent R, Dussault S (2012) Finite element analysis of ductile fuses for W-shape steel bracing members. In: 15th World conference of earthquake engineering, Lisboa

FEMA 274 (1997) NEHRP commentary on the guidelines for the seismic rehabilitation of buildings. Report no. FEMA 274. Federal Emergency Management Agency, Washington, DC

FEMA 351 (2000) Recommended seismic evaluation and upgrade criteria for existing welded steel moment frame buildings. Report No. FEMA 351. Federal Emergency Management Agency, Washington, DC

FEMA 356 (2000) Prestandard and commentary for the seismic rehabilitation of buildings. In: Report No. FEMA 356. Federal Emergency Management Agency, Washington, DC

FEMA 547 (2006) Techniques for the seismic rehabilitation of existing buildings. Report No. FEMA 547. Federal Emergency Management Agency, Washington, DC

Garlock M, Ricles JM, Sauce R (2004) Experimental studies on full-scale post-tensioned steel moment connections. In: 13th World conference of earthquake engineering, Vancouver

Ghobadi MS, Ghassemieh M, Mazroi A, Abolmaali A (2009) Seismic performance of ductile welded connections using T-stiffener. J Construct Steel Res 65:766–775

Hedayat A, Celikag M (2009) Post-Northridge connection with modified beam end configuration to enhance strength and ductility. J Construct Steel Res 65 (7):1413–1430

Khabit IF, Mahin SA, Pister KS (1988) Seismic behavior of concentrically braced steel frames. Report no. UCB/EERC-88/01. Earthquake Engineering Research Center, University of California, Berkeley

Kiamanesh R, Abolmaali A, Ghassemieh M (2010) The effect of stiffeners on the strain patterns of the welded connection zone. J Construct Steel Res 66(1):19–27

Leelataviwat S, Goel SC, Stojadinovic B (1998) Drift and yield mechanism based seismic design and upgrading of steel moment frames. Research report no. UMCEE 98–29. Department of civil and environmental engineering, The University of Michigan, Ann Arbor

Lenwari A, Thepchatri T, Albrecht P (2006) Debonding strength of steel beams strengthened with CFRP plates. J Comp Constr 10(1):69–78

Leon RT, Yang CS (2003) Special inverted-V-braced frames with suspended zipper struts. Georgia Institute of Technology

Marian L, Giaralis A (2014) Optimal design of a novel tuned mass-damper-inerter (TMDI) passive vibration control configuration for stochastically support-excited structural systems. Prob Eng Mech. http://www.sciencedirect.com/science/article/pii/S026689014000216

Pachoumis DT, Galousis EG, Kalfas CN, Christitsa AD (2008) Reduced beam section moment connections subjected to cyclic loading: experimental analysis and FEM simulation. Eng Struct 31:216–223

Plumier A (1997) The dog-bone: back to the future. Eng J 34(2):61–67

Qingshan Y, Bo LA, Na Y (2009) Aseismic behaviors of steel moment resisting frames with opening in beam web. J Construct Steel Res 65:1323–1336

Sarhosis V, Tsavdaridis KD, Giannopoulos I (2014) Discrete element modelling of masonry infilled steel frames with multiple window openings subjected to lateral load variations. The Open Construction and Building Technology Journal 8:93–103

Schnerch DA (2005) Strengthening of steel structures with high module carbon fiber reinforced polymer (CFRP) materials. PhD dissertation, North Carolina State University, p 265

Schwinger C (2014) Quality assurance for structural engineering firms. Online: available in May 2014 http://www.aisc.org/store/p-1795-quality-assurance-for-structural-engineering-firms.aspx

Shen Y (2009) Seismic performance of steel moment-resisting frames with nonlinear replaceable links. Master Thesis, Department of Civil Engineering, University of Toronto

Slotboom M, Robbemont A, Habraken A, Teuffel P (2014) Safety and redundancy of adaptive buildings structures. In: Civil engineering for sustainability and resilience international conference, CESARE'14, Amman

Smith MC (2002) Synthesis of mechanical networks: the inerter. IEEE Trans Autom Contr 47(10):1648–1662

Thorburn LJ, Kulak GL, Montgomery CJ (1983) Analysis of steel plate shear walls. In: Structural engineering report no. 107. Department of Civil Engineering, University of Alberta, Edmonton

Tsavdaridis KD, D'Mello C, Huo BY (2013) Experimental and computational study of vertical shear behaviour of partially encased perforated steel beams. J Eng Struct 56:805–822

Tsavdaridis KD, Faghih F, Nikitas N (2014) Assessment of perforated steel beam-to-column connections subjected to cyclic loading. J Earthquake Eng 18(8):1302–1325

Uang CM, Yu QS, Noel SA, Gross JL (2000) Cyclic testing of steel moment connections rehabilitated with RBS or welded haunch. J Struct Eng 126(1):57–68

Vasdravellis G, Karavasilis TL, Uy B (2013) Large-scale experimental validation of steel post-tensioned connections with web hourglass pins. J Struct Eng ASCE 139(6):1033–1042

Vian D, Bruneau M (2004) Testing of special LYS steel plate shear walls. In: 13th World conference on earthquake engineering, Vancouver, Paper no. 978

Vincent RB (2008) Minimizing the strength of bracing connections. In: 6th international workshop connections in steel structures VI. Chicago, pp 127–141

Yu QS, Uang CM, Gross JL (2000) Seismic rehabilitation design of steel moment connection with welded haunch. J Struct Eng ASCE 126(1):57–68

S

Strengthening Techniques: Masonry and Heritage Structures

Elizabeth Vintzileou
Department of Structural Engineering, Faculty of
Civil Engineering, National Technical University
of Athens, Athens, Greece

Synonyms

Arch; Box action; Deep rejointing; Diaphragm;
Dome; Foundation; Grouting; Masonry; Mortar;
Repair and strengthening; Stitching; Ties; Vault

Introduction

Masonry structures constitute a large part of the
existing structures throughout the world.
A significant part of the existing masonry struc-
tures belong to the built cultural heritage. Fur-
thermore, many countries with rich cultural
heritage are situated in earthquake-prone areas
(in Europe, in Asia, in South America). The
built cultural heritage includes unique monu-
ments (temples, churches, castles, palaces, brid-
ges, etc.) and (inhabited and leaving) historic
centers. Therefore, there is a continuous need
for the preservation of this wealth, whose cul-
tural, historical, social, and economic values can-
not be overestimated.

On the other hand, the international scientific
community is working quite intensively in the
field of preservation of the existing stock of struc-
tures, and, thus, the knowledge acquired in the
last decades allows for interventions to be more
rationally selected and better applied.

It is obvious that the vast variety of existing
masonry structures, in terms of typology, of
materials used for their construction, of their cul-
tural significance, etc., does not allow this text to
be of general applicability and to cover all possi-
ble cases. Thus, to make this text to be of some
use, one has to limit its scope to a part of the
existing masonry structures. Unique monuments
are usually the subject of specific studies

(exhaustive surveys, in situ and laboratory tests,
monitoring, sophisticated analyses, etc.) leading
to the selection of adequate intervention tech-
niques by multidisciplinary national or interna-
tional teams. Thus, they are left out of the scope
of this text. Furthermore, structures other than
buildings (e.g., towers, bridges, etc.) present
peculiarities that need to be treated separately.
Therefore, those structures as well are not cov-
ered by this text. However, the repair and
strengthening techniques, concisely presented
and treated herein, are applicable also to the cat-
egories of structures that are not explicitly
covered.

Prerequisites for the Selection of Adequate Intervention Techniques: Intervention Strategies

Interventions to individual structural members or
to parts of an existing masonry structure should
be part of a global intervention strategy. The
selection of a strategy and, hence, of appropriate
intervention techniques (for the repair or
strengthening of a structure) is based on a series
of investigations and studies that should precede
the redesign of the structure, as briefly described
herein:

(a) **Documentation of the history of the struc-
    ture**: It comprises the date of construction,
    information on construction phases, events
    that may have affected the structural system
    (earthquakes, floods, fires, wars, etc.), previ-
    ous damages and interventions, etc. This
    phase is of paramount significance especially
    for cultural heritage structures.

(b) **Documentation of the structural system**: It
    comprises the documentation of both the
    structural system and its current state. The
    documentation of the structural system
    includes the survey of the bearing system
    (bearing elements, their interconnections,
    path of forces from the region of their appli-
    cation to the Earth), as well as type and prop-
    erties of materials, construction type of
    masonry, etc. The documentation of the

current state of the structure includes the survey of its pathology (survey of damage and decay), as well as qualitative interpretation of damage and decay. This step may require in situ and laboratory investigations (destructive and/or nondestructive, depending on the significance of the structure, its state, its future use, the characteristic to be detected, etc.).

(c) **Diagnosis and assessment**: The input from steps (a) and (b) serves the needs of diagnosis (i.e., the identification of the causes of the observed pathological image) and assessment (i.e., the margins of safety of the structure in its current state). At this step, the structure is adequately modeled and subjected to normal (self weight, live loads, etc.) and to accidental actions (e.g., seismic action) with the purpose to reproduce the main damages of the structure. This is a step of major significance, as it allows for identification of the causes of damage and of the regions of the structure that are vulnerable, for checking the adequacy of modeling, and for assessing the margins of safety of the existing structure.

The output of steps (a) to (c) allows the engineer to select the strategy of interventions, to choose among the available intervention techniques those techniques that serve better the purpose of repair or strengthening of the structure, and to proceed with dimensioning of the interventions.

Depending on the results of the assessment of the existing structure and taking into account the actions that are expected to act on the structure in the future (e.g., change of use, required seismic upgrading of the structure, etc.), one of the following strategies of interventions may be applied:

(i) *Local interventions to bearing elements*: This strategy is applicable when the bearing system is in principle adequate, whereas local inadequacies are detected in a limited number of bearing elements. The interventions consist in enhancing the bearing capacity or the deformation capacity or the

stiffness or the interconnection of limited number of elements, without major alteration of the bearing system.

(ii) *Alleviation of irregularities*: This strategy is applicable, when the irregular distribution of masses, stiffnesses, and/or strengths leads to respective excessive requirements in some regions of the structure. In such a case, the alleviation of irregularities (by adding shear walls, by modifying the arrangement of existing bearing walls, etc.) may constitute the best strategy. It should be noted, however, that this strategy may be of limited applicability in the case of cultural heritage structures.

(iii) *Global enhancement of bearing capacity*: This strategy is adequate when (a) a large number of bearing elements are severely damaged or (b) large inelastic deformations are expected to be imposed to the structure. In such a case, a global enhancement of the bearing capacity of the structure is advisable (strengthening of existing elements and/or addition of new bearing elements). Needless to say, the resulting enhancement of the global stiffness of the structure has to be taken into account.

(iv) *Global enhancement of stiffness*: This strategy may be selected when the calculations show that the expected deformations are larger than those the (brittle) masonry elements can sustain. If this is the case, global enhancement of stiffness may be achieved by strengthening of the bearing elements and/or by improving the box action of the structure, e.g., by enhancing the diaphragm action of floors. Thus, the deformations to be imposed to the structure are reduced to a level the masonry elements can sustain.

(v) *Reduction of masses*: This is a strategy that is desirable for structures to be subjected to seismic actions. In case of masonry structures, it may consist in using lighter materials for roof and floor pavements, as well as by reducing live loads (selection of appropriate use of the structure). It should be noted, however, that this strategy does not

S

lead to spectacular results, as – in the case of masonry structures – their self weight constitutes the most significant part of the vertical loads.

(vi) *New bearing system*: When the existing bearing system is inadequate for seismic actions, a new bearing system may be constructed inside the existing structure. The new bearing system (e.g., made of reinforced concrete (RC) or steel) is connected to the existing structure, thus contributing to the resistance to seismic actions. This is a strategy applied in some cases to cultural heritage structures, where extensive interventions to the existing bearing system are invasive and, thus, undesirable. It should be noted that, the dimensioning of the new bearing system will be governed by the compatibility of deformations with those of the existing one, rather than by the required additional bearing capacity.

Repair and Strengthening Techniques

The information included in this chapter regards the most commonly applied techniques for the repair and strengthening of existing masonry structures. It should be noted that some of the techniques presented here, although of frequent use, are not sufficiently documented through experimental work or through monitoring of structures after intervention. The research on intervention techniques, as well as on the production, testing, and use of new materials, is ongoing.

The material included in this chapter is organized as follows: (a) Repair and strengthening techniques for masonry as well as for the connection between masonry elements are first presented. The aim of those techniques is to reinstate the pre-damage bearing capacity of individual elements or to enhance their bearing capacity, so that they may sustain seismic actions in the future. Finally, the repair of damage in the connection between masonry elements as well as its strengthening is examined, since the so-called box action of masonry structures strongly depends on those connections. Subsequently, (b) techniques for the enhancement of the diaphragm action of floors and roofs as well as for their connection with the vertical elements are presented. The purpose of those techniques is, in most cases, not to enhance the bearing capacity of floors and roofs, but to ensure almost equal displacements of all vertical elements and to prevent diaphragms from moving independently of the vertical elements during an earthquake, thus, causing severe damage to the walls. Finally, (c) some indicative solutions frequently applied to cope with problems such as sliding soil or insufficient dimensions of the foundation are briefly presented.

Repair and Strengthening Techniques for Masonry Walls

Deep Rejointing

Deep rejointing consists in the removal of the mortar from the joints to a depth at least equal to the thickness of the joints and filling with a mortar of improved properties. The technique should preferably be applied on both faces of masonry. In order to reach a significant enhancement of the mechanical properties of masonry, and taking into account that the contribution of the mortar to the strength of masonry is rather limited, a large part of the existing mortar has to be replaced. Therefore, this technique may be efficient only when applied to masonry of limited thickness (say 300 mm max.). This limits the applicability of deep rejointing to brick or block masonry, excluding stone masonry walls (of minimum thickness around 400 mm). It has to be noted that rejointing may be needed for reasons other than strengthening (e.g., to avoid water permeability of masonry or to prepare masonry for the application of grouts). In those cases, usually, there is no need for deep rejointing.

The application of the new mortar has to be made very carefully, so that the voids created due to the removal of the existing mortar are fully filled. Furthermore, the new mortar should not be excessively strong compared to the existing one. Otherwise, the compressive stresses on the

Strengthening Techniques: Masonry and Heritage Structures, Fig. 1 (**a**) Qualitative interpretation of spalling of masonry units due to defective application of deep rejointing, (**b**) A brick wallette tested in compression after a defective application of deep rejointing (Vintzileou 2001)
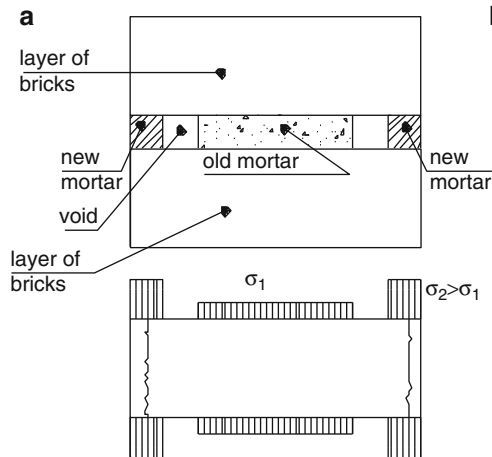
exterior (limited in thickness) part of masonry may be excessive (Fig. 1a) and lead to spalling of the masonry units (Fig. 1b). As a result, the compressive strength of masonry may be reduced, instead of being increased.

In conclusion, (a) deep rejointing may be efficient in case of brick masonries, 300 mm thick at maximum, and (b) provided that the technique is applied with due care and using adequate materials, (c) in case of stone masonries, it is not advisable to apply deep rejointing both because it is not expected to contribute to the improvement of the mechanical properties of masonry and because it is a quite expensive technique (mainly in workmanship). Even if applied to stone masonries, deep rejointing is considered as a repair technique, reinstating the initial mechanical properties of masonry.

Grouting of Cracks

This repair technique is applicable to cracks having a width that does not exceed 10 mm approximately:

(a) Preparation of masonry: When masonry is plastered, plaster is removed in the vicinity of the crack. In cases where plaster cannot be removed (e.g., when masonry is covered with paintings or frescoes or mosaics or other decorative elements), special care has to be taken for the protection of the decoration of masonry. For example, it should be

checked whether the frescoes are fixed on masonry. If not, their adequate fixing should precede the application of grout. Furthermore, during the application of grout, special care should be taken for immediate cleaning of the face of masonry, in case of leaking of the grout.

(b) Holes are drilled to masonry at distances along the crack. The distance of consecutive holes should not exceed the thickness of masonry. Transparent plastic tubes are installed into the holes. Part of the length of the tubes is protruding.

(c) The crack is sealed along its entire length, using an adequate mortar.

(d) The materials for the grout (see also section "Generalities") are mixed and the grout is introduced to the crack from the lower tube until it leaks from the immediately higher up placed tube and so on, until the crack is filled over its entire length. The grout is applied with a pressure not exceeding 1 bar. Depending on the state of the masonry, in order to avoid excessive hydrostatic pressure that might cause further damage to masonry, the crack length to be grouted should be limited within 1 day. For damaged, low strength masonries, this length should be limited to 1 m per day.

(e) After hardening of the grout, the protruding parts of the tubes are cut and masonry is plastered, if this is the case.
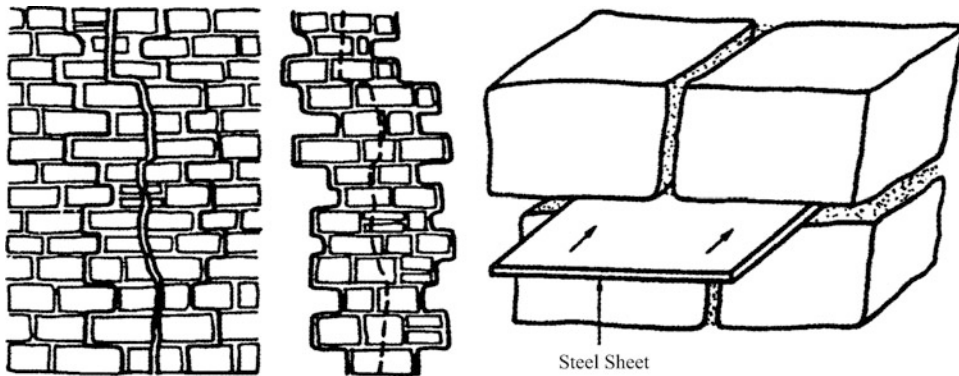
Strengthening Techniques: Masonry and Heritage Structures, Fig. 2 Stitching of wide cracks

Strengthening Techniques: Masonry and Heritage Structures, Fig. 3 Repair of wide cracks by means of reinforced zones (indicative reinforcement)
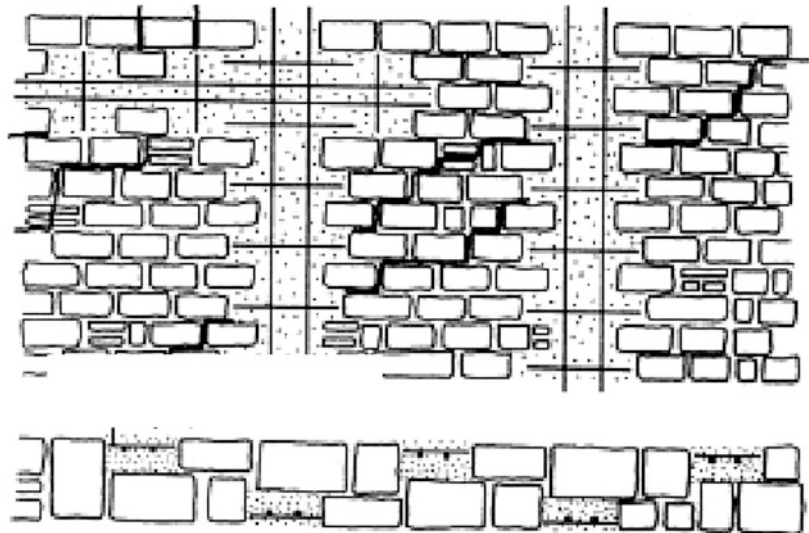
Stitching of Cracks

In case of almost vertical wide cracks (width exceeding 10 mm), grouting alone cannot lead to reinstatement of the initial strength of masonry. In such a case, stitching of the cracks, prior to grouting, is needed.

Stitching consists in the removal of cracked masonry units along the crack and replacement with new sound masonry units laid in mortar of adequate composition (Fig. 2). If necessary, some of the bed joints along the crack may be locally reinforced. To this purpose, steel plates are inserted in the mortar joints (Fig. 2). The plates are smooth and, hence, of limited bond with the mortar. It is therefore advisable to artificially roughen them.

This technique cannot be applied to masonry walls that are cracked due to shear (diagonal or bi-diagonal cracks), as the removal of cracked masonry units along the cracks is not feasible. In those cases, the wall should be either demolished (and reconstructed) or it should be carefully grouted, provided that its construction type and its state allow grouting (see section "Grouting").

An alternative, applicable exclusively to buildings without any heritage value, would be that of vertical and horizontal reinforced zones (Fig. 3). Grooves are created in the wall, both horizontally and vertically on both sides of masonry. Steel reinforcement is placed and a mortar of adequate composition or concrete is poured.

Stitching of Vertical Cracks Between Perpendicular Walls

The same technique is applicable to large cracks in the connection of walls. In that case, stones or bricks are removed and replaced by intact ones positioned diagonally (Fig. 4), to connect the two walls.

Alternatively, steel plates can be introduced to mortar joints (Fig. 5), provided that the thickness thereof is large enough for accommodating the plates.
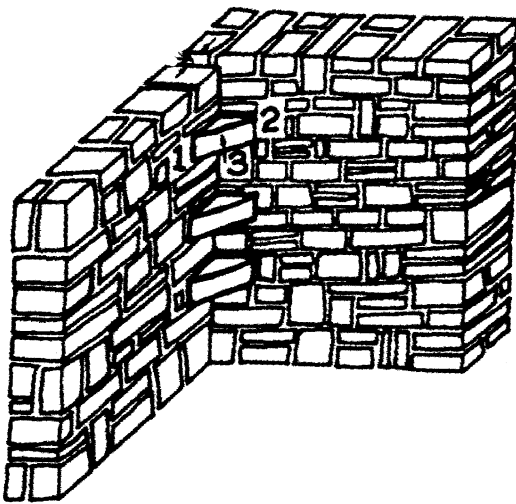
Strengthening Techniques: Masonry and Heritage Structures, Fig. 4 Stitching of vertical cracks between perpendicular walls. Stones 1 and 2 are substituted by stone 3

It should be noted that (a) the separation between the walls is not reversible and (b) the steel plates inserted into mortar joints are passive (i.e., they are mobilized in case further relative displacement of the walls takes place). Therefore, the cracks – after stitching – have to be grouted.

Grouting

Generalities Grouting of masonry (termed also as "homogenization" of masonry) consists in filling all voids and cracks within the mass of masonry using a hydraulic mixture. Organic binders (mainly epoxy resin) are not in use for masonry; the exothermal reaction of their hardening, combined with the (quite often) large voids within the mass of masonry may lead to disintegration of the epoxy resin-based grout before the application of any actions to the building. Furthermore, due to the dust present in the mass of masonry, the bond between the grout and the in situ materials may be defective. Finally, due to durability considerations, organic binders are excluded. It should also be noted that hydraulic grouts, even of low to medium mechanical properties can ensure significant enhancement of the mechanical properties of masonry (e.g., Binda et al. 1994; Valluzzi 2004; Vintzileou and Miltiadou-Fezans 2008), thanks to their high bond with the in situ materials.

A special mention needs to be made here of cement-based grouts (with a cement content exceeding 50 % per weight). Although they

Strengthening Techniques: Masonry and Heritage Structures, Fig. 5 Steel plates accommodated in horizontal mortar joints for stitching cracks

a   steel plate for stitching of walls

crack

b   steel plate for stitching of walls

crack
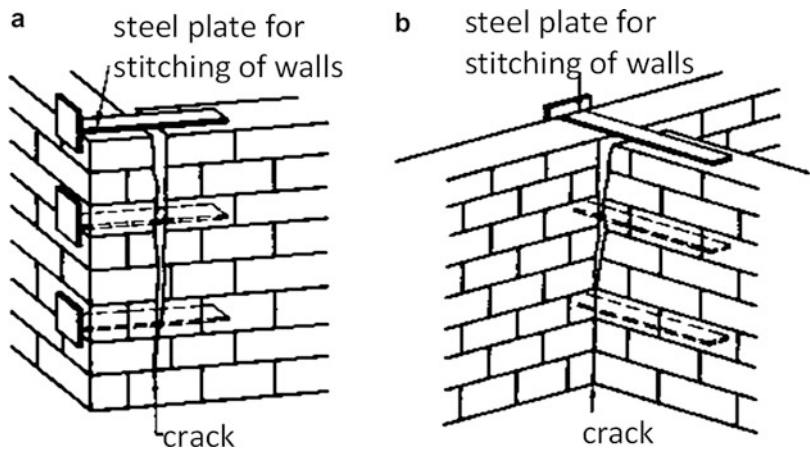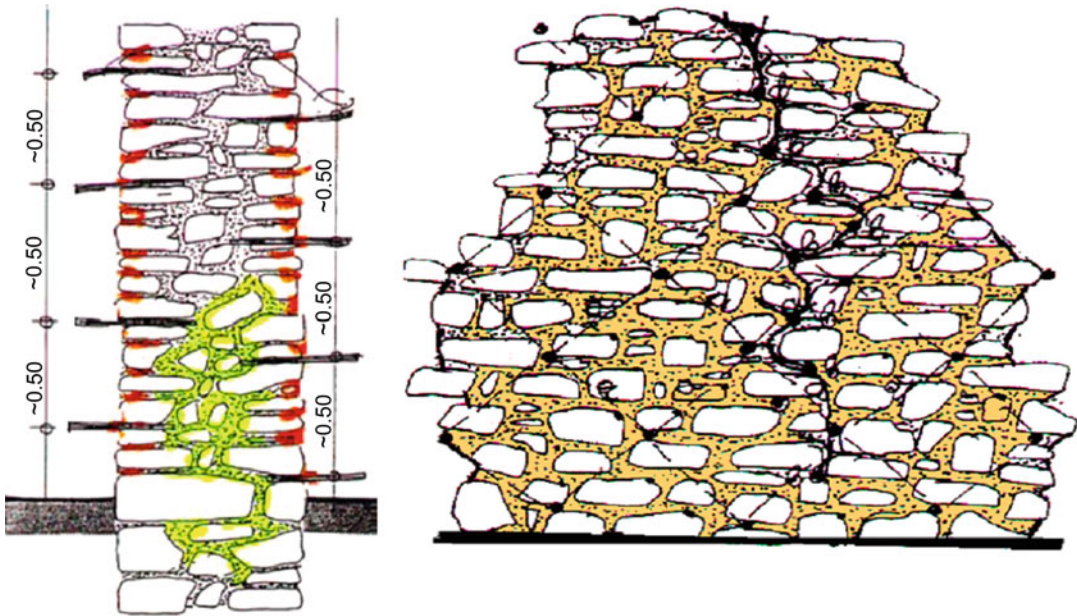
Strengthening Techniques: Masonry and Heritage Structures, Fig. 6 Grid of holes in a masonry wall to be grouted

constitute the first application of grouts to masonry, the experimental work carried out in several countries (Vintzileou and Tassios 1995; Toumbakari 2002) has proven that cement-based grouts are not more efficient than ternary grout (cement content of approximately 30 % per weight, lime (and additives)) or hydraulic lime-based grouts. Furthermore, several cases of durability problems due to the use of cement are reported (e.g., efflorescence). Thus, in the recent years, the use of cement-based grouts is very limited, if not fully abandoned.

The Design of the Grout Hydraulic grouts injected into masonry have to comply with a set of requirements, namely, rheological (injectability, i.e., penetrability into fine cracks and voids, according to the design, and sufficient fluidity for the grout to be diffused into masonry, as well as stability), physical (low hydration heat, limited shrinkage, adequate hardening time, and hygroscopic properties), chemical (e.g., resistance to expansion and chemical stability of the products of chemical reactions), and mechanical requirements that are related to the desirable mechanical properties of the grouted masonry (i.e., strength and deformability characteristics), depending on the actual state of masonry, the actions to be imposed, the overall scheme of interventions, etc.

Therefore, the design of a grout requires sufficient knowledge on the construction type of masonry, on its state, on the in situ materials, and on the aim of the intervention, all these combined with knowledge on material science.

Preparation of Masonry, Mixing of Grout, and Application The preparation of masonry for grouting is similar to that described in section "Grouting of Cracks." Holes are drilled to masonry at distances, forming a grid. In the general case of application of grouts from both faces of masonry elements, holes are drilled on both faces of masonry. As shown in Fig. 6, the holes on the interior face of masonry are drilled halfway between consecutive holes drilled on the exterior face of masonry. Thus, the resulting grid of holes is dense and it ensures better filling of the internal voids and cracks. The holes should be drilled deep enough (Fig. 7) to reach the core of

Strengthening Techniques: Masonry and Heritage Structures, Fig. 7 Drilling holes in masonry

masonry, thus ensuring that the grout will fill all the internal voids and cracks. Transparent plastic tubes are installed into the holes (Fig. 8) with an inclination toward the bottom of the element (to facilitate the flow of the grout). The (horizontal and vertical) distance between consecutive tubes may vary between 0.5 m and 1.0 m. However, it should in no case exceed the thickness of the elements to be injected.

The plastic tubes are numbered and reported on a drawing (Fig. 8). This is of major importance, because it allows for full control of the application procedure (tubes from which the grout is introduced and those from which the grout overflows are marked during the injection). All cracks are sealed (Fig. 9).

The materials to be used for the production of the grout should be available at the vicinity of the region to be injected, together with the mixer (Fig. 10). The use of high turbulence mixer is recommended, to provide adequate mix of the materials. After mixing of the constituent materials, the grout is transferred to a transparent

tube (Fig. 11), equipped with a device that provides a mechanical stirring of the grout during application (to avoid segregation).

When the preparation of masonry is completed (Fig. 12) and the grout mix is ready, the grout is applied with low pressure from the bottom to the top of masonry. The consumed quantity of grout is recorded. This is a very significant piece of information; it is taken into account when the enhanced mechanical properties of masonry are estimated.

Masonry is cured (kept wet) during the entire procedure that may last for several days, depending on the size and state of the injected element.

Reinforced Plaster

This is a technique applicable to masonry of limited thickness (not exceeding 0.40 m). It is applied after repair of masonry (by grouting or by stitching of cracks, by rejointing masonry, etc.). Reinforced plaster should preferably be applied on both sides of masonry. Actually, it is beneficial especially against out-of-plane bending, and, therefore, either the interior plaster or the exterior one is mobilized (in tension) depending on the direction of the seismic action.

The plaster is removed and masonry is cleaned from any loose material. The (metallic or nonmetallic) grid of reinforcement is placed on the faces of masonry and fixed on it (using anchors). In order to achieve better bond of the reinforced plaster, it is advisable to remove the mortar from the joints as deep as possible and fill them during the application of the plaster. In case of metallic reinforcement (steel grid), the plaster is usually a cement-based one. In case of nonmetallic reinforcement (e.g., fiber-reinforced wires or bars), the plaster can be a hydraulic lime + pozzolan mortar.

Recently, fiber-reinforced mortars are also applied to masonry.

It should be noted that the use of reinforced plasters may not be advisable in the case of cultural heritage assets, where masonry is quite thick, and the application of the reinforced plaster may be possible only on the interior face of masonry (therefore, in its least efficient position).

S
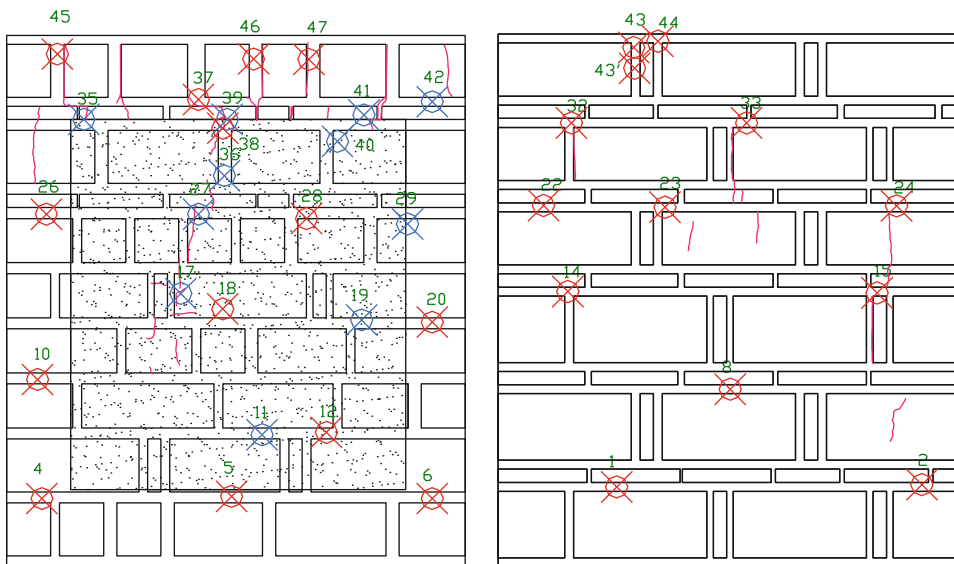
Strengthening Techniques: Masonry and Heritage Structures, Fig. 8 Plastic tubes installed into drilled holes. Tubes are numbered and reported to drawings

Local Demolition and Reconstruction
In case of local collapse of walls or in case of partial collapse of the exterior leaf of a double- or three-leaf masonry or in case of significant out-of-plane displacement of a wall (Fig. 13), the only solution is the reconstruction of the part that has collapsed or undergone significant deformations.

Needless to say, reconstruction is usually combined with repair or strengthening techniques described in the previous paragraphs (e.g., grouting of three-leaf masonry, stitching of

large cracks in other than the collapsed regions of the building, etc.).

Strengthening of Walls Using FRP Rods, Laminates, or Sheets
The use of FRPs in strengthening (bearing and infill) masonry is quite extensive in the recent years. FRPs with various types of fibers and in various forms (rods, laminates, or sheets) are used for strengthening masonry walls (Fig. 14) either against out-of-plane (bending) or in-plane (shear) actions.

The use of FRP materials is supported by extensive experimental research (collected, e.g., in Morbin 2001; Rashadul 2008), whereas models for the dimensioning of the intervention are also proposed (e.g., Faella et al. 2004;

Strengthening Techniques: Masonry and Heritage Structures, Fig. 9 Sealing of cracks

ElGawady 2005). The available experimental data prove that the technique is quite efficient in enhancing the bearing capacity of masonry, as well as its ductility, provided that a good bond with the substrate (masonry) is ensured.

The technology of FRP applications is not presented in detail in this text for several reasons, namely, (a) most of the tests available in the Literature refer to monotonic loading of masonry specimens; in a limited number of cases, also repeated loading (without change of sign) or cyclic loading (e.g., Grillo 2003, on concrete masonry) was applied to the specimens. On the other hand, (b) in laboratory tests, masonry walls are subjected either to out-of-plane bending or to in-plane shear. However, during an earthquake, masonry walls are subjected to cyclic actions, as well as to simultaneous in-plane and out-of-plane actions. This is a very adverse condition which is not sufficiently covered by the international Literature. The problem that may arise is

Strengthening Techniques: Masonry and Heritage Structures, Fig. 10 Materials for preparation of the grout mix. High turbulence mixer

a premature de-bonding of the FRPs from the substrate with subsequent loss of the strengthening of the walls. Furthermore, (c) fire protection of the intervention is an issue. Another issue that

Strengthening Techniques: Masonry and Heritage Structures, Fig. 11 Device for mechanical mixing of the grout before application

needs to be very carefully examined is (d) whether, after the occurrence of damage in FRP strengthened masonry, its repair or strengthening is still possible. Actually, as proven by tests, the failure of FRP strengthened walls – even under monotonic actions – may be associated with extensive cracking and disintegration of masonry. Especially in the case of historic masonry, for which the possibility to re-intervene after a seismic event is a prerequisite, this aspect may be governing. Last but not least, (e) there are limited data regarding strengthening of historic masonry using FRP materials. It is questionable whether poor-quality (double-leaf or three-leaf) stone masonry, having very low elastic properties, can be strengthened using high-strength and low-deformability materials. In such a case, prior strengthening of masonry using "traditional" techniques would be required, whereas the issue of compatibility of the in situ and added materials should be carefully examined. For example, the application of FRP sheets, covering large portions of masonry elements, may create problems of building physics (e.g., entrapment of humidity in the mass of masonry walls), thus leading to decay of the covered masonry. Another example would be that of inserting FRP rods in (horizontal and vertical) mortar joints. Even if the practical difficulties in finding

Strengthening Techniques: Masonry and Heritage Structures, Fig. 12 Masonry walls ready for the application of grouts

Strengthening Techniques: Masonry and Heritage Structures, Fig. 13 (**a**) collapse of the corner of a building, (**b**) partial collapse of the exterior leaf of a three-leaf masonry wall, (**c**) significant out-of-plane displacement of a masonry wall

Strengthening Techniques: Masonry and Heritage Structures, Fig. 14 Alternative configurations of FRPs for masonry strengthening (schematic)

FRP rods inserted in mortar joints   FRP laminates   FRP sheets
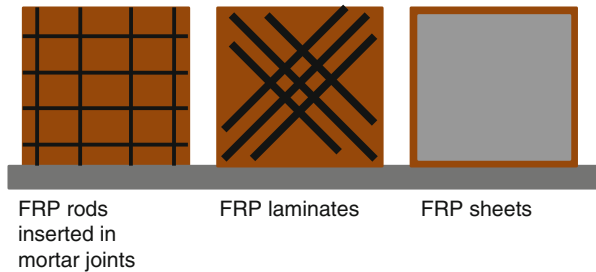
continuous bed joints in rubble stone masonries, the fact that grooves should be created to insert vertical FRP rods, and the difficulties in anchoring the rods at the top and at the base of a wall, as well as into transverse walls, are neglected, the problem remains of the behavior of rather soft original mortar joints that are partially filled with a (by orders of magnitude) stronger material. In real buildings, where the wall will be subjected to simultaneous cyclic shear and cyclic out-of-plane bending, there is a danger of stress concentration in the reinforced portions of the masonry section. The stress concentration may lead to spalling of the masonry units and to subsequent adverse effects on the bearing capacity of the masonry element (comp. with Fig. 1). Moreover, (f) in case of historic buildings, the application of FRPs may be not be acceptable, as it alters the appearance of (unplastered) masonry. Thus, although the application of FRPs seems to be a technically

Strengthening Techniques: Masonry and Heritage Structures, Fig. 15 Typical damages due to out-of-plane bending of walls

promising intervention, it is believed that this technique should be applied with caution for strengthening masonry buildings that are to be subjected to seismic actions. It is believed that the (under investigation) use of inorganic bonding material (mortar) instead of an organic one (resins) may be a sensible alternative.

Enhancement of the Box Action of Masonry Buildings

Introductory Remarks

It is well known that masonry is a brittle material; it reaches its maximum resistance at small deformation value, whereas an abrupt decrease of its resistance takes place for imposed deformations exceeding the deformation corresponding to the maximum resistance. The vulnerability of existing masonry buildings to large deformations imposed by an earthquake is due to various typical construction features, namely, the defective connection of walls among them, the flexibility of floors and roofs, the large openings (especially when located close to the corners of the building), the large distance between transverse walls, etc. All measures taken with the purpose of alleviating those sources of vulnerability contribute to an improved seismic behavior of masonry. Last but not least, all the techniques that improve the box action of masonry buildings lead to an increase of their stiffness and to subsequent reduction of their natural period. In this way, the building "moves" to the left region of the response spectrum and, hence, to lower seismic actions.

Enhancement of the Diaphragm Action of Floors and Roofs

In existing masonry buildings, typically, timber floors and roofs are simply resting on vertical elements (walls or piers). Furthermore, due to the flexibility of floors and roofs in their plane, masonry walls are quite vulnerable to out-of-plane bending (Fig. 15). Under such actions, vertical cracks close to the connections between longitudinal and transverse walls, as well as out-of-plane collapse of portions of walls, constitute typical damages of masonry buildings.

The enhancement of the diaphragm action of floors and roofs leads to almost equal deformations of all vertical elements, thus limiting the out-of-plane vulnerability. Thanks to the overall increase of the stiffness of the building, the imposed deformations are also significantly reduced.

There are several techniques for the enhancement of the diaphragm action of floors and roofs. The Engineer has to select those most appropriate for each specific case and document (through analysis) the efficiency of the selected techniques.
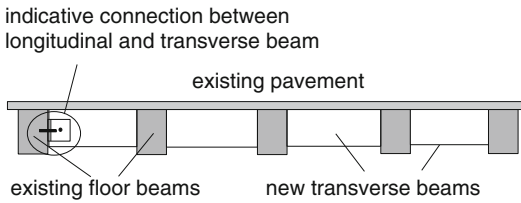
Strengthening Techniques: Masonry and Heritage Structures, Fig. 16 Enhancement of the diaphragm action of a floor, arrangement of transverse beams connected to the existing ones (schematic)

Some of the possible alternatives are briefly presented herein:

(a) Providing transverse beams to the existing floor: Typically, in existing structures, floors consist of beams spanning in one direction (normally, the smaller of the two). The in-plane flexibility of the floor can be reduced by arranging (timber or steel) beams perpendicular to the existing ones (Fig. 16), thus forming a grid. It should be noted that the cross-sectional dimensions of the transverse beams, their spacing, and the dimensioning of their connections to the existing members are to be determined for each specific case. Similarly, in case of a trussed roof, beams perpendicular to the existing trusses can be arranged.

(b) Providing a system of bi-diagonal steel ties: Significant enhancement of the in-plane stiffness of a floor (or roof) can be achieved by arranging steel ties along the two diagonals of the floor (Fig. 17). A system of multiple ties may be needed in the case of rectangular floors, in which one in-plan dimension is significantly larger than the other. As in the previous case, the ties (adequately connected to the existing floor beams or to the existing trusses of the roof) are to be adequately dimensioned. It should be noted that, typically, ties should be dimensioned to develop a tensile stress not exceeding 50 % of their yield strength, in order to avoid excessive strains of the steel elements and, hence, avoid significant in-plane deformations of the floor.
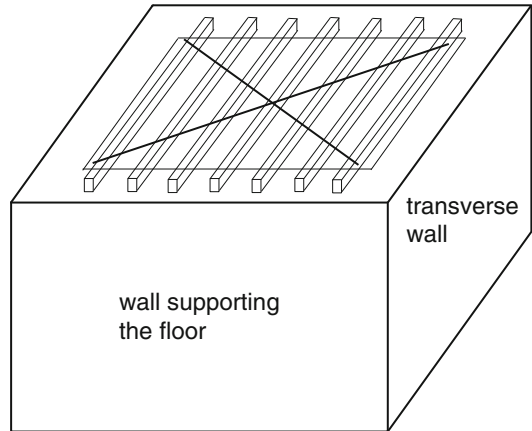
Strengthening Techniques: Masonry and Heritage Structures, Fig. 17 Arrangement of steel ties for the enhancement of the in-plane stiffness of a floor (schematic)

(c) Providing a second pavement: In the frequent case of a timber pavement of floors and roof, significant enhancement of the in-plane stiffness can be reached through a second pavement. This second pavement can be located either on top of the existing one or at the bottom of the beams (to avoid changes in the height of the story). The second pavement should be either perpendicular to the existing one or, preferably, arranged at an angle of 45° (Fig. 18). In this case too, the Designer should check the efficiency of the solution through modeling and analysis of the structure.

(d) Alternatively, instead of providing a second pavement, the in-plane stiffness of the floor may be enhanced by arranging bi-diagonal FRP laminates (Fig. 19). This alternative solution should be applied with caution: The efficiency of the FRP plates depends on their bond with the existing pavement. The simultaneous flexural (due to vertical loads) and in-plane (due to seismic actions) deformations of the floor may lead to detachment of FRP laminates from the substrate. In such a case, the efficiency of the intervention may be significantly reduced.

(e) Replacement of the existing floors and roof by reinforced concrete slabs: This alternative, quite popular in previous decades, is
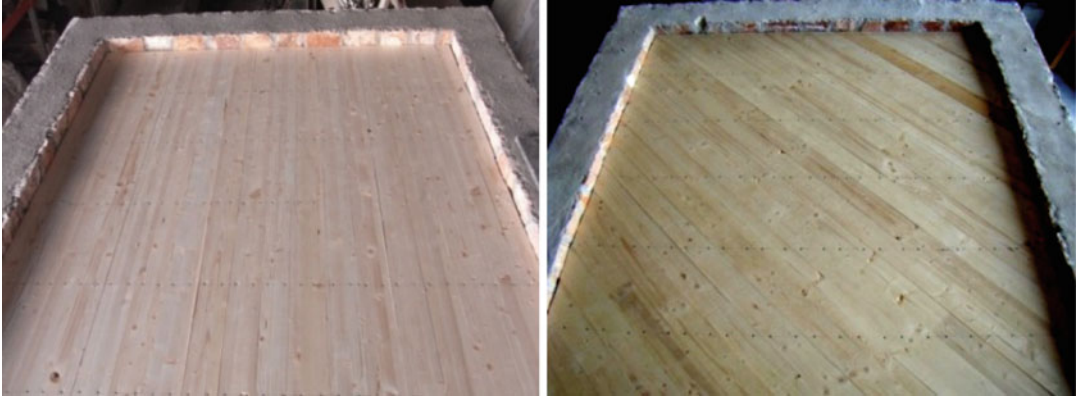
Strengthening Techniques: Masonry and Heritage Structures, Fig. 18 Existing timber pavement and additional timber pavement at an angle of 45°
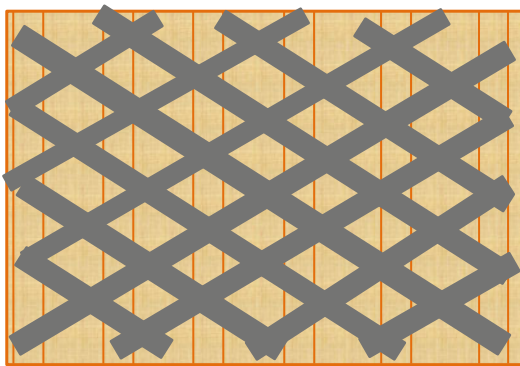
Strengthening Techniques: Masonry and Heritage Structures, Fig. 19 FRP laminates glued on the existing pavement (schematic)

nowadays of rather limited application. Actually, it is rather invasive in case of cultural heritage buildings and, thus, not accepted by the competent authorities. Furthermore, recent research (Zaopo 2011; Mazzon et al. 2009; Magenes et al. 2012; Mouzakis et al. 2012) on floors stiffened using several alternative techniques, as well as shaking table tests on model buildings, has proven that solutions (a) to (d) provide sufficient in-plane stiffness to the floor, and, hence, the replacement of the existing floors and roofs with RC slabs is not inevitable.

When this alternative is selected, adequate connection of the RC slab with the walls should be ensured. To this purpose, one of the following two techniques may be applied: (i) Grooves are made in masonry walls at distances (Fig. 20), steel reinforcement is positioned, and the groves are filled with concrete, during the construction of the slab. The spacing, the dimensions, and the reinforcement of the grooves are determined so that they are able to transfer the shear (due to the design seismic action) from the slab to the walls. The respective resistance of masonry should also be checked, (ii) steel beams are positioned along the masonry walls (Fig. 20), and they are fixed onto masonry. The RC slab is supported by the steel beams. In this case too, the connection of the steel beams to masonry walls should be adequately dimensioned.

Independently of the alternative selected for the enhancement of the in-plane stiffness of floors and roofs, there are two prerequisites for the efficiency of the solution, namely, (i) adequate connection of the diaphragm to masonry walls should be ensured and (ii) the supporting masonry walls should be adequately repaired or strengthened. Actually, a very stiff diaphragm not connected to masonry walls may be displaced as a rigid body during the earthquake, causing damage to the masonry walls. If the stiff diaphragm is connected to masonry walls made of unstrengthened poor-quality masonry, catastrophic damage may occur (Fig. 21).

Figure 22 shows the connection between floor and walls provided to a building model tested on

Strengthening Techniques: Masonry and Heritage Structures, Fig. 20 Alternative connections of RC slabs to masonry walls (schematic)
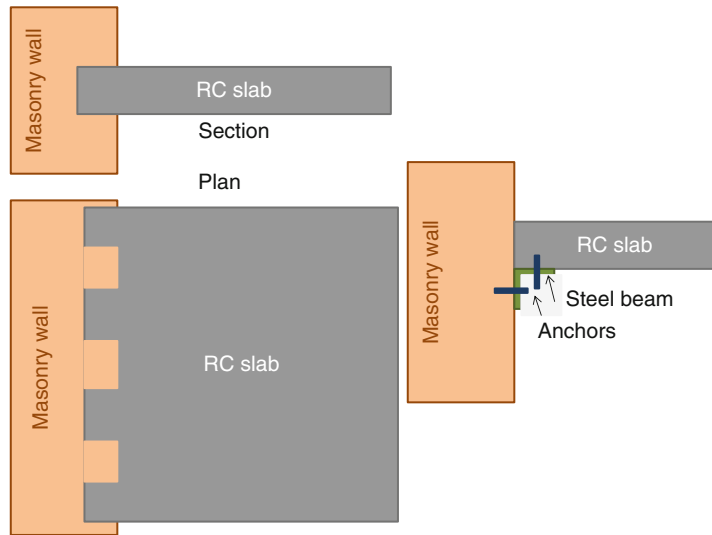
Strengthening Techniques: Masonry and Heritage Structures, Fig. 21 Catastrophic effects of RC slabs supported by poor-quality masonry (Courtesy of Prof. C. Modena)

the earthquake simulator (Mouzakis et al. 2012). The diaphragm action of the floor was enhanced as shown in Fig. 18, and the walls (made of three-leaf masonry) were grouted. The connection was proven to be efficient, and no damage was caused to masonry walls even for an input acceleration exceeding 1g.

Improving the Connection Between Floors/Roofs and Vertical Elements

Frequently, in existing masonry buildings, roof and floors are simply resting on masonry walls. Thus, the vertical loads (the self weight of floors and roof, the weight of pavements, as well as the live loads of the building) are concentrated at intervals. In case of poor-quality masonry, this may lead to local damage. Furthermore, during an earthquake, the floors are moving rather independently from the vertical masonry elements, thus, causing damage to the walls at their top (Fig. 23). Finally, the out-of-plane deformation of the walls is not prevented, due to loose connection with the horizontal elements. For all those reasons, it is desirable to provide better connection between horizontal and vertical elements.

Strengthening Techniques: Masonry and Heritage Structures, Fig. 22 Detail of the connection between diaphragm and masonry walls. The building model after strengthening
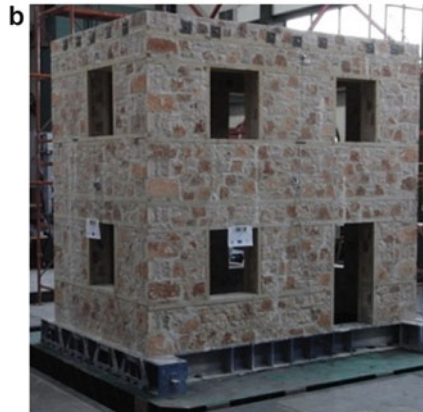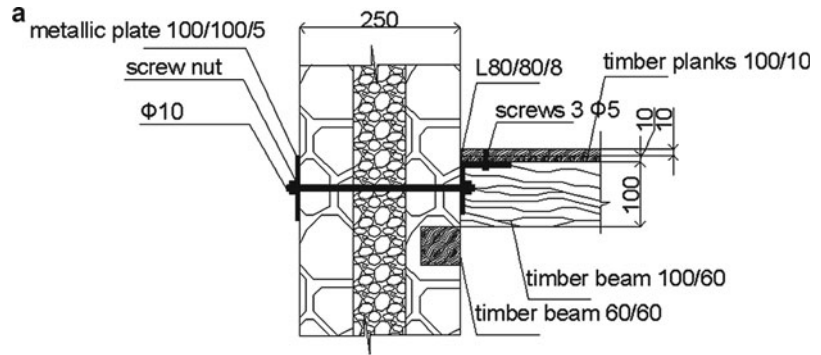
Strengthening Techniques: Masonry and Heritage Structures, Fig. 23 Damaged masonry walls due to the movement of simply supported roof

The improvement of the connection between horizontal and vertical elements can be achieved as follows:

(1) Arrangement of a timber or a steel beam in the perimeter of floors and roof (Fig. 24) and connection of the (timber or steel) beams of the floor with the collector beam. Although the beam shown in Fig. 24 was positioned during construction, it is feasible to provide a collector beam during interventions as well, by creating a recess in masonry walls. Floor

beams or roof trusses have to be connected to the collector beam, which should be connected to the walls (Fig. 22).

(2) When the roof is replaced or temporarily removed during interventions, a reinforced concrete tie beam can be constructed at the top of the upper floor walls. Usually, the width of the RC tie beam is equal to the thickness of the walls. However, in case there are decorative elements on the facade (to be restored after interventions) or the accommodation of insulation is needed, the width of the RC tie beam can be adequately reduced. The construction of the RC tie beam is also possible after temporary shoring of the roof, if the roof is not to be replaced. The construction of the tie beam with the roof in its place is not encouraged, although it is technically feasible.

The tie beam may be connected to masonry using anchors. The diameter, the length, and the spacing of the anchors should be adequately designed. It should be noted that the construction of RC tie beams in intermediate floors is not recommended because (a) it is a quite invasive intervention and (b) the RC tie beam will inevitably be of limited width. Thus, its connection to the beams of the floor and to the walls should be ensured in a way similar to that described in the previous paragraph (1), where all the application is dry and (c) durability problems may arise due to the continuous contact of timber elements and concrete.

Improving the Connection Between Walls

The significance of the adequate connection between walls was known to the constructors of masonry buildings. This is proven by the use of better quality, large-size cut stones in the corners of the buildings, even in the case of rather poor general construction quality (Fig. 25).

Nevertheless, the insufficiently stiff in their plane floors and roofs that are simply supported on the walls do not prevent damage due mainly to out-of-plane bending of walls and piers (Fig. 16).

It should be noted that, if enhancement of the diaphragm action of floors and roofs and adequate connection thereof with the walls is provided, further improvement of the connection between walls may not be needed. Actually, in such a case, the walls do behave as vertical slabs fixed along their perimeter, and their vulnerability to out-of-plane actions is significantly reduced. Nevertheless, if further improvement of the connection between walls is needed, steel ties can be used. The use of ties for the connection between walls was quite common during the construction of the existing masonry buildings.

Strengthening Techniques: Masonry and Heritage Structures, Fig. 25 Elaborate connection between walls at the corners of masonry buildings

Strengthening Techniques: Masonry and Heritage Structures, Fig. 26 The metallic tie (of limited length) was not able to prevent the out-of-plane collapse of the wall

In that case, the ties (of limited length) were arranged within the thickness of one wall and anchored on its transverse wall (Fig. 26).

Steel ties provided as an intervention to improve the connection between walls are not embedded in masonry. They are (preferably) arranged at both sides of a wall or only along the interior face of a wall. They may be arranged in one single level or in more levels, provided that the height of the story allows for more ties to be positioned. Ties may be provided along one or both main axes of the building.

The ties are usually passive (i.e., non-prestressed). Therefore, they are expected to be mobilized only in case differential movement of the two connected walls takes place. In case wall A (Fig. 27) tends to separate from walls C and D, the ties should be able to develop a tensile force sufficient to sustain wall A in its position. The ties are dimensioned so that their tensile stress does not exceed 50 % of their yield strength, in order to avoid excessive deformations when they are mobilized. Special care should be taken for their anchorage on masonry. The dimensions of the anchoring plates should allow (a) for almost uniform distribution of the compressive stresses on masonry and (b) for sufficiently small compressive stresses on masonry, to avoid local crushing.

The use of FRP ties as an alternative to steel ties is not encouraged. Although FRPs have a very high tensile strength (compared to normal steel) and they may have very high modulus of elasticity (which could prevent the occurrence of large deformations), they are characterized by complete lack of ductility. Therefore, if deformations larger than their deformation at failure occur, their fracture will be imminent.

C

B

Total force to be taken by the four
ties= (Self-weight of wall
"A")x(seismic coefficient)

D

A
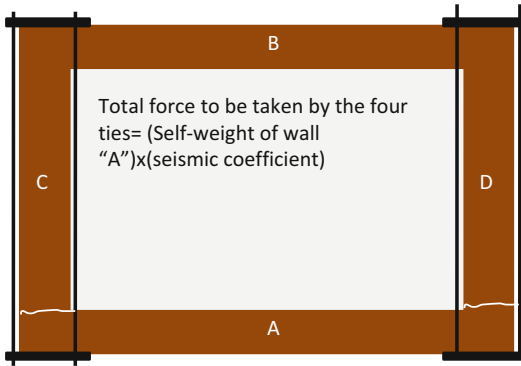
Strengthening Techniques: Masonry and Heritage Structures, Fig. 27 Arrangement of steel ties to improve the connection between walls (schematic)

Arches, Vaults, and Domes

In old masonry structures, arches (Fig. 28) were constructed to bridge openings in buildings (Fig. 28 a, b, f) and river bridges (Fig. 28g), to transfer loads from domes to masonry piers, and so on, whereas vaults and domes (Fig. 28c, d, e, f) are frequently used to bridge large spans. In most of cases, stone or brick curved elements can safely bear the vertical loads for which they were conceived and constructed. Damage may be due to excessive vertical loading resulting, for example, from inadequate use (Fig. 29) or due to the lack of maintenance and repair (Fig. 28e).

In cases of damage due to excessive loads or to decay, simple repair is sufficient. Repair may include local reconstruction, grouting, or stitching of cracks, as described in section "Repair and Strengthening Techniques for Masonry Walls."

It should be noted, however, that most of the damages of curved elements are due to the effects of the thrust. It is well known that the loads are transferred from the curved elements to their supports following a thrust line; the thrust line is oblique when it reaches the supports. Therefore, there is a horizontal component of the thrust that causes horizontal displacements to the supporting elements. In the (quite frequent) case of flexible supports, outwards moving vertical elements may be damaged and induce damage to the curved

elements as well. This phenomenon is accentuated in the case of seismic events.

A typical example of the aforementioned damage is presented in Fig. 30: The cupola itself remained undamaged. However, all piers of the drum failed (some of them in shear, others in out-of-plane bending). Due to the out-of-plane deformations of the piers, the supporting arches have failed. The horizontal component of the thrust of the arches has pushed the vertical elements apart, resulting to out-of-plane displacements of the walls exceeding 0.10 m (observe the deformed shape of the monument in Fig. 30).

It is, therefore, of major importance to take measures for the reduction of the effects of thrust, by applying the techniques described in the following.

Ties and Struts

This is a typical efficient measure to reduce the effect of thrust. The technique was applied in masonry structures either during construction or as a remedial measure during previous interventions (Fig. 31). The technique consists in providing at the base of the curved element horizontal ties that can sustain the horizontal component of the thrust. The ties (made of steel or stainless steel or titanium, depending on the case) are anchored to transverse elements. For the maximum acceptable stress in the ties, as well as for their anchorage to masonry, the rules presented in section "Improving the Connection Between Walls" are valid.

When the structure is subjected to seismic action, there is a cyclic (inwards/outwards) movement of the vertical supporting elements. Therefore, in addition to the ties, elements functioning in compression (struts) should also be provided. In old structures, quite frequently, timber elements were used, able to play the double role of strut and tie (due to their sufficient resistance in buckling). This solution can be applied in modern interventions as well. Alternatively, the solution shown in Fig. 32 (adopted for interventions to Dafni Monastery, Miltiadou et al. 2012) can be chosen: A timber compression element is provided. A hole is drilled at the core of the

S

Strengthening Techniques: Masonry and Heritage Structures, Fig. 28 Arches, vaults, and domes

timber element and a steel tie is accommodated. The tie is anchored on masonry.

Steel or FRP Ties in the Perimeter of the Curved Elements

This is a technique frequently applied at the base of cupolas, as well as at the base of their drum (Fig. 33). The ties (made either of steel or FRP) are designed to sustain the horizontal component of the thrust. It has to be noted, however, that they do not lead to any reduction in the flexibility of the system. Therefore, their application is useful but, quite often, not sufficient for the protection of the curved elements.

Strengthening Techniques: Masonry and Heritage Structures, Fig. 29 Partial collapse of the bridge after the passage of a heavy track
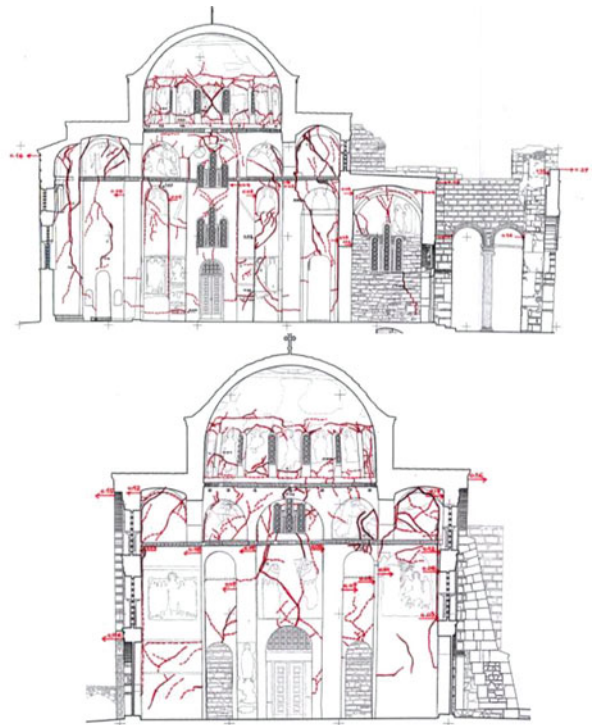
Strengthening Techniques: Masonry and Heritage Structures, Fig. 30 The main church of Dafni Monastery (Greece) damaged due to the 1999 earthquake

Buttresses

This is a technique applied also in the past (Fig. 34): The flexible vertical elements supporting a vault or a dome are stiffened by an external buttress. Thus, the out-of-plane deformations of the supporting elements are

Strengthening Techniques: Masonry and Heritage Structures, Fig. 31 Arches with timber ties

Strengthening Techniques: Masonry and Heritage Structures, Fig. 32 The timber strut/ steel tie solution was tested in-laboratory and adopted for application to the monument (Dafni Monastery)
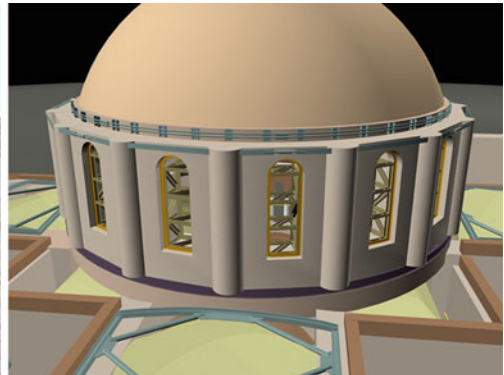
Strengthening Techniques: Masonry and Heritage Structures, Fig. 33 Dafni Monastery: existing steel tie at the base of the cupola, to be replaced by a stainless steel tie

Strengthening Techniques: Masonry and Heritage Structures, Fig. 34 (**a**) Ossios Lucas Monastery: external stone masonry buttresses were added against the southern facade. Due to architectural constraints, the buttresses do not reach the critical region of the system of arches that support the cupola. (**b**) Dafni Monastery: during the extensive interventions of the early twentieth century, a pair of stone masonry buttresses was constructed against the northern wall of the church. This technique was not repeated against the southern wall (where the courtyard is located)

Strengthening Techniques: Masonry and Heritage Structures, Fig. 35 Dafni Monastery: steel diaphragms to be constructed on top of the system of vaults; they will be covered by a timber pavement
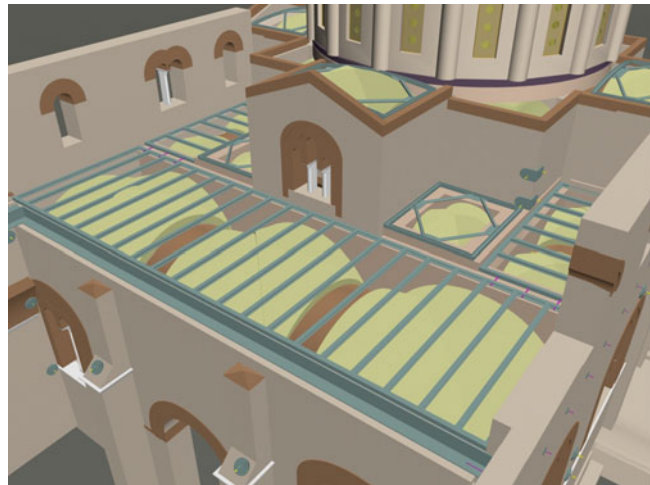
significantly reduced, and damage is avoided or significantly reduced.

It is obvious that in many cases, the construction of external buttresses, although a mechanically efficient solution, cannot be adopted either because there is no space for them to be accommodated or because the appearance of the structure is significantly altered.

Enhancing the Diaphragm Action of Vaults and Domes

This is an alternative solution applicable when diaphragms (metallic or wooden ones) can be arranged either under the curved elements or on top of them, when the curved elements are visible from the interior of the building. In this way, the effects of the thrust are minimized and both the curved elements and their supporting vertical elements are protected. This solution was adopted also in the case of Dafni Monastery (Fig. 35).

Modification of the Structural System

If the existing bearing system proves to be inadequate to resist seismic actions or if the new use of the structure requires higher performance (e.g., a residential building is given the use of a cultural
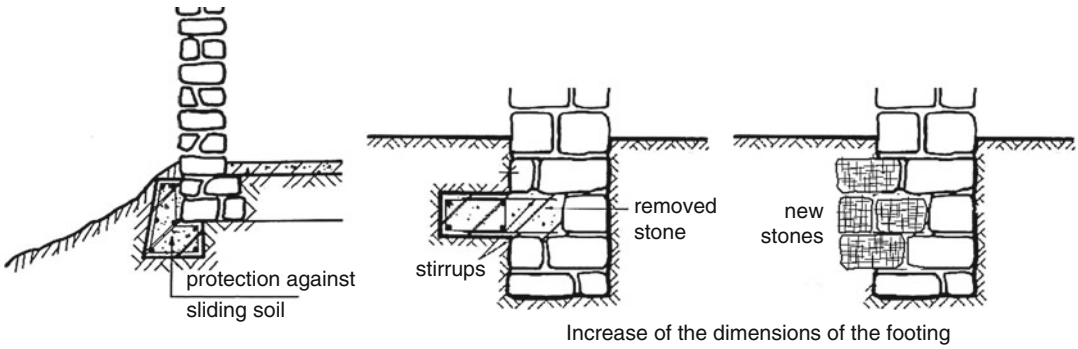
Strengthening Techniques: Masonry and Heritage Structures, Fig. 36 Interventions to foundation

Strengthening Techniques: Masonry and Heritage Structures, Fig. 37 Micro-piles
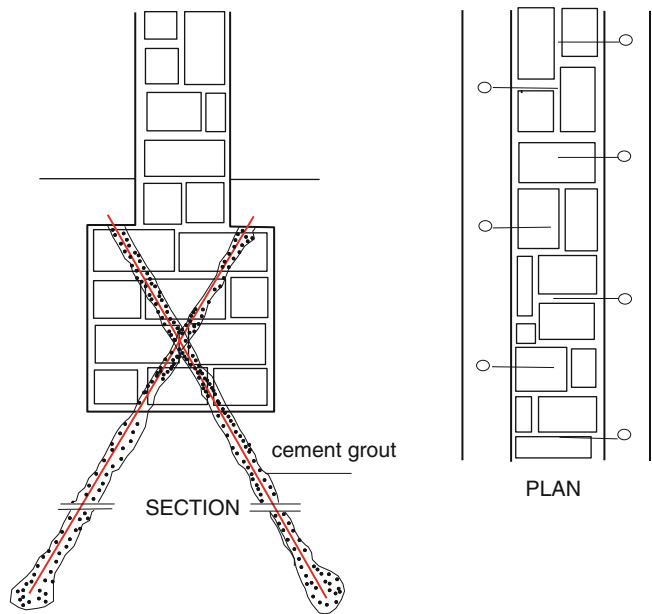
center), there may be a need for drastic interventions, such as construction of additional bearing elements, closing of openings with masonry, etc. In such cases, the construction of a new bearing system in the interior of the existing structure may be envisaged. When this alternative is selected, usually the new bearing system is made of steel. It should be noted, however, that for the new added system to assist the existing one in sustaining seismic actions, the former has to be connected with the original structural members. Moreover, the two systems have to be compatible from the deformation point of view; otherwise, the new added system may cause damage to the original one (due to pounding). Therefore, in the design of the new system, deformations (rather than forces) may be the governing parameter.

Strengthening of the Foundation

Interventions to the foundation are needed when the mass of the structure is significantly increased due to the interventions or due to the change of use (higher live loads or heavier pavements, etc.). There are also cases of interventions needed when there are damages in the building due to sliding or differential settlements. The interventions to the foundation consist usually in increasing the dimensions of foundation walls using

masonry or reinforced concrete (Fig. 36), depending on the case.

When the properties of the foundation soil are such that a deep foundation is needed, micropiles (Fig. 37) constitute a good solution. In the case of buildings in regions of archaeological interest, it should be made sure – before the application of micro-piles – that there are no ruins of older structures under the foundation of the building under examination.

Summary

Techniques that are commonly applied either to reinstate or to enhance the bearing capacity of masonry, and masonry components and masonry buildings are briefly presented and commented upon. This text comprises techniques aiming at (a) reinstating the mechanical properties of masonry (such as rejointing, grouting and stitching of cracks, etc.), (b) enhancing the mechanical properties of masonry (such as grouting of masonry, reinforced plaster), (c) improving the overall behavior of a masonry building (such as enhancement of the diaphragm action of floors and roof, improved connection between horizontal and vertical elements, etc.), and (d) improving the foundation conditions of the building. The techniques presented herein are, in principle, applicable to existing masonry buildings independently of their architectural, historical, and cultural values. There are, however, some cases – indicated in the text – where intervention techniques have to be applied with caution, when a heritage building is to be repaired or strengthened.

Cross-References

References

Binda L, Fontana A, Mirabella Roberti G (1994) Mechanical and stress distribution in multiple-leaf walls. In: Proceedings, 10th international brick block masonry conference, Calgary, Canada, pp 51–59

ElGawady MA (2005) Shear model for URM walls retrofitted with FRP, Annual NZSEE Conference, Wairakei, New Zealand (electronic source)

Faella C, Martinelli E, Nigro E, Paciello S (2004) Tuff masonry walls strengthened with a new kind of C-FRP sheet: experimental tests and analysis. In: 13th WCEE, Vancouver (Paper no, 923)

Grillo VE (2003) FRP/steel strengthening of unreinforced concrete masonry piers. MSc thesis. University of Florida. 222 pp

Magenes G, Penna A, Rota M, Galasco A, Senaldi I (2012) Shaking table teat of a full scale stone masonry building with stiffened floor and roof diaphragms. In: 15th WCEE, Lisbon

Mazzon N, Valluzzi MR, Aoki T, Garbin E, De Canio G, Ranieri N, Modena C (2009) Shaking table tests on two multi-leaf stone masonry buildings. In: 11th Canadian masonry symposium, Toronto

Miltiadou A, Delinicolas N, Vintzileou E, Mouzakis CH, Dourakopoulos J, Yannopoulos P (2012) Retrofitting of the main church of Dafni Monastery. Hellenic Ministry of Culture (in Greek)

Morbin A (2001) Strengthening of masonry elements with FRP composites. Diploma Dissertation. Dipartimento di Ingegneria Edile, Università di Padova, Italy, 190 pp

Mouzakis C, Vintzileou E, Adami C-E, Karapitta L (2012) Dynamic tests on three leaf stone masonry building without timber ties before and after interventions. SAHC, Wroclaw

Rashadul I (2008) Inventory of FRP strengthening methods in masonry structures. Master thesis, Technical University of Catalunya. 131 pp

Toumbakari E-E (2002) Lime-pozzolan-cement grouts and their structural effects on composite masonry walls. Doctor thesis, Katholieke Universiteit Leuven

Valluzzi M-R (2004) Consolidamento di murature in pietra. Iniezioni di calce idraulica naturale. Collana Scientifica REFICERE, GruppoEditoriale Faenza EditriceS.p.a., 128 pp. (in Italian)

Vintzileou E (2001) The effect of deep rejointing on the compressive strength of brick masonry. Masonry Int 15(1):8–12

S

Vintzileou E, Miltiadou-Fezans A (2008) Mechanical properties of three-leaf stone masonry grouted with ternary or hydraulic lime based grouts. Eng Struct 30(8):2265–2276

Vintzileou E, Tassios TP (1995) Three-leaf stone masonry strengthened by injecting cement-grouts. J Struct Eng Struct Division ASCE 121(5):848–856

Zaopo N (2011) Valutazione sperimentale dell'efficacia di interventi di miglioramento sismico di solai in legno riforzati nel piano, Diploma dissertation (supervisor MR Calluzzi), University of Padua, 212 pp. (in Italian)

Structural Design Codes of Australia and New Zealand: Seismic Actions

George P. Kouretzis, Mark J. Masia and Clive Allen
Faculty of Engineering and Built Environment, The University of Newcastle, Callaghan, NSW, Australia

Synonyms

Australia; Design; Earthquake; New Zealand; Seismic actions; Structural

Introduction

The design of structures to resist earthquake effects in Australia and New Zealand follows the pertinent Australian Standard AS 1170.4:2007 and New Zealand Standard NZS 1170.5:2004 provisions. Both Standards are based on the common Australian/New Zealand Standard AS/NZS 1170:2002 on structural design actions; and although they share elements such as the site subsoil classification system, they incorporate certain provisions such as the near-fault factor in NZS 1170.5:2004 to account for the different seismotectonic regime of the two countries. Indeed, Australia is an area of generally low seismicity, with the most catastrophic recent event being the 1989 Newcastle earthquake of magnitude $M = 5.6$, which resulted in 13 casualties and significant damages in the wider Newcastle area. On the other hand, New Zealand is located on the boundary between the Indo-Australian and Pacific Plate and suffers from frequent strong earthquakes. In fact, New Zealand is one of the first countries to account for seismic actions in a building standard; as early as 1935, and following the 1931 Hawke's Bay earthquake that claimed 256 lives, the NZS 95 provided seismic loads to be considered in design (McRae et al. 2011). Despite constant advances and amendments in the seismic standards over the years, the recent 22 February 2011 Christchurch earthquake of magnitude $M = 6.3$ resulted in 185 casualties, numerous injuries, and extensive damages in the central business district of Christchurch and its eastern suburbs.

The purpose of this entry is to provide an outline of the seismic design actions currently considered in the abovementioned Standards and a short discussion on the rationale behind their adaptation, in the light of other widely used modern seismic codes provisions, such as the Eurocode 8 EN-1998-1 and the American ASCE/SEI 7-10. It covers the determination of elastic and inelastic design response spectra to be used together with static and dynamic analysis methods, as well as scaling of strong motion recordings according to NZS 1170.5:2004 to perform time history analyses. Emphasis is put on the determination of the input rather than on the provisions about the performance of equivalent static, modal, and time-domain analyses, which can be found elsewhere, including the Standards themselves.

Elastic Response Spectra

The elastic response spectra for horizontal loading, $C(T)$, can be described by the following generic equation, compatible with both Standards:

$$C(T) = C_h(T) \cdot (Z \cdot R) \cdot N(T, D) \qquad (1)$$

where $C_h(T)$ is the spectral shape factor, $Z \cdot R$ is the product of the hazard factor Z times the factor R that depends on the annual probability of exceedance of the design earthquake and the
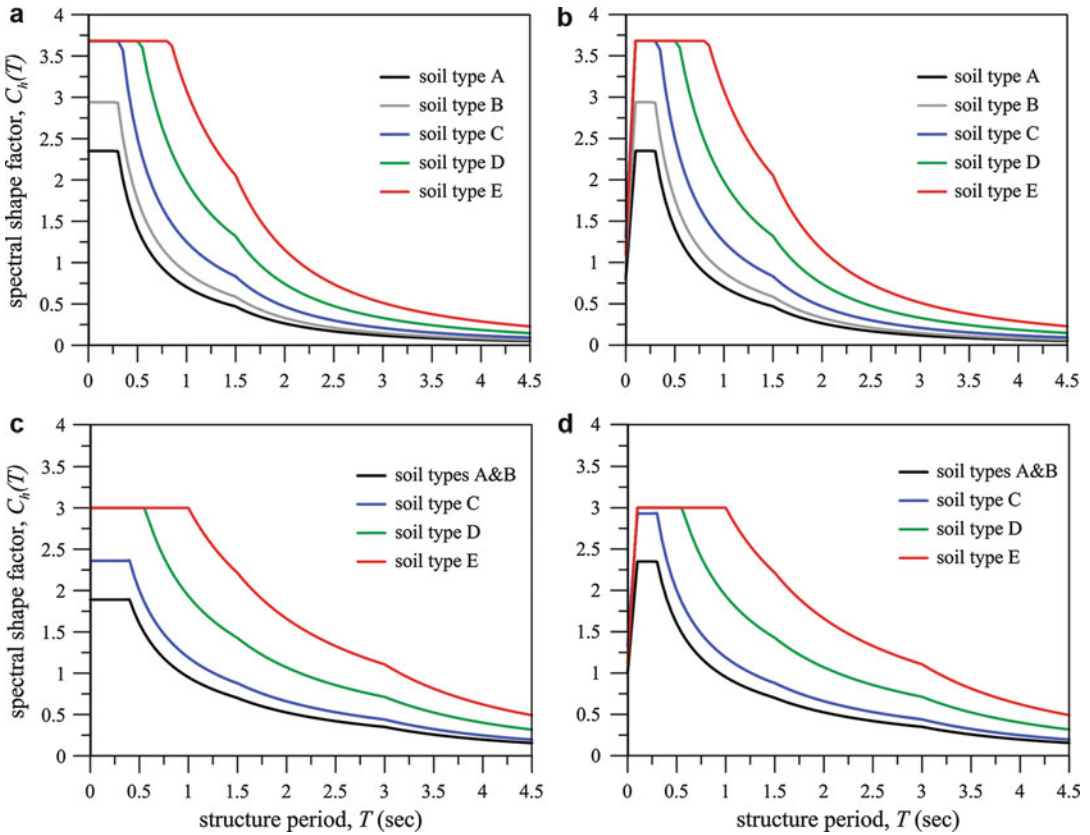
Structural Design Codes of Australia and New Zealand: Seismic Actions, Fig. 1 Spectral shape factor (5 % damping) plots for (**a**) AS 1170.4:2007/equivalent static analyses, (**b**) AS 1170.4:2007/dynamic analyses, (**c**) NZS 1170.5:2004/equivalent static analyses, (**d**) NZS 1170.5:2004/dynamic analyses

design limit state (ultimate or serviceability) under consideration, and $N(T, D)$ is the near-fault factor, applicable only to the NZS 1170.5:2004. The parameters comprising the elastic response spectra equation are presented in the following paragraphs.

Spectral Shape Factor for Different Subsoil Classes

The spectral shape factors $C_h(T)$ to be used for equivalent static analyses and for dynamic (modal and time history) analyses are plotted separately in Fig. 1, while the relevant functions are provided in Table 1 for the different site subsoil classes. These spectral shape factors correspond to 5 % structural damping, and a correction factor to convert them to other damping values is not provided as, e.g., in EN-1998-1. However, such

a provision would be of value only for special structures such as tanks containing liquids, which are not within the scope of the specific Australian and New Zealand Standards.

As mentioned in the Introduction, Australian and New Zealand Standards share a common site classification scheme, which is preferably based on the predominant site period (or low amplitude natural period) of the site T_s, estimated as:

$$T_s = 4H_s/V_s \qquad (2)$$

where H_s is the depth to the seismic bedrock of the site and V_s is the shear wave velocity of the overlying soil layer. In the absence of sufficient data to estimate the predominant site period, the site can be classified to one of the five subsoil classes (A to E) using (in the order of most preferred to least

Structural Design Codes of Australia and New Zealand: Seismic Actions, Table 1 Spectral shape factor (5 % damping) for different site subsoil classes

Site subsoil class	Equivalent static method			Dynamic analyses methods		
	Structure period, T (sec)	AS 1170.4:2007	NZS 1170.5:2004	Structure period, T (sec)	AS 1170.4:2007	NZS 1170.5:2004
A	$0 < T \leq 0.1$	2.35	1.89	$0 < T \leq 0.1$	$0.8 + 15.5\,T$	$1.0 + 1.35(T/0.1)$
	$0.1 < T < 0.3$	$0.704/T \leq 2.35$		$0.1 < T < 0.3$	$0.704/T \leq 2.35$	2.35
	$0.3 \leq T < 0.4$			$0.3 \leq T \leq 1.5$		$1.60(0.5/T)^{0.75}$
	$0.4 \leq T \leq 1.5$		$1.60(0.5/T)^{0.75}$			
	$1.5 < T \leq 3.0$	$1.056/T^2$	$1.05/T$	$1.5 < T \leq 3.0$	$1.056/T^2$	$1.05/T$
	$3 < T$		$3.15/T^2$	$3 < T$		$3.15/T^2$
B	$0 < T \leq 0.1$	2.94	1.89	$0 < T \leq 0.1$	$1.0 + 19.4\,T$	$1.0 + 1.35(T/0.1)$
	$0.1 < T < 0.3$	$0.88/T \leq 2.94$		$0.1 < T < 0.3$	$0.88/T \leq 2.94$	2.35
	$0.3 \leq T < 0.4$					
	$0.4 \leq T \leq 1.5$		$1.60(0.5/T)^{0.75}$	$0.3 \leq T \leq 1.5$		$1.60(0.5/T)^{0.75}$
	$1.5 < T \leq 3.0$	$1.32/T^2$	$1.05/T$	$1.5 < T \leq 3.0$	$1.32/T^2$	$1.05/T$
	$3 < T$		$3.15/T^2$	$3 < T$		$3.15/T^2$
C	$0 < T \leq 0.1$	3.68	2.36	$0 < T \leq 0.1$	$1.3 + 23.8\,T$	$1.33 + 1.60(T/0.1)$
	$0.1 < T < 0.3$	$1.25/T \leq 3.68$		$0.1 < T < 0.3$	$1.25/T \leq 3.68$	2.93
	$0.3 \leq T < 0.4$					
	$0.4 \leq T \leq 1.5$		$2.0(0.5/T)^{0.75}$	$0.3 \leq T \leq 1.5$		$2.0(0.5/T)^{0.75}$
	$1.5 < T \leq 3.0$	$1.874/T^2$	$1.32/T$	$1.5 < T \leq 3.0$	$1.874/T^2$	$1.32/T$
	$3 < T$		$3.96/T^2$	$3 < T$		$3.96/T^2$
D	$0 < T \leq 0.1$	3.68	3.0	$0 < T \leq 0.1$	$1.1 + 25.8\,T$	$1.12 + 1.88(T/0.1)$
	$0.1 < T < 0.3$	$1.98/T \leq 3.68$		$0.1 < T < 0.3$	$1.98/T \leq 3.68$	3.0
	$0.3 \leq T < 0.56$			$0.3 \leq T < 0.56$		
	$0.56 \leq T \leq 1.5$		$2.4(0.75/T)^{0.75}$	$0.56 \leq T \leq 1.5$		$2.4(0.75/T)^{0.75}$
	$1.5 < T \leq 3.0$	$2.97/T^2$	$2.14/T$	$1.5 < T \leq 3.0$	$2.97/T^2$	$2.14/T$
	$3 < T$		$6.42/T^2$	$3 < T$		$6.42/T^2$
E	$0 < T \leq 0.1$	3.68	3.0	$0 < T \leq 0.1$	$1.1 + 25.8\,T$	$1.12 + 1.88(T/0.1)$
	$0.1 < T < 0.3$	$3.08/T \leq 3.68$		$0.1 < T < 0.3$	$3.08/T \leq 3.68$	3.0
	$0.3 \leq T < 1.0$			$0.3 \leq T < 1.0$		
	$1.0 \leq T \leq 1.5$		$3.0/T^{0.75}$	$1.0 \leq T \leq 1.5$		$3.0/T^{0.75}$
	$1.5 < T \leq 3.0$	$4.62/T^2$	$3.32/T$	$1.5 < T \leq 3.0$	$4.62/T^2$	$3.32/T$
	$3 < T$		$9.96/T^2$	$3 < T$		$9.96/T^2$

preferred): borehole log data together with in situ and laboratory tests, seismological methods including strong motion recordings (Nakamura 1989; Bouckovalas et al. 2002), qualitative borehole log descriptions, and geological information for the site. For layered sites, the site period can be estimated via an averaging method described in AS 1170.4:2007 and NZS 1170.5:2004.

The five different subsoil classes are defined in the Standards as follows (AS 1170.4:2007, NZS 1170.5:2004):

Class A: Strong rock. Strong to extremely strong rock, with:

(a) Unconfined compressive strength greater than 50 MPa, and

Structural Design Codes of Australia and New Zealand: Seismic Actions, Table 2 Maximum depth limits for site subsoil class C (After AS 1170.4:2007, NZS 1170.5:2004)

Soil type and description		Representative undrained shear strength, S_u (KPa)	Representative SPT N-values	Maximum depth of soil (m)
Cohesive soils	Very soft	<12.5	–	0
	Soft	12.5–25	–	20
	Firm	25–50	–	25
	Stiff	50–100	–	40
	Very stiff to hard	100–200	–	60
Cohesionless soils	Very loose	–	<6	0
	Loose dry	–	6–10	40
	Medium dense	–	10–30	45
	Dense	–	30–50	55
	Very dense	–	>50	60
	Gravels	–	>50	100

(b) An average shear wave velocity over the top 30 m $V_{s,30} > 1500$ m/s, and

(c) Not underlain by materials having a compressive strength less than 18 MPa or a shear wave velocity less than 600 m/s

Class B: Rock. Rock, with:

(a) Unconfined compressive strength between 1 MPa and 50 MPa, and

(b) An average shear wave velocity over the top 30 m $V_{s,30} > 360$ m/s, and

(c) Not underlain by materials having a compressive strength less than 0.8 MPa or a shear wave velocity less than 300 m/s

A surface layer of no more than 3 m depth of highly weathered or completely weathered rock or soil material may be present in a class B site. It should be mentioned here that, based on the average shear wave velocity over the top 30 m criterion, other seismic codes such as EN-1998-1 or ASCE/SEI 7-10 would classify a site with $V_{s,30} > 360$ m/s to dense/stiff soil deposits, rather than rock (class B according to EN-1998-1, which corresponds to $360 < V_{s,30} < 800$ m/s, and class C according to ASCE/SEI 7-10, which corresponds to $366 < V_{s,30} < 762$ m/s). This suggests that class B in AS 1170.4:2007 and NZS 1170.5:2004 is rather broad, covering a range of subsoil conditions that may not exhibit similar behavior during an earthquake rich in high- to medium-frequency content (Bouckovalas and kouretzis 2001).

Class C: Shallow soil sites. Sites that are not classified as class A, class B, or class E:

(a) The predominant site period estimated as above is $T_s \leq 0.6$ s, or

(b) Soil depth does not exceed the maximum values listed in Table 2.

Class D: Deep or soft soil sites. Sites that are not classified as class A, class B, or class E; and:

(a) The predominant site period estimated as above is $T_s > 0.6$ s, or

(b) Soil depth exceeds the maximum values listed in Table 2, or

(c) Sites that are underlain by less than 10 m of cohesive soil with undrained shear strength $S_u < 12.5$ KPa or cohesionless soil soils with SPT values $N < 6$.

Class E: Very soft soil sites:

(a) More than 10 m of very soft cohesive soil with undrained shear strength $S_u < 12.5$ KPa, or

(b) More than 10 m of very loose cohesionless soils with SPT values $N < 6$, or

(c) More than 10 m depth of soft-loose soils with shear wave velocity values $V_s < 150$ m/s, or

(d) More than 10 m combined depth of soils with properties described in (a), (b), and (c) above.

A comparison between the spectral shape factors of AS 1170.4:2007 and NZS 1170.5:2004 is
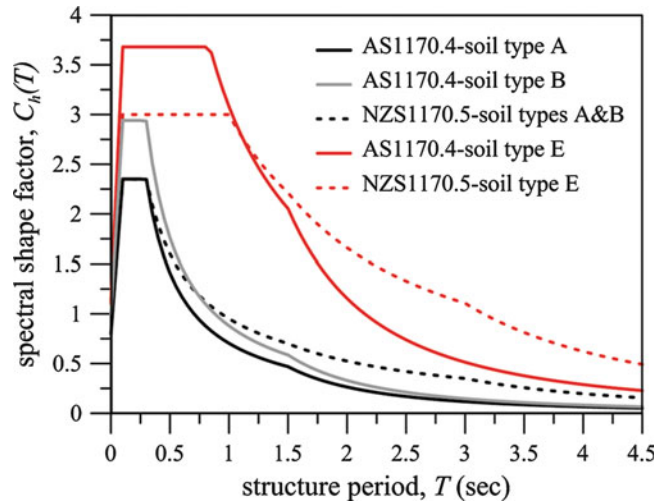
attempted in Fig. 2, for the two extreme cases of rock (A and) and soft soil (E) sites. The New Zealand Standard, which generally applies to earthquakes with higher magnitudes ($M \geq 6.5$), features lower spectral factor values in the low structural period range ($T < 1$ s) and higher spectral factor values in the high period range ($T > 1$ s). This is compatible with the fact that soil nonlinearity effects alter the spectral content of stronger earthquakes with magnitude $M > 5.5$, which in addition have a richer low-frequency content attributed also to near-fault effects.

Hazard Factor and Annual Probability of Exceedance

The hazard factor Z of a particular location corresponds to the peak ground acceleration (in g's) for site class B (AS 1170.4:2007) or classes A and B (NZS 1170.5:2004), considering a design earthquake with return period of 500 years, i.e., a 10 % probability of exceedance during a 50-year design life. The hazard factors in the Australian Standard have not been updated since its previous version AS 1170.4:1993 and are based on the work of Gaull et al. (1990), who used data of the Australian Geological Survey Organisation dated back to 1856. All areas in Australia are assumed to be seismically active for design purposes, with the hazard factor generally ranging between $Z = 0.05$ and $Z = 0.13$, with the exception of the Meckering region in

Western Australia, where the hazard factor ranges between $Z = 0.14$ and $Z = 0.22$, and the Macquarie Island, located halfway between New Zealand and Antarctica, where the hazard factor is equal to $Z = 0.60$.

In NZS 1170.5:2004, the hazard factor Z corresponds to 0.5 times the spectral acceleration (5 % damping) for a structure period $T = 0.5$ s and soil subclass C, considering a design earthquake with a return period of 500 years (Fig. 1c, d). The minimum value across New Zealand is $Z = 0.13$, to ensure no collapse of structures even in areas of low seismicity, i.e., every structure is designed to survive the 84-percentile strong motion of a magnitude $M = 6.5$ normal-faulting earthquake at source-to-site distance of 20 km, a seismic scenario corresponding to the low-seismicity areas of New Zealand. The hazard factor at the high-seismicity major cities (e.g., Wellington, Napier, Hastings) is of the order of $Z = 0.40$, with the maximum value considered in the Standard being $Z = 0.60$.

The probability factor (AS 1170.4:2007) or return period factor (NZS 1170.5:2004) R, listed in Table 3, is used to scale the hazard factor to the required annual probability of exceedance, for the design limit state under consideration. In AS 1170.4:2007, the annual probability of exceedance is correlated with the design working life and the importance level of the structure, with the

Structural Design Codes of Australia and New Zealand: Seismic Actions, Table 3 Probability factor (AS 1170.4:2007) or return period factor (NZS 1170.5:2004)

Annual probability of exceedance	R
1/2,500	1.8
1/2,000	1.7
1/1,500	1.5
1/1,000	1.3
1/800	1.25
1/500	1.0
1/250	0.75
1/200	0.70
1/100	0.50
1/50	0.35
1/25	0.25
1/20	0.20

latter provided in AS/NZS 1170.0:2002 (Appendix F). The design annual probability of exceedance applies to the ultimate limit state only: According to AS 1170.4:2007, structures of importance levels 1–3 (i.e., all structures except ones carrying critical post-disaster functions and exceptional structures) de facto satisfy the serviceability limit state requirements when designed in accordance to the Standard. A special study needs to be carried out for critical post-disaster structures only (importance level 4), to ensure that they remain operational after a seismic event equivalent to the ultimate limit state design event for ordinary importance level 2 structures. The design annual probability of exceedance for structures and buildings (importance level 2) with a working life of 50 years is 1/500 and corresponds to a 10 % probability of exceedance during the design life of the structure. For major structures affecting crowds and associated with "very great" economic, social, and environmental consequences of failure (importance level 3), the annual probability of exceedance is reduced to 1/1,000 for structures with a design working life of 50 years and to 1/2,500 for structures with a design working life of 100 years.

The required annual probability of exceedance for structures designed according to NZS 1170.5:2004 is provided again in AS/NZS

1170.0:2002 (Table 3.3 of AS/NZS 1170.0:2002); however, in the New Zealand Standard, it is correlated with the design (serviceability or ultimate) limit state: The annual probability of exceedance for the common serviceability limit state SLS1 (requirement for no repairs on structural and nonstructural components) is 1/25 for all structures. The special serviceability limit state SLS2 applies to structures carrying critical post-disaster functions only (importance level 4) and is associated with an annual probability of exceedance of 1/500 for design working life of 50 years. A special hazard study is required to define the SLS2 actions for structures of importance level 4 with design working life of 100 years or more.

As far as the ultimate limit state is concerned, the design annual probability of exceedance is the same as in AS 1170.4:2007, for structures classified to importance levels 2 and 3. Note that the product $Z \cdot R$ in NZS 1170.5:2004 need not exceed $Z \cdot R = 0.70$ for ultimate limit state analyses but must be higher than $Z \cdot R = 0.20$ when an annual probability of exceedance 1/2,500 is considered. These $Z \cdot R$ bounds correspond to the higher and lower seismicity regions of New Zealand, respectively: The upper bound matches the 84-percentile near-fault strong motion due to a $M = 8.1$ event from the major Alpine Fault that runs along the South Island, divided by a margin of safety equal to 1.5 likely to result from applying the code design provisions.

Near-Fault Factor

The near-fault factor $N(T,D)$ is introduced in NZS 1170.5:2004 to account for near-source effects on the strong motion during the activation of a strike-slip fault: a) forward directivity, resulting in high peak velocities and displacements, and b) polarization of the long-period motions in the near-source region, where medium- to long-period pulses tend to be stronger in the direction perpendicular to the strike of the fault. Near-fault effects in NZS 1170.5:2004 are considered for eleven (11) major strike-slip faults classified as "class A" faults according to the Californian Category A fault criteria (Petersen et al. 2000), capable of producing earthquakes with magnitude $M = 7.0$

S

Structural Design Codes of Australia and New Zealand: Seismic Actions, Table 4 Maximum values of the near-fault factor $N_{max}(T)$ for different structure periods

NZS 1170.5:2004		Linear interpolation functions	
Structure period, T (sec)	$N_{max}(T)$	Structure period, T (sec)	$N_{max}(T)$
≤ 1.5	1.0	≤ 1.5	1.0
2	1.12	$1.5 < T < 4$	0.24 T + 0.64
3	1.36		
4	1.60	$4 \leq T < 5$	0.12 T + 1.12
≥ 5	1.72	≥ 5	1.72

or greater and having slip rates of 5 mm/year of greater. The near-fault factor $N(T,D)$ must be considered during the derivation of the medium- to long-period band of the elastic response spectra for sites located within a distance D less than 20 km from the traces of the major strike-slip faults listed in NZS 1170.5:2004 and applies to annual probability of exceedance less that 1/250:

$$N(T,D) = N_{max}(T) \qquad \text{for} \quad D \leq 2\,\text{km} \quad (3.1)$$

$$N(T,D) = 1 + (N_{max}(T)-1) \cdot (20-D)/18$$
$$\text{for} \quad 2\,\text{km} < D \leq 20\,\text{km}$$
$$(3.2)$$

$$N(T,D) = 1.0 \qquad \text{for} \quad D > 20\,\text{km} \quad (3.3)$$

where the factor $N_{max}(T)$ is provided as a function of the structure period T in Table 4 and Fig. 3a. Linear interpolation functions are provided for intermediate period values, as proposed in NZS 1170.5:2004. In addition, a comparison of the effect of near-fault correction (Eqs. 3.1, 3.2, and 3.3) on the spectral shape factors for dynamic analyses is illustrated in Fig. 3b, for different distances from the fault.

Vertical Design Actions

According to AS 1170.4:2007, vertical earthquake actions generally need not to be considered in design. For the design of mechanical and electrical components, the vertical earthquake forces are taken equal to 50 % of the horizontal forces.

However, a provision is included stating that when the dynamic analysis requires the consideration of vertical earthquake forces, both upward and downward motions need to be considered and the vertical spectral shape factor shall be taken equal to

$$C_v(T) = 0.5 \cdot C_h(T_v) \quad (\text{AS } 1170.4 : 2007) \quad (4)$$

where T_v is the vertical period of vibration. In NZS 1170.5:2007 on the other hand, when vertical design actions need to be considered in the analyses (e.g., time history analyses), the elastic response spectrum for vertical loading is taken equal to 70 % of the corresponding horizontal spectrum, as

$$C_v(T) = 0.7 \cdot C_h(T_v) \quad (\text{NZS } 1170.5 : 2004)$$
$$(5)$$

where $T_v = 0$ for the design of the structure as a whole, or equal to the vertical period of the element under consideration, for the design of parts and components. Note that the vertical component of strong earthquake motion is usually rich in high-frequency content; thus, peak ground acceleration and spectral values in the near-source region may exceed the horizontal values (NZS 1170.5 Supp 1:2004, Niazi and Bozorgnia 1992; Ambrasseys and Simpson 1996). In light of this, NZS 1170.5 Supp 1:2004 recommends that when the distance of the site from the fault trace is less than $D = 10$ km, the vertical design actions be considered equal to the horizontal design actions for structure period $T \leq 0.30$ s. Strong motion recordings from recent earthquakes, such as the Christchurch 2011 earthquake, provided further evidence in support of the above, as will be discussed later in this entry.

Structural Ductility and Structural Performance Factor-Inelastic Design Spectrum

Derivation of the (inelastic) design spectrum requires the correction of the elastic response spectrum to account for the ability of the
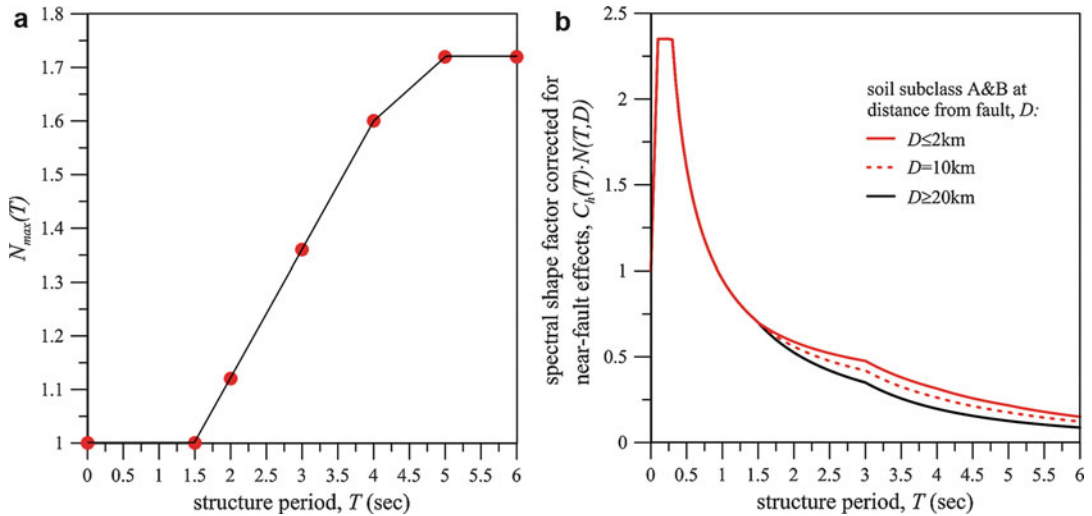
Structural Design Codes of Australia and New Zealand: Seismic Actions, Fig. 3 (**a**) Variation of the near-fault factor $N_{\max}(T)$ with the structure period, T, and (**b**) comparison of spectral shape factors corrected for near-fault effects for a rock site (soil subclass A and B) at a distance $D \leq 2$ km, $D = 10$ km, and $D > 20$ km from a major active fault (5 % damping)

particular structure to dissipate seismic energy via nonlinear response. Apart from the structural ductility factor, μ, which is related to the level of inelastic demand that can be reliably sustained by the structure, AS 1170.4:2007 and NZS 1170.5:2004 introduce the structural performance factor S_p to reduce the elastic design seismic actions. The structural performance factor is employed to quantify a number of effects that are not explicitly taken into account in state-of-practice structural analysis procedures by simply scaling the design loads. Such effects include (NZS 1170.5 Supp 1:2004):

(a) Excitation effects: The estimated seismic loads correspond to a peak ground acceleration value, which may be reached only during a single loading cycle and therefore is unlikely to lead to significant damage ("effective" ground acceleration concept).
(b) Individual structural elements typically feature a higher capacity than modeled during the analysis of the structure, due to higher material strength, strain hardening, strain rate effects, etc.
(c) The total structural capacity is typically higher than predicted, due to redundancy

effects or the contribution of nonstructural elements (e.g., in fill walls) which is not taken directly into account in typical analysis models.
(d) The energy dissipation of the structure is typically higher than assumed, due to damping introduced from nonstructural elements and soil-foundation interaction effects.

The ductility factor μ for the ultimate limit state is calculated in accordance with the appropriate material standard, where such data are provided. Contrary to NZS 1170.5:2004, where reference is made to material standards only and a special study is required otherwise, maximum ductility factors for typical structural systems are provided directly in AS 1170.4:2007 (Table 5). Alternatively, for a specific structure, μ and S_p can be determined via a nonlinear static pushover analysis. Note also that AS 1170.4:2007 applies only to structures with ductility factor $\mu \leq 3$, and when a higher ductility factor is considered, the design must be performed in compliance with NZS 1170.5:2004 provisions (Table 5).

While NZS 1170.5:2004 does not refer explicitly to ductility factor values for the ultimate limit state, for serviceability limit state analyses of

Structural Design Codes of Australia and New Zealand: Seismic Actions, Table 5 Ductility factor μ and structural performance factor S_p for different structural systems and specific structure types (After AS 1170.4:2007)

Structural system	Description	μ	S_p
Steel structures	Special moment-resisting frames (fully ductile)[a]	4	0.67
	Intermediate moment-resisting frames (moderately ductile)	3	0.67
	Ordinary moment-resisting frames (limited ductile)	2	0.77
	Moderately ductile concentrically braced frames	3	0.67
	Limited ductile concentrically braced frames	2	0.77
	Fully ductile concentrically braced frames[a]	4	0.67
	Other steel structures not defined above	2	0.77
Concrete structures	Special moment-resisting frames (fully ductile)[a]	4	0.67
	Intermediate moment-resisting frames (moderately ductile)	3	0.67
	Ordinary moment-resisting frames	2	0.77
	Ductile coupled walls (fully ductile)[a]	4	0.67
	Ductile partially coupled walls[a]	4	0.67
	Ductile shear walls	3	0.67
	Limited ductile shear walls	2	0.77
	Ordinary moment-resisting frames in a combination with limited ductile shear walls	2	0.77
	Other concrete structures not listed above	2	0.77
Timber structures	Shear walls	3	0.67
	Braced frames (with ductile connections)	2	0.77
	Moment-resisting frames	2	0.77
	Other wood- or gypsum-based seismic-force-resisting systems not listed above	2	0.77
Masonry structures	Close-spaced-reinforced masonry[b]	2	0.77
	Wide-spaced-reinforced masonry[b]	1.5	0.77
	Unreinforced masonry[b]	1.25	0.77
	Other masonry structures not complying with AS 3700:2001	1	0.77
Specific structure types	Tanks, vessels, or pressurized spheres on braced on unbraced legs	2	1.0
	Cast-in-place concrete silos and chimneys having walls continuous to the foundation	3	1.0
	Distributed mass cantilever structures, such as stacks, chimneys, silos, and skirt-supported vertical vessels	3	1.0
	Trussed towers (freestanding or guyed), guyed stacks, and chimneys	3	1.0
	Inverted pendulum-type structures	2	1.0
	Cooling towers	3	1.0
	Bins and hoppers on braced or unbraced walls	3	1.0
	Storage racking	3	1.0
	Signs and billboards	3	1.0
	Amusement structures and monuments	2	1.0
	All other self-supporting structures not otherwise covered	3	1.0

[a]The design of structures with $\mu > 3$ should be in accordance to NZS 1170.5:2004
[b]Values from AS 3700:2001

common structures (SLS1), it is provisioned that $1.0 \leq \mu \leq 1.25$, and for critical post-disaster structures (SLS2), it is provisioned that $1.0 \leq \mu \leq 2.0$. Furthermore, the ductility factor for vertical actions is always considered to be $\mu = 1.0$.

Note that in NZS 1170.5:2004, the inelastic horizontal design spectra are not derived by directly multiplying the elastic spectra by $(1/\mu)$. Instead, to account for the transition between equal displacement theory (which is valid for longer structure periods) and equal energy theory

(valid for shorter structure periods), a transition point is defined at $T_1 = 0.7$ s for soil subclasses A to D and at $T_1 = 1.0$ s for soil subclass E, where T_1 is the largest translation period of vibration along the direction under consideration. So, the factor to derive the inelastic response spectra is estimated as:

For subsoil classes A, B, C, and D:

$$k_\mu = \mu \qquad \text{for} \quad T_1 \geq 0.7\,\text{s} \qquad (6.1)$$

$$k_\mu = 1 + (\mu - 1)T_1/0.7 \qquad \text{for} \quad T_1 < 0.7\,\text{s} \tag{6.2}$$

For subsoil class E:

$$k_\mu = \mu \qquad \text{for} \quad T_1 \geq 1.0\,\text{s or } \mu < 1.5 \quad (6.3)$$

$$k_\mu = 1.5 + (\mu - 1.5)T_1 \quad \text{for } T_1 < 1.0\text{s and } \mu \geq 1.5 \tag{6.4}$$

For the purpose of calculating k_μ, the largest translation period of vibration T_1 is not taken less than 0.4 s.

The structural performance factor S_p in AS 1170.4:2007 is provided in Table 5. In NZS 1170.5:2004, the structural performance factor is taken equal to $S_p = 0.7$ for ultimate limit state analyses, except for low-ductility structures with $1.0 < \mu < 2.0$, where S_p is calculated as:

$$S_p = 1.3 - 0.3\mu \tag{7}$$

For ultimate limit state verification of structures against sliding or overturning, S_p is taken equal to $S_p = 1.0$. For serviceability limit state SLS1 and SLS2 analyses, S_p is taken equal to $S_p = 0.7$, except otherwise dictated by the relevant material standard.

In accordance to the above, the inelastic design response spectrum in AS 1170.4:2007 is defined as

$$C_d(T) = \frac{C(T)S_p}{\mu} \tag{8}$$

and in NZS 1170.5:2004 as

$$C_d(T) = \frac{C(T)S_p}{k_\mu} \text{ (for horizontal actions)} \quad (9.1)$$

$$C_d(T) = C_v(T)S_p \text{ (for vertical actions)} \quad (9.2)$$

Actions for Dynamic Time History Analyses

In addition to actions applicable to common equivalent static and modal dynamic analyses, AS 1170.4:2007 and, mainly, NZS 1170.5:2004 include certain provisions for the determination of the appropriate design actions when dynamic time-domain analyses are to be performed. While in AS 1170.4:2007 there is a vague requirement that the response spectra of the actual acceleration time histories used shall "approximate" the design spectrum of Eq. 8, NZS 1170.5:2004 prescribes a more elaborate and explicit procedure for the selection and scaling of ground motion records.

Ground motion accelerographs to be used for time-domain analyses according to NZS 1170.5:2004 shall consist of both horizontal components of the recording; the vertical component may have to be considered too, for the analysis of structures sensitive to vertical strong ground motion. The above imply that the procedure refers to three-dimensional analyses, yet an adaption to two-dimensional models is feasible, by considering only one horizontal component of each ground motion record as, e.g., in ASCE/SEI 7-10. A "family" of no less than three records must be employed in each time history analysis.

The selected acceleration records must be consistent with the local seismotectonic regime (expected magnitude, fault rupture mechanism, source-to-site distance, etc.). When the site is located near a major fault, i.e., $N(T, D) > 1$ (Eqs. 3.1, 3.2, and 3.3), then one set of the records must have a forward directivity component. Simulated, artificial accelerographs may be used too, when appropriate real acceleration recordings are not available. Each record shall be scaled by a "record scale factor" k_1 to
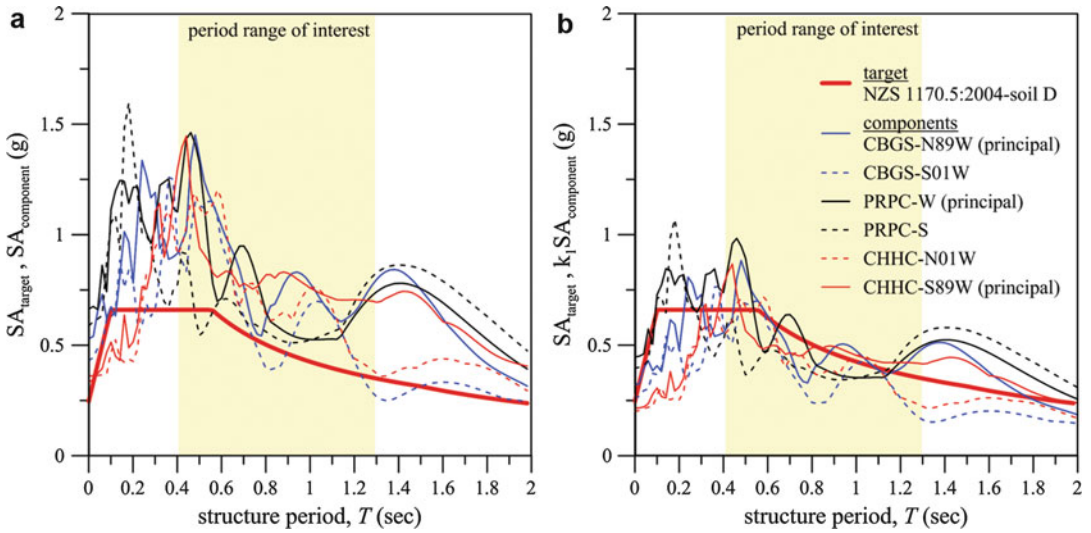
Structural Design Codes of Australia and New Zealand: Seismic Actions, Fig. 4 (**a**) Comparison of the 2011 Christchurch strong motion recordings elastic response spectra (5 % damping, horizontal components) with the design NZS 1170.5:2004 elastic design spectra for the city of Christchurch and soil subclass D, (**b**) scaled elastic response spectra of the same strong motion recordings to match the NZS 1170.5:2004 design spectra for largest structure translation period $T_1 = 1$ s and $S_p = 1$

match the target design spectrum over the period range of interest and by a "family scale factor" k_2 to ensure that the energy content of at least one record in the family exceeds that of the design spectrum over the period range of interest. The target design spectrum is defined as

$$SA_{\text{target}} = \left(\frac{1 + S_p}{2}\right) C(T) \qquad (10)$$

where the spectral shape factor $C(T)$ is calculated from Eq. 1 and S_p is the structural performance factor defined in the section "Structural Ductility and Structural Performance Factor-Inelastic Design Spectrum."

In order to present the procedure of scaling ground motion records according to NZS 1170.5:2004, the horizontal components of three typical strong ground motion recordings from the 2011 Christchurch $M = 6.3$ earthquake were used. These records were scaled to match the target ultimate limit state spectrum of NZS 1170.5:2004 for a common structure (importance level 2) located in Christchurch, founded on subsoil class D (soft soil). According to NZS 1170.5:2004, for a structure of importance level 2 in Christchurch with a design life of 50 years, $Z \cdot R = 0.22$ g and $N(T,D) = 1$, i.e., no near-fault effects, need to be taken into account due to possible rupture of the Port Hills Fault that is believed to have caused the 2011 earthquake. Assuming for simplicity that $S_p = 1$, the target spectrum is plotted in Fig. 4.

Strong motion recordings during the Christchurch 2011 earthquake exceeded by far the design peak ground acceleration at the ground surface of class D soft soils implied by NZS 1170.5:2004 (0.246 g) and reached maximum values of 1.67 g in the horizontal direction and 2.20 g in the vertical direction at the Heathcote Valley Primary School Station (HVSC) (www. geonet.org.nz). Here a family of three near-field recordings on soft to very soft soil are used, with peak horizontal ground accelerations ranging from 0.34 to 0.67 g (Table 6).

The elastic response spectra of the horizontal components of these three recordings are plotted in Fig. 4a, in comparison with the design response spectra of NZS 1170.5:2004. A band-pass filter in the frequency ranges of 0.10–0.25 Hz and 24.5–25.5 Hz was applied to the time histories

Structural Design Codes of Australia and New Zealand: Seismic Actions, Table 6 Characteristics of the typical strong motion recordings from the Christchurch 22 February 2011 $M = 6.3$ earthquake (Source: www.geonet.org.nz)

Site code	Subsoil class acc. to NZS 1170.5:2004	Epicentral distance	PGA-horiz.1 (g)	PGA-horiz.2 (g)	PGA-vert. (g)
CBGS	D	7	0.553	0.452	0.360
PRPC	E	6	0.669	0.595	1.88
CHHC	D	6	0.345	0.364	0.601

used to derive the spectra. It is clear that the spectral values of these typical records are above the design spectra across practically the whole range of important structural periods, an indication of the severity of the Christchurch earthquake.

Scaling of ground motion records to match the target design spectrum via the scale factor k_1 was performed while taking into account a structure period range that depends on the largest translational period of the structure in the direction of interest, T_1. The period range of interest is defined as $0.4\,T_1 \leq T \leq 1.3\,T_1$, with the product $0.4\,T_1$ not taken less than 0.4 s. Notice that this band of significant periods is considerably narrower compared to the corresponding one prescribed in EN-1998-1, which is $0.2\,T_1 \leq T \leq 2.0\,T_1$, or even ASCE/SEI 7-10, where the corresponding range is $0.2\,T_1 \leq T \leq 1.5\,T_1$. For the case at hand, it was assumed that $T_1 = 1$ s, and the period range of interest is denoted by the yellow band in Fig. 4.

The best-fit scale factor k_1 of each horizontal ground motion component was determined as the value minimizing the function $\log(k_1 SA_{\text{component}}/SA_{\text{target}})$ over the period range of interest, in a least mean square sense. In other words, the aim was to find via an interactive procedure the k_1 value that minimizes the sum:

$$\int_{0.4T_1}^{1.3T_1} \left[\log\left(\frac{k_1 SA_{\text{component}}}{SA_{\text{target}}} \right) \right]^2 dT = \min \quad (11)$$

The selected recordings must satisfy a "similarity" criterion $0.33 < k_1 < 3.0$. Moreover, in order to verify that each selected record reasonably matches the target spectrum over the period range of interest, the following inequality quantifying the mean (over the period of interest) square difference of the scaled over the target spectral values must be satisfied:

$$D_1 = \sqrt{\frac{1}{(1.3 - 0.4)T_1} \cdot \int_{0.4T_1}^{1.3T_1} \left[\log\left(\frac{k_1 SA_{\text{component}}}{SA_{\text{target}}} \right) \right]^2 dT} \leq \log(1.5) \quad (12)$$

Although not a normative criterion, a stricter fit $D_1 \leq \log(1.3)$ is proposed for most cases, according to NZS 1170.5 Supp 1:2004.

The k_1 and D_1 factors for the records at hand are listed in Table 7, where it is depicted that the specific records are compatible with NZS 1170.5:2004 requirements. The component of each record with the lower k_1 value was nominated as "principal" (Fig. 4, Table 7), and this value of k_1 was used as the record scale factor.

The record family scale factor, k_2, ensures that the principal component of at least one record spectrum (after being scaled by k_1) exceeds the target spectrum over the period range of interest. It is estimated as the maximum value of the ratio $SA_{\text{target}}/\max(SA_{\text{principal}}) \geq 1.0$ within the period range of interest, where $\max(SA_{\text{principal}})$ is the maximum spectral value between all the principal components of the family, at each period step considered for the derivation of the spectra. To confirm the selection of the principal and

S

Structural Design Codes of Australia and New Zealand: Seismic Actions, Table 7 Scaling factors for the considered typical strong motion records

Site code	Component	k_1	D_1	k_2
CBGS	**N89W (principal)**	**0.608**	0.070	1.05
	S01W	0.789	0.105	
PRPC	**W (principal)**	**0.672**	0.081	
	S	0.785	0.097	
CHHC	N01W	0.683	0.064	
	S89W (principal)	**0.598**	0.058	

secondary components, the record family scale factor must be within the range $1.0 < k_2 \le 1.3$. If $k_2 > 1.3$, then either a different record must be selected or the selection of the principal/secondary components may be reversed, aiming to minimize the product $k_1 k_2$. This may be the case when all three principal components in a family feature low spectral values within a particular period band, while one of the secondary components is rich in frequency content within the same band; thus, although the scale factor k_1 may be greater for that particular strong motion component, the product $k_1 k_2$ may be lower. In the case at hand however, the family factor k_2 is within the desirable range (Table 7), a fact that can be confirmed visually from Fig. 4.

The above scaling procedure must be followed for all directions of interest, corresponding to different translational periods of the structure, T_1, and thus a different period range of interest. Although more elaborate than the analogous procedure of EN-1998-1 or ASCE/SEI 7-10 for the time history representation of the seismic action, where it is simply required that no value of the mean elastic spectrum of all time histories used is less than 90 % (EN-1998-1) or 100 % (ASCE/SEI 7-10) of the corresponding value of the elastic response spectrum, the procedure prescribed in NZS 1170.5:2004 results in a more rational scaling of strong motion records, and in a closer match to the design spectrum, as it does not impose restrictions related only to lower-than-target spectral acceleration values found in EN-1998-1 and ASCE/SEI 7-10. As a result of
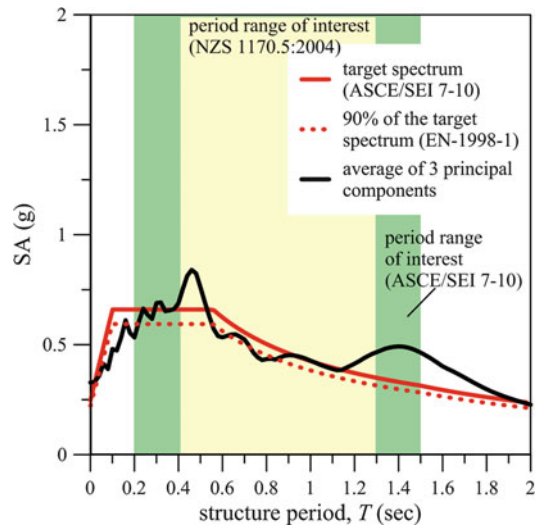
Structural Design Codes of Australia and New Zealand: Seismic Actions, Fig. 5 Comparison of the mean response spectrum (5 % damping) of the three principal components listed in Table 7 with the NZS 1170.5:2004 target spectrum for soil subclass D and $S_p = 1$. The period of interest band according to NZS 1170.5:2004 ($0.4\ T_1 \le T \le 1.3\ T_1$) and according to ASCE/SEI 7-10 ($0.2\ T_1 \le T \le 1.5\ T_1$) is also drawn

the above, if one embraced the more conservative EN-1998-1 and ASCE/SEI 7-10 provisions regarding the average response spectra from the three principal components, the specific scaling factors applied would not be acceptable, as their average spectrum is below 100 % (ASCE/SEI 7-10) and marginally below 90 % (EN-1998-1) of the target spectrum (Fig. 5).

Finally, note that NZS 1170.5:2004 (as also ASCE/SEI 7-10) bases the accelerograph scaling procedure on spectral values within the range of interest only and does not impose restrictions on the zero period spectral response acceleration values of the strong motion, as, e.g., EN-1998-1.

Summary

Seismic design action definitions in the Australian AS 1170.4:2007 and New Zealand NZS 1170.5:2004 Standards were presented in a concise way, focusing on the determination of the elastic and inelastic design response spectra to be used together with static and dynamic methods

of analysis. A short discussion against provisions of other modern seismic codes, such as EN-1998-1 and ASCE/SEI 7-10, was attempted to point out certain key differences in particular clauses. Although Australian and New Zealand Standards embrace most of the recent developments on earthquake engineering, such as accounting for near-source effects on strong motion (but not topography effects as, e.g., EN-1998-1), any seismic code is not a "static" document, but rather a "dynamic" one, and evolves with lessons learned from recent major earthquakes that are included in future revisions as, e.g., the increase of the hazard factor Z for the Christchurch area from $Z = 0.22$ g to $Z = 0.30$ g in the latest compliance document of the New Zealand Building Code (Department of Building and Housing 2011).

Cross-References

▶ Earthquake Response Spectra and Design Spectra
▶ Earthquake Return Period and Its Incorporation into Seismic Actions
▶ European Structural Design Codes: Seismic Actions
▶ Response-Spectrum-Compatible Ground Motion Processes
▶ Seismic Actions due to Near-Fault Ground Motion
▶ Selection of Ground Motions for Response History Analysis

References

Ambraseys NN, Simpson KA (1996) Prediction of vertical response spectra in Europe. Earthq Eng Struct Dyn 25:401–412

AS 1170.4:2007. Structural design actions. Part 4: earthquake actions in Australia. Standards Australia

AS 3700:2001. Masonry structures. Standards Australia

AS/NZS 1170:2002. Structural design actions. Standards Australia/standards New Zealand

ASCE/SEI 7-10. Minimum design loads for buildings and other structures. ASCE Standard, American Society of Civil Engineers, Structural Engineering Institute

Bouckovalas GD, Kouretzis GP, Kalogeras IS (2002) Site-specific analysis of strong motion data from the September 7, 1999 Athens, Greece Earthquake. Nat Hazards 27:105–131

Bouckovalas G, Kouretzis GP (2001) Stiff soil amplification effects in the 7 september 1999 Athens (Greece) earthquake. Soil Dyn Earthq Eng 21(8): 671–687

Department of Building and Housing (2011) Compliance document for New Zealand building code, clause B1, structures, Amendment 11, August. New Zealand Government

EN-1998-1. Eurocode 8: design of structures for earthquake resistance – Part 1: general rules, seismic actions and rules for buildings. European Committee for Standardization

Gaull BA, Michael-Leiba MO, Rynn JMW (1990) Probabilistic earthquake risk maps of Australia. Aust J Earth Sci 37:169–187

GeoNet-Earthquake Commission-GNS Science. website: www.geonet.org.nz. Accessed Sept 2013

McRae G, Clifton C, Megget L (2011) Review of NZ building codes of practice. Report to the royal commission of inquiry into the building failure caused by the Christchurch earthquakes ENG.ACA.0016.1

Nakamura Y (1989) A method for dynamic characteristics estimation of subsurface using microtremors on the ground surface. RTRI Q Rep 30(1):25–33

Niazi M, Bozorgnia Y (1992) Behaviour of near-source vertical and horizontal response spectra at SMART-1 array. Earthq Eng Struct Dyn 21:37–50

NZS 1170.5:2004. Structural design actions. Part 5: earthquake actions-New Zealand. Standards New Zealand

NZS 1170.5 Supp 1:2004. Structural design actions. Part 5: earthquake actions. New Zealand-commentary. Standards New Zealand

NZS 95. New Zealand standard code of building by-laws. Wellington, Standards Association of NZ

Petersen MD, Toppozada TR, Cao T, Cramer CH, Reichle MS, Bryant WA (2000) Active fault near-source zones within and bordering the state of California for the 1997 Uniform Building Code. Earthq Spectra 16(1):69–83

Structural Optimization Under Random Dynamic Seismic Excitation

Christian Bucher
Institute of Building Construction and Technology, Vienna University of Technology, Wien, Austria

Synonyms

Optimal design; Random vibration; Response surface; Seismic isolation; Structural reliability

Introduction

In civil engineering, the design of structures has always been governed by fulfilling the needs of both safety and economy. Ideally, the designer would want to minimize cost while simultaneously maximizing safety. To a large extent, these objectives are in conflict; therefore, suitable compromises need to be found. A well-established tool for finding this set of best compromises is Pareto optimization. Figure 1 illustrates the trade-off between two conflicting objectives, both of which should be minimized. Reducing one objective automatically implies increasing the other one. Good compromises are to be found on the Pareto front.

Mathematically, this can be expressed as follows. Given a multi-objective minimization problem with m scalar-valued objective functions $f_k(\mathbf{x})$

$$f_k : D_{\mathbf{x}} \subset \mathbb{R}^n \to \mathbb{R} \tag{1}$$

a point $\mathbf{x}^*$ is called nondominated if (Ngatchou et al. 2008)

$$\begin{aligned} f_k(\mathbf{x}^*) \leq f_k(\mathbf{x}) \forall k \ \forall \mathbf{x} \in D_{\mathbf{x}} \text{ and} \\ \exists \ell : f_\ell(\mathbf{x}^*) < f_\ell(\mathbf{x}) \forall \mathbf{x} \in D_{\mathbf{x}} \end{aligned} \tag{2}$$

The Pareto set, i.e., the set of all nondominated solutions of an m objective problem can be constructed by assembling all solutions $\mathbf{x}_\alpha$ of all aggregated optimization problems of the form

$$f_\alpha = \sum_{k=1}^m \alpha_k f_k(\mathbf{x}) \to \text{Min.!} \tag{3}$$

in which α_k are positive coefficients. For more involved problems, alternative method based on multi-objective genetic algorithms (MOGA) such as non-dominated sorting genetic algorithms (NSGA) (Ngatchou et al. 2008) may be more appropriate. Applications of these approaches can be found, e.g., in Liu and Frangopol (2005) and Furuta and Kameda (2006).

Traditionally, however, in structural engineering, the abovementioned compromise has been realized by setting minimum target safety levels
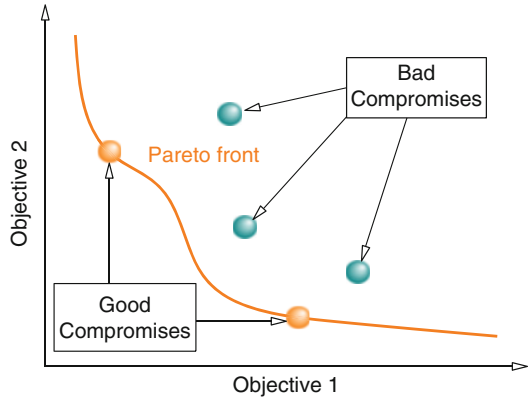
Structural Optimization Under Random Dynamic Seismic Excitation, Fig. 1 Pareto front for the case of two conflicting objectives to be minimized

(as specified, e.g., by design codes) and aiming for the economically best solution satisfying the safety requirements. Mathematically, this corresponds to a *constrained* optimization problem. Due to many uncertainties involved in the construction process as well as in the environmental loads acting on a structure, the optimization problem cannot be treated deterministically. Figure 2 shows a simple example of a stochastic optimization problem in which the design variables contain uncertainties (e.g., due to the manufacturing or construction tolerances) and the objective (cost) function contains uncertainties (e.g., due to price fluctuations on the markets), and finally, the safety constraint function contains uncertainties (e.g., due to randomness of environmental loads such as earthquakes or storms).

Optimization Methodology

In a deterministic setting, a constraint optimization problem involves an objective function $f(\mathbf{x})$ and several constraint functions $g_k(\mathbf{x})$, $k = 1 \dots m$. Then the problem can be formulated as

$$\mathbf{x}^* = \text{argmin} f(\mathbf{x}); g_k(\mathbf{x}) \leq 0; k = 1 \dots m \tag{4}$$

Since the cost function f will typically be random, any optimization will have to be formulated on the basis of expected values or probabilities
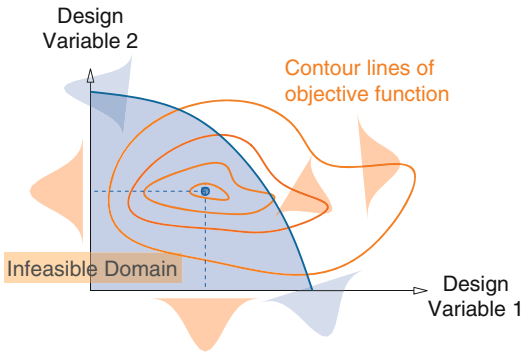
Structural Optimization Under Random Dynamic Seismic Excitation, Fig. 2 Simple example for a stochastic optimization problem

Structural Optimization Under Random Dynamic Seismic Excitation, Fig. 3 Safety margin for constraints in a stochastic optimization problem

derived from the cost function. In this way, the objective function R for the optimization can be formulated, e.g., in terms of the mean value $\overline{f}$ and standard deviation σ_f, as

$$R = \overline{f} + k\sigma_f \qquad (5)$$

In this form, a multiple of standard deviation of the cost function is added to the mean values to form the objective function. The case $k = 0$ corresponds to game theory (suitable for, e.g., large portfolio management, fair gambling), whereas $k > 0$ implies a so-called *risk-aversive* strategy.

Such a strategy is commonly employed when large losses cannot be compensated by long-term averaging (strong asymmetry).

Also, the constraint functions $g_k(\mathbf{x})$ will be random. Therefore, it is helpful to use the probability of constraint violation as new constraint, i.e., subject to the minimization of R to

$$P[g_k(\mathbf{x}^*)] > 0] \leq \varepsilon \qquad (6)$$

In safety-related constraints, the probability level ε may be very small, e.g., $\varepsilon = 10^{-5}$. Hence, it is absolutely mandatory to provide computationally efficient methods for computing such small probabilities. This need is even more pronounced as the application of optimization algorithm required the repeated computation of small failure probabilities (whenever the algorithm changes the design variables).

In traditional structural design, this is realized by introducing a safety margin (usually expressed in terms of partial safety factors) into the otherwise deterministic formulation of the constraints (cf. Fig. 3).

Computational Tools

The reduction of the large computational effort can be tackled in various ways, such as:

- Reduce and simplify the structural model, e.g., by modal reduction (Clough and Penzien 1993) or proper orthogonal decomposition (Bamer and Bucher 2012).
- Reduce the probabilistic description, e.g., by equivalent linearization (Roberts and Spanos 2003) or tail-equivalent linearization (Fujimura and Kiureghian 2007).
- Approximate the input-output relation of the system by a meta-model (Bucher and Macke 2005).
- Approximate the input-failure probability relation by a meta-model (Gasser and Schuëller 1997).
- Compute the input-failure probability relation by efficient Monte Carlo methods (Bucher 2009a).

In the following, the response surface method and a specific Monte Carlo method (asymptotic sampling) will be explained in some detail.

S

Response Surface Method

The response surface method has been a topic of extensive research in many different application areas since the influential paper by Box and Wilson in 1951 (Box and Wilson 1951). A fairly complete review on existing techniques and research directions of the response surface methodology can be found in Hill and Hunter (1966), Mead and Pike (1975), Myers et al. (1989), and Myers (1999). One of the earliest suggestions to utilize the response surface method for structural reliability assessment was made in Rackwitz (1982). Here Lagrangian interpolation surfaces and second-order polynomials are rated as useful response surfaces. Moreover, the importance of reducing the number of basic variables and error checking is emphasized. Polynomials of different orders in combination with regression analysis are proposed in Faravelli (1989), in which fractional factorial designs are utilized to obtain a sufficiently large number of support points. The absolutely essential validation of the chosen response surface model is done by means of analysis of variance.

In Ouypornprasert et al. (1989), it has been pointed out that for reliability analysis, it is most important to obtain support points for the response surface very close to or exactly at the limit state $g(\mathbf{x}) = 0$. This finding has been further extended in Kim and Na (1997) and Zheng and Das (2000). In Brenner and Bucher (1995), the response surface concept has been applied to problems involving random fields and nonlinear structural dynamics. Besides polynomials of different orders, piecewise continuous functions such as hyperplanes or simplexes can also be utilized as response surface models.

Assume that an appropriate response surface model $q(\cdot)$ has been chosen to represent the experimental data. Then, for estimating the values of the parameters $\boldsymbol{\theta}$ in the model, the method of maximum likelihood can be utilized. Under the assumptions of a Gaussian distribution of the random error terms ε, the method of maximum likelihood can be replaced by the more common method of least squares (Box and Draper 1987). In the latter case, the parameters $\boldsymbol{\theta}$ are determined in such a way that the sum of squares of the differences between the value of
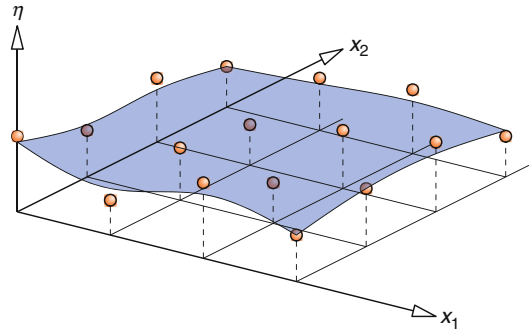
Structural Optimization Under Random Dynamic Seismic Excitation, Fig. 4 Response surface and data points

the response surface $q(\boldsymbol{\theta}; \mathbf{x}^{(k)})$ and the measured response $z^{(k)}$ at the m points of experiment

$$\mathbf{x}^{(k)} = \left(x_1^{(k)}, \ldots, x_n^{(k)}\right)', \quad k = 1, 2, \ldots, m \quad (7)$$

becomes as small as possible. In other words, the sum of squares function

$$s(\boldsymbol{\theta}) = \sum_{k=1}^{m} \left(z^{(k)} - q\left(\boldsymbol{\theta}; \mathbf{x}^{(k)}\right)\right)^2 \quad (8)$$

has to be minimized. This corresponds to a minimization of the variance of the random error terms ε. The minimizing choice of $\boldsymbol{\theta}$ is called a *least squares estimate* and is denoted by $\hat{\boldsymbol{\theta}}$.

The above regression problem becomes more simple to deal with when the response surface model is linear in its parameters $\boldsymbol{\theta}$. Let us assume that the response surface is given by (Fig. 4)

$$\eta = \theta_1 q_1(\mathbf{x}) + \theta_2 q_2(\mathbf{x}) + \cdots + \theta_p q_p(\mathbf{x}) \quad (9)$$

The least squares estimate of the parameter vector is given by

$$\hat{\boldsymbol{\theta}} = (\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'\,\mathbf{z} \quad (10)$$

This estimator is unbiased, i.e.,

$$\mathrm{E}\left[\hat{\boldsymbol{\theta}}\right] = \boldsymbol{\theta} \quad (11)$$

A detailed discussion is given by (Bucher and Macke 2005).

Structural Optimization Under Random Dynamic Seismic Excitation, Fig. 5 Schematic sketch of first-passage problem
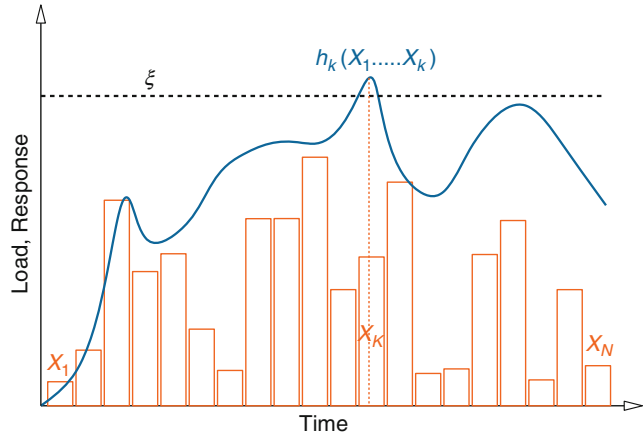
Monte Carlo Simulation

First-Passage Problem
Generally, the probability of failure P_F in an n-dimensional space of random variables $X_1, \ldots X_n$ can be computed as

$$P_F = \int \cdots \int_{D_F} f_{X_1, \ldots X_n}(x_1, \ldots x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n \quad (12)$$

In this equation, $f_{X_1, \ldots X_n}(x_1, \ldots x_n)$ denotes the joint probability function of the random variables $X_1, \ldots X_n$, and D_F denotes the failure domain, i.e., the region of the n-dimensional random variable space in which failure occurs. Typically, this is denoted in terms of a scalar limit state function $g(.)$ attaining negative values, i.e., $DF = \{(X_1, \ldots X_n) \mid g(X_1, \ldots X_n) \leq 0\}$. For applications in time-dependent problems as those arising structural dynamics, the appropriate formulation leads to the first-passage problem. Here the failure domain is typically written as

$$D_F = \left\{ (X_1, \ldots X_n) \mid \max_{1 \leq k \leq N} h_k(X_1, \ldots X_n) \geq \xi \right\} \quad (13)$$

In this equation, $h(.)$ denotes a response quantity of interest, and ξ is a critical threshold value of this response. The random variables X_i usually denote the random excitation (e.g., earthquake or wind) which is discretized in time. A schematic sketch is shown in Fig. 5. Note that due to the

principle of causality, the values of h at point k in time can only depend on the basic variables $X_1 \ldots X_k$ (i.e., on those in the past put to the present as expressed by the index k) and not on those in the future.

The generalized safety index (or reliability index) β is defined by

$$\beta = \Phi^{-1}(1 - P_F) = \Phi^{-1}[F_H(h)] \quad (14)$$

Here $\Phi^{-1}(.)$ is the inverse standardized Gaussian distribution function. Without loss of generality, it is assumed that the random variables X_i are Gaussian and that they are independent and identically distributed (i.i.d.) with zero mean and unit standard deviation.

Since the failure probabilities to be computed are usually very small (e.g., $PF = 10^{-5}$), it is not feasible to evaluate the integral in Eq. 12 using standard numerical integration procedures. This is due to the fact that the number of integration points required to perform, e.g., Gaussian integration in dimension n, grows exponentially with n. Monte Carlo methods do not have this dependence on the dimension; however, crude or plain Monte Carlo simulation requires a number of samples roughly larger than $\frac{10}{P_F}$. This, again, may lead to prohibitively large computational effort.

Asymptotic Sampling Method
The asymptotic sampling method as described in Bucher (2009a, b) exploits the asymptotic

behavior of the failure probability expressed in terms of the reliability index β when changing the standard deviation of the basic random variables. If the original standard Gaussian random variables are replaced by variables with non-unit standard deviations $\sigma = \frac{1}{f}$, then the computed reliability index will depend on the choice of f. As a first simple case, consider a linear function of the basic random variables X_k, say

$$g(\mathbf{X}) = \sum_{k=1}^{N} a_k X_k \qquad (15)$$

with arbitrary real-valued coefficients a_k. The random variable $Y = g(\mathbf{X})$ will then be Gaussian with a zero mean and a variance

$$\sigma_Y^2 = \sum_{k=1}^{N} a_k^2 \qquad (16)$$

The distribution function $F_Y(y)$ of this variable will therefore be given by

$$F_Y(\xi) = \Phi\left(\frac{\xi}{\sigma_Y}\right) = \Phi\left(\frac{\xi}{\sqrt{\sum_{k=1}^{N} a_k^2}}\right) \qquad (17)$$

Upon changing the standard deviation of all basic variables from unity to a value of $1/f$, the standard deviation of Y changes by the same amount. Thus, the distribution function changes to

$$F_Y(\xi) = \Phi\left(\frac{\xi f}{\sqrt{\sum_{k=1}^{N} a_k^2}}\right) \qquad (18)$$

from which the generalized reliability index according to Eq. 14 is immediately found as

$$\beta = \frac{\xi f}{\sqrt{\sum_{k=1}^{N} a_k^2}} \rightarrow \frac{\beta}{f} = \frac{\xi}{\sqrt{\sum_{k=1}^{N} a_k^2}} \qquad (19)$$

This means that for a linear function of Gaussian variables, the scaled reliability index β/f is invariant with respect to the choice of f.

For the general case, this is not true. There is, however, an asymptotic property which ensures similar behavior for many nonlinear cases. As stated by Breitung (1984) and Gollwitzer and Rackwitz (1988), the reliability index asymptotically depends linearly on f or in scaled notation

$$\lim_{f \to \infty} \frac{\beta(f)}{f} = \text{const.} \qquad (20)$$

In order to exploit this asymptotic relation, Bucher (2009b) suggested to utilize the formulation

$$\frac{\beta}{f} = A + \frac{B}{f^2} \qquad (21)$$

in which the coefficients A and B are determined from a regression analysis using the values of β as estimated from several Monte Carlo runs with different values of f. It should be mentioned that the asymptotic sampling method is a representative of a very recent family of methods exploiting the dependency of the failure probability on a scaling parameter of the standard deviation of the input variables. A different approach to this is given by Naess and Gaidai (2008) and Naess et al. (2010). For a comparative review and synthesis of these two approaches, it is referred to Qin et al. (2012).

Optimal Design of Seismic Isolation Systems

Definition of Problem

In order to ensure structural safety and integrity in earthquake conditions, it may be useful or even necessary to equip structures with protective devices. One possible choice is seismic isolation devices.

The basic scenario in which such devices are utilized is shown in Fig. 6.

In this scenario, a structure is subjected to an earthquake described by the ground acceleration $a(t)$. Its effect on the structure is to be mitigated

Structural Optimization Under Random Dynamic Seismic Excitation, Fig. 6 Structure and friction-based seismic isolation device
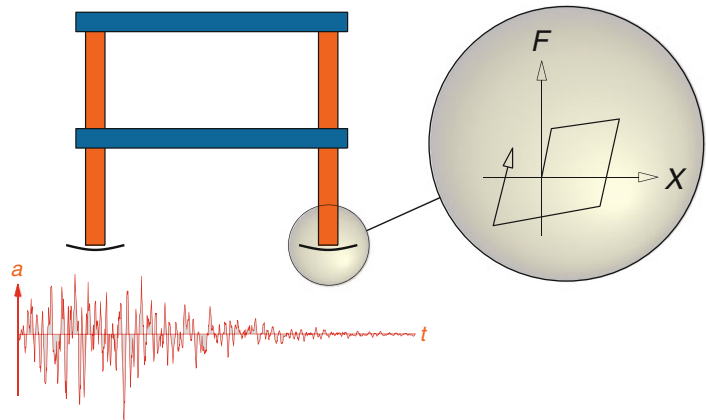
by a device which limits the transfer of forces from the ground to the structure. One such device consists of a combined friction and spring element, in which the spring can also be replaced by a re-centering force due to gravity effects (so-called friction pendulum systems, see, e.g., Roussis and Constantinou (2006)). The device has two characteristic parameters, one is the maximum force transmitted by friction (described by the friction coefficient μ) and the other is the re-centering (or restoring) spring stiffness constant (described by the effective radius of curvature R) (Fig. 7). Both parameters should be chosen such as to ensure the desired protective effect for the structure and the integrity of the isolation device itself.

The question of optimal design for stiffness and friction of seismic isolation systems has been addressed previously in the literature (e.g., Constantinou and Tadjbakhsh (1983), Iemura et al. (2007), Jangid (2005)). The method of equivalent linearization in conjunction with a random process model for the earthquake excitation tuned to the El Centro record was used in Jangid (1996). Here, a full nonlinear dynamic analysis will be used as a basis for the computation of the structural response. The ground motion is modeled as a nonstationary (i.e., evolutionary) random process; hence, the response should to be characterized in suitable probabilistic terms.

As the main function, if the seismic isolation is to keep the effect of the ground acceleration from the superstructure, it is obvious that a "soft"
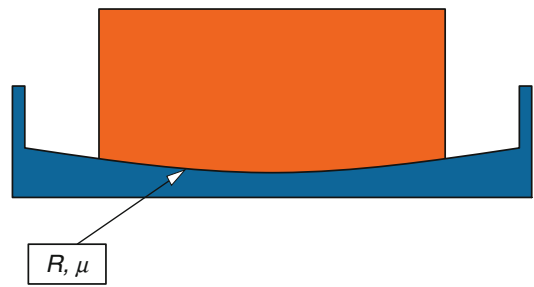
Structural Optimization Under Random Dynamic Seismic Excitation, Fig. 7 Simple sketch for friction pendulum system

isolator will be good for that purpose. On the other hand, the isolator itself can have only limited displacements due to constructive reasons. This means that the isolator should be "stiff" enough. Hence, the targets (objectives) of the optimal design of a seismic isolation device are in conflict with each other. As mentioned above, the determination of the Pareto set allows the preselection of suitable candidates for the final decision-making process.

The relevant quantities for the Pareto optimization are the maximum displacement x_D of the friction device (this should be small in order to ensure safety of the supports) and the maximum structural displacement x_S (measured relative to the device, this should be small to ensure seismic isolation). As the quantities are random, characteristic values (i.e., mean value plus three times the respective standard deviation) are chosen for the objectives.
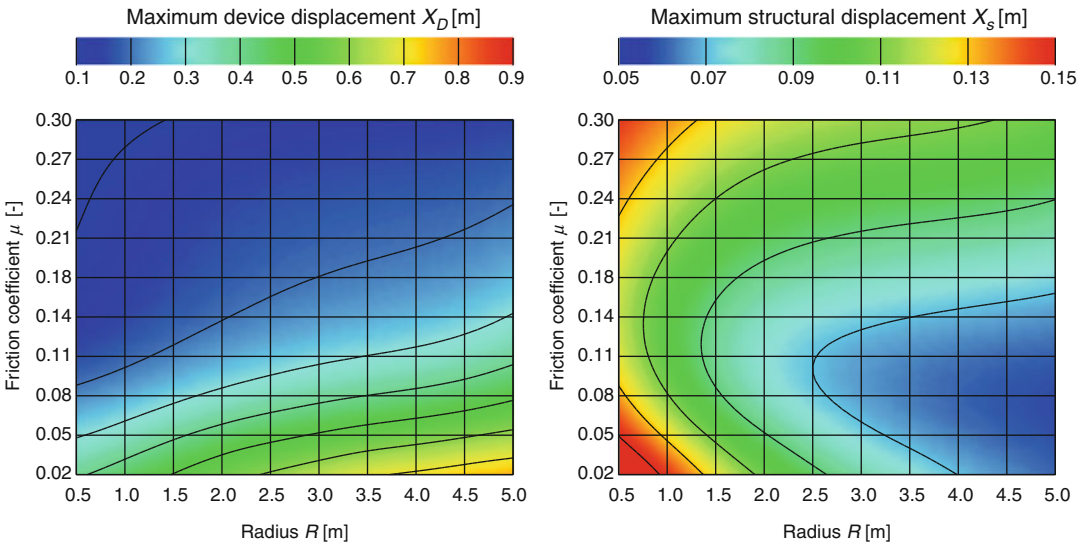
Structural Optimization Under Random Dynamic Seismic Excitation, Fig. 8 Characteristic value of maximum device displacement depending (*left*) and maximum structural displacement (*right*) depending on design parameters R and μ

Structural Optimization Under Random Dynamic Seismic Excitation, Fig. 9 Convergence of particle swarm optimization method to one point in the Pareto set
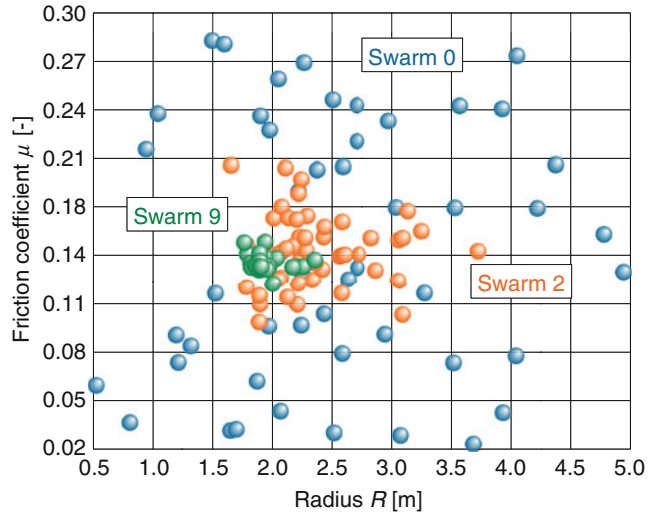
Selected Results

In order to prepare the optimization, the characteristic values for x_D and x_D are computed for 100 randomly selected combinations of the design variables R and μ. At each sample point, the structural system is analyzed for 100 randomly generated earthquakes (using a nonstationary Kanai-Tajimi model (Bucher 2009c)), and the characteristic values for x_D and x_S are computed using sample statistics. These discrete values are utilized to form response surface approximations which by inherent smoothing largely eliminate the sampling uncertainty inherent in Monte Carlo methods. The resulting response surfaces are shown in Fig. 8.

An aggregated objective function R is then generated as

$$R = \alpha x_D + (1 - \alpha)x_S \qquad (22)$$

**Structural Optimization
Under Random Dynamic
Seismic Excitation,
Fig. 10** Final Pareto set
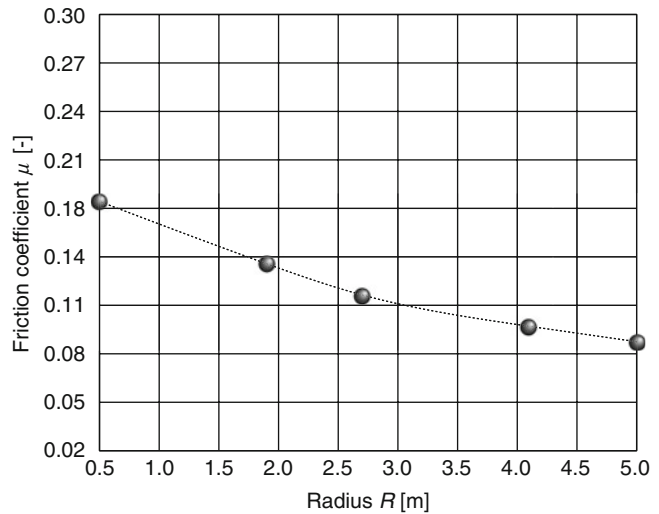
in which α is a real number in the interval $[0, 1]$. Solving the aggregated problem results in one point of the Pareto set. The convergence of this optimization using a particle swarm optimizer is shown in Fig. 9.

Finally, the results of aggregated optimization runs for different values of α yield the Pareto set as shown in Fig. 10.

It turns out that the Pareto set contains very large friction coefficients which are very difficult to realize in practice. Hence, the optimal solutions may not be accessible.

Cross-References

▶ Code-Based Design: Seismic Isolation of Buildings
▶ Nonlinear Dynamic Seismic Analysis
▶ Reliability Estimation and Analysis for Dynamical Systems
▶ Robust Design Optimization for Earthquake Loads
▶ Stochastic Analysis of Nonlinear Systems

References

Bamer F, Bucher C (2012) Application of the proper orthogonal decomposition for linear and nonlinear structures under transient excitations. Acta Mech 223:2549–2563

Box GEP, Draper NR (1987) Empirical model-building and response surfaces. Wiley, New York

Box GEP, Wilson KB (1951) On the experimental attainment of optimum conditions. J R Stat Soc Ser B 13:1–45

Breitung KW (1984) Asymptotic approximations for multinormal integrals. J Eng Mech 110(3):357–366

Brenner CE, Bucher C (1995) A contribution to the SFE-based reliability assessment of nonlinear structures under dynamic loading. Probabilist Eng Mech 10:265–273

Bucher C (2009a) Computational analysis of randomness in structural mechanics. In: Frangopol DM (ed) Structures and infrastructures book series, vol 3. Taylor & Francis, London

Bucher C (2009b) Asymptotic sampling for high-dimensional reliability analysis. Probabilist Eng Mech 24:504–510

Bucher C (2009c) Probability-based optimization of friction damping devices. Struct Saf 31:500–507

Bucher C, Macke M (2005) Response surface methodology. In: Nikolaidis E, Ghiocel DM, Singhal S (eds) Structural reliability handbook. CRC Press, Boca Raton, p 19, 1–19, 23

Clough RW, Penzien J (1993) Dynamics of structures, 2nd edn. McGraw-Hill, New York

Constantinou MC, Tadjbakhsh IG (1983) Probabilistic optimum base isolation of structures. J Struct Eng 109(3):676–689

Faravelli L (1989) Response-surface approach for reliability analysis. J Eng Mech 115:2763–2781

Fujimura K, Kiureghian AD (2007) Tail-equivalent linearization method for nonlinear random vibration. Probabilist Eng Mech 22:63–76

Furuta H, Kameda T (2006) Application of multi-objective genetic algorithm to bridge maintenance. In: Ceragioli F, Dontchev A, Futura H, Marti K, Pandolfi L (eds) System modeling and optimization.

S

IFIP International Federation for Information Processing, Boston, pp 139–148

Gasser M, Schuëller GI (1997) Reliability-based optimization of structural systems. Math Method Oper Res 46:287–307

Gollwitzer R, Rackwitz R (1988) An efficient numerical solution to the multinormal integral. Probabilist Eng Mech 3(2):98–101

Hill WJ, Hunter WG (1966) A review of response surface methodology: a literature survey. Technometrics 8:571–590

Iemura H, Taghikhany T, Jain S (2007) Optimum design of resilient sliding isolation system for seismic protection of equipments. Bull Earthquake Eng 5(1):85–103

Jangid RS (1996) Optimum damping in a non-linear base isolation system. J Sound Vibration 189(4):477–487

Jangid RS (2005) Optimum friction pendulum system for near-fault motions. Eng Struct 27(3):349–359

Kim S-H, Na S-W (1997) Response surface method using vector projected sampling points. Struct Saf 19:3–19

Liu M, Frangopol DM (2005) Bridge annual maintenance prioritization under uncertainty by multiobjective combinatorial optimization. Comp-Aided Civil Inf Eng 20(5):343–353

Mead R, Pike DJ (1975) A review of response surface methodology from a biometric view-point. Biometrics 31:803–851

Myers RH (1999) Response surface methodology – current status and future directions. J Qual Technol 31:30–44

Myers RH, Khuri AI, Carter JWH (1989) Response surface methodology: 1966–1988. Technometrics 31:137–157

Naess A, Gaidai O (2008) Monte Carlo methods for estimating the extreme response of dynamical systems. J Eng Mech 134(8):628–636

Naess A, Gaidai O, Batsevych O (2010) Prediction of extreme response statistics of narrow-band random vibrations. J Eng Mech 136(3):290–298

Ngatchou PN, Zarei A, Fox WLJ, El-Sharkawi MA (2008) Pareto multiobjective optimization. In: Lee KY, El-Sharkawi MA (eds) Modern heuristic optimization techniques. Wiley-Interscience, Piscataway, pp 189–207

Ouypornprasert W, Bucher C, Schuëller GI (1989) On the application of conditional integration in structural reliability analysis. In: Ang AHS, Shinozuka M, Schuëller GI (eds) Proceeding 5th international conference on structural safety and reliability. San Francisco, pp 1683–1689

Qin J, Nishijima K, Faber MH (2012) Extrapolation method for system reliability assessment: a new scheme. Adv Struct Eng 15(11):1893–1909

Rackwitz R (1982) Response surfaces in structural reliability. Berichte zur Zuverlässigkeitstheorie der Bauwerke. Heft 67, München

Roberts JB, Spanos P (2003) Random vibration and statistical linearization. Dover Publications, Mineola

Roussis PC, Constantinou MC (2006) Uplift-restraining friction pendulum seismic isolation system. Earthquake Eng Struct Dyn 35(5):577–593

Zheng Y, Das PK (2000) Improved response surface method and its application to stiffened plate reliability analysis. Eng Struct 22:544–551

Structural Reliability Estimation for Seismic Loading

Achintya Haldar
Department of Civil Engineering and Engineering Mechanics, University of Arizona, Tucson, AZ, USA

Synonyms

Energy dissipation in dynamic analysis; Finite element analysis; Implicit performance function; Nonlinear analysis; Risk management; Seismic loading; Structural reliability; Time-domain dynamic analysis; Time-domain reliability analysis

Introduction

The estimation of structural reliability has matured over the years. It has been accepted by the profession that if the risk in engineering design cannot completely be eliminated, it needs to be managed appropriately. The issue has attracted added significance since the basic design philosophy has changed from human safety to structural safety. This is in response to limit enormous amount of damage caused to infrastructures during recent earthquakes in China, Chile, Haiti, India, Iran, Japan, the USA, and other parts of the world. The word "structure" is used here in a generic sense. It represents real engineered systems including structures in a nuclear power plant, multistory buildings, onshore and offshore structures, soil–pile interaction problems in offshore mooring systems, aerospace structures, computer

packaging systems, systems represented by finite elements, etc. Most of the discussions made here are appropriate for different forms of dynamic including seismic loadings.

The estimation of the probability of failure, expressed in terms of risk or reliability, implied that it needs to be estimated just before failure indicating a very complicated process a structure has to go through before failure. The appropriate risk management requires an acceptable reliability method considering all major loads and load combinations that may act on the structures during their lifetime, modeling the structures appropriately satisfying the underlying physics at the time of failure and analyzing their nonlinear behavior as realistically as possible just before failure.

If the performance functions or limit state functions are explicit, i.e., if they can be expressed as explicit functions of all the random variables present in the formulation and the performance requirements, there are several methods with various degrees of sophistication that are available (Haldar and Mahadevan 2000a). The most commonly used procedures are the first-order and second-order reliability method commonly known as FORM/SORM. They are also referred to as the Rackwitz-Fiessler (1976) algorithm. Although not always necessary, for the most efficient implementation of the procedures require the performance functions are available in explicit forms. The algorithms are iterative in nature and the gradients of performance or limit state functions are required to estimate the coordinates of the most probable failure point (MPFP) and the corresponding reliability index and failure probability. However, their applicability in estimating the reliability of real structural systems could be very limited. For routine applications of FORM/SORM, the derivatives of the performance functions are difficult to obtain even for a simple structure like a two-dimensional small frame. Furthermore, risk analysis methods currently available were developed for relatively simple systems with numerous assumptions which cannot be satisfied for large realistic structural systems. For such systems, the limit states are expected to be implicit, and for

dynamic nonlinear problems, they are also to be a function of time. In response to questions on future research directions in the risk evaluation techniques, Rackwitz (2000) commented that the use of FORM/SORM complemented by Monte Carlo simulations (MCS) would be the next frontier. Several recent studies reflected this idea and simulation has become an integral part of reliability evaluation studies.

Challenges in the Reliability Estimation for Seismic Loading

Reliability estimation for dynamic or seismic loading is evolving. The probability of failure implies that it needs to be estimated just before failure developing various sources of nonlinearities. The finite element method (FEM) is commonly used by the deterministic community to study nonlinear problems where dynamic loadings are applied in time domain. This clearly indicates that an FEM-based general purpose reliability evaluation method, parallel to the deterministic analysis, is necessary.

The most sophisticated deterministic dynamic analysis of structures requires that the load should be applied in time domain. This is one of the major challenges in the reliability analysis for seismic loading. The classical random vibration-based approaches were used in the past for this purpose; however, they did not provide information acceptable to the deterministic community. The classical random vibration-based approaches have numerous limitations including the loads which are applied in the form of power spectral density functions essentially appropriate for linear structural behavior, the uncertainty in the linear or nonlinear structural behavior which may need to be incorporated in approximate ways, several performance-enhancing features currently introduced in structures which cannot be incorporated in appropriate ways, etc. The most severe weakness is that the seismic loading cannot be applied in time domain.

Structural elements (beams, columns, connections, etc.) are generally designed first for

strength according to design guidelines given in codes satisfying some underlying reliability requirements, although unknown to most designers. However, the overall system reliability in strength considering the reliabilities of all the elements remains unknown. It is not addressed in codes since it is difficult to estimate and can only be assessed based on numerous idealistic assumptions (brittle or ductile behavior, dependency of failure of elements, consideration of nonlinear limit states, etc.) (Haldar and Mahadevan 2000a). Moreover, the design satisfying strength requirements may not satisfy the global performance or serviceability requirement, e.g., excessive lateral defection, a major cause of structural failure for seismic loading. Serviceability requirements also may control the seismic design in most cases and should not be ignored (Huh and Haldar 2002).

Physics-Based Dynamic Behavior: Connection Conditions

Structural members are connected to each other using various types and forms of connections, specifically for steel structures. These connections are routinely modeled as fully restrained (FR). However, extensive experimental studies indicate that they are essentially partially restrained (PR) connections with different rigidities. In a deterministic analysis, PR connections add a major source of nonlinearity by decreasing the overall stiffness of a frame. In a dynamic analysis, it changes the dynamic properties of the structures and adds a major source of energy dissipation. In reliability analysis, it also adds a major source of uncertainty. Thus, the reliability of steel frames with PR connections and subjected to seismic excitation is expected to be quite different than that of frames with FR connections. The implications of the presence of PR connections and the uncertainty in modeling them are essential for the reliability analysis.

Static Versus Dynamic or Seismic Analyses: Improved FEM-Based Formulation

The modeling of structures to extract response information under the static and dynamic application of loadings is expected to be very different.

Under dynamic or seismic loading, the structural response will depend on many factors including the changes in the structural dynamic properties (stiffness, frequency, damping, etc.) and the corresponding response amplification as the loading progresses from linear to nonlinear stages; the sophistication in the physics-based structural modeling, i.e., FR versus PR connections and support conditions; the various sources of energy dissipation intentionally introduced in the system to improve its behavior; etc.

One basic limitation of reliability analysis under seismic loading is the incapability of handling structural behavior appropriately. Because the analysis is based on tracking the uncertainty propagating through the steps of deterministic analysis (Haldar and Mahadevan 2000b), the efficiency of the deterministic FEM is extremely important. It has been reported in the literature (Kondoh and Atluri 1987; Shi and Atluri 1988; Haldar and Nee 1989) that the assumed stress-based FEM has many advantages over the commonly used displacement-based FEM, particularly for nonlinear analysis of frame structures capable of producing large deformation. In the assumed stress-based FEM, the tangent stiffness can be expressed in an explicit form, the stresses of an element can be obtained directly, fewer elements are required to describe a large deformation configuration, and integration is not required to obtain the tangent stiffness. Different sources of nonlinearity, especially due to the presence of partially restrained (PR) connections, can be incorporated without losing the basic simplicity. It is very accurate and efficient in analyzing nonlinear responses. Some of its essential features in the context of the dynamic problems are discussed very briefly in the following sections.

Unified Stochastic Finite Element Method (SFEM)

As discussed earlier, to capture a realistic structural behavior, an FEM-based formulation is desirable. When uncertainty in the variables in the formulation is considered, it leads to the concept of stochastic finite element method (SFEM) (Haldar and Mahadevan 2000b).

Without losing any generality, the limit state function can be expressed in terms of the set of basic random variables $\mathbf{x}$ (e.g., loads, material properties, and structural geometry), the set of displacements $\mathbf{u}$, and the set of load effects $\mathbf{s}$ (except the displacements) such as internal forces, stresses, etc. The displacement $\mathbf{u} = \mathbf{QD}$, where $\mathbf{D}$ is the global displacement vector and $\mathbf{Q}$ is a transformation matrix. The limit state function can be expressed as $g(\mathbf{x}, \mathbf{u}, \mathbf{s}) = 0$. For reliability computation, it is convenient to transform $\mathbf{x}$ into the standard normal space $\mathbf{y} = \mathbf{y}(\mathbf{x})$ such that the elements of $\mathbf{y}$ are statistically independent and have a standard normal distribution. An iterative algorithm can be used to locate the design point or MPFP on the limit state function using the first-order approximation. During each iteration, the structural response and the response gradient vectors are calculated using finite element models. The following iteration scheme can be used for finding the coordinates of the design point:

$$\mathbf{y}_{i+1} = \left[\mathbf{y}_i^t \alpha_i + \frac{g(\mathbf{y}_i)}{|\nabla g(\mathbf{y}_i)|} \right] \alpha_i \tag{1}$$

where

$$\nabla g(\mathbf{y}) = \left[\frac{\partial g(\mathbf{y})}{\partial y_1}, \ldots, \frac{\partial g(\mathbf{y})}{\partial y_n} \right]^t \quad \text{and} \quad \alpha_i = -\frac{\nabla g(\mathbf{y}_i)}{|\nabla g(\mathbf{y}_i)|} \tag{2}$$

To implement the algorithm, the gradient $\nabla g(\mathbf{y})$ of the limit state function in the standard normal space can be derived as (Haldar and Mahadevan 2000b):

$$\nabla g(\mathbf{y}) = \left[\frac{\partial g(\mathbf{y})}{\partial \mathbf{s}} \mathbf{J}_{s,x} + \left(\mathbf{Q} \frac{\partial g(\mathbf{y})}{\partial \mathbf{u}} + \frac{\partial g(\mathbf{y})}{\partial \mathbf{s}} \mathbf{J}_{s,D} \right) \mathbf{J}_{D,x} + \frac{\partial g(\mathbf{y})}{\partial \mathbf{x}} \right] \mathbf{J}_{y,x}^{-1} \tag{3}$$

where $\mathbf{J}_{i,j}$'s are the Jacobians of transformation (e.g., $\mathbf{J}_{s,x} = \partial \mathbf{s}/\partial \mathbf{x}$) and y_i's are statistically independent random variables in the standard normal space. The evaluation of the quantities in Eq. 3 will depend on the problem under consideration (linear or nonlinear, two or three dimensional, etc.) and the performance functions used. The essential numerical aspect of SFEM is the evaluation of three partial derivatives, $\partial g/\partial \mathbf{s}$, $\partial g/\partial \mathbf{u}$, and $\partial g/\partial \mathbf{x}$ and four Jacobians, $\mathbf{J}_{s,x}$, $\mathbf{J}_{s,D}$, $\mathbf{J}_{D,x}$, and $\mathbf{J}_{y,x}$. They can be evaluated by procedures suggested by Haldar and Mahadevan (2000b) for linear and nonlinear, two- or three-dimensional structures. Once the coordinates of the design point $\mathbf{y}^*$ are evaluated with a preselected convergence criterion, the reliability index β can be evaluated as:

$$\beta = \sqrt{(\mathbf{y}^*)^t (\mathbf{y}^*)} \tag{4}$$

The probability of failure, P_f, can be calculated as:

$$P_f = \Phi(-\beta) = 1.0 - \Phi(\beta) \tag{5}$$

where Φ is the standard normal cumulative distribution function. Equation 5 can be considered as a notational failure probability. When the reliability index is larger, the probability of failure will be smaller. The author and his team published numerous papers to validate the above procedure for static application of the loading.

Dynamic Governing Equation for SFEM

The nonlinear dynamic equilibrium equation can be expressed at time $t + \Delta t$ as:

$$\mathbf{M}^{t+\Delta t}\ddot{\mathbf{D}}^{(k)} + {}^t\mathbf{C}^{t+\Delta t}\dot{\mathbf{D}}^{(k)} + {}^t\mathbf{K}^{(k)}{}^{t+\Delta t}\Delta\mathbf{D}^{(k)}$$
$$= {}^{t+\Delta t}\mathbf{F}^{(k)} - {}^{t+\Delta t}\mathbf{R}^{(k-1)} - \mathbf{M}^{t+\Delta t}\ddot{\mathbf{D}}_g^{(k)} \tag{6}$$

where $\mathbf{M}$ is the mass matrix, ${}^t\mathbf{C}$ is the viscous damping coefficient matrix at time t, ${}^t\mathbf{K}^{(k)}$ is the global tangent stiffness matrix of the kth iteration

at time t, $^{t+\Delta t}\Delta\mathbf{D}^{(k)}$ is the incremental displacement vector of the kth iteration at time $t + \Delta t$, $^{t+\Delta t}\mathbf{F}^{(k)}$ is the external load vector of the kth iteration at time $t + \Delta t$, $^{t+\Delta t}\mathbf{R}^{(k-1)}$ is the internal force vector of the $(k-1)^{\text{th}}$ iteration at time $t + \Delta t$, and $^{t+\Delta t}\ddot{\mathbf{D}}_g^{(k)}$ is the seismic ground acceleration vector of the kth iteration at time $t + \Delta t$.

The damping matrix $^t\mathbf{C}$ in Eq. 6 is considered to be viscous. In a realistic seismic analysis of frames, the amount of damping energy to be generated will depend on the non-yielding and yielding state of the material and the hysteretic behavior if the material yields. For mathematical simplicity, the effect of non-yielding energy dissipation is generally represented by equivalent viscous damping varying between 0.1 % and 7 % of the critical damping (Leger and Dussault 1992). Among many alternatives, Rayleigh-type damping is utilized in this study. It can be represented as:

$$^t\mathbf{C} = \alpha\mathbf{M} + \gamma\,^t\mathbf{K} \qquad (7)$$

where $^t\mathbf{K}$ is the tangent stiffness matrix and α and γ are proportional constants which can be evaluated from the natural frequencies of the structure. Since Rayleigh-type damping is proportional to mass and stiffness matrixes, it is more appropriate for nonlinear behavior as it incorporates information on changes in the stiffness values.

Numerical Procedures

The following numerical procedures can be followed to implement SFEM concept for seismic or dynamic loading. The step-by-step direct integration numerical procedure using the Newmark-β method, with parameters $\eta = 1/2$ and $\beta_N = 1/4$, can be used to solve Eq. 6. The iterative scheme can be expressed as:

$$^t\mathbf{K}_D\,^{t+\Delta t}\Delta\mathbf{D}^{(k)} = {}^{t+\Delta t}\mathbf{F}^{(k)} - {}^{t+\Delta t}\mathbf{R}^{(k-1)} \qquad (8)$$

where

$$^{t+\Delta t}\mathbf{F}^{(k)} = {}^{t+\Delta t}\mathbf{F}^{(k-1)} + {}^{t+\Delta t}\Delta\mathbf{F}_D^{(k)} \qquad (9)$$

and $^t\mathbf{K}_D$ is the dynamic tangent stiffness matrix and can be shown to be:

$$^t\mathbf{K}_D = f_1\mathbf{M} + f_2\,^t\mathbf{K} \qquad (10)$$

$^{t+\Delta t}\mathbf{F}^{(k-1)}$ and $^{t+\Delta t}\Delta\mathbf{F}_D^{(k)}$ are the modified external force and its incremental vector, respectively. The modified external force vector can be expressed as:

$$^{t+\Delta t}\mathbf{F}^{(k-1)} = {}^{t+\Delta t}\mathbf{F}^{(k-1)} + {}^{t+\Delta t}\mathbf{P}^{(k-1)} - \mathbf{M}\,^{t+\Delta t}\ddot{\mathbf{D}}^{(k-1)} \qquad (11)$$

And $^{t+\Delta t}\mathbf{R}^{(k-1)}$ is the internal force vector of the system. The term $^{t+\Delta t}\mathbf{P}^{(k-1)}$ in Eq. 11 is the modified force vector contributed by the displacement, velocity, and acceleration at time t and the displacement at time $t + \Delta t$ and can be written as

$$^{t+\Delta t}\mathbf{P}^{(k-1)} = \mathbf{M}\left[f_1\,^t\mathbf{D} + f_3\,^t\dot{\mathbf{D}} + f_4\,^t\ddot{\mathbf{D}} - f_1\,^{t+\Delta t}\mathbf{D}^{(k-1)}\right]$$
$$+\,^t\mathbf{K}\left[f_5\,^t\mathbf{D} + f_6\,^t\dot{\mathbf{D}} + f_7\,^t\ddot{\mathbf{D}} - f_5\,^{t+\Delta t}\mathbf{D}^{(k-1)}\right] \qquad (12)$$

The incremental external force term $^{t+\Delta t}\Delta\mathbf{F}_D^{(k)}$ can be shown to be:

$$^{t+\Delta t}\Delta\mathbf{F}_D^{(k)} = {}^{t+\Delta t}\Delta\mathbf{F}^{(k)} - \mathbf{M}\,^{t+\Delta t}\Delta\ddot{\mathbf{D}}_g^{(k)} \qquad (13)$$

The coefficients, f_i's, are constants and can be evaluated (Haldar and Nee 1989) in terms of η, α, β_N, γ, and Δt as:

$$f_1 = \frac{1}{\beta_N\Delta t^2} + \frac{\eta\alpha}{\beta_N\Delta t}, \quad f_2 = \frac{\eta\gamma}{\beta_N\Delta t} + 1,$$
$$f_3 = \frac{1}{\beta_N\Delta t} + \frac{\eta\alpha}{\beta_N} - \alpha,$$
$$f_4 = \left(\frac{1}{2\beta_N} - 1\right) + \eta\alpha\left(\frac{1}{2\beta_N} - \frac{1}{\eta}\right)\Delta t,$$
$$f_5 = \frac{\eta\alpha}{\beta_N\Delta t}, \quad f_6 = \frac{\eta\gamma}{\beta_N} - \gamma, \quad f_7 = \left(\frac{\eta\gamma}{2\beta_N} - \gamma\right)\Delta t \qquad (14)$$

Equation 8 now can be solved using the modified Newton–Raphson method. Once the displacements are obtained, the member forces can be calculated accordingly.

Consideration of PR Connections in the Finite Element Formulation

The presence of partial connection rigidity needs to be incorporated into the deterministic analysis of structures to capture their behavior. In general, the relationship between the moment M, transmitted by the connection, and the relative rotation angle θ is used to represent the flexible behavior. Among the many alternatives (Richard model, piecewise linear model, polynomial model, exponential model, B-spline model, etc.), the Richard four-parameter moment–rotation model is chosen here to represent the flexible behavior of a connection. It can be expressed as (Richard and Abbott 1975):

$$M = \frac{(k - k_p)\theta}{\left(1 + \left|\frac{(k-k_p)\theta}{M_0}\right|^N\right)^{1/N}} + k_p\theta \qquad (15)$$

where M is the connection moment, θ is the relative rotation between the connecting elements, k is the initial stiffness, k_p is the plastic stiffness, M_0 is the reference moment, and N is the curve shape parameter. These parameters are identified in Fig. 1.

Although an ordinary beam-column element is used to represent a PR connection element for numerical analyses, its stiffness needs to be updated at each iteration since the stiffness representing the partial rigidity depends on θ. This can be accomplished by updating the Young's modulus as:

$$E_C(\theta) = \frac{l_C}{I_C}K_C(\theta) = \frac{l_C}{I_C}\frac{\partial M(\theta)}{\partial \theta} \qquad (16)$$

where l_C, I_C, and $K_C(\theta)$ are the length, the moment of inertia, and the tangent stiffness of the connection element, respectively. $K_C(\theta)$ is calculated using Eq. 15 and can be shown to be:

$$K_C(\theta) = \frac{dM}{d\theta} = \frac{(k - k_p)}{\left(1 + \left|\frac{(k-k_p)\theta}{M_0}\right|^N\right)^{\frac{N+1}{N}}} + k_p \quad (17)$$

The Richard model discussed up to now represents only the monotonically increasing loading portion of the M-θ curves. However, the unloading and reloading behavior of the M-θ curves is also essential for any nonlinear seismic analysis. This subject was extensively addressed in the literature (Colson 1991; El-Salti 1992). They theoretically developed the unloading and reloading parts of the M-θ curves using the Masing rule. A general class of Masing models can be defined with a virgin loading curve expressed as:

$$f(M, \theta) = 0 \qquad (18)$$

and its unloading and reloading curve can be described by the following equation:

$$f\left(\frac{M - M_a}{2}, \frac{\theta - \theta_a}{2}\right) = 0 \qquad (19)$$

where (M_a, θ_a) is the load reversal point as shown in Fig. 1.

Using the Masing rule and the Richard model represented by Eqs. 15 and 17, the unloading and reloading behavior of a PR connection can be generated as:

$$M = M_a - \frac{(k - k_p)(\theta_a - \theta)}{\left(1 + \left|\frac{(k-k_p)(\theta_a - \theta)}{2M_0}\right|^N\right)^{1/N}} - k_p(\theta_a - \theta) \qquad (20)$$

and

$$K_C(\theta) = \frac{dM}{d\theta} = \frac{(k - k_p)}{\left(1 + \left|\frac{(k-k_p)(\theta_a - \theta)}{2M_0}\right|^N\right)^{\frac{N+1}{N}}} + k_p \qquad (21)$$

If (M_b, θ_b) is the next load reversal point, as shown in Fig. 1, the reloading relation between M and θ can be obtained by simply replacing (M_a, θ_a) with (M_b, θ_b) in Eqs. 20 and 21. Thus, the proposed method uses Eqs. 15 and 17 when the connection is loading and Eqs. 20 and 21 when
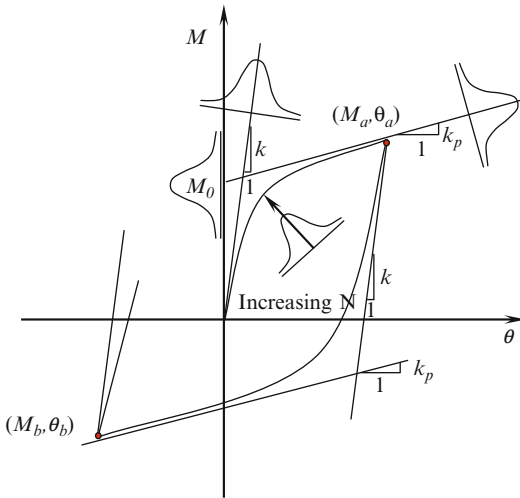
Structural Reliability Estimation for Seismic Loading, Fig. 1 M-θ curve using the Richard model, Masing rule, and uncertainty (Huh and Haldar 2002)

the connection is unloading and reloading. This represents hysteretic behavior at the PR connections.

The basic FEM formulation of the structure remains unchanged, and thus the incorporation of the PR connection can be successfully accomplished.

Pre- and Post-Northridge PR Connections

During the Northridge earthquake of 1994, several connections in steel frames fractured in a brittle and premature manner. A typical connection, shown in Fig. 2, was fabricated with the beam flanges attached to the column flanges by full penetration welds (field welded) and with the beam web bolted (field bolted) to single plate shear tabs (Richard and Radau 1998), denoted hereafter as the pre-NC.

In the post-Northridge design practices, the thrusts were to make the connections more flexible than the pre-NC and to move the location of formation of any plastic hinge away from the connection and to provide more ductility to increase the energy absorption capacity. Several improved connections can be found in the literature including cover-plated connections, spliced beam

connections, side-plated connections, bottom haunch connections, connections with vertical ribs, and connections with a reduced beam sections (RBS) or dog boned (FEMA 350–3 2000). Seismic Structural Design Associates, Inc. (SSDA) proposed a unique proprietary slotted web (SSDA SlottedWeb[TM]) moment connection (Richard et al. 1997), as shown in Fig. 3, denoted hereafter as the post-NC. The author was given access to some of the actual SSDA full-scale test results. Using the four parameters Richard model discussed earlier, Mehrabian et al. (2005) proposed a mathematical model to represent moment-relative rotation (M-θ) curves for this type of connections. Once (M-θ) curves are available, they can be used for the reliability evaluation, as discussed in the previous section. This is very important that this type of performance-enhancing feature is incorporated in any reliability evaluation.

Uncertainties in the Connection Model

A considerable amount of the uncertainty in the connection behavior comes from the uncertainties in the manufacturing and assembling processes and from the mathematical modeling (Haldar and Mahadevan 2000a). Deterministic prediction of connection behavior, based on either empirical formulations or single test data, is likely to overestimate the strength and stiffness. In practice, parameters in a typical M-θ curve are estimated from experimental results using a curve-fitting technique. Therefore, deterministic curves do not account for the scatter in the connection behavior. To consider the uncertainty in modeling the behavior of PR connections, the four parameters in the Richard model can be considered to be the basic random variables as shown in Fig. 1, along with other load- and resistance-related random parameters.

Seismic Reliability Evaluation for Large Structures

Considering all the challenges discussed above, a robust reliability evaluation technique for large

Structural Reliability Estimation for Seismic Loading, Fig. 2 A typical pre-NC (Mehrabian et al. 2005)
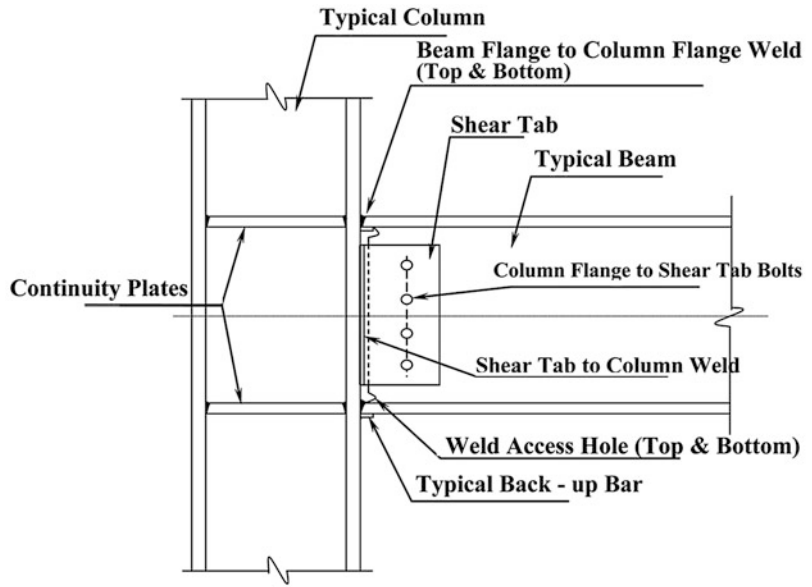
Structural Reliability Estimation for Seismic Loading, Fig. 3 A typical post-NC (Mehrabian et al. 2005)
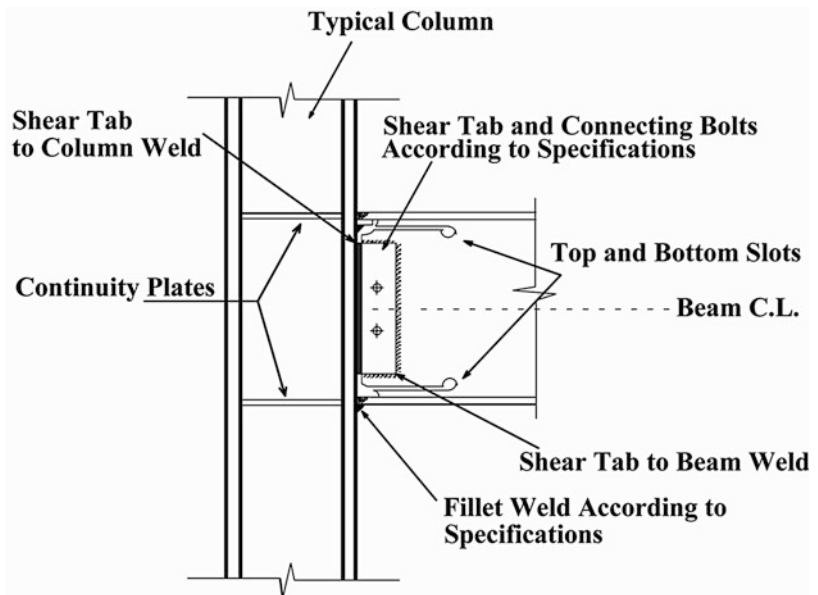
structural systems may not be available at this time. In the presence of such vacuum, one may decide a simple Monte Carlo simulation (MCS) to address the presence of uncertainty. It is observed that one deterministic analysis of a large structure excited by seismic loading applied in time domain satisfying different sources of energy dissipation and satisfying underlying physics may take over 10 h of computer time.

If one decides only say 10,000 cycles of MCS, it will take about 100,000 h or about 11.4 years of continuous running of a computer. Obviously, it will be impractical.

Eliminating the basic MCS as a realistic alternative for large structural systems with relatively moderate probability of failure events, the available computational approaches can be broadly divided into two categories: (i) the

sensitivity-based stochastic finite element method (SFEM) analysis and (ii) the response surface method (RSM). This led to the development of sensitivity-based SFEM formulation (Haldar and Mahadevan 2000b). The other alternative is RSM (Box et al. 1978). The primary purpose of applying RSM in reliability analysis is to approximate the original complex and implicit limit state function using a simple and explicit polynomial (Bucher and Bourgund 1990; Yao and Wen 1996; Khuri and Cornell 1996). The three basic weaknesses of RSM that need to be addressed before applying it for structural reliability evaluations are the following: (1) it cannot incorporate distribution information of random variables even when it is available; (2) if the response surface is not generated in the failure region, it may not be directly applicable or robust; and (3) for large systems, it may not give the optimal sampling points. Thus, a basic RSM-based reliability method may not be acceptable.

At present, a second-order polynomial without and with cross terms are generally used to generate response surfaces. Recently, Li et al. (2001) proposed high-order response surface method (HORSM). The method employs Hermite polynomials and the one-dimensional Gaussian points as sampling points to determine the highest power of each variables. In recent past, several methods with the general objective of approximately developing multivariate expressions for response surface for mechanical engineering applications were proposed. One such method is the high-dimensional model representation (HDMR) (Alis and Rabitz 2001; Rao and Chowdhury 2009; Sobol 2003; Wei and Rahman 2007). It is also referred to as "decomposition method," "univariate approximation," "bivariate approximation," "S-variate approximation," etc. HDMR captures the high-dimensional relationships between sets of input and output model variables in such a way that the component functions of the approximation are ordered starting from a constant and adding terms such as first order, second order, and so on. The concept appears to be reasonable if higher-order variable correlations are weak,

allowing the physical model to be captured by the first few lower-order terms.

Another major work is known as the explicit design space decomposition (EDSD). It can be used when responses can be classified into two classes, e.g., safe and unsafe. The classification is performed using explicitly defined boundaries in space. A machine learning technique known as support vector machines (SVM) (Basudhar et al. 2008; Layman et al. 2007) is used to construct the boundaries separating distinct classes. The failure regions corresponding to different modes of failure are represented with a single SVM boundary, which is refined through adaptive sampling.

The HORSM, HDMR, and EDSD-SVM approaches use MCS to estimate the underlying reliability. They may not be suitable for reliability evaluation of large structural systems where dynamic loadings need to be applied in time domain and several important features like PR connections, several sources of nonlinearities and energy dissipation features must be explicitly analytically incorporated in the formulation satisfying the underlying physics.

Proposed New Method

The proposed reliability evaluation method for large structural systems is developed in two stages. In the first stage, the two weaknesses of RSM, i.e., the consideration of distributional information of the random variables present in the formulation and identifying the location of the failure region, are addressed by integrating it with FORM/SORM. This approach will lead to a hybrid approach consisting of SFEM, FORM/SORM, and RSM. In the second stage, the efficiency of the method is improved by using several advanced factorial schemes so that the response surface can be generated with fewer sampling points.

Proposed Method: Stage 1
Considering the fact that a higher-order polynomial may result in an ill-conditional system of equations for unknown coefficients and exhibit

an irregular behavior outside of the domain of samples, their utilization in generating RSM has received relatively little attention (Gavin and Yau 2008; Rajashekhar and Ellingwood 1993). For complicated problems considered by the research team of the author, a second-order polynomial, without and with cross terms, is considered to be appropriate. They can be expressed as:

$$\hat{g}(\mathbf{X}) = b_0 + \sum_{i=1}^{k} b_i X_i + \sum_{i=1}^{k} b_{ii} X_i^2 \qquad (22)$$

$$\hat{g}(\mathbf{X}) = b_0 + \sum_{i=1}^{k} b_i X_i + \sum_{i=1}^{k} b_{ii} X_i^2 + \sum_{i=1}^{k-1} \sum_{j>i}^{k} b_{ij} X_i X_j \qquad (23)$$

where X_i ($i = 1, 2, \ldots, k$) is the ith random variable and b_0, b_i, b_{ii}, and b_{ij} are unknown coefficients to be determined. The numbers of coefficient necessary to define Eqs. 22 and 23 are $p = 2k + 1$ and $p = (k + 1)(k + 2)/2$, respectively. The coefficients can be fully defined by estimating deterministic responses at intelligently selected data points called experimental sampling points. The concept behind a sampling scheme can be expressed as:

$$X_i = X_i^C \pm h_i \sigma_{x_i} x_i \qquad (24)$$

where X_i^C and σ_{x_i} are the coordinates of the center point and the standard deviation of a random variable X_i, respectively, and h_i is an arbitrary factor that defines the experimental region.

Sampling points are selected around the center point. The selection of the center point and experimental sampling points around it are crucial factors in establishing the efficiency and accuracy of the proposed iterative method. The center point is selected to be the coordinates of the checking points as the iteration continues. In the context of iterative scheme of FORM/SORM, the initial center point $\mathbf{x}_{C_1}$ is selected to be the mean values of the random variable X_i's. Then, using the responses obtained from the deterministic FEM evaluations for all the experimental sampling points around the center point, the response surface $\hat{g_1}(\mathbf{X})$ can be generated explicitly in terms of the random variables $\mathbf{X}$. Once a closed-form expression for the limit state function is obtained, the coordinates of the checking point $\mathbf{x}_{D_1}$ can be estimated using FORM/SORM, using all the statistical information on the X_i's, eliminating one major deficiency of RSM. The response can be evaluated again at the checking point $\mathbf{x}_{D_1}$, and a new center point $\mathbf{x}_{C_2}$ can be selected using linear interpolation from the center point $\mathbf{x}_{C_1}$ to $\mathbf{x}_{D_1}$ such that $g(\mathbf{X}) = 0$, i.e.,

$$\mathbf{x}_{C_2} = \mathbf{x}_{C_1} + (\mathbf{x}_{D_1} - \mathbf{x}_{C_1}) \frac{g(\mathbf{x}_{C_1})}{g(\mathbf{x}_{C_1}) - g(\mathbf{x}_{D_1})} \quad \text{if} \quad g(\mathbf{x}_{D_1}) \geq g(\mathbf{x}_{C_1}) \qquad (25)$$

$$\mathbf{x}_{C_2} = \mathbf{x}_{D_1} + (\mathbf{x}_{C_1} - \mathbf{x}_{D_1}) \frac{g(\mathbf{x}_{D_1})}{g(\mathbf{x}_{D_1}) - g(\mathbf{x}_{C_1})} \quad \text{if} \quad g(\mathbf{x}_{D_1}) < g(\mathbf{x}_{C_1}) \qquad (26)$$

A new center point $\mathbf{x}_{C_2}$ then can be used to develop an explicit performance function for the next iteration. This iteration scheme can be repeated until a preselected convergence criterion of $(\mathbf{x}_{C_{i+1}} - \mathbf{x}_{C_i})/\mathbf{x}_{C_i} \leq \varepsilon$ is satisfied. ε can be considered to be $|0.05|$. In the final iteration, the information on the most recent center point

can be used to formulate the final response surface. FORM/SORM can then be used to calculate the reliability index and the corresponding coordinates of the MPFP.

It is to be noted that since increasing efficiency is a major objective, random variables with low-sensitivity indexes (Haldar and Mahadevan 2000a)

can be considered as deterministic at their mean values without compromising the accuracy of the reliability estimation.

To select experimental sampling points around the center point, saturated design (SD) and central composite design (CCD) are the two most promising schemes. SD is less accurate but more efficient since it requires only as many sampling points as the total number of unknown coefficients to define the response surface. CCD is more accurate but less efficient since a regression analysis is needed to evaluate the unknown coefficients for the response surface. Also, a second-order polynomial with cross terms (Eq. 23) must be used for CCD.

To illustrate the computational effort required for the reliability evaluation of a large structural system, suppose the total number of significant random variables, after making some of them deterministic at their mean values, present in the formulation is $k = 50$. The total number of coefficients necessary to define Eq. 22 will be $2 \times 50 + 1 = 101$ and to define Eq. 23 will be $(50 + 1)(50 + 2)/2 = 1326$. It can also be shown that if Eq. 22 and SD scheme are used to generate the response surface, the total number of sampling points, essentially the total number of deterministic FEM-based time-domain nonlinear response analyses will be 101. However, if Eq. 23 and a full SD scheme are used, the corresponding deterministic analyses will be 1326. If Eq. 23 and CCD scheme are used, the corresponding deterministic analyses will be $2^{50} + 2 \times 50 + 1 > 1.1258999 \times 10^{15}$. Obviously, some of these alternatives may not be meaningful.

Since the proposed algorithm is iterative and the basic SD and CCD require different amounts of computational effort, considering efficiency without compromising accuracy, several schemes can be followed. Among the numerous schemes considered by the research team, one basic and two promising schemes are:

Scheme 0 – SD using second-order polynomial without the cross terms throughout all the iterations.

Scheme 1 – Implement SD using Eq. 22 for the intermediate iterations and SD using Eq. 23 for the final iteration.

Scheme 2 – Implement SD using Eq. 22 for the intermediate iterations and CCD using Eq. 23 for the final iteration.

The efficiency and accuracy of the above schemes will be discussed with the help of an example.

Proposed Method: Stage 2

The above two schemes may be major improvements but still may not be able to estimate the reliability of large structural systems. They may require fewer deterministic evaluations but still they could be in thousands. Their efficiency can be improved significantly by reducing the deterministic evaluations in hundreds in the following way.

Scheme M1: To improve the efficiency of Scheme 1, the cross terms (edge points), $k(k - 1)$, are suggested to be added only for the most important variables in the last iteration. Since the proposed algorithm is an integral part of FORM/SORM, all the random variables in the formulation can be arranged in descending order of their sensitivity indexes $\alpha(X_i)$, i.e., $\alpha(X_1) > \alpha(X_2) > \alpha(X_3)\ldots\ldots > \alpha(X_k)$. The sensitivity of a variable X, $\alpha(X)$, is the directional cosines of the unit normal vector at the design point. In the last iteration, the cross terms are added only for the most sensitive random variables, m, and the corresponding reliability index is calculated. The total number of FEM analyses required for Scheme 1 and $M1$ are $(k + 1)(k + 2)/2$ and $2k + 1 + m(2k - m - 1)/2$, respectively. For an example, suppose for a large structural system, $k = 50$ and $m = 3$. The total number of required FEM analyses will be 1326 and 245, respectively, for the two schemes; a significant improvement in the efficiency without compromising the accuracy.

Illustrative Examples

To illustrate the different procedures discussed here for the seismic risk evaluation of a structure, the following illustrative examples are presented.

Structural Reliability Estimation for Seismic Loading, Fig. 4 A three-story three-bay SMRF structure (Haldar et al. 2012)
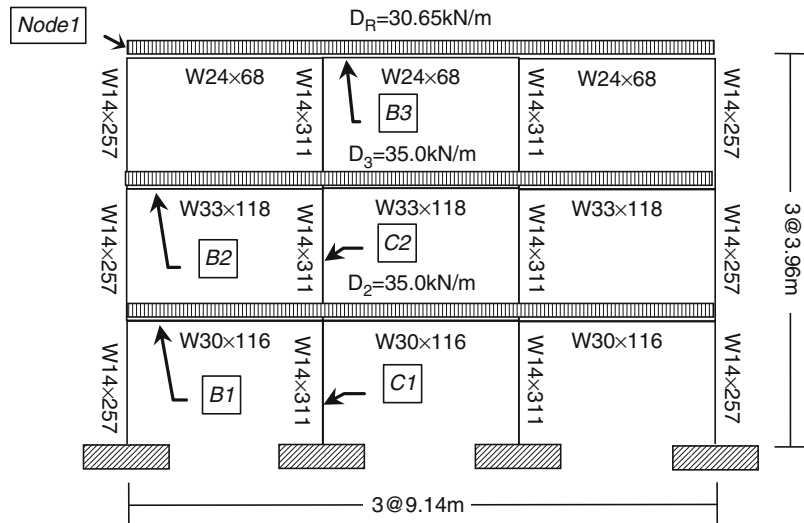
Structural Reliability Estimation for Seismic Loading, Fig. 5 Earthquake time history used in this study
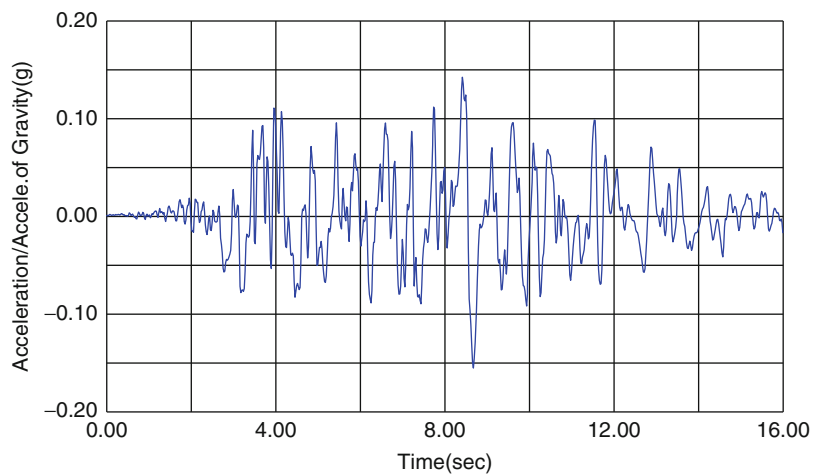
Reliability of Steel Frames in the Presence of Reinforced Concrete Shear Wall: Seismic Loading is Applied in Time Domain

A three-story three-bay steel frame, as shown in Fig. 4, is considered (Haldar et al. 2012). Section sizes of beams and columns, using A36 steel, are also shown in the figure. The gravity loads acting on the second, third, and roof are given in the figure. A building is supposed to consist of several such frames. It was excited by a seismic time history shown in Fig. 5.

The four parameters of the Richard model are calculated by PRCONN (Richard 1993), a commercially available computer program for both pre-NC and post-NC connections. For the example under consideration, considering the sizes of columns and beams, three types of connection are necessary. They are denoted as types A, B, and C, hereafter. Four Richard parameters for both pre-NC and post-NC connections are summarized in Table 1.

Limit States of Performance Functions

Both strength and serviceability limit states are used for structural reliability estimation.

Strength limit states: They mainly depend on the failure mode to be considered. Most of the elements in the structural system considered are beam columns, i.e., they are subjected to both axial load and bending moment at the same time.

Structural Reliability Estimation for Seismic Loading, Table 1 Parameters of Richard equation for M-θ curves

Connection assembly type			Connection parameters			
ID		Beam	Column	k^a	$k_p{}^a$	$M_0{}^b$	N
Pre-NC	A	W24 × 68	W14 × 257	2.51×10^7	5.56×10^5	4.16×10^4	1.1
		W24 × 68	W14 × 311				
	B	W33 × 118	W14 × 257	5.08×10^7	1.14×10^5	6.79×10^4	1.1
		W33 × 118	W14 × 311				
	C	W30 × 116	W14 × 257	3.95×10^7	9.19×10^5	5.65×10^4	1.1
		W30 × 116	W14 × 311				
Post-NC	A	W24 × 68	W14 × 257	1.00×10^9	4.52×10^5	9.64×10^4	1.0
		W24 × 68	W14 × 311				
	B	W33 × 118	W14 × 257	2.34×10^9	4.52×10^5	2.44×10^5	1.0
		W33 × 118	W14 × 311				
	C	W30 × 116	W14 × 257	2.14×10^9	4.52×10^5	2.21×10^5	1.0
		W30 × 116	W14 × 311				

[a] kN · cm/rad
[b] kN · cm

For design purposes, interaction equations are generally used to consider the combined effect of axial load and bending moment. The interaction equations suggested by the American Institute of Steel Construction's *Load and Resistance Factor Design* manual (AISC 2005) for two-dimensional structures are used for this illustrative example. They can be expressed as:

$$\frac{Pu}{\phi p_n} + \frac{8}{9}\left(\frac{M_{ux}}{\phi_b M_{nx}}\right) \leq 1.0; \quad \text{if} \quad \frac{p_u}{\phi p_n} \geq 0.2 \tag{27}$$

$$\frac{Pu}{2\phi p_n} + \left(\frac{M_{ux}}{\phi_b M_{nx}}\right) \leq 1.0; \quad \text{if} \quad \frac{p_u}{\phi p_n} < 0.2 \tag{28}$$

where ϕ and ϕ_b are the resistance factors, P_u is the required tensile/compressive strength, P_n is the nominal tensile/compressive strength, M_{ux} is the required flexural strength, and M_n is the nominal flexural strength. P_n and M_{nx} can be calculated using AISC's LRFD code (AISC 2005).

Serviceability limit states: For seismic loading, the design may be controlled by the serviceability, e.g., inter-story drift or the overall lateral displacement. The general form of a serviceability limit state can be represented as:

$$g(\mathbf{X}) = \delta_{allow} - y_{max}(\mathbf{X}) = \delta_{allow} - \hat{g}(\mathbf{X}) \tag{29}$$

where δ_{allow} is the allowable inter-story drift or overall lateral displacement specified in codes and $y_{max}(\mathbf{X})$ is the corresponding maximum inter-story drift or overall lateral displacement estimated from the analysis.

Reliability Evaluations of Frames with Different Connection Conditions and Without RC Shear Wall

The statistical information on all the random variables for the illustrative example is summarized in Table 2. The allowable displacement at the top of the frame, δ_{allow}, is considered to be $h/400$, where h is the height of the frame. For this example, δ_{allow} is 2.97 cm. The corresponding serviceability limit state is defined by Eq. 29 (Huh and Haldar 2011). The probabilities of failure of the frame for the lateral deflection at the top of the frame for serviceability limit state and the strength limit state of the weakest members are first estimated assuming all the connections are of FR type, considered routinely in the profession. The results are summarized in Table 3 using the proposed method. To verify the results, 20,000 MCS for the serviceability and 30,000 MCS for the strength limit states were carried out.

The results clearly indicate that the bare steel frame will not satisfy the serviceability requirement. Then, the reliabilities of the frame are estimated assuming all the connections are post-NC

Structural Reliability Estimation for Seismic Loading, Table 2 Statistical information on the design variables

	Item	Random variable	Mean value	COV	Dist.
Member	All	$E(kN/m^2)$	1.999×10^8	0.06	Ln
		$F_y(kN/m^2)$	2.482×10^5	0.10	Ln
	Column W14 × 257	$I_x^{C1}(m^4)$	1.415×10^{-3}	0.05	Ln
		$Z_x^{C1}(m^3)$	7.981×10^{-3}	0.05	Ln
	Column W14 × 311	$A^{C2}(m^2)$	5.897×10^{-2}	0.05	Ln
		$I_x^{C2}(m^4)$	1.802×10^{-3}	0.05	Ln
	Beam W33 × 118	$I_x^{B2}(m^4)$	2.456×10^{-3}	0.05	Ln
	Beam W30 × 116	$I_x^{B3}(m^4)$	2.052×10^{-3}	0.05	Ln
		$Z_x^{B3}(m^3)$	6.194×10^{-3}	0.05	Ln
Seismic load		ξ	0.05	0.15	Type I
		g_e	1.0	0.20	Type I
Connection	Richard model parameter	K^a	Refer to values in Table 3	0.15	N
		k_p^a		0.15	N
		M_0^b		0.15	N
		N		0.05	N
Shear wall[c]		$E_C(kN/m^2)$	2.137×10^7	0.18	Ln
		ν	2.137×10^7	0.18	Ln

Ln lognormal distribution
[a]kN · cm/rad
[b]kN cm
[c]$f_C' = 2.068 \times 10^4$ (kN/m^2)

Structural Reliability Estimation for Seismic Loading, Table 3 Reliability evaluation of FR frame

		Serviceability	Strength limit state	
Limit state		Node at 1	Beam ($B1$)	Column ($C1$)
MCS	P_f	0.08740	N/A[a]	N/A[a]
	$\beta \approx \Phi^{-1}(1 - P_f)$	1.357	N/A	N/A
	NOS[b]	20,000	30,000	30,000
Proposed algorithm	No. of RV	8	6	6
	Scheme	1	2	2
	β	1.330	4.724	5.402
	Error w.r.t β	1.99 %	N/A	N/A
	TNSP[c]	79	103	103

[a]Not a single failure observed for 30,000 cycles of simulation since large reliability indexes are expected in the strength limit state
[b]Number of simulation for deterministic FEM analyses
[c]Total number of sampling points (total number of deterministic FEM analyses)

and pre-NC types and the results are summarized in Table 4. The behavior of the frame in the presence of FR and post-NC for both serviceability and strength limit states are very similar. This was also observed during the full-scale experimental investigations establishing the advanced features of the proposed method. In any case, the lateral stiffness of the frame needs to be increased.

Reliability Evaluations of Frames with Different Connection Conditions with RC Shear Wall

To increase the lateral stiffness, the steel frame is strengthened with a reinforced concrete (RC) shear wall at the first floor level, as shown in Fig. 6. Obviously, the RC shear wall will increase the complexity in the problem.

Structural Reliability Estimation for Seismic Loading, Table 4 Reliability evaluations of frame without and with shear wall

3 × 3 SMRF without shear wall			Connection type		
			FR	Post-NC	Pre-NC
Serviceability limit state (Node 1)		β	1.330	1.329	0.463
		$P_f \approx \Phi(-\beta)$	0.09176	0.09192	0.32168
		No. of RV	8	20	20
		TNSP	79	313	313
Strength limit state	Beam	β	4.724	4.756	3.681
		$P_f \approx \Phi(-\beta)$	1.156×10^{-6}	9.873×10^{-7}	1.162×10^{-4}
		No. of RV	6	18	18
		TNSP	103	264	264
	Column	β	5.402	5.376	4.154
		$P_f \approx \Phi(-\beta)$	3.295×10^{-8}	3.808×10^{-8}	1.634×10^{-5}
		No. of RV	6	18	18
		TNSP	103	264	264
3 × 3 SMRF with shear wall					
Serviceability limit state (Node 1)		β	3.667	3.534	1.685
		$P_f \approx \Phi(-\beta)$	0.00012	0.00020	0.04599
		No. of RV	10	22	22
		TNSP	108	366	366
Strength limit state	Beam	β	6.879	6.714	4.467
		$P_f \approx \Phi(-\beta)$	3.014×10^{-12}	9.468×10^{-12}	3.966×10^{-6}
		No. of RV	8	20	20
		TNSP	79	313	313
	Column	β	6.879	6.714	4.467
		$P_f \approx \Phi(-\beta)$	3.014×10^{-12}	9.468×10^{-12}	3.966×10^{-6}
		No. of RV	8	20	20
		TNSP	79	313	313

Structural Reliability Estimation for Seismic Loading, Fig. 6 Steel frame with RC shear wall in the first floor (Haldar et al. 2012)
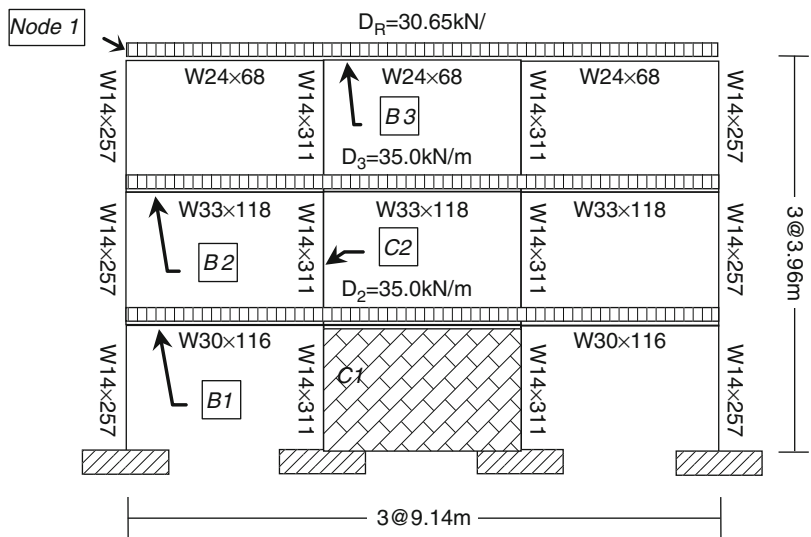
However, if it is placed properly, it will increase the lateral stiffness of the frame.

For the steel and concrete dual system, all the steel elements in the frame are modeled as beam-column elements. A four-node plane stress element is introduced for the shear wall in the frame. To consider the presence of RC shear wall, two additional parameters, namely, the modulus of elasticity, E_C, and the Poisson ratio of concrete, v, are necessary in the deterministic formulation. The tensile strength of concrete is small compared to its compressive strength and cracking may develop at a very early stage of loading. The behavior of an RC shear wall is expected to be significantly different before and after cracking. It was observed that the degradation of the stiffness of the shear walls occurs after cracking and can be considered effectively by reducing the modulus of elasticity of the shear walls (Lefas et al. 1990). The same concept is used in this study. The shear wall is assumed to develop cracks when the tensile stress in concrete exceeds the prescribed value. The rupture strength of concrete, f_r, is assumed to be $f_r = 7.5 \times \sqrt{f_c'}$, where f_c' is the compressive strength of concrete. After the tensile stress of each shear wall exceeds the prescribed tensile stress of concrete, the degradation of the shear wall stiffness is assumed to be reduced to 40 % of the original stiffness (Lefas et al. 1990). The uncertainty in all the variables considered for the bare steel frame will remain the same. However, two additional sources of uncertainty, namely, in E_C and v, need to be considered (Lee and Haldar 2003), as given in Table 2.

The frame is again excited by the same earthquake time history as shown in Fig. 2. The probabilities of failure for the combined dual system in the presence of FR, post-NC, and pre-NC connections are calculated using the proposed algorithm for the strength and serviceability limit states. The results are summarized in Table 4. The results indicate that the presence of shear wall at the first floor level significantly improves both the serviceability and strength behavior of the steel frame. If the probabilities of failure need to be reduced further, RC shear walls can be added in the second and/or third floor. Again, this improved behavior can be observed and quantified by carrying out about a hundred deterministic evaluations instead of thousands of MCS. The improved behavior of the frame in the presence of RC shear wall is expected; however, the proposed algorithm can quantify the amount of improvement in terms of probability of failure considering all major sources of uncertainty. This is a significant development. The information will help to make decision as what to do next in the design.

Summary

Structural reliability estimation for seismic loading, specifically for large structural systems, is evolving. An overview of the existing state of the art is given. For wider acceptance, it is necessary to estimate reliability by applying the seismic loading in time domain, considering special features recently being introduced to improve seismic response behavior and explicitly considering all major sources of nonlinearity and uncertainty. A new method is suggested incorporating all these features. Several seismic risk evaluation procedures are illustrated with examples.

Cross-References

▶ Nonlinear Dynamic Seismic Analysis
▶ Nonlinear Finite Element Analysis
▶ Probabilistic Seismic Hazard Models
▶ Reliability Estimation and Analysis for Dynamical Systems
▶ Response Variability and Reliability of Structures
▶ Seismic Reliability Assessment, Alternative Methods for
▶ Seismic Risk Assessment, Cascading Effects
▶ Stochastic Analysis of Nonlinear Systems

S

References

Alis OF, Rabitz H (2001) Efficient implementation of high dimensional model representations. J Math Chem 29(2):127–142

American Institute of Steel Construction (2005) Manual of steel construction: load and resistance factor design. American Institute of Steel Construction, Chicago

Basudhar A, Missoum S, Harrison Sanchez A (2008) Limit state function identification using support vector machines for discontinuous responses and disjoint failure domains. Probab Eng Mech 23(1):1–11

Box GP, William GH, Hunter JS (1978) Statistics for experimenters: an introduction to design, data analysis and modeling building. Wiley, New York

Bucher CG, Bourgund U (1990) A fast and efficient response surface approach for structural reliability problems. Struct Saf 7:57–66

Colson A (1991) Theoretical modeling of semirigid connections behavior. J Constr Steel Res 19:213–224

El-Salti MK (1992) Design of frames with partially restrained connections. PhD dissertation, Department of Civil Engineering and Engineering Mechanics, University of Arizona

Federal Emergency Management Agency (FEMA) (2000) Seismic design criteria for steel moment-frame structures, FEMA 350–353 and 355A-F

Gavin HP, Yau SY (2008) High-order limit state function in the response surface method for structural reliability analysis. Struct Saf 30(2):162–179

Haldar A, Mahadevan S (2000a) Probability, reliability, and statistical methods in engineering design. Wiley, New York

Haldar A, Mahadevan S (2000b) Reliability assessment using stochastic finite element analysis. Wiley, New York

Haldar A, Nee KM (1989) Elasto-plastic large deformation analysis of PR steel frames for LRFD. Comput Struct 31(5):811–823

Haldar A, Farag R, Huh J (2012) A novel concept for the reliability evaluation of large systems. Adv Struct Eng 15(11):1879–1892

Huh J, Haldar A (2002) Seismic reliability of nonlinear frames with PR connections using systematic RSM. Probab Eng Mech 17(2):177–190

Huh J, Haldar A (2011) A novel risk assessment method for complex structural systems. IEEE Trans Reliab 60(1):210–218

Khuri AI, Cornell JA (1996) Response surfaces designs and analyses. Marcel Dekker, New York

Kondoh K, Atluri SN (1987) Large-deformation, elasto-plastic analysis of frames under nonconservative loading, using explicitly derived tangent stiffnesses based on assumed stresses. Comput Mech 2(1):1–25

Layman R, Missoum S, Vande Geest J (2007) Failure prediction and robust design of grafts for aortic aneurysms. In: 48th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and material conference, Honolulu

Lee SY, Haldar A (2003) Reliability analysis of frame and shear wall structural systems – dynamic loading. J Struct Eng ASCE 129(2):233–240

Lefas D, Kotsovos D, Ambraseys N (1990) Behavior of reinforced concrete structural walls: strength, deformation characteristics, and failure mechanism. ACI Struct J 87(1):23–31

Leger P, Dussault S (1992) Seismic-energy dissipation in MDOF structures. J Struct Eng ASCE 118(5):1251–1269

Li G, Rosenthal C, Rabitz H (2001) High dimensional model representations. J Phys Chem A 105:7765–7777

Mehrabian A, Haldar A, Reyes AS (2005) Seismic response analysis of steel frames with post-Northridge connection. Steel Compos Struct 5(4):271–287

Rackwitz R (2000) Reliability analysis: past, present and future, Keynote Lecture. In: 8th ASCE joint specialty conference on probabilistic mechanics and structural reliability, University of Notre Dame, South Bend

Rackwitz R, Fiessler B (1976) Note on discrete safety checking when using non-normal stochastic models for basic variables. Load Project Working Session, MIT, Cambridge

Rajashekhar MR, Ellingwood BR (1993) A new look at the response surface approach for reliability analysis. Struct Saf 12(3):205–220

Rao BN, Chowdhury R (2009) Enhanced high dimensional model representation for reliability analysis. Int J Numer Methods Eng 77(5):719–750

Richard RM (1993) PRCONN manual. RMR Design Group, Tucson

Richard RM, Abbott BJ (1975) Versatile elastic–plastic stress–strain formula. J Eng Mech ASCE 101(EM4):511–515

Richard RM, Radau RE (1998) Force, stress and strain distribution in FR bolted welded connections. In: Proceedings of structural engineering worldwide, San Francisco, CA

Richard RM, Allen CJ, Partridge JE (1997) Proprietary slotted beam connection designs. Modern Steel Construction, Chicago

Shi G, Atluri SN (1988) Elasto-plastic large deformation analysis of space frames. Int J Numer Methods Eng 26:589–615

Sobol IM (2003) Theorems and examples on high dimensional model representations. Reliab Eng Syst Saf 79(2):187–193

Wei D, Rahman S (2007) Structural reliability analysis by univariate decomposition and numerical integration. Probab Eng Mech 22:27–38

Yao TH-J, Wen YK (1996) Response surface method for time-variant reliability analysis. J Struct Eng ASCE 122(2):193–201

Structural Seismic Reliability Analysis

Shankar Sankararaman
Intelligent Systems Division, SGT Inc., NASA
Ames Research Center, Moffett Field, CA, USA

Synonyms

First-order reliability method; Monte Carlo sampling; Probability of failure; Reliability analysis; Uncertainty

Introduction

Perhaps, no other discipline within engineering has to deal with as much uncertainty as the field of earthquake engineering (Der Kiureghian 1996). To begin with, the occurrence of earthquakes in time and space is completely random in nature, and this leads to a large amount of uncertainty while predicting the intensities of ground motions resulting from earthquakes. Further, it is challenging to precisely assess the load-carrying capacity of the structural system of interest, due to the inherent variability across different structural members that constitute the overall structural system. It is necessary to analyze all of these different sources of uncertainty and assess the safety of the structure by accounting for such sources of uncertainty. Structural safety assessment is important both during the design of the structural system and for analyzing its performance while the system is under operation, particularly before and after earthquakes.

Engineering design is usually a trade-off between maximizing safety levels and minimizing cost. It is necessary to account for the uncertainty in anticipated loading conditions and the different sources of uncertainty regarding the structural system while assessing safety during system design. Traditional design approaches simplify the problem by treating the uncertain quantities to be deterministic and account for the inherent uncertainty through the use of empirical safety factors, also referred to as deterministic safety factors. These empirical safety factors do not provide any information on how the different uncertain quantities influence the overall structural safety. Therefore, it is difficult to design a system with a uniform distribution of safety levels among the different components using empirical safety factors (Haldar and Mahadevan 2000). Further, during the design stage, deterministic safety factors do not provide adequate information to achieve optimal use of the available resources to maximize safety. For these reasons, it is necessary to use probabilistic approaches that rigorously account for the different sources of uncertainty and directly aid in the design of the structural system. Probabilistic approaches directly calculate the probability that the structural system may fail by probabilistically analyzing the load-carrying capacity of the structure and the actual loading on the structure. The safety of the structure is defined in terms of the converse of the probability of failure, which in turn is used to compute the so-called reliability metric that measures the probabilistic reliability of the structure. In fact, probabilistic methods provide a systematic framework to analyze the different sources of uncertainty, quantify their individual contributions to the overall structural safety, and aid in efficient design by choosing design parameters that maximize the reliability of the structure.

During the operation of the structural system, particularly before and after earthquakes, it is important to assess the safety of the structure and make decisions regarding repairs and replacements that are necessary for maintenance, rehabilitation, and structural retrofitting. The quality of the structural components may have degraded over the course of the operation of the structure, and therefore, the load-carrying capacity of the structures may have changed over the course of time. Therefore, it is important to reassess structural safety by recalculating the structural demands and load-carrying limits and reestimating the reliability of the structure.

Since the advantages of probabilistic assessment of structural safety have become evident

during the past two decades, several design guidelines and codes are being revised to incorporate probabilistic analysis. Examples of such revision include the American Institute of Steel Construction Load and Resistance Factor Design (1994) specifications and the European and Canadian structural design specifications. The use of probabilistic analysis in these codes is expected to provide more information about system behavior, the influence of the various uncertain quantities on system performance, and the interaction between the different system components. Therefore, the engineering design community has been increasingly resorting to the use of probabilistic reliability assessment techniques over the past few years. This has further been aided by the advent of high power computing technology that supports the use of advanced computational techniques for probabilistic analysis.

This entry discusses the fundamental concepts of structural reliability analysis in the context of earthquake engineering and provides an overview of the various mathematical tools that may be used to assess the safety and reliability of different types of structural systems. The rest of this entry is organized as follows. The fundamental problem of reliability analysis is formally defined and mathematically presented in section "Structural Reliability Analysis." The various sources of uncertainty are discussed, and the challenges involved in reliability calculation are outlined. Section "Simulation-Based Methods" discusses different types of simulation techniques for structural reliability analysis; while simulation techniques rely on high computational power, they are generally accurate and serve as benchmark solutions for comparison against other computational methods. Section "First-Order Reliability Methods" discusses a class of analytical techniques, popularly known as first-order reliability methods, to calculate structural reliability. While analytical methods are based approximations, they are computationally far cheaper than simulation-based approaches. The aforementioned methods are first illustrated using a structural beam in section "Numerical Example I: Structural Beam" and using a structural

frame in section "Numerical Example 2: Structural Frame." Section "Advanced Concepts in Structural Reliability" reviews a few advanced methods for reliability assessment, thereby providing an overview of the state of the art in the topic of structural reliability analysis.

Structural Reliability Analysis

The fundamental formulation of structural reliability analysis is based on the fact that a structural failure occurs when the loading demand on the structure is greater than its load-carrying capacity. Conventionally, the load-carrying capacity of the structure is equivalent to the resistance offered by the structure and is denoted by R. The loading applied on the structure is denoted by S. The classical formulation of reliability analysis is based on the following equation:

$$G \equiv R - S \qquad (1)$$

If the load-carrying capacity is larger than the applied loading, $G = R - S > 0$, and the structure is considered to be safe. If the load-carrying capacity is smaller than the applied loading, $G = R - S < 0$, and the structure is considered to have failed. The equation $G = 0$ represents the condition when the applied loading is equal to the resistance offered, and this equation is popularly known as the limit state equation in structural reliability analysis. Further, the function G is also referred to as the performance function.

If the loading on the structure (S) were to be known precisely and if the load-carrying capacity (R) of a structure could be estimated precisely, then R and S would become completely deterministic. It would then be possible to design a structure whose load-carrying capacity would be greater than the anticipated loading, and design would become a trivial problem. However, since R and S are mostly uncertain, particularly in earthquake engineering, it becomes challenging to ascertain whether the structure is safe or not.

Sources of Uncertainty

It is important to understand the reason why R and S are uncertain in practical structural systems. The uncertainty in S, i.e., the applied/anticipated loading conditions, can be explained from the fact a given structure does not experience the same amount of loading at all possible time instants. In the context of earthquake engineering, the occurrences of earthquakes and their intensities are random, and therefore, the structural loading that is resultant of these earthquakes is also random.

The resistance offered by a structural member, i.e., R, is also random due to several reasons. The resistance offered by a structural member depends on several quantities, including its material and geometric properties. For example, the maximum bending moment that can be applied on a beam is a function of its yield stress (material property) and the sectional modulus (geometric property). There are two types of uncertainty that contribute to the overall uncertainty in R. First, the material and geometric properties that govern the value of R may themselves be uncertain due to natural variability across nominally identical members of the same type. Second, the functional dependence between these quantities and R may not be precisely known; though a physics-based model may be developed to represent this functional relationship, it may not be accurately representative of the actual carrying capacity of the structure. The latter issue is commonly known as model uncertainty and has been an important topic of research during the past few years.

Some researchers classify the different types of uncertainty into aleatory and epistemic (Der Kiureghian and Ditlevsen 2009). Aleatory uncertainty refers to those types of uncertainty that are inherent in nature and, therefore, irreducible by definition. For example, the variability in material properties, variability in earthquake loading, etc. are all irreducible in nature. Epistemic uncertainty refers to those types of uncertainty that may be reduced when more information is available. For example, when an improved model can be used to predict the response of a structure, then the uncertainty regarding the model prediction would decrease, thereby decreasing the overall uncertainty.
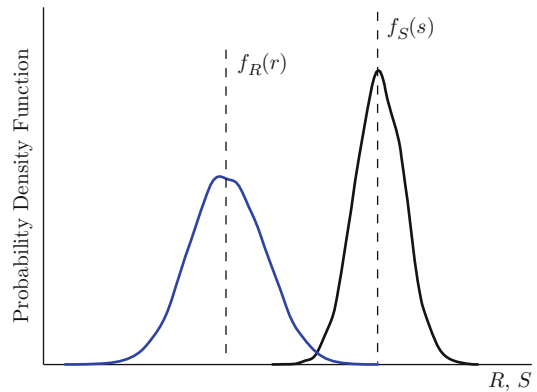
Structural Seismic Reliability Analysis, Fig. 1 Region of overlap between R and S

Nevertheless, it is clear as to why R and S need to be treated as uncertain variables. Therefore, they need to be represented as probability distributions. Let $f_R(r)$ and $f_S(s)$ denote the probability density functions (PDF) of R and S, respectively, as shown in Fig. 1. Note that r and s are generic realizations of the random variables R and S, respectively. While some realizations of R and S may render the structure safe (whenever $r > s$), some other realizations of R and S may correspond to structural failure. The most important aspect of structural reliability analysis is to calculate the probability of structural failure based on the probability distributions of R and S.

Calculation of Failure Probability

From the preceding discussion, it can be easily seen that the probability of the event $G < 0$ is directly equal to the probability of structural failure, and this can be mathematically written as

$$P_f = P(R < S) = P(G < 0) \qquad (2)$$

Using principles of probability, Eq. 2 can be rewritten as (Ditlevsen and Madsen 1996):

$$P_f = \int_{R < S} f_R(r) f_S(S) ds dr \qquad (3)$$

The reliability of the structure is the converse of failure probability and is equal to $1 - P_f$. Typically, good design practices ensure that

$\mu_R > \mu_S$, where μ_R and μ_S denote the means of R and S, respectively. However, due to the uncertainty in R and S, there is small region of overlap between the densities of R and S, and in this region of overlap, structural failure will occur because $R < S$. Therefore, the probability content in this region of overlap is equal to P_f, and in fact, this region is the domain of integration in Eq. 3. The goal in design is to minimize this region of overlap, and this may be accomplished either by moving μ_R and μ_S apart from each other or by reducing σ_R and σ_S, where σ_R and σ_S denote the standard deviations of R and S, respectively. There is a design cost related with each of these activities; while it may be cost-wise prohibitive to eliminate this region of overlap, good design practices advocate minimizing the probability of failure subject to certain cost constraints. That is why it is important to be able to accurately estimate the value of failure probability.

Analytical, closed-form expressions are available for P_f only in the special case when both R and S are normally distributed, i.e., when R and S follow Gaussian distributions as N(μ_R, σ_R) and N(μ_S, σ_S), respectively. In this case, G follows the distribution N$\left(\mu_R - \mu_S, \sqrt{\sigma_R^2 + \sigma_S^2}\right)$, and P_f can be evaluated as

$$P_f = P(G < 0) = \Phi\left(-\frac{\mu_R - \mu_S}{\sigma_R^2 + \sigma_S^2}\right) \qquad (4)$$

In order to delve deeper into this computation, consider the transformation of the Gaussian variables R and S into standard normal distributions (with zero mean and unit standard deviation) as follows:

$$R' = \frac{R - \mu_R}{\sigma_R} \qquad (5)$$

$$S' = \frac{S - \mu_S}{\sigma_S} \qquad (6)$$

Now, both R' and S' follow the Gaussian distribution N(0, 1). This coordinate space is referred to as the reduce coordinate system or the standard normal coordinate system. Rewrite the limit
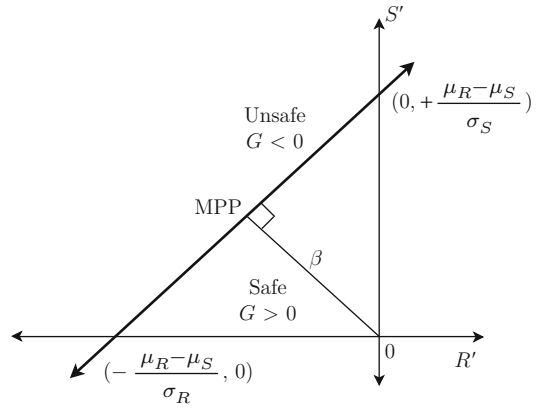
Structural Seismic Reliability Analysis, Fig. 2 Linear performance function

state equation in Eq. 1 in terms of the transformed variables as

$$G \equiv \sigma_R R' - \sigma_S S' + \mu_R - \mu_S = 0 \qquad (7)$$

Since the performance function G is linear, the limit state equation corresponds to a straight line in the $R' - S'$ coordinate axis, as shown in Fig. 2. This straight line divides the plane into two regions. The first region corresponds to $G > 0$ and is known as "safe region," while the second region corresponds to $G < 0$ and is known as the "unsafe region."

It is apparent from Fig. 2 that if the limit state line is closer to the origin in the reduced coordinate system, the failure region is larger, and if it is away from the origin, the failure region is smaller. Hence, it is intuitive that the distance of this straight line to the origin is of importance and this can be computed using basic trigonometry as

$$\beta = \frac{\mu_R - \mu_S}{\sigma_R^2 + \sigma_S^2} \qquad (8)$$

Comparing Eq. 8 and Eq. 4, it can be easily seen that

$$\beta = \Phi^{-1}\left(-P_f\right), \qquad (9)$$

where Φ^{-1} denotes the inverse of the standard normal cumulative distribution function (Haldar and Mahadevan 2000).

The minimum distance point on the limit state is referred to as the checking point or the design point or the most probable point (MPP); since this coordinate axis consists of standard normal variables, the closer a point is to the origin, the higher is the probability of the maximum, and hence, the most probable point is called so. The minimum distance β is popularly known as the Hasofer-Lind Reliability Metric (Hasofer and Lind 1974). The lower the value of P_f, the higher is the value of β and vice versa. Sometimes, in design applications, a target reliability value is specified, and the component or the structure is designed to meet this target value.

Having discussed the computation of failure probability by considering R and S to be Gaussian variables, it is necessary to understand that the probability distributions of R and S may not be readily available, in many structural applications. In fact, they may not even follow Gaussian distributions. It may be necessary to estimate their probability distributions based on the probability distributions of those quantities that influence R and S. An alternate approach is to redefine the limit state directly in terms of these quantities, as explained in the following subsection.

Generalized Definition of the Limit State

Consider that the performance of a structure is defined by generic set of quantities, given by X, where $X = \{X_i, i = 1 \text{ to } n\}$. The loading on the structure (S) and load-carrying capacity (R) can be expressed in terms of these generic set of quantities, and without loss of generality, the limit state in Eq. 1 can be rewritten as

$$G(X) = 0 \qquad (10)$$

The structure is said to be safe when $G(X) > 0$, and the structure is said to have failed when $G(X) > 0$. Limit states in earthquake engineering are of different types and correspond to different types of failures. Limit state related to the serviceability of the structure is used to check whether the deflections of structural

components are within acceptable limits, whether the stresses are lower than yield stress, etc. Limit states related to the safety of the structure are also known as ultimate limit states and are used to check whether accelerations are smaller than acceptable thresholds, whether the stress on the structure is lower than the ultimate stress, etc. Depending upon the criteria of design, a single limit state or a combination of states may be chosen. The majority of this entry deals with situations consisting of single limit states. When multiple limit states need to be simultaneously evaluated, it is necessary to resort to system reliability techniques (Mahadevan et al. 2001; Mahadevan and Raghothamachar 2000), as later explained in section "Advanced Concepts in Structural Reliability" of this entry.

Consider a limit state equation that is based on a specific functional performance of the structure. Using the corresponding limit state equation, the structural failure probability can be calculated as

$$P_f = \int_{G(X)<0} f_X(x)dx, \qquad (11)$$

where $f_X(x)$ denotes the joint probability density function of X. If X consists of independent quantities, then the joint probability density function can be expressed as the product of the individual probability densities of X_i ($i = 1$ to m).

Note that, in general, G may be nonlinear unlike Eq. 1, and therefore, the concepts explained in section "Calculation of Failure Probability" cannot be readily extended to this case. Further, underlying random variables, i.e., X, may not follow Gaussian distributions at all. In order to overcome these challenges, researchers have developed a suite of computational methods to efficiently evaluate the integral in Eq. 11 and compute the failure probability. While some of these methods are based on simulation, other methods are based on linearizing G and, hence, popularly known as first-order reliability methods. These computational methods are discussed in sections "Simulation-Based Methods" and "First-Order Reliability Methods," respectively.

Simulation-Based Methods

The most intuitive approach to structural reliability analysis is to make use of simulation, i.e., repeatedly evaluate the performance function G for multiple realizations of the random variable X, and use the results of these multiple evaluations to calculate the failure probability. Simulation techniques rely on the ability to generate pseudorandom numbers using computer programs and may require several thousand evaluations of G in order to produce meaningful results.

In any simulation-based approach, there are two steps. The first step is to randomly generate samples of X, based on the joint probability density function $f_X(x)$. The second step is to evaluate the function G for each of the randomly generated samples. While the former is a statistical problem, the latter is a modeling problem. Mathematical models (which may either be based on physics or which may be constructed using observed data) can be constructed to represent G, but it is necessary to resort to advanced statistical methods to generate random samples.

Generation of random samples is particularly easy when the various quantities in X are independent of each other. Then, random samples can be generated for each individual X_i. Several computer programs are capable of generating pseudorandom numbers that are uniformly distributed on the interval [0, 1]. Once a pseudorandom number (u_i) can be generated using a computer code, then the corresponding random sample of X_i may be generated by inverting the cumulative distribution function of X_i as

$$x_i = F_{X_i}^{-1}(u_i) \qquad (12)$$

This procedure is repeated multiple times to obtain random samples of X_i. Then, the entire procedure is repeated for all the quantities in X to obtain multiple samples of X.

When the quantities in X are dependent on one another, one approach is to use the full conditional distributions of the quantities in X, i.e., probability densities such as $f_{X_i|X_j}(x_i|x_j)$, where $i \neq j$. Then, the conditional cumulative distribution functions are used in Eq. 12 to successively generate one sample for each X_i. However, such information on conditional distributions may be difficult to obtain, in practice. Therefore, it is necessary to rely on correlation structure of X to generate correlated samples. The most popular approach to generate correlated samples using the covariance matrix of X is, perhaps, using the Nataf model (Liu and Der Kiureghian 1986). Alternatively, Cholesky decomposition (Haldar and Mahadevan 2000) or principal component analysis (Cox and Hinkley 1974) may also be used for such transformation. All of these methods are based on transforming the random variable X from the original, correlated coordinate space to an uncorrelated coordinate space using linear transformation techniques. An appropriate transformation matrix is computed, and using this matrix, it is possible to uniquely transform the variables from the correlated space to the uncorrelated space and vice versa. While generating correlated normal variables, it must be noted that it is challenging to accurately satisfy both the marginal distributions of the individual variables and the covariance matrix. In fact, it is not possible to uniquely define the joint probability density function using information on marginal distributions and the covariance matrix (Bickel and Doksum 1977). It must also be cautioned that correlation is able to capture only one type of dependence among variables, i.e., linear dependence. Other types of dependence cannot be captured using the aforementioned transformation techniques.

Having developed a mechanism to generate samples of X, there are several types of simulation techniques that may be used to evaluate the failure probability. Some of the important techniques are briefly explained in this section.

Monte Carlo Sampling

Monte Carlo sampling (MCS) is the most straightforward approach to evaluate failure probability P_f. First, it is necessary to generate N random samples of X, and then, the function G needs to be evaluated for every random sample X. Recall that structural failure occurs when $G < 0$. Out of N evaluations of G, let N_f denote

the number of simulations in which $G < 0$. Then, an estimate of the probability of failure can be calculated as

$$P_f = \frac{N_f}{N} \qquad (13)$$

Theoretically, an infinite number of samples are necessary to accurately estimate the probability of failure. Since only a finite number of samples are being used here, there is an uncertainty regarding the estimated value of P_f. This uncertainty can be expressed in terms of the variance of P_f as

$$\text{Variance of } P_f = \sqrt{\frac{P_f(1 - P_f)}{N}} \qquad (14)$$

Thus, it can be seen that, for a given number of samples, the accuracy in the estimation of P_f increases proportional to the square root of the number of samples; for example, if the number of samples is quadrupled, then the accuracy is doubled.

The choice of the value of N is very critical in Monte Carlo analysis. As it can be seen from Eq. 14, the accuracy depends on the number of samples used. Further, in many engineering applications, the true probability of failure could be smaller than 10^{-5}. Therefore, on average, only 1 out of 100,000 samples would correspond to failure. Thus, at least 100,000 samples are necessary to observe failure. For a reliable estimate of failure probability, at least 10 times, this minimum (one million samples, in this case) is usually recommended. Sometimes, if G consists of a complex computer code (e.g., a finite element analysis), it may not be computationally feasible to evaluate G a million times. In such cases, it is necessary to resort to other alternative techniques for structural reliability estimation.

Importance Sampling

The basic idea of importance sampling is to deliberately select those samples of X that result in structural failure, instead of sampling them as per their original density $f_X(x)$. For this purpose, a new sampling density function $h_X(x)$ is defined so

that most samples derived out of this density function will lead to failure. Samples are drawn from $h_X(x)$ and corrected based on their original densities, while computing the failure probability. Therefore, $h_X(x)$ is popularly known as the proposal density function.

In order to mathematically derive the expression for failure probability, consider the definition of P_f as per Eq. 11. Note that the domain of integration is specifically over the failure region, defined by $G < 0$. Consider a mathematical indicator function as follows:

$$I_g(x) = \begin{cases} 0, & \text{if} \quad G(x) > 0 \\ 1, & \text{if} \quad G(x) < 0 \end{cases} \qquad (15)$$

Using this indicator function, the expression for failure probability can be rewritten as

$$P_f = \int I_g(x) f_X(x) dx \qquad (16)$$

Note that the domain of integration has now been explicitly rolled into the indicator function and need not be specified as integral limits. Now, multiplying and dividing by $h_X(x)$,

$$P_f = \int \left[I_g(x) \frac{f_X(x)}{h_X(x)} \right] h_X(x) dx \qquad (17)$$

If samples of X are drawn from $h_X(x)$, then the right-hand side of the above equation is simply the expectation of $I_g(x) \frac{f_X(x)}{h_X(x)}$. Such an expectation over N samples can be computed numerically as

$$P_f = \frac{1}{N} \sum_{i=1}^{i=N} I_g(x^i) \frac{f_X(x^i)}{h_X(x^i)} \qquad (18)$$

The accuracy of the estimate of P_f obtained through importance sampling depends on the choice of $h_X(x)$. Several researchers have explored importance sampling in detail and studied its accuracy (Melchers 1989). In fact, the concept of importance sampling is also used in several other engineering applications such as particle filtering (Arulampalam et al. 2002), while tracking the behavior of dynamic systems.

S

Stratified Sampling

In this sampling approach, the overall domain of X is divided into multiple sub-domains, and samples are drawn from each sub-domain independently. The process of dividing the overall domain into multiple sub-domains is referred to as stratification. This method is applicable when subpopulations within the overall population are significantly different and when certain subpopulations contribute more to the failure probability than others.

Let the overall failure domain be divided into m mutually exclusive regions, denoted by Ω_i, where $i = 1$ to m. Each of these regions has its own probability, denoted by $P(\Omega_i)$. The unique feature of this method is that different number of samples may be chosen for each region; let N_i denote the number of samples in each region. Then, the failure probability can be expressed as

$$P_f = \sum_{i=1}^{i=m} \left[P(\Omega_i) \frac{1}{N} \sum_{j=1}^{j=N_i} I_g(x^j) \right], \quad (19)$$

where x^j is the jth sample. An advantage of stratified sampling is that no particular region of the overall domain would be neglected during sampling.

Other Sampling Techniques

While it almost impossible to explain all sampling approaches in detail, a few important techniques are briefly explained in this section, without delving into the details of implementation. Adequate references are provided for all of these techniques to instigate further reading on these topics.

While importance sampling requires far fewer number of samples than Monte Carlo sampling to estimate the failure probability, it is challenging to select an appropriate proposal density function because the region of "importance" (where more samples are likely to indicate structural failure) is usually unknown. **Adaptive sampling** is an advanced sampling technique where the efficiency of importance sampling is continuously improved by updating the proposal density function based on the information obtained after

evaluating G for a few samples. Two classes of adaptive sampling methods are multimodal sampling (Karamchandani et al. 1989) and curvature-based sampling (Wu 1992). It has been reported by researchers that adaptive sampling techniques can accurately estimate the failure probability using 100–400 samples, while traditional Monte Carlo techniques may require several hundreds of thousands of samples.

Recall that Eq. 14 provided an estimate for the variance P_f. Using the so-called variance reduction techniques, it is possible to reduce this variance and thereby obtain an improved estimate of P_f. Such techniques are called variance reduction techniques (Kalos and Whitlock 2008) and are commonly used while estimating P_f. One such technique is popularly called the **conditional expectation method**; in this method, a control variable is selected and the variance of P_f is reduced by removing the random fluctuations of this control variable which was not conditioned. In another technique, popularly known as the technique of **antithetic variates**, negative correlation is purposefully induced between successive samples to decrease the variance of the estimated mean value. It is also common to use the technique of antithetic variates in combination with the conditional expectation method (Haldar and Mahadevan 2000).

There are other sampling methods such as the Latin hypercube sampling (Iman 2008) method or the unscented transform sampling methods (Daigle and Sankararaman 2013). These sampling techniques are more well suited to compute the overall statistics of G, rather than to estimate $P_f = P(G < 0)$. While the Latin hypercube sampling approach focuses on covering the entire domain of X accurately, the unscented transform sampling method focuses on estimating the central moments of G.

Summary

This section discussed a variety of simulation techniques to estimate structural reliability. Topics such as Monte Carlo sampling, importance sampling, and stratified sampling were discussed in detail, and some other advanced topics such as adaptive sampling, variance

reduction techniques in sampling, etc. were briefly explained. The topic of structural reliability estimation using sampling has always been an important research topic, and researchers are constantly developing new methods to mitigate the computational challenges and estimate the failure probability as accurately as possible, using as fewer evaluations of G as possible.

First-Order Reliability Methods

This section discusses a class of methods known as the first-order reliability methods to compute the probability of failure of structural systems. These methods are based on the first-order Taylor's series expansion of the performance function $G(X)$. The first-method, known as the first-order second-moment (FOSM) method, focuses on approximating the mean and standard deviation of G and uses this information to compute P_f. Then, the FOSM method is extended to the advanced FOSM method in two steps: first, the methodology is developed for the case where all the variables in X are Gaussian (normal) and, second, the methodology is extended to the general case of non-normal variables.

First-Order Second-Moment Method

First, consider the generic performance function $G(X)$, and let $f_X(x)$ denote the joint probability density function of X. Recall $X = \{X_i, i = 1 \text{ to } n\}$, and let μ_{X_i} and σ_{X_i} denote the mean and standard deviation of X_i, respectively. Further, the covariance of X_i and X_j is denoted by $Cov(X_i, X_j)$. The first-order second-moment (shortly, referred to as FOSM) method approximates G to be a Gaussian distribution, using only the mean (μ_X) and covariance of X.

Consider the first-order Taylor series expansion of $G(X)$ around μ_X as

$$G = G(\mu_X) + \sum_{i=1}^{i=n} (X_i - \mu_{X_i}) \left(\frac{\partial G}{\partial X_i}\right)_{\mu_X} \quad (20)$$

Note that G is now a linear function of X with the partial derivatives as coefficients, and

therefore, it is straightforward to approximate its mean and variance as

$$\mu_G = G(\mu_X) \quad (21)$$

$$\sigma_G^2 = \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} \left(\frac{\partial G}{\partial X_i}\right)_{\mu_X} \left(\frac{\partial G}{\partial X_j}\right)_{\mu_X} Cov(X_i, X_j) \quad (22)$$

When the inputs to G are uncorrelated, then the expression for variance in Eq. 22 simplifies to

$$\sigma_G^2 = \sum_{i=1}^{i=n} \left(\frac{\partial G}{\partial X_i}\right)_{\mu_X}^2 \sigma_{X_i}^2 \quad (23)$$

Once the mean and standard deviation of G is calculated, then G is approximated as a Gaussian distribution, with mean and standard deviation σ_G. Then, the failure probability can be calculated as

$$P_f = P(G < 0) = \Phi\left(-\frac{\mu_G}{\sigma_G}\right) \quad (24)$$

It is obvious that the failure probability calculated using the FOSM approach is accurate only in two cases: (1) when G is a linear function of X and the quantities in X are statistically independent normal random variables or (2) when G is a multiplicative function of X and the quantities in X are statistically independent lognormal random variables. In many practical examples, it is unlikely that all the variables are statistically independent normals or lognormals. Nor it is likely that G is an additive or multiplicative function of X. In such cases, Eq. 24 is only approximate; however, it can be used to provide a rough idea of the level of risk or reliability.

Nevertheless, the FOSM approach has two important deficiencies. First, the method does not use the information regarding the variable X when available. The function G in Eq. 20 is linearized at the mean of X. When G is nonlinear, significant error may be introduced by neglecting higher order terms. Second, and more importantly, the P_f calculated using Eq. 24 does not remain invariant for

different but mechanically equivalent formulations of the same performance function. For example, the following performance functions G_1 and G_2 are mechanically equivalent:

$$G_1 \equiv X_1 - X_2 \qquad (25)$$

$$G_2 \equiv \frac{X_1}{X_2} < 1 \qquad (26)$$

However, the FOSM approach, when implemented using these two performance functions, yields different values for $\frac{\mu_G}{\sigma_G}$ and hence different values of P_f. In other words, the same engineering problem can be formulated either in terms of strength or stress, but the FOSM approach should not lead to different results for different formations. The lack of invariance is a serious problem and was recognized by the researchers in 1970s, and in order to overcome this problem, Hasofer and Lind (Hasofer and Lind 1974) proposed the advanced FOSM method.

Advanced FOSM for Normal Variables

First, consider the case where the quantities in X are independent normal variables. While it may be meaningful to use Taylor's series expansion to linearize G, there is no rationale behind choosing the point of linearization to be the mean of X. It is necessary to choose a point of linearization such that the estimate P_f obtained after such linearization is directly related to the probability of failure of the structure.

Point of Linearization: Most Probable Point
Any point lying on the curve of demarcation would satisfy the equation $G(X) = 0$. Since (1) this curve serves as the demarcation between the two zones given by $G > 0$ and $G < 0$ and (2) it is of interest to calculate the probability $P(G < 0)$, it is intuitive that it is important to identify a linear function which closely resembles the contour $G(x) = 0$. Hence, the point of linearization must lie on this curve of demarcation; in other words, the point of linearization must satisfy the equation $G(x) = 0$. This is

clearly different from the FOSM approach, where the mean μ_X was chosen as the point of linearization; in general, the mean μ_X will not satisfy this equation.

Therefore, the point of linearization should be located on the curve of demarcation. However, there are infinite points that satisfy this criterion, and it is important to select the appropriate one. Each of these infinite points has a likelihood of occurrence, and intuitively, the *point of maximum likelihood* is chosen as the point of linearization. This likelihood can be calculated using the probability density function of the underlying random variables. For a single normal random variable X with mean μ and standard deviation σ, the PDF is given by

$$f_X(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad (27)$$

For example, when $\mu = 10$ and $\sigma = 1, x = 10$ is 1.65 times more likely to occur than $x = 9$. The maximum value of the likelihood function occurs at $x = \mu$; therefore, the farther x is away from the mean μ, the lower the likelihood of occurrence of x. If there is only one random variable X instead of X, the mean μ_X cannot be chosen as the point of linearization since $G(\mu_X) \neq 0$, in general. Therefore, if there is a single input variable, the point of linearization is chosen in such a way that it satisfies the equation $G(x) = 0$ and the value of $\|x - \mu_X\|$ is minimum.

However, in a general structural reliability analysis problem, the input to G is a vector, i.e., $X = \{X_1, X_2, \ldots X_i, \ldots X_n\}$, and each X_i has its own mean μ_{X_i} and standard deviation σ_{X_i}. The objective is to identify the point of maximum likelihood, which can be calculated by maximizing the joint probability density function of all the input random variables. If the variables are independent, then the joint density function of X is expressed as (when the variables are independent)

$$f_X(x) = \prod_{i=1}^{i=n} \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left[-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}\right] \qquad (28)$$

It can be verified (by taking the logarithm) that the maximizer of the above function simultaneously minimizes

$$\delta = \sum_{i=1}^{i=n} \left(\frac{x_i - \mu_i}{\sigma_i}\right)^2 \qquad (29)$$

Eq. 29 can rewritten as

$$\delta = \sum_{i=1}^{i=n} u_i^2, \qquad (30)$$

where

$$u_i = \frac{x_i - \mu_i}{\sigma_i} \qquad (31)$$

If the above computation were performed for every realization x_i of the random variable X_i, then the corresponding u_i's would be realizations of the standard normal variable U_i, i.e., $U_i \sim N(0,1)$. Therefore, Eq. 31 is referred to as the standard normal transformation, similar to that in Eq. 5 and Eq. 6. In the space containing standard normal variables, maximizing the likelihood of occurrence is equivalent to minimizing Eq. 30, which implies that the point of linearization is that point on the curve of demarcation, whose distance (measured in the standard normal space) from the origin is minimum. Since the point of linearization has the maximum likelihood of occurrence, it is popularly known as the Maximum Probable Point (MPP), as indicated in Fig. 3.

Computing Failure Probability

The next task is to expand $G(X)$ around the MPP and make use of the linear approximation. Let $x^* = \{x_i^*, i = 1 \text{ to } n\}$ denote the most probable point in the original coordinate space and $u^* = \{u_i^*, i = 1 \text{ to } n\}$ denote the corresponding point in the standard normal space. Then,

$$G = G(x^*) + \sum_{i=1}^{i=n} \left(X_i - x_i^*\right) \left(\frac{\partial G}{\partial X_i}\right)_{x^*} \qquad (32)$$

Now. G is a linear combination of Gaussian variable X and, therefore, can be approximated to
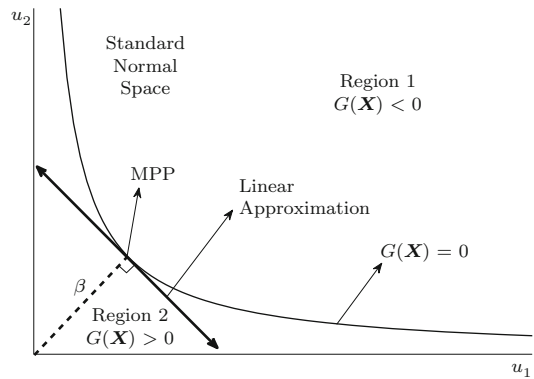
Structural   Seismic   Reliability   Analysis, Fig. 3 Estimating MPP in FORM

be Gaussian with mean μ_G^* and standard deviation σ_G^*. The mean can be calculated as

$$\mu_G^* = \sum_{i=1}^{i=n} \left(\mu_{X_i} - x_i^*\right) \left(\frac{\partial G}{\partial X_i}\right)_{x^*} \qquad (33)$$

By definition of the standard normal transformation, it can be easily seen that (for all $i = 1$ to n)

$$x_i^* = \mu_{X_i} + u_i^* \sigma_{X_i} \qquad (34)$$

$$\frac{\partial u_i}{\partial x_i} = \frac{1}{\sigma_{X_i}} \qquad (35)$$

$$\left(\frac{\partial G}{\partial X_i}\right)_{x^*} = \left(\frac{\partial G}{\partial u_i}\right)_{u^*} \frac{\partial u_i}{\partial x_i} = \frac{1}{\sigma_{X_i}} \left(\frac{\partial G}{\partial u_i}\right)_{u^*} \qquad (36)$$

Substituting Eq. 34 and 36 in Eq. 33,

$$\mu_G^* = \sum_{i=1}^{i=n} \left(\mu_{X_i} - x_i^*\right) \left(\frac{\partial G}{\partial X_i}\right)_{x^*} \qquad (37)$$

Substituting Eq. 34 and 36 in Eq. 37,

$$\mu_G^* = \sum_{i=1}^{i=n} - u_i^* \times \left(\frac{\partial G}{\partial u_i}\right)_{u^*} \qquad (38)$$

Note that $\left(\frac{\partial G}{\partial u_i}\right)_{u^*}$ refers to the ith term of the gradient vector ($\alpha^* = \{\alpha_i^* - i = 1 \text{ to } n\}$) of G in

the standard normal space. Therefore, Eq. 38 can be rewritten as:

$$\mu_G^* = -(\boldsymbol{u}^*)^T(\boldsymbol{\alpha}^*) \qquad (39)$$

Note that the vectors represented by $\boldsymbol{u}^*$ and $\boldsymbol{\alpha}^*$ are both perpendicular to the limit state equation and hence parallel to each other. However, they are of opposite directions; while $\boldsymbol{u}^*$ is directed away from the origin, the gradient vector $\boldsymbol{\alpha}^*$ is directed towards the origin, i.e., in the direction of increasing value of G. Therefore, the above vector multiplication is basically a dot product of collinear but opposite vectors, and hence,

$$\mu_G^* = \|\boldsymbol{u}^*\| \times \|\boldsymbol{\alpha}^*\| \qquad (40)$$

Similarly, based on Eq. 32, the variance of G, denoted by $(\sigma_G^*)^2$, is calculated as

$$(\sigma_G^*)^2 = \sum_{i=1}^n \sigma_{X_i}^2 \times \left[\left(\frac{\partial G}{\partial X_i}\right)_{x^*}\right]^2 \qquad (41)$$

Substituting Eq. 36 in Eq. 41,

$$\begin{aligned}(\sigma_G^*)^2 &= \sum_{i=1}^n \sigma_{X_i}^2 \times \left[\left(\frac{\partial G}{\partial X_i}\right)_{x^*}\right]^2 \left[\frac{\partial u_i}{\partial x_i}\right]^2 \\ &= \sum_{i=1}^n \left[\left(\frac{\partial G}{\partial u_i}\right)_{x^*}\right]^2 \end{aligned} \qquad (42)$$

Therefore,

$$\sigma_G^* = \sqrt{\sum_{i=1}^n \left[\left(\frac{\partial G}{\partial u_i}\right)_{x*}\right]^2} = \|\boldsymbol{\alpha}^*\| \qquad (43)$$

Having estimated μ_G^* and σ_G^*, the reliability index and the value of P_f can be calculated as

$$\beta = \frac{\mu_G^*}{\sigma_G^*} = \|\boldsymbol{u}^*\| \qquad (44)$$

$$P_f = \Phi(-\beta) \qquad (45)$$

Therefore, the reliability index is simply equal to the minimum distance measured from the

> Given PDFs of $\boldsymbol{X}$
> Minimize $\beta = \boldsymbol{u}^T \boldsymbol{u}$
> such that $G(\boldsymbol{x}) = 0$
> where standard normal $\boldsymbol{u} = T(\boldsymbol{x})$
> $P(G < 0) = \Phi(-\beta)$

Structural   Seismic   Reliability   Analysis, Fig. 4 Optimization in advanced FOSM

origin to the curve represented by the limit state equation, in the standard normal space. Note that Eq. 39 and Eq. 44 are correct only when $\beta > 0$, i.e., $P_f < 0.5$. Otherwise, the right-hand sides of these two equations need to be negated. Since the failure probability of structures is usually less than 0.1, the above equations hold in normal circumstances. The most general expression (appropriate sign incorporated) for β can be obtained by dividing the right-hand side of Eq. 39 by that of Eq. 43.

Algorithm for Structural Reliability Analysis

Having discussed the theory behind the advanced FOSM, the method is now presented as an algorithm to aid practitioners. The key of the advanced FOSM method is to identify the MPP by solving an optimization problem, as shown in Fig. 4.

In this optimization problem, the focus is to select that point on the limit state equation that is closest to the origin, in the standard normal space. In Fig. 4, T represents the standard normal transformation function from the original space ($\boldsymbol{x}$) to the standard normal space ($\boldsymbol{u}$). This optimization is solved using the Rackwitz-Fiessler (Fiessler et al. 1979) algorithm, an iterative procedure, as follows:

1. Initialize counter $j = 0$ and start with an initial guess for the most probable point (MPP), i.e., $\boldsymbol{x}^j = \left\{x_1^j, x_2^j, \ldots x_i^j, \ldots x_n^j\right\}$, a column vector.
2. Transform into standard normal space and calculate $\boldsymbol{u}^j = \left\{u_1^j, u_2^j, \ldots u_i^j, \ldots u_n^j\right\}$ using Eq. 31, a column vector.

3. Compute the gradient vector in the standard normal space, i.e., $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \ldots \alpha_n\}$, another column vector where

$$\alpha_i = \frac{\partial G}{\partial u_i} = \frac{\partial G}{\partial x_i} \times \frac{\partial x_i}{\partial u_i} = \frac{\partial G}{\partial x_i} \times \sigma_i \qquad (46)$$

4. In the iterative procedure, the next point u^{j+1} is calculated using the following equation:

$$\boldsymbol{u}^{j+1} = \frac{1}{\|\boldsymbol{\alpha}\|} \left[\boldsymbol{\alpha}^T \boldsymbol{u}^j - G(\boldsymbol{x}^j) \right] \frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|} \qquad (47)$$

5. Transform back into original space, i.e., compute $\boldsymbol{x}^{j+1}$, and continue starting from Step 3 until the iterative procedure converges. Using tolerance limits δ_1 and δ_2, convergence can be checked if the following two criteria are satisfied: (i) the point lies on the curve of demarcation, i.e., $\left| G(\boldsymbol{x}^j) \right| \leq \delta_1$, and (2) the solution does not change between two iterations, i.e., $\left| \boldsymbol{x}^{j+1} - \boldsymbol{x}^j \right| \leq \delta_2$.

Note that, since this approach is gradient based, the gradient vector at MPP is an indicator of which sources of uncertainty are the strongest contributors to structural failure. The higher the magnitude of the gradient in a direction that corresponds to a particular uncertain variable, the more important is that variable in the context of structural safety.

Until now, the discussion did not account for statistical among variables. If there is any statistical dependence, then it is necessary to transform the variables into uncorrelated standard normal space. The same type of transformations discussed earlier in section "Simulation-Based Methods" may be used for this purpose.

Advanced FOSM for Non-normal Variables

Now consider the case where the inputs X_i ($i = 1$ to n) have arbitrary probability distributions given by their CDFs as $F_{X_i}(x_i)$ ($i = 1$ to n). Now that X_i is not normally distributed, Eq. 31 cannot be used for standard normal transformation. Therefore, it is necessary to calculate u_i from a given x_i meaningfully, so that u_i represents

a realization of the standard normal variable. The only difference from the implementation of the advanced FOSM algorithm is the transformation step and the gradient computation which is dependent on the choice of transformation.

One simple transformation is based on probability integral transform concept as

$$u_i = \Phi^{-1}(F_{X_i}(X_i = x_i)), \qquad (48)$$

where $\Phi^{-1}(.)$ refers to the inverse of the standard normal distribution function (Haldar and Mahadevan 2000). Now, the calculation of the gradient in the standard normal space is different from Eq. 46 and can be derived directly using Eq. 48. First, decompose Eq. 48 into two parts as

$$v_i = F_{X_i}(X_i = x_i) \qquad (49)$$

$$u_i = \Phi^{-1}(v_i) \qquad (50)$$

Then, each element of the gradient vector $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \ldots \alpha_n\}$ can be calculated as

$$\alpha_i = \frac{\partial G}{\partial u_i} = \frac{\partial G}{\partial x_i} \times \frac{\partial x_i}{\partial v_i} \times \frac{\partial v_i}{\partial u_i} = \frac{\partial G}{\partial x_i} \times \frac{\phi(u_i)}{f_{X_i}(x_i)}, \qquad (51)$$

where $\phi(.)$ refers to the standard normal density function and $f_{X_i}(x_i)$ is the PDF of the ith input variable X_i.

In addition to the above procedure, there are also other transformation techniques. For example, a two-parameter transformation procedure estimates the mean μ_i and standard deviation σ_i of the normal distribution by equating the CDF and PDF values of the distribution of X and the normal distribution. Then, Eq. 31 can be used to calculate u_i from x_i. Note that the mean μ_{X_i} and standard deviation σ_{X_i} are dependent on the value of x_i. Similarly, Chen and Lind (Chen and Lind 1983) proposed a three-parameter transformation procedure by introducing a third parameter, a scale factor which is estimated by matching the slope of the probability density function in addition to the PDF and CDF values. Further, when

S

the inputs are correlated or statistically dependent, it is necessary to transform them to *uncorrelated* standard normal space. Haldar and Mahadevan (Haldar and Mahadevan 2000) describe methods for such transformation. It must be noted that any transformation must be accompanied by suitably computing the derivatives in the standard normal space, and Eq. 46 must be appropriately replaced.

Sometimes, the variable X may follow arbitrary distributions and be correlated. In that case, it is still necessary to transform these variables to uncorrelated standard normal space (Liu and Der Kiureghian 1986; Haldar and Mahadevan 2000). This is usually performed in two steps: first, the variables are converted to uncorrelated space and then transformed to standard normal space where optimization is performed to estimate MPP.

Summary

The section discussed the use of first-order reliability methods in order to estimate the reliability of structures. First, the first-order second-moment (FOSM) method was presented and then extended to the advanced FOSM method. The concept of most probable point (MPP) was introduced. It was derived that the distance from the origin to the MPP, in standard normal space, is equal to the safety index or reliability index, denoted by ft. Information regarding the gradient at the MPP can be used to identify the sources of uncertainty that are significant contributors to the failure of the structure.

The inverse of the advanced FOSM algorithm is commonly employed in design. This procedure is known as inverse-FORM and is used to select a design parameter (which is input as unknown quantity to the performance function G) so that a target reliability index may be attained. Details of the inverse-FORM methodology can be found in several research articles (Der Kiureghian et al. 1994) and textbooks (Haldar and Mahadevan 2000).

Numerical Example I: Structural Beam

This section illustrates the estimation of structural reliability through an illustrative example that has been extended from a case study discussed by Haldar and Mahadevan (Haldar and Mahadevan 2000). This numerical example consists of a steel beam that is suggested to carry an applied deterministic bending moment M that follows a Gaussian distribution whose mean is equal to 1,500 kip-in and standard deviation is equal to 75 kip-in. The yield stress (Y) of steel is considered to be a lognormal variable with mean equal to 38 ksi and standard deviation equal to 3.8 ksi. The plastic section modulus (Z) is known to be a normal random variable with mean equal to 54 in^3 and standard deviation equal to 2.7 in^3. The goal is to compute the structural reliability of this beam.

The resistance offered by this beam (maximum loading possible) can be expressed as the product of yield stress (Y) and plastic section modulus (Z). The structure will fail if the resistance is smaller than the applied bending moment (M). Therefore, the limit state equation is

$$G = YZ - M \qquad (52)$$

The structure is said to fail when $G < 0$.

First, the problem is solved using Monte Carlo simulation. It is possible to directly generate samples of M and Z since they follow Gaussian distributions. However, in order to generate samples from the lognormally distributed Y, its distribution parameters (the mean and standard deviation of the corresponding normal distribution) need to be estimated. The location parameter λ_Y is calculated to be equal to 3.632611 and the scale parameter ζ_Y is calculated to be equal to 0.0997513. It is trivial to code Monte Carlo simulation is a programming environment such as MATLAB. For this numerical example, the MATLAB codes would be

1. N = 50,000,000
2. M = randn(N,1)*75 + 1,500
3. Z = randn(N,1)*2.7 + 54
4. Y = icdf('logn',rand (N,1),3.632611,0.0997513)
5. G = Y.*Z-M
6. Nf = length(find(G < 0))
7. Pf = Nf/N

Structural Seismic Reliability Analysis, Table 1 Advanced FOSM method: implementation

Quantity	Iteration I	Iteration II	Iteration III
Y	38	30.6	30.81
Z	54	51.4	51.1
M	1,500	1,552.5	1,573.7
G	552	19.21	−0.08
μ_Y	37.81	37.07	37.11
μ_Z	54	54	54
μ_M	1,500	1,500	1,500
σ_Y	3.79	3.05	3.07
σ_Z	2.7	2.7	2.7
σ_M	75	75	75
u_Y	0.05	−2.13	−2.05
u_Z	0	−0.96	−1.08
u_M	0	0.7	0.98
α_Y	54	51.41	51.08
α_Z	38	30.57	30.81
α_M	−1	−1	−1
α_{u_Y}	204.69	156.78	156.97
α_{u_Z}	102.06	82.53	83.18
α_{u_M}	−75	−75	−75
New u_Y	−1.9	−2.05	−2.05
New u_Z	−0.96	−1.08	−1.09
New u_M	0.70	0.98	0.98
β	2.24	2.52	2.52

As seen from the above code, a very large number of samples have been used and the resultant P_f is calculated to be equal to 0.0059. Since the number of samples is high, this serves a benchmark solution to verify the solutions from FOSM and advanced FOSM methods.

The same numerical example can also be solved using the first-order second-moment method, by linearizing the limit state equation at the mean of the variables. The distribution information of the variables is not used. In this approach, $\mu_G = 38 \times 54 - 1,500 = 552$, and $\sigma_G^2 = (3.8 \times 54)^2 + (2.7 \times 38)^2 + (75)^2$ which yields $\sigma_G = 241.37$. Therefore, $\beta = 2.287$, and hence, $P_f = 0.011$. Evidently, this is extremely erroneous because of the inaccuracy of the FOSM method.

Finally, the numerical example is solved using the advanced FOSM method, by estimating the most probable point (MPP), as shown in Table 1. The initial guess for the MPP is $Y = 38$, $Z = 54$, and $M = 1,500$. In this table, μ_Y, μ_Y, μ_M, and σ_Y, σ_Y, σ_M denote the equivalent normal mean and standard deviation in each iteration. Obviously, since Z and M are originally normal variables, their means and standard deviations do not change. The coordinate values in the standard normal space are denoted by u_Y, u_Z, and u_M. The derivatives in the original space are denoted by σ_Y, σ_Z, and σ_M, and these are multiplied by the equivalent normal standard deviation to calculate the derivatives in the standard normal space, denoted by σ_{u_Y}, σ_{u_Z}, and σ_{u_M}, respectively. The coordinate values after the derivative-based update is given by "New u_Y," "New u_Z," and "New u_M." These are transformed back to the original space using the previously computed equivalent normal mean and standard deviation, and the new values of Y, Z, and M are the starting values for the next iteration. This procedure is continued until the optimization converges.

It is seen that, after three iterations, the value of G is close to zero, and there is no significant change in the value of $\beta = 2.52$ that corresponds to $P_f = 0.0059$, which is in excellent agreement with the Monte Carlo simulation approach. Note that the FOSM approach required only 12 evaluations of G to estimate this failure probability. As explained earlier, the gradient in the standard normal space is an indicator of those sources of uncertainty that are the strongest contributors to structural failure. In this case, the yield stress (Y) is found to be the most important contributor since it has the highest gradient.

Numerical Example 2: Structural Frame

This section illustrates the calculation of structural reliability using the model of a simple two-story frame, as shown in Fig. 5.

This two-story frame has six parameters, m_1, m_2, k_1, k_2, D_1, D_2, that represent the masses, stiffness, and damping parameters of the two stories of the structural frame. Except the masses ($m_1 = 136$ and $m_2 = 66$ kg), all of the other quantities are uncertain. Further, the inputs to the system are forces (lateral excitations) at the two levels given by $A_1\sin(\omega_1 t)$ and $A_2\sin(\omega_1 t)$.
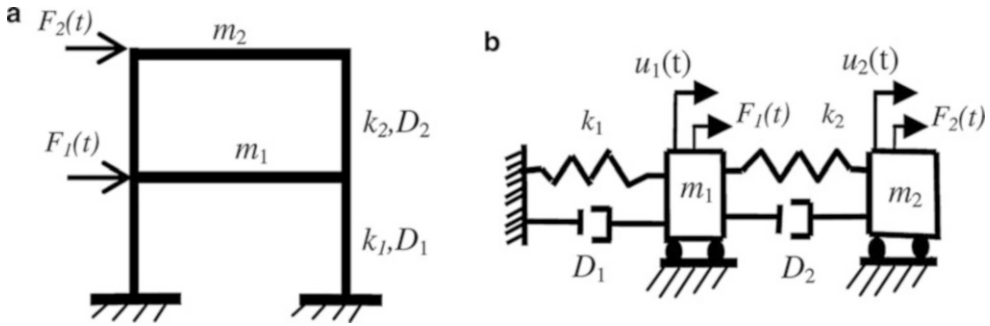
Structural Seismic Reliability Analysis, Fig. 5 (**a**) Structural frame, (**b**) mechanical model

Structural Seismic Reliability Analysis, Table 2 Statistics of uncertain quantities

Parameter	Value	Std. Dev.	Unit
k_1	30,700	1,500	N/m
k_2	44,300	2,000	N/m
D_1	307	30	Ns/m
D_2	443	40	Ns/m
A_1	75	5	N
A_2	100	5	N
ω_1	9	0.5	s^{-1}
ω_2	9	0.5	s^{-1}

All the uncertain quantities are assumed to follow Gaussian distributions, and their statistics are provided in Table 2.

In this example, the maximum displacement of the second story during the first 5 s of loading is the quantity of interest, and the structure is said to have failed due to serviceability reasons if this maximum displacement exceeds 35 mm.

First, the response of the system (u_2, in this case) needs to be computed based on the fundamental equations:

$$m_1 \ddot{u}_1 + (D_1 + D_2)\dot{u}_1 - D_2 \dot{u}_2 + (k_1 + k_2)u_1 \\ - k_2 u_2 = A_1 \sin(\Omega_1 t)$$

(53)

$$m_2 \ddot{u}_2 - D_2 \dot{u}_1 + D_2 \dot{u}_2 - k_2 u_1 - k_2 u_2 = A_2 \sin(\Omega_2 t)$$

(54)

Once u_2 is computed, its maximum value is computed, and the limit state G is defined as

$$G = \max(u_2) - 0.035$$

(55)

The probability of failure is directly evaluated using Monte Carlo sampling (10,000 samples), and P_f is observed to be equal to 0.0028.

Note that sinusoidal loading functions were used in this example so that the response of structural frame may be easily computed using well-known equations structural dynamics. In order to study the response of a structure to an earthquake, the appropriate earthquake loading needs to be used, and the equations of structural dynamics need to be solved using numerical techniques. Then the maximum deflection would be calculated and used to construct the limit state for reliability analysis. Since the scope of this article is to familiarize the readers with reliability methods, simpler numerical examples were considered for illustrative purposes.

Advanced Concepts in Structural Reliability

In addition to the first-order reliability methods and simulation-based techniques, there are other types of techniques and methods that have become popular for structural reliability analysis, over the past two decades. The purpose of this section is to provide an overview of some of these approaches and list appropriate references that would aid in-depth understanding of these methods.

Second-Order Reliability Methods

Recall that $G(X)$ is a generic function and the class of first-order reliability methods focus on linearizing G using the first-order Taylor's series expansion. Several researchers have developed computational methods to improve the estimates provided by the first-order reliability methods. In particular, second-order reliability methods focus on quadratic approximations of the limit state (Fiessler et al. 1979). Breitung (1984) estimated closed-form analytical expressions for P_f, using the principal curvatures of the limit state at the MPP. While Breitung's method was based on a parabolic approximation, Tvedt (Tvedt 1990) developed a generalized second-order approximation to the limit state to compute the failure probability. Der Kiureghian et al. (1987) approximated the limit state using two different semi-parabolas around the MPP and used the analytical expressions developed by Breitung (1984).

Subset Simulation

The topic of subset simulation addresses structural reliability calculation using sampling. Recall that, since failure probabilities are generally small, it is necessary to consider a large number of samples to estimate the failure probability accurately. The basic idea of subset simulation is to express the failure probability as a product of larger conditional failure probabilities by introducing intermediate failure events (Au and Beck 2001). As a result, the original problem of calculating a small failure probability, which is computationally demanding, is reduced to calculating a sequence of conditional probabilities, which can be readily and efficiently estimated by means of simulation. This approach has been applied to structural reliability analysis of frames subjected to seismic excitation (Au and Beck 2001).

Surrogate Modeling

Another class of methods for structural reliability analysis relies on approximating the performance function $G(X)$ using different types of mathematical tools. This class of methods is referred to as surrogate modeling techniques (sometimes, as response surface methods) since they use a few evaluations of $G(X)$ (referred to as training points) to construct a mathematical function that approximates the original G. Obviously, the surrogate model will not be able to match the value of G at values of X, and therefore, this imparts additional uncertainty to the problem. Once a surrogate model is constructed, then Monte Carlo simulation may be used to compute the failure probability. Since the surrogate model is simple to evaluate, it is easy to use a million samples of X during Monte Carlo sampling.

Commonly used surrogate modeling approaches include regression techniques (Haldar and Mahadevan 2000), polynomial chaos expansion (Najm 2009), kriging (Stein 1999), etc. Each of these methods uses different types of basis functions, and one may approximate G better than the other, and it is necessary to choose a suitable surrogate model based on the application of interest.

Efficient Global Reliability Analysis

While conventional surrogate modeling approaches focus on approximating the performance function $G(X)$ over the entire domain of X, the technique of efficient global reliability analysis (Bichon et al. 2008) argues that it is not necessary for such approximation. It is sufficient to approximate the function $G(X)$ near the limit state equation, i.e., around the region where $G(X) = 0$. This method uses a few training points that lie near the curve represented by the limit state equation to construct a Gaussian process surrogate model (Rasmussen 2004) and continues updating this surrogate model using additional training points until the resultant surrogate model sufficiently approximates the performance function $G(X)$ around the curve represented by the limit state equation. Finally, Monte Carlo simulation can be used along with the final surrogate model; it is sufficient to know whether $G > 0$ or $G < 0$, and the numerical value of G is not significant. Since this method approximates G only near the limit state curve, it is well suited to find the sign of G and hence provides estimates of P_f with reasonable accuracy.

S

System Reliability Methods

Sometimes, a structural system may consist of multiple structural components, each of which has its own limit state. In a series system, the failure of any one component implies the failure of the entire structural system. In a parallel system, it is imperative that all the components fail individually to imply that the system has failed. The probability of system failure can be expressed as a union of component-level failures in the former case, while it is expressed as an intersection of component-level failures in the latter case. Methods for predicting system reliability have been studied by several researchers and documented in several research articles (Hohenbichler and Rackwitz 1983; Cruse et al. 1994) and textbooks (Ditlevsen and Madsen 1996; Haldar and Mahadevan 2000). Sometimes, even a single structural component may have multiple limit states; methods for system reliability methods are applicable even to such situations since it is necessary to evaluate probability of union or intersection of different events that correspond to failure across multiple limit states.

Summary

This entry introduced the concept of structural reliability analysis, in the context of earthquake engineering, and reviewed several fundamental concepts that may be used to assess the safety of structural systems. There are several sources of uncertainty that affect the performance of structural systems, and therefore, the safety of structural systems is uncertain. It is important to constantly perform structural safety assessment for the purposes of analysis and design and estimate the probability that the structure may fail due to the applied loading. In general, structural failure occurs when the applied loading is greater than the load-carrying capacity (resistance) of the structure. Since both the load-carrying capacity and the actual loading may be uncertain, the safety of the structure also becomes uncertain. This concept was explored in detail by considering Gaussian distributions for the loading and the resistance quantities, mathematical expressions for failure probability were derived, and the concepts of limit state and reliability index were introduced.

In many structural systems, the loading and the resistance may be functions of various quantities such as material and geometric properties, and this led to the explanation of the generalized limit state function. The quantification of failure probability was explained in detail using simulation-based approaches and first-order reliability methods. These methods were also illustrated using two different numerical examples, one consisting of a beam and the other consisting of a structural system. Finally, an overview of advanced reliability concepts such as the second-order reliability method, efficient global reliability analysis, system reliability techniques, etc. was provided, thereby explaining the state of the art in the field of structural reliability analysis.

Cross-References

▶ Reliability Analysis of Nonlinear Vibrating Systems-Spectral Approach
▶ Reliability Estimation and Analysis
▶ Response Variability and Reliability of Structures
▶ Seismic Reliability Assessment, Alternative Methods for
▶ Structural Reliability Estimation for Seismic Loading

References

Arulampalam MS, Maskell S, Gordon N, Clapp T (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans Signal Process 50(2):174–188

Au S, Beck JL (2001) Estimation of small failure probabilities in high dimensions by subset simulation. Probab Eng Mech 16(4):263–277

Bichon BJ, Eldred MS, Swiler LP, Mahadevan S, McFarland JM (2008) Efficient global reliability analysis for nonlinear implicit performance functions. AIAA J 46(10):2459–2468

Bickel PJ, Doksum KA (1977) Mathematical statistics: basic ideas and selected topics. Holden Day, San Francisco

Breitung K (1984) Asymptotic approximations for multinormal integrals. J Eng Mech 110(3):357–366

Chen X, Lind NC (1983) Fast probability integration by three-parameter normal tail approximation. Struct Saf 1(4):269–276

Cox DR, Hinkley DV (1974) Theoretical statistics. Chapman and Hall, London

Cruse TA, Mahadevan S, Huang Q, Mehta S (1994) Mechanical system reliability and risk assessment. AIAA J 32(11):2249–2259

Daigle M, Sankararaman S (2013) Advanced methods for determining prediction uncertainty in model-based prognostics with application to planetary rovers. In: Annual conference of the Prognostics and Health Management Society, New Orleans, pp 262–274

Der Kiureghian A (1996) Structural reliability methods for seismic safety assessment: a review. Eng Struct 18(6):412–424

Der Kiureghian A, Ditlevsen OD (2009) Aleatory or epistemic? Does it matter? Struct Saf 31(2):105–112

Der Kiureghian A, Lin HZ, Hwang SJ (1987) Second-order reliability approximations. J Eng Mech 113(8):1208–1225

Der Kiureghian A, Zhang Y, Li CC (1994) Inverse reliability problem. J Eng Mech 120(5):1154–1159

Ditlevsen OD, Madsen HO (1996) Structural reliability methods. Wiley, Chichester

Fiessler B, Rackwitz R, Neumann HJ (1979) Quadratic limit states in structural reliability. J Eng Mech Div 105(4):661–676

Haldar A, Mahadevan S (2000) Probability, reliability, and statistical methods in engineering design. Wiley, New York

Hasofer AM, Lind NC (1974) Exact and invariant second-moment code format. J Eng Mech Div 100(1):111–121

Hohenbichler M, Rackwitz R (1983) First-order concepts in system reliability. Struct Saf 1(3):177–188

Iman RL (2008) Latin hypercube sampling. Wiley Online Library

Kalos MH, Whitlock PA (2008) Monte carlo methods. Wiley, Weinheim

Karamchandani A, Bjerager P, and Cornell, AC (1989) Adaptive Importance Sampling, Proceedings, International Conference on Structural Safety and Reliability (ICOSSAR), San Francisco, pp 855–862

Liu P, Der Kiureghian A (1986) Multivariate distribution models with prescribed marginals and covariances. Probab Eng Mech 1(2):105–112

Mahadevan S, Raghothamachar P (2000) Adaptive simulation for system reliability analysis of large structures. Comput Struct 77(6):725–734

Mahadevan S, Zhang R, Smith N (2001) Bayesian networks for system reliability reassessment. Struct Saf 23(3):231–251

Melchers RE (1989) Importance sampling in structural systems. Struct Saf 6(1):3–10

Najm HN (2009) Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. Annu Rev Fluid Mech 41:35–52

Rasmussen CE (2004) Gaussian processes in machine learning. In: Advanced lectures on achine learning. Springer, pp 63–71

Stein ML (1999) Interpolation of spatial data: some theory for kriging. Springer, New York

Tvedt L (1990) Distribution of quadratic forms in normal space-application to structural reliability. J Eng Mech 116(6):1183–1197

Wu, YT (1992) An adaptive importance sampling method for structural system reliability analysis. In: Cruse TA (ed) Reliability technology 1992, ASME winter annual meeting, vol AD-28), Anaheim, pp 217–231

Structures with Nonviscous Damping, Modeling, and Analysis

Sondipon Adhikari
College of Engineering, Swansea University, Swansea, UK

Introduction

The role of damping is vitally important in predicting dynamic response of structures, such as building and bridges subjected to earthquake loads. Noise and vibration are not only uncomfortable to the users of these complex dynamical systems but also may lead to fatigue, fracture, and even failure of such systems. Increasing use of composite structural materials, active control, and damage-tolerant systems in the aerospace and automotive industries has led to renewed demand for energy absorbing and high damping materials. Effective applications of such materials in complex engineering dynamical systems require robust and efficient analytical and numerical methods. Due to the superior damping characteristics, the dynamics of viscoelastic materials and structures have received significant attention over the past two decades. This chapter is aimed at developing computationally efficient and physically insightful approximate numerical methods for linear dynamical systems with nonviscous damping.

A key feature of nonviscously damped systems is the incorporation of the time history of the state variables in the equation of motion. Here we use the Biot model (Biot 1958) which

S

allows one to incorporate a wide range of functions in the frequency domain by means of summation of simple "pole residue forms." Several authors have considered this model due to its simplicity and generality (see, e.g., Adhikari 2013a, b; Muravyov 1998; Muravyov and Hutton 1997; Zhang and Zheng 2007). The equation of motion of an n-degree-of-freedom linear viscously damped system can be expressed by coupled differential equations as

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \int_0^t \mathscr{G}(t-\tau)\dot{\mathbf{u}}(\tau) + \mathbf{K}\mathbf{u}(t) = \mathbf{f}(t). \quad (1)$$

Here $\mathbf{u}(t) \in \mathbb{R}^n$ is the displacement vector; $\mathbf{f}(t) \in \mathbb{R}^n$ is the forcing vector; $\mathbf{M}$, $\mathbf{K} \in \mathbb{R}^{n \times n}$ are respectively the mass matrix and stiffness; and $\mathscr{G}(t)$ is the matrix of damping kernel functions. In general $\mathbf{M}$ is a positive definite symmetric matrix and $\mathbf{K}$ is a nonnegative definite symmetric matrix. In the special case when $\mathscr{G}(t) = \mathbf{C}\delta(t)$, where $\delta(t)$ is the Dirac delta function, it reduces to the classical viscous damping case with a damping matrix $\mathbf{C}$. Therefore, Eq. 1 can be viewed as the generalization of the conventional viscously damped systems.

The natural frequencies $(\omega_j \in \mathbb{R})$ and the mode shapes $(\mathbf{x}_j \in \mathbb{R}^n)$ of the corresponding undamped system can be obtained (Meirovitch 1997) by solving the matrix eigenvalue problem

$$\mathbf{K}\mathbf{x}_j = \omega_j^2 \mathbf{M}\mathbf{x}_j, \quad \forall j = 1, 2, \ldots, n. \quad (2)$$

The undamped eigenvectors satisfy an orthogonality relationship over the mass and stiffness matrices, that is,

$$\mathbf{x}_k^T \mathbf{M}\mathbf{x}_j = \delta_{kj} \quad (3)$$

and

$$\mathbf{x}_k^T \mathbf{K}\mathbf{x}_j = \omega_j^2 \delta_{kj}, \quad \forall k, j = 1, 2, \ldots, n \quad (4)$$

where δ_{kj} is the Kronecker delta function. We construct the modal matrix

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] \in \mathbb{R}^n. \quad (5)$$

The modal matrix can be used to diagonalize system (1) provided the damping matrix $\mathscr{G}(t)$ is simultaneously diagonalizable with $\mathbf{M}$ and $\mathbf{K}$. This condition, known as the proportional damping, originally introduced by Lord Rayleigh (1877) in 1877 in the context of viscous damping, is still in wide use today. The mathematical condition for proportional damping can be obtained from the commutative behavior of the system matrices (Adhikari 2001; Caughey and O'Kelly 1965). This can be expressed as $\mathscr{G}(t)\mathbf{M}^{-1}\mathbf{K} = \mathbf{K}\mathbf{M}^{-1}\mathscr{G}(t)$. The concern of this chapter is when this condition is not met, the most likely case for many practical applications. In particular, due to the recent developments in actively controlled structures and the increasing use of composite and smart materials, the need to consider general nonproportionally damped linear dynamic systems is more than ever before.

For nonproportionally damped systems, the nonviscous modal damping matrix

$$\mathscr{G}'(t) = \mathbf{X}^T \mathscr{G}(t)\mathbf{X} \quad (6)$$

is not a diagonal matrix. Such problems can be solved using a spectral approach similar to the undamped or proportionally damped system by transforming Eq. 1 into a state-space form (Wagner and Adhikari 2003). The state-space approach is not only computationally more expensive; it also lacks the physical insight provided by the classical normal mode-based approach. Therefore, many authors have developed approximate methods in the original space (Adhikari 1999a, b).

The eigenvalue problem corresponding to system (1) can be expressed as

$$\left[s_j^2 \mathbf{M} + s_j \mathbf{G}(s_j) + \mathbf{K} \right] \mathbf{u}_j = 0, \quad \forall j = 1, 2, \ldots, m \quad (7)$$

where $s_j \in \mathbb{C}$ are the eigenvalues and $\mathbf{u}_j \in \mathbb{C}^n$ are the eigenvectors. The matrix $\mathbf{G}(s)$ is the Laplace transfer of $\mathscr{G}(t)$.

In general m is more than $2n$, that is, $m = 2n + p$; $p \geq 0$. Thus, although the system has n degrees of freedom, the number of eigenvalues is

more than $2n$. This is a major difference between the nonviscously damped systems and the viscously damped systems where the number of eigenvalues is exactly $2n$, including any multiplicities. When the eigenvalue s_j appears in complex conjugate pairs, $\mathbf{u}_j$ also appears in complex conjugate pairs, and when s_j is real, $\mathbf{u}_j$ is also real. Corresponding to the $2N$ complex conjugate pairs of eigenvalues, the n eigenvectors together with their complex conjugates will be called *elastic modes*. These modes are related to the n modes of vibration of the structural system. Physically, the assumption of "$2N$ complex conjugate pairs of eigenvalues" implies that all the elastic modes are oscillatory in nature, that is, they are subcritically damped. The modes corresponding to the "additional" p eigenvalues will be called *nonviscous modes*. These modes are induced by the nonviscous effect of the damping mechanism. For stable passive systems, the nonviscous modes are overcritically damped (i.e., negative real eigenvalues) and not oscillatory in nature. Nonviscous modes, or similar to these, are known by different names in the literature of different subjects, for example, "wet modes" in the context of ship dynamics (Bishop and Price 1979) and "damping modes" in the context of viscoelastic structures (McTavish and Hughes 1993). Determination of the eigenvectors is considered next.

In this work we consider complex conjugate eigensolutions only as for stable systems such eigenvalues are of great practical importance. Using the eigensolutions, the frequency response function (FRF) can be obtained (see, for example, Adhikari (2002)) as

$$\mathbf{H}(i\omega) = \sum_{j=1}^{n} \left[\frac{\gamma_j \mathbf{u}_j \mathbf{u}_j^T}{i\omega - s_j} + \frac{\gamma_j^* \mathbf{u}_j^* \mathbf{u}_j^{*T}}{i\omega - s_j^*} \right] + \sum_{j=2n+1}^{m} \frac{\gamma_j \mathbf{u}_j \mathbf{u}_j^T}{i\omega - s_j}$$
$$\text{where} \quad \gamma_j = \frac{1}{\mathbf{u}_j^T \left[2s_j \mathbf{M} + d\mathbf{G}/ds \big|_{s=s_j} \right] \mathbf{u}_j}.$$
$$(8)$$

Here $(\bullet)^*$ denotes complex conjugation, $(\bullet)^T$ denotes matrix transposition, and $(\bullet)'$ denotes differentiation with respect to s. This equation shows that if the complex eigensolutions s_j and $\mathbf{u}_j$ can be

obtained efficiently, the dynamic response can be obtained exactly using Eq. 8. In this chapter an iterative approach is developed to obtain the complex eigensolutions of nonproportionally damped systems from the undamped eigensolutions.

Iterative Approach for the Elastic Modes

Considering the proportional damping assumption, recently few methods (Adhikari and Pascual 2009, 2011) have been proposed to obtain the eigenvalues of nonviscously damped systems. So far, only perturbation type of approaches (Adhikari 2002) is available to obtain the eigenvectors for the general nonproportionally damped systems. This type of approaches may be suitable for the case on small nonproportionality. For the general case, only computationally expensive state-space approach (Wagner and Adhikari 2003) is currently available. Here a novel iterative method is proposed as an alternative to the state-space approach to obtain the elastic modes of general nonviscously damped systems.

For distinct undamped eigenvalues (ω_l^2), $\mathbf{x}_l$, $\forall l = 1, \ldots, n$, form a complete set of vectors. For this reason, $\mathbf{u}_j$ can be expanded as a complex linear combination of $\mathbf{x}_l$. Thus, an expansion of the form

$$\mathbf{u}_j = \sum_{l=1}^{n} \alpha_l^{(j)} \mathbf{x}_l \qquad (9)$$

may be considered. Without any loss of generality, we can assume that $\alpha_j^{(j)} = 1$ (normalization) which leaves us to determine $\alpha_j^{(j)}$, $\forall l \neq j$. Substituting the expansion of $\mathbf{u}_j$ into the eigenvalue Eq. 7, one obtains the approximation error for the j-th mode as

$$\varepsilon_j = \sum_{l=1}^{n} s_j^2 \alpha_l^{(j)} \mathbf{M} \mathbf{x}_l + s_j \sigma_l^{(j)} \mathbf{G}(s_j) \mathbf{x}_l + \alpha_l^{(j)} \mathbf{K} \mathbf{x}_l.$$
$$(10)$$

We use a Galerkin approach to minimize this error by viewing the expansion (9) as a projection in the basis functions $\mathbf{x}_l \in \mathbb{R}^n$, $\forall l = 1, 2, \ldots n$.

Therefore, we make the error orthogonal to the basis functions, that is,

$$\varepsilon_j \perp \mathbf{x}_l \quad \text{or} \quad \mathbf{x}_k^T \varepsilon_j = 0 \quad \forall k = 1, 2, \ldots, n. \quad (11)$$

Using the orthogonality property of the undamped eigenvectors described by (3) and (4), one obtains

$$s_j^2 \alpha_k^{(j)} + s_j \sum_{l=1}^{n} \alpha_l^{(j)} G'_{kl}(s_j) + \omega_k^2 \alpha_k^{(j)} = 0, \quad \forall k = 1, \ldots, n \quad (12)$$

where $G'_{kl}(s_j) = \mathbf{x}_k^T \mathbf{G}(s_j) \mathbf{x}_l$ are the elements of the modal damping matrix $\mathbf{G}'(s_j)$ defined in Eq. 6. The j-th equation of this set obtained by setting $k = j$ can be written as

$$\left(s_j^2 + s_j G'_{jj}(s_j) + \omega_j^2 \right) \alpha_j^{(j)} + s_j \sum_{l \neq j}^{n} \alpha_l^{(j)} G'_{jl}(s_j) = 0. \quad (13)$$

Recalling that $\alpha_j^{(j)} = 1$ and $\mathbf{G}'(s_j)$ is a symmetric matrix, this equation can be rewritten as

$$s_j^2 + s_j \underbrace{\left(G'_{jj}(s_j) + \sum_{l \neq j}^{n} \alpha_l^{(j)} G'_{lj}(s_j) \right)}_{\gamma_j} + \omega_j^2 = 0 \quad (14)$$

where

$$\gamma_j = G'_{jj}(s_j) + \mathbf{b}_j^T \mathbf{a}_j \quad (15)$$

$$\mathbf{b}_j = \left\{ G'_{1j}(s_j), G'_{2j}(s_j), \ldots, {}^{\{j\text{-th term deleted}\}}, \right. \\ \left. \ldots, G'_{nj}(s_j) \right\}^T \in \mathbb{R}^{(n-1)} \quad (16)$$

and

$$\mathbf{a}_j = \left\{ \alpha_1^{(j)}, \alpha_2^{(j)}, \ldots, {}^{\{j\text{-th term deleted}\}}, \ldots, \alpha_n^{(j)} \right\}^T \in \mathbb{C}^{(n-1)}. \quad (17)$$

The vector $\mathbf{a}_j$ is unknown and can be obtained by excluding the $j = k$ case in Eq. 12. Excluding this case, one has

$$s_j^2 \alpha_k^{(j)} + s_j \left(G'_{kj}(s_j) + \alpha_k^{(j)} G'_{kk}(s_j) + \sum_{l \neq k \neq j}^{n} \alpha_l^{(j)} G'_{kl}(s_j) \right) + \omega_k^2 \alpha_k^{(j)} = 0$$

$$\text{or} \quad \left(s_j^2 + \omega_k^2 + G'_{kk}(s_j) \right) \alpha_k^{(j)} + s_j \sum_{l \neq k \neq j}^{n} G'_{kl}(s_j) \alpha_l^{(j)} = -s_j G'_{kj}(s_j), \quad \forall k = 1, \ldots, n; \neq j. \quad (18)$$

These equations can be combined into a matrix form as

$$\left[\mathbf{P}_j - \mathbf{Q}_j \right] \mathbf{a}_j = \mathbf{b}_j. \quad (19)$$

In the above equation, the vectors $\mathbf{a}_j$ and $\mathbf{b}_j$ have been defined before. The matrices $\mathbf{P}_j$ and $\mathbf{Q}_j$ are defined as

$$\mathbf{P}_j = \text{diag}\left[\frac{s_j^2 + s_j G'_{11}(s_j) + \omega_1^2}{-s_j}, \ldots, {}^{\{j\text{-th term deleted}\}}, \ldots, \frac{s_j^2 + s_j G'_{nn}(s_j) + \omega_n^2}{-s_j} \right] \in \mathbb{C}^{(n-1) \times (n-1)}, \quad (20)$$

and

$$
\mathbf{Q}_j = \begin{bmatrix} 0 & G'_{12}(s_j) & \dots & \{j\text{-th term deleted}\} & \dots & G'_{1n}(s_j) \\ G'_{21}(s_j) & 0 & \vdots & \vdots & \vdots & G'_{2n}(s_j) \\ \vdots & \vdots & \vdots & \{j\text{-th term deleted}\} & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ G'_{n1}(s_j) & G'_{n2}(s_j) & \dots & \{j\text{-th term deleted}\} & \dots & 0 \end{bmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}. \qquad (21)
$$

From Eq. 19, $\mathbf{a}_j$ should be obtained by solving the set of linear equations. Because $\mathbf{P}_j$ is a diagonal matrix, one way to do this is by using the Neumann expansion method (Adhikari 1999a). Using the Neumann expansion, we have

$$
\mathbf{a}_j = \left[\mathbf{I}_{n-1} - \mathbf{P}_j^{-1}\mathbf{Q}_j \right]^{-1} \left\{ \mathbf{P}_j^{-1}\mathbf{b}_j \right\}
$$
$$
= \left[\mathbf{I}_{n-1} + \mathbf{R}_j + \mathbf{R}_j^2 + \mathbf{R}_j^3 + \dots \right] \mathbf{a}_{j0} \qquad (22)
$$

where $\mathbf{I}_{n-1}$ is a $(n-1) \times (n-1)$ identity matrix:

$$
\mathbf{R}_j = \mathbf{P}_j^{-1}\mathbf{Q}_j \in \mathbb{C}^{(n-1)\times(n-1)} \text{ and } \mathbf{a}_{j0} = \mathbf{P}_j^{-1}\mathbf{b}_j \in \mathbb{C}^{(n-1)}. \qquad (23)
$$

Because $\mathbf{P}_j$ is a diagonal matrix, its inversion can be carried out analytically, and subsequently, the closed-form expressions of the elements of $\mathbf{a}_j$ can be obtained. Keeping one term in the series (22), the first-order expression of the elements of $\mathbf{a}_j$ can be obtained as

$$
\mathbf{a}_j \equiv \left\{ \alpha_k^{(j)} \right\}_{\forall k \neq j} = -\frac{s_j G'_{kj}(s_j)}{\omega_k^2 + s_j^2 + s_j G'_{kk}(s_j)}. \qquad (24)
$$

Similarly, the second-order expression of the elements of $\mathbf{a}_j$ can be obtained as

$$
\mathbf{a}_j \equiv \left\{ \alpha_k^{(j)} \right\}_{\forall k \neq j} = -\frac{s_j G'_{kj}(s_j)}{\omega_k^2 + s_j^2 + s_j G'_{kk}(s_j)} + \sum_{\substack{l=1 \\ l\neq j\neq k}}^{n} \frac{s_j^2 G'_{kl}(s_j) G'_{lj}(s_j)}{\left(\omega_k^2 + s_j^2 + s_j G'_{kk}(s_j)\right)\left(\omega_l^2 + s_j^2 + s_j G'_{ll}(s_j)\right)}. \qquad (25)
$$

The vector $\mathbf{a}_j$ obtained using this way can be substituted back in the expression of the eigenvalues in (14), which in turn can be solved for s_j as

$$
s_j = -\left(\gamma_j \pm \mathrm{i}\sqrt{4\omega_j^2 - \gamma_j^2}\right)/2. \qquad (26)
$$

However, the vectors $\mathbf{a}_j$ and $\mathbf{b}_j$ are also a function of s_j. As a result γ_j in Eq. 15 becomes a

function of s_j. This forms the basics of the iterative approach as from Eq. 26, one can write

$$
s_j^{(r+1)} = -\gamma_j\left(s_j^{(r)}\right)/2 \pm \mathrm{i}\sqrt{\omega_j^2 - \gamma_j^2\left(s_j^{(r)}\right)/4};
$$
$$
r = 0, 1, 2, \dots \qquad (27)
$$

For every iteration step, the vectors $\mathbf{a}_j$ and $\mathbf{b}_j$ get updated based on new values of Sj using

S

Eq. 24 or Eq. 25 depending on the order of terms retained in the series (22). The iteration can be started with the equivalent proportional viscous damping assumption (Udwadia 2009), namely,

$$s_j^{(0)} = -G'_{jj}(i\omega_j)/2 \pm i\sqrt{\omega_j^2 - G'^2_{jj}(i\omega_j)/4}.$$

(28)

The iteration can be stopped when the successive values of s_j or $\mathbf{a}_j$ do not change significantly. Once the final values of $\alpha_k^{(j)}, \forall k$ are obtained, the j-th complex mode $\mathbf{u}_j$ can be obtained from the series (9).

The necessary and sufficient conditions for the convergence of the proposed method are difficult to obtain. Below we give a sufficient condition.

Proposition 1 *A sufficient condition for the convergence of the proposed iterative method is that $|\mathbf{G}'(s_j)|$ is a diagonally dominant matrix $\forall j \in [1, 2n]$.*

Proof During the iteration process, the value of s_j changes for different iteration steps. We aim to derive the condition for the convergence of series (22) for an arbitrary value of s_j. This will guarantee the convergence of the iterative method, no matter what the value of s_j. The complex matrix power series (22) converges if, and only if, for all the eigenvalues $\sigma_l^{(j)}$ of the matrix $\mathbf{R}_j$, the inequality $\left|\sigma_l^{(j)}\right| < 1$ holds. Although this condition is both necessary and sufficient, checking convergence for all $j = 1, \cdots, n$ is not feasible for every iteration step. So we look for a sufficient condition which is relatively easy to check and which ensures convergence for all $j = 1, \cdots, n$.

For an arbitrary r-th iteration, let us denote the matrix $\mathbf{R}_j$ defined in Eq. 23 as $\mathbf{R}_j^{(r)}$. Suppose the value of s_j for the r-th iteration step is $s_j^{(r)}$. The kl-th element of the matrix $\mathbf{R}_j^{(r)}$ can be obtained as

$$R_{jkl}^{(r)} = \frac{-s_j^{(r)} G'_{kl}(s_j)(1 - \delta_{kl})}{\omega_k^2 + s_j^{(r)^2} + s_j^{(r)} G'_{kk}(s_j)}, \quad \forall k, l \neq j.$$

(29)

Since a matrix norm is always greater than or equal to its maximum eigenvalue, it follows from the inequality $\left|\sigma_l^{(j)}\right| < 1$ that the convergence of the series is guaranteed if $\left\|\mathbf{R}_j^{(r)}\right\| < 1$. Writing the sum of absolute values of entries of $\mathbf{R}_j^{(r)}$ results in the following inequality as the required sufficient condition for the convergence

$$\sum_{\substack{k=1 \\ k \neq j}}^{n} \sum_{\substack{l=1 \\ l \neq j}}^{n} \left| \frac{s_j^{(r)} G'_{kl}(s_j)}{\omega_k^2 + s_j^{(r)^2} + s_j^{(r)} G'_{kk}(s_j)} \right| (1 - \delta_{lk}) < 1.$$

(30)

Dividing both the numerator and denominator by $s_j^{(r)}$, the above inequality can be written as

$$\sum_{\substack{k=1 \\ k \neq j}}^{n} \sum_{\substack{l=1 \\ l \neq i \neq k}}^{n} \frac{\left|G'_{kl}(s_j)\right|}{\left|1/s_j^{(r)}\left(\omega_k^2 + s_j^{(r)^2}\right) + G'_{kk}(s_j)\right|} < 1.$$

(31)

Taking the maximum for all $k \neq j$, this condition can further be represented as

$$\max_{k \neq j} \frac{\sum_{\substack{l=1 \\ l \neq j, k}}^{n} \left|G'_{kl}(s_j)\right|}{\left|1/s_j^{(r)}\left(\omega_k^2 + s_j^{(r)^2}\right) + G'_{kk}(s_j)\right|} < 1. \quad (32)$$

It is clear that (32) always holds if

$$\sum_{\substack{l=1 \\ l \neq i \neq k}}^{n} \left|G'_{kl}(s_j)\right| < \left|G'_{kk}(s_j)\right|, \quad \forall k \neq j$$

(33)

which in turn implies that for all $j = 1 \cdots, n$, the inequality $\left\|\mathbf{R}_j^{(r)}\right\| < 1$ holds if $|\mathbf{G}'(s_j)|$ is a diagonally dominant matrix. It is important to note that the diagonal dominance of $|\mathbf{G}'(s_j)|$ is only a sufficient condition and the lack of it does not necessarily prevent convergence of the proposed iterative method.

Summary of the Algorithm

We give a simple iterative algorithm to implement the idea developed in the previous section. We select a tolerance between the differences of the successive values of s_j, denoted by ε_m. A small value, say $\varepsilon_m = 0.001$, can be selected for numerical calculations. Considering that the undamped eigensolutions (ω_j and x_j) and the modal damping matrix $\mathbf{G}'(s_j)$ are known, the complex eigensolutions (s_j and $\mathbf{u}_j$) can be obtained using the following iterative algorithm:

> **for** $j = 1, 2, \ldots n$ **do**
>   Initialize $\varepsilon = 100, r = 0$
> $s_j^{(r)} = -G'_{jj}(i\omega_j)/2 \pm i\sqrt{\omega_j^2 - G'^2_{jj}(i\omega_j)/4}$
>   **while** $\varepsilon > \varepsilon_m$ **do**
> $\mathbf{b}_j = \left\{ G'_{1j}\left(s_j^{(r)}\right), G'_{2j}\left(s_j^{(r)}\right), \ldots, \{j\text{–th term deleted}\}, \ldots, G'_{nj}\left(s_j^{(r)}\right) \right\}^T$
> $\mathbf{a}_j \equiv \left\{ \alpha_k^{(j)} \right\}_{\forall k \neq j} = -\dfrac{s_j^{(r)} G'_{kj}\left(s_j^{(r)}\right)}{\omega_k^2 + s_j^{(r)^2} + s_j^{(r)} G'_{kk}\left(s_j^{(r)}\right)}$
> $\gamma_j = G'_{jj}(s_j) + \mathbf{b}_j^T \mathbf{a}_j$
> $s_j^{(r+1)} = -\gamma_j\left(s_j^{(r)}\right)/2 \pm i\sqrt{\omega_j^2 - \gamma_j^2\left(s_j^{(r)}\right)/4}$
> $\varepsilon = \left| s_j^{(r+1)} - s_j^{(r)} \right| / \left| s_j^{(r)} \right|$
> $r = r + 1$
>   **end while**
> $\mathbf{u}_j = \sum_{k=1}^{n} \alpha_k^{(j)} \mathbf{x}_k$
> **end for**

The algorithm is outlined for the first-order expression of $\alpha_k^{(j)}$ given by Eq. 24. However, the extension to the second- or higher-order expressions is straightforward. One simply needs to change the expression of $\mathbf{a}_j$ in this algorithm. If the higher-order terms are used, then less number of steps in the iteration is needed. Once the complex eigensolutions s_j and $\mathbf{u}_j$ are obtained using this method for all j, the dynamic response such as the frequency response function can be obtained exactly using Eq. 8. Next we illustrate this new method using a numerical example.

Numerical Illustration: A 3-DOF System with Exponential Nonviscous Damping

We consider a three-degree-of-freedom nonviscous and nonproportionally damped system from Adhikari (2013a). The system is shown in Fig. 1. Three masses, each of mass m_u, are connected by springs of stiffness k_u. The nonviscous damping elements of the system are shown in the figure. The equation of motion of this model system can be represented by Eq. 1. The mass and the stiffness matrices of the system are given by

$$\mathbf{M}_1 = \begin{bmatrix} m_u & 0 & 0 \\ 0 & 2m_u & 0 \\ 0 & 0 & m_u \end{bmatrix} \quad \text{and}$$

$$\mathbf{K} = \begin{bmatrix} 2k_u & -k_u & 0 \\ -k_u & 2k_u & -k_u \\ 0 & -k_u & 2k_u \end{bmatrix}. \tag{34}$$

The matrix of damping kernel functions can be expressed in the time domain as

$$\mathscr{G}(t) = \mathbf{C}_1 \mu_1 e^{-\mu_1 t} + \mathbf{C}_2 \mu_2 e^{-\mu_2 t}. \tag{35}$$

The coefficient matrices of this double-exponential model are given by
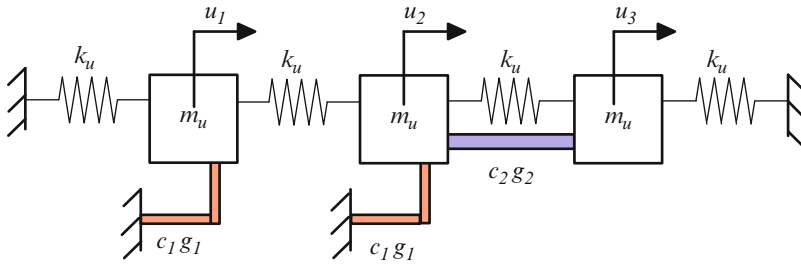
Structures with Nonviscous Damping, Modeling, and Analysis, Fig. 1 A three-DOF model system with nonviscous damping, the shaded bars represent the nonviscous damping elements. c_i, $i = 1, 2$ are the damping constants and $g_i(t - \tau) = \mu_i e^{-\mu_i(t-\tau)}$, $i = 1, 2$ are the nonviscous damping functions

$$\mathbf{C}_1 = \begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and}$$

$$\mathbf{C}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & c_2 & -c_2 \\ 0 & -c_2 & c_2 \end{bmatrix}. \quad (36)$$

Both the matrices have rank deficiency because one can easily verify that

$$r_1 = \text{rank}(\mathbf{C}_1) = 2 \le 3 \quad (37)$$

and

$$r_2 = \text{rank}(\mathbf{C}_2) = 1 \le 3. \quad (38)$$

The order of the system, that is, the number of eigenvalues

$$m = 2N + \sum \text{rank}(\mathbf{C}_i) = 6 + 3 = 9 \quad (39)$$

For the numerical calculation, we have assumed $m_u = 3.0$ kg, $k_u = 2.0$ N/m, $c_1 = 0.6$ Ns/m, $c_2 = 0.2$ Ns/m, $\mu_1 = 1.0$ s^{-1}, and $\mu_2 = 5.0$ s^{-1}. The matrix $\mathbf{G}(s)$, necessary for the implementation of the iterative method, can be obtained from (35) as

$$\mathbf{G}(s) = \mathbf{C}_1 \frac{\mu_1}{s + \mu_1} + \mathbf{C}_2 \frac{\mu_2}{s + \mu_2}. \quad (40)$$

The undamped eigenvalues and eigenvectors are obtained as

$$\{\omega_1, \omega_2, \omega_3\} = \{0.4315, 1.1547, 1.2616\} \quad (41)$$

and

$$[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3] = \begin{bmatrix} 0.1750 & -0.4082 & -0.3688 \\ 0.3012 & -0.0000 & 0.1429 \\ 0.1750 & 0.4082 & -0.3688 \end{bmatrix}. \quad (42)$$

Note that the last two undamped eigenvalues are very close and therefore one would expect significant modal coupling. The complex conjugate eigenvalues obtained using the proposed iterative method is compared with the exact state-space method and the first-order perturbation method in Table 1. We have used the first-order expression of $\alpha_k^{(j)}$ given by Eq. 24 and considered the error tolerance to be $\varepsilon_m = 0.001$. For all the three eigenvalues, 3 iterations are used. The first-order perturbation results are obtained from Eq. 28. The percentage errors are calculated with respect to the exact state-space results as

$$\varepsilon = 100 \times \frac{|\text{exact} - \text{approximate}|}{|\text{exact}|}. \quad (43)$$

Using the proposed iterative method, errors corresponding to all the three modes are reduced compared to the first-order perturbation results.

Next we consider the complex eigenvectors of the system. The exact eigenvectors obtained using the state-space approach (Wagner and Adhikari 2003) are given by

Structures with Nonviscous Damping, Modeling, and Analysis, Table 1 The complex eigenvalues of the system obtained using the proposed method are compared with the exact state-space method and the first-order perturbation method. The numbers in the parenthesis represent the percentage error

Eigenvalue number	State space (exact)	First-order perturbation	Proposed iterative (three iterations)	Method
1	$-0.0335 \pm 0.4453i$	$-0.0323 \pm 0.4448i$ (0.2918)	$-0.0335 \pm 0.4453i$	(0.0003)
2	$-0.0386 \pm 1.1806i$	$-0.0373 \pm 1.1831i$ (0.2445)	$-0.0383 \pm 1.1810i$	(0.0450)
3	$-0.0420 \pm 1.2941i$	$-0.0427 \pm 1.2907i$ (0.2695)	$-0.0424 \pm 1.2911i$	(0.2321)

$$\mathbf{U}_{\text{Exact}} = \begin{bmatrix} 0.1721 - 0.0073i - 0.3519 - 0.0115i - 0.4397 + 0.0163i \\ 0.3014 - 0.0002i - 0.0305 - 0.0066i \ 0.1410 \ + 0.0006i \\ 0.1769 + 0.0078i \ \ 0.4632 - 0.0096i - 0.3019 - 0.0124i \end{bmatrix}. \tag{44}$$

The eigenvectors calculated using the expression of the first-order perturbation in (24) are given by

$$\mathbf{U}_{\text{1st}} = \begin{bmatrix} 0.1721 - 0.0073i - 0.3519 - 0.0115i - 0.4397 + 0.0163i \\ 0.3014 - 0.0001i - 0.0309 - 0.0058i \ 0.1420 \ + 0.0020i \\ 0.1770 + 0.0077i \ \ 0.4646 - 0.0115i - 0.2989 - 0.0139i \end{bmatrix}. \tag{45}$$

Using the propose iterative approach, the matrix of complex eigenvectors can be obtained as

$$\mathbf{U}_{\text{iter}} = \begin{bmatrix} 0.1721 - 0.0073i - 0.3519 - 0.0115i - 0.4397 + 0.0163i \\ 0.3013 - 0.0000i \ - 0.0305 - 0.0062i \ \ 0.1421 + 0.0019i \\ 0.1769 + 0.0078i \ \ 0.4623 \ - 0.0094i - 0.2994 - 0.0135i \end{bmatrix}. \tag{46}$$

The eigenvectors are normalized such that the first element is identical for all the three computational method. Therefore, the eigenvectors only differ in the second and the third elements. The absolute value of the errors for the first-order perturbation method and the iterative method is given by

$$\varepsilon_{\text{1st}} = \begin{bmatrix} 0.0593 & 3.0320 & 1.2494 \\ 0.0340 & 0.5051 & 1.1030 \end{bmatrix} \quad \text{and}$$

$$\varepsilon_{\text{iter}} = \begin{bmatrix} 0.0756 & 1.2961 & 1.2448 \\ 0.0085 & 0.1864 & 0.9141 \end{bmatrix}. \tag{47}$$

In general, the iterative method gives lower error in the eigenvectors. In Fig. 2, errors in two typical frequency response functions of the system calculated from Eq. 8 using the first-order perturbation and the iterative method are shown. The first-order perturbation method performs poorly across the frequency range. From these results, the relative accuracy of the proposed iterative method can be observed.

Conclusions

Due to the recent focus in sustainable earthquake-resistant design, there is a renewed interest to consider general nonviscous and nonproportionally damped linear dynamic systems as new

Error in the cross-FRF $H_{13}(\omega)$

Error in the driving-point FRF $H_{33}(\omega)$

generation of materials are being used. State-space-based methods were normally applied to address such problems. These methods are computationally more expensive and often do not give the physical insight compared to the classical normal mode-based method. In this work a new iterative method has been proposed to obtain the complex eigensolutions of a general nonviscous nonproportionally damped system from the undamped eigensolutions. It is assumed that all the eigenvalues are distinct and are real or appear in complex conjugate pairs. The proposed method exploits a mathematical construction where complex eigenvalues and eigenvectors

can be updated from their previous values in an iterative manner. A sufficient condition for the convergence of the proposed iterative method is derived. A simple algorithm is proposed to implement this method.

The applicability of the proposed method is investigated using an example with two nonviscous damping kernels. Acceptable accuracy has been observed. Using the iterative method developed here, it is possible to obtain the eigenvalues, eigenvectors, and consequently the dynamic response of nonproportionally damped systems by post-processing of the undamped eigenvalues and eigenvectors, which in turn can be obtained using a general-purpose finite element software. Future work is necessary to extend this method to systems with repeated eigenvalues.

References

Adhikari S (1999a) Modal analysis of linear asymmetric non-conservative systems. ASCE J Eng Mech 125(12):1372–1379

Adhikari S (1999b) Rates of change of eigenvalues and eigenvectors in damped dynamic systems. AIAA J 37(11):1452–1458

Adhikari S (2001) Classical normal modes in non-viscously damped linear systems. AIAA J 39(5):978–980

Adhikari S (2002) Dynamics of non-viscously damped linear systems. ASCE J Eng Mech 128(3):328–339

Adhikari S (2013) Structural dynamic analysis with generalized damping models: analysis. Wiley ISTE, UK, 368 pp. http://eu.wiley.com/WileyCDA/WileyTitle/productCd–1848215215.html

Adhikari S (2013) Structural dynamic analysis with generalized damping models: identification. Wiley ISTE, UK, 272 pp. http://eu.wiley.com/WileyCDA/WileyTitle/productCd–184821670X.html

Adhikari S, Pascual B (2009) Eigenvalues of linear viscoelastic systems. J Sound Vib 325(4–5):1000–1011

Adhikari S, Pascual B (2011) Iterative methods for eigenvalues of viscoelastic systems. Trans ASME J Vib Acoust 133(2):021002-1–7

Biot MA (1958) Linear thermodynamics and the mechanics of solids. In: Proceedings of the third U. S. National Congress on applied mechanics. ASME, New York, pp 1–18

Bishop RED, Price WG (1979) An investigation into the linear theory of ship response to waves. J Sound Vib 62(3):353–363

Caughey TK, O'Kelly MEJ (1965) Classical normal modes in damped linear dynamic systems. Trans ASME J Appl Mech 32:583–588

McTavish DJ, Hughes PC (1993) Modeling of linear viscoelastic space structures. Tran ASME J Vib Acoust 115:103–110

Meirovitch L (1997) Principles and techniques of vibrations. Prentice-Hall International, New Jersey

Muravyov A (1998) Forced vibration responses of a viscoelastic structure. J Sound Vib 218(5):892–907

Muravyov A, Hutton SG (1997) Closed-form solutions and the eigenvalue problem for vibration of discrete viscoelastic systems. Trans ASME J Appl Mech 64:684–691

Rayleigh JW (1877) Theory of sound (two volumes), 1945 re-issue, 2nd edn. Dover Publications, New York

Udwadia FE (2009) A note on nonproportional damping. J Eng Mech-ASCE 135(11):1248–1256

Wagner N, Adhikari S (2003) Symmetric state-space formulation for a class of non-viscously damped systems. AIAA J 41(5):951–956

Zhang J, Zheng GT (2007) The biot model and its application in viscoelastic composite structures. J Vib Acoust 129:533–540

Subset Simulation Method for Rare Event Estimation: An Introduction

Konstantin M. Zuev
Department of Computing and Mathematical Sciences, California Institute of Technology, USA

S

Synonyms

Engineering reliability; Failure probability; Markov chain Monte Carlo; Monte Carlo simulation; Rare events; Subset Simulation

Introduction

This entry provides a detailed introductory description of Subset Simulation, an advanced stochastic simulation method for estimation of small probabilities of rare failure events. A simple and intuitive derivation of the method is given along with the discussion on its

implementation. The method is illustrated with several easy-to-understand examples. The reader is assumed to be familiar only with elementary probability theory and statistics.

Subset Simulation (SS) is an efficient and elegant method for simulating rare events and estimating the corresponding small tail probabilities. The method was originally developed by Siu-Kui Au and James Beck in the already classical paper (Au and Beck 2001a) for estimation of structural reliability of complex civil engineering systems such as tall buildings and bridges at risk from earthquakes. The method turned out to be so powerful and general that over the last decade, SS has been successfully applied to reliability problems in geotechnical, aerospace, fire, and nuclear engineering. Moreover, the idea of SS proved to be useful not only in reliability analysis but also in other problems associated with general engineering systems, such as sensitivity analysis, design optimization, and uncertainty quantification. As of October 2014, according to the Web of Science (ISI) database and Google Scholar, the original SS paper (Au and Beck 2001a) received 315 and 572 citations respectively, that indicates the high impact of the Subset Simulation method on the engineering research community.

Subset Simulation is essentially based on two different ideas: conceptual and technical. The conceptual idea is to decompose the rare event F into a sequence of progressively "less-rare" nested events,

$$F = F_m \subset F_{m-1} \subset \ldots \subset F_1, \qquad (1)$$

where F_1 is a relatively frequent event. For example, suppose that F represents the event of getting exactly m heads when flipping a fair coin m times. If m is large, then F is a rare event. To decompose F into a sequence (Eq. 1), let us define F_k to be the event of getting exactly k heads in the first k flips, where $k = 1,\ldots,m$. The smaller the k, the less rare the corresponding event F_k, and F_1 – getting heads in the first flip – is relatively frequent.

Given a sequence of subsets (Eq. 1), the small probability $\mathbb{P}(F)$ of the rare event F can then be represented as a product of larger probabilities as follows:

$$\begin{aligned} \mathbb{P}(F) &= \mathbb{P}(F_m) \\ &= \mathbb{P}(F_1)\frac{\mathbb{P}(F_2)}{\mathbb{P}(F_1)}\frac{\mathbb{P}(F_3)}{\mathbb{P}(F_2)}\cdots\frac{\mathbb{P}(F_{m-1})}{\mathbb{P}(F_{m-2})}\frac{\mathbb{P}(F_m)}{\mathbb{P}(F_{m-1})} \\ &= \mathbb{P}(F_1) \cdot \mathbb{P}(F_2|F_1) \cdot \ldots \cdot \mathbb{P}(F_m|F_{m-1}), \end{aligned}$$
$$(2)$$

where $\mathbb{P}(F_k|F_{k-1}) = \mathbb{P}(F_k)/\mathbb{P}(F_{k-1})$ denotes the conditional probability of event F_k given the occurrence of event F_{k-1}, for $k = 2,\ldots,m$. In the coin example, $\mathbb{P}(F_1) = 1/2$, all conditional probabilities $\mathbb{P}(F_k|F_{k-1}) = 1/2$, and the probability of the rare event $\mathbb{P}(F) = 1/2^m$.

Unlike the coin example, in real applications, it is often not obvious how to decompose the rare event into a sequence in Eq. 1 and how to compute all conditional probabilities in Eq. 2. In Subset Simulation, the "sequencing" of the rare event is done adaptively as the algorithm proceeds. This is achieved by employing Markov chain Monte Carlo, an advanced simulation technique, which constitutes the second – technical – idea behind SS. Finally, all conditional probabilities are automatically obtained as a by-product of the adaptive sequencing.

The main goals of this entry are (a) to provide a detailed exposition of Subset Simulation at an introductory level, (b) to give a simple derivation of the method and discuss its implementation, and (c) to illustrate SS with intuitive examples. Although the scope of SS is much wider, in this entry the method is described in the context of engineering reliability estimation; the problem SS was originally developed for in Au and Beck (2001a).

The rest of the entry is organized as follows: section "Engineering Reliability Problem" describes the engineering reliability problem and explains why this problem is computationally challenging. Section "The Direct Monte Carlo Method" discusses how the Direct Monte Carlo method can be used for engineering reliability estimation and why it is often inefficient. In section "Preprocessing: Transformation of Input Variables," a necessary preprocessing step which is

often used by many reliability methods is briefly discussed. Section "The Subset Simulation Method" is the core of the entry, where the SS method is explained. Illustrative examples are considered in section "Illustrative Examples." For demonstration purposes, the MATLAB code for the considered examples is provided in section "MATLAB code." Section "Summary" concludes the entry with a brief summary.

Engineering Reliability Problem

One of the most important and computationally challenging problems in reliability engineering is to estimate the probability of failure for a system, that is, the probability of unacceptable system performance. The behavior of the system can be described by a *response variable y*, which may represent, for example, the roof displacement or the largest interstory drift. The response variable depends on *input variables* $x = (x_1,\ldots,x_d)$, also called *basic variables*, which may represent geometry, material properties, and loads,

$$y = g(x_1, \ldots, x_d), \tag{3}$$

where $g(x)$ is called the *performance function*. The performance of the system is measured by comparison of the response y with a specified critical value y^*: if $y \leq y^*$, then the system is safe; if $y > y^*$, then the system has failed. This failure criterion allows to define the *failure domain F* in the input x-space as follows:

$$F = \{x : g(x) > y^*\}. \tag{4}$$

In other words, the failure domain is a set of values of input variables that lead to unacceptance system performance, namely, to the exceedance of some prescribed critical threshold y^*, which may represent the maximum permissible roof displacement, maximum permissible interstory drift, etc.

Engineering systems are complex systems, where complexity, in particular, means that the information about the system (its geometric and material properties) and its environment (loads)

is never complete. Therefore, there are always uncertainties in the values of input variables x. To account for these uncertainties, the input variables are modeled as random variables whose marginal distributions are usually obtained from test data, from expert opinion, or from literature. Let $\pi(x)$ denote the joint probability density function (PDF) for x. The uncertainty in the input variables is propagated through Eq. 3 into the response variable y, which makes the *failure event $\{x \in F\} = \{y > y^*\}$* also uncertain. The *engineering reliability problem* is then to compute the probability of failure pF, given by the following expression:

$$p_F = \mathbb{P}(x \in F) = \int_F \pi(x)dx. \tag{5}$$

The behavior of complex systems, such as tall buildings and bridges, is represented by a complex model (3). In this context, complexity means that the performance function $g(x)$, which defines the integration region F in Eq. 5, is not explicitly known. The evaluation of $g(x)$ for any x is often time-consuming and usually done by the finite element method (FEM), one of the most important numerical tools for computation of the response of engineering systems. Thus, it is usually impossible to evaluate the integral in Eq. 5 analytically because the integration region, the failure domain F, is not known explicitly.

Moreover, traditional numerical integration is also generally not applicable. In this approach, the d-dimensional input x-space is partitioned into a union of disjoint hypercubes, $\square_1,\ldots,\square_N$. For each hypercube $\square_i$, a "representative" point $x^{(i)}$ is chosen inside that hypercube, $x^{(i)} \in \square_i$. The integral in Eq. 5 is then approximated by the following sum:

$$p_F \approx \sum_{x^{(i)} \in F} \pi\left(x^{(i)}\right)\text{vol}(\square_i), \tag{6}$$

where $\text{vol}(\square_i)$ denotes the volume of $\square_i$ and summation is taken over all failure points $x^{(i)}$. Since it is not known in advance whether

a given point is a failure point or not (the failure domain F is not known explicitly), to compute the sum in Eq. 6, the failure criterion in Eq. 4 must be checked for all $x(i)$. Therefore, the approximation in Eq. 6 becomes

$$p_F \approx \sum_{i=1}^{N} I_F\left(x^{(i)}\right) \pi\left(x^{(i)}\right) \text{vol}(\square_i), \qquad (7)$$

where $I_F(x)$ stands for the indicator function, i.e.,

$$I_F(x) = \begin{cases} 1, & \text{if } x \in F, \\ 0, & \text{if } x \notin F. \end{cases} \qquad (8)$$

If n denotes the number of intervals each dimension of the input space is partitioned into, then the total number of terms in Eq. 7 is $N = n^d$. Therefore, the computational effort of numerical integration grows exponentially with the number of dimensions d. In engineering reliability problems, the dimension of the input space is typically very large (e.g., when the stochastic load time history is discretized in time). For example, $d \sim 10^3$ is not unusual in the reliability literature. This makes numerical integration computationally infeasible.

Over the past few decades, many different methods for solving the engineering reliability problem (5) have been developed. In general, the proposed reliability methods can be classified into three categories, namely:

(a) *Analytic methods* are based on the Taylor-series expansion of the performance function, e.g., the first-order reliability method (FORM) and the second-order reliability method (SORM) (Ditlevsen and Madsen 1996; Madsen et al. 2006; Melchers 1999).

(b) *Surrogate methods* are based on a functional surrogate of the performance function, e.g., the response surface method (RSM) (Faravelli 1989; Schuëller et al. 1989; Bucher 1990), Neural Networks (Papadrakakis et al. 1996), support vector machines (Hurtado and Alvarez 2003), and other methods (Hurtado 2004).

(c) *Monte Carlo simulation methods*, among which are Importance Sampling (Engelund and Rackwitz 1993), Importance Sampling using Elementary Events (Au and Beck 2001b), Radial-based Importance Sampling (Grooteman 2008), Adaptive Linked Importance Sampling (Katafygiotis and Zuev 2007), Directional Simulation (Ditlevsen and Madsen 1996), Line Sampling (Koutsourelakis et al. 2004), Auxiliary Domain Method (Katafygiotis et al. 2007), Horseracing Simulation (Zuev and Katafygiotis 2011), and *Subset Simulation* (Au and Beck 2001a).

Subset Simulation is thus a reliability method which is based on (advanced) Monte Carlo simulation.

The Direct Monte Carlo Method

The Monte Carlo method, referred in this entry as *Direct Monte Carlo* (DMC), is a statistical sampling technique that has been originally developed by Stan Ulam, John von Neumann, Nick Metropolis (who actually suggested the name "Monte Carlo" (Metropolis 1987)), and their collaborators for solving the problem of neutron diffusion and other problems in mathematical physics (Metropolis and Ulam 1949). From a mathematical point of view, DMC allows to estimate the expected value of a quantity of interest. More specifically, suppose the goal is to evaluate $\mathbb{E}_\pi[h(x)]$, that is, an expectation of a function $h : \chi \to \mathbb{R}$ with respect to the PDF $\pi(x)$,

$$\mathbb{E}_\pi[h(x)] = \int_\chi h(x)\pi(x)dx. \qquad (9)$$

The idea behind DMC is a straightforward application of the *law of large numbers* that states that if $x^{(1)}, x^{(2)}, \ldots$ are i.i.d. (independent and identically distributed) from the PDF $\pi(x)$, then the empirical average $\frac{1}{N}\sum_{i=1}^{N} h\left(x^{(i)}\right)$ converges to the true value $\mathbb{E}_\pi[h(x)]$ as N goes to $+\infty$. Therefore, if the number of samples N is large enough,

then $\mathbb{E}_\pi[h(x)]$ can be accurately estimated by the corresponding empirical average:

$$\mathbb{E}_\pi[h(x)] \approx \frac{1}{N}\sum_{i=1}^{N} h\left(x^{(i)}\right). \qquad (10)$$

The relevance of DMC to the reliability problem (5) follows from a simple observation that the failure probability p_F can be written as an expectation of the indicator function (8), namely,

$$p_F = \int_F \pi(x)dx = \int_\chi I_F(x)\pi(x)dx = \mathbb{E}_\pi[I_F(x)], \qquad (11)$$

where χ denotes the entire input x-space. Therefore, the failure probability can be estimated using the DMC method (10) as follows:

$$p_F \approx \hat{p}_F^{DMC} = \frac{1}{N}\sum_{i=1}^{N} I_F\left(x^{(i)}\right), \qquad (12)$$

where $x^{(1)},\ldots.0x^{(N)}$ are i.i.d. samples from $\pi(x)$.

The DMC estimate of p_F is thus just the ratio of the total number of *failure samples* $\sum_{i=1}^{N} I_F(x^{(i)})$, i.e., samples that produce system failure according to the system model, to the total number of samples, N. Note that $\hat{p}_F^{DMC}$ is an *unbiased* random estimate of the failure probability, that is, on average, $\hat{p}_F^{DMC}$ equals to p_F. Mathematically, this means that $\mathbb{E}[\hat{p}_F^{DMC}] = p_F$. Indeed, using the fact that $x^{(i)} \sim \pi(x)$ and in Eq. 11,

$$\mathbb{E}[\hat{p}_F^{DMC}] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} I_F\left(x^{(i)}\right)\right]$$

$$= \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\left[I_F\left(x^{(i)}\right)\right]$$

$$= \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}_\pi[I_F(x)] = p_F. \qquad (13)$$

The main advantage of DMC over numerical integration is that its accuracy does not depend on the dimension d of the input space. In reliability analysis, the standard measure of accuracy

of an unbiased estimate $\hat{p}_F$ of the failure probability is its *coefficient of variation* (c.o.v.) $\delta(\hat{p}_F)$, which is defined as the ratio of the standard deviation to the expected value of $\hat{p}_F$, i.e., $\delta(\hat{p}_F) = \sqrt{\mathbb{V}[\hat{p}_F]}/\mathbb{E}[\hat{p}_F]$, where $\mathbb{V}$ denotes the variance. The smaller the c.o.v. $\delta(\hat{p}_F)$, the more accurate the estimate $\hat{p}_F$ is. It is straightforward to calculate the variance of the DMC estimate:

$$\mathbb{V}[\hat{p}_F^{DMC}] = \mathbb{V}\left[\frac{1}{N}\sum_{i=1}^{N}I_F\left(x^{(i)}\right)\right] = \frac{1}{N^2}\sum_{i=1}^{N}\mathbb{V}\left[I_F\left(x^{(i)}\right)\right]$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\left(\mathbb{E}\left[I_F\left(x^{(i)}\right)^2\right] - \mathbb{E}\left[I_F\left(x^{(i)}\right)\right]^2\right)$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}(p_F - p_F^2) = \frac{p_F(1-p_F)}{N}. \qquad (14)$$

Here, the identity $I_F(x)^2 = I_F(x)$ was used. Using Eqs. 13 and 14, the c.o.v. of the DMC estimate can be calculated:

$$\delta(\hat{p}_F^{DMC}) = \frac{\sqrt{\mathbb{V}[\hat{p}_F^{DMC}]}}{\mathbb{E}[\hat{p}_F^{DMC}]} = \sqrt{\frac{1-p_F}{Np_F}}. \qquad (15)$$

This result shows that $\delta(\hat{p}_F^{DMC})$ depends only on the failure probability p_F and the total number of samples N and does not depend on the dimension d of the input space. Therefore, unlike numerical integration, the DMC method does not suffer from the "curse of dimensionality," i.e., from an exponential increase in volume associated with adding extra dimensions, and is able to handle problems of high dimension.

Nevertheless, DMC has a serious drawback: it is inefficient in estimating small failure probabilities. For typical engineering reliability problems, the failure probability p_F is very small, $p_F \ll 1$. In other words, the system is usually assumed to be designed properly, so that its failure is a *rare event*. In the reliability literature, $p_F \sim 10^{-2} - 10^{-9}$ have been considered. If p_F is very small, then it follows from Eq. 15 that

$$\delta(\hat{p}_F^{DMC}) \approx \frac{1}{\sqrt{Np_F}}. \qquad (16)$$

This means that the number of samples N needed to achieve an acceptable level of accuracy is inverse proportional to p_F, and therefore very large, $N \propto 1/p_F \gg 1$. For example, if $p_F = 10^{-4}$ and the c.o.v. of 10 % is desirable, then $N = 10^6$ samples are required. Note, however, that each evaluation of $I_F(x^{(i)})$, $i = 1,\ldots,N$, in Eq. 12 requires a system analysis to be performed to check whether the sample $x^{(i)}$ is a failure sample. As it has been already mentioned in section "Engineering Reliability Problem," the computation effort for the system analysis, i.e., computation of the performance function $g(x)$, is significant (usually involves the FEM method). As a result, the DMC method becomes excessively costly and practically inapplicable for reliability analysis. This deficiency of DMC has motivated research to develop more advanced simulation algorithms for efficient estimation of small failure probabilities in high dimensions.

Remark 1 It is important to highlight, however, that even though DMC cannot be routinely used for reliability problems (too expensive), it is a very robust method, and it is often used as a check on other reliability methods.

Preprocessing: Transformation of Input Variables

Many reliability methods, including Subset Simulation, assume that the input variables x are independent. This assumption, however, is not a limitation, since in simulation one always starts from independent variables to generate the dependent input variables. Furthermore, for convenience, it is often assumed that x are i.i.d. Gaussian. If this is not the case, a "preprocessing" step that transforms x to i.i.d. Gaussian variables z must be undertaken. The transformation from x to z can be performed in several ways depending on the available information about the input variables. In the simplest case, when x are independent Gaussians, $x_k \sim \mathcal{N}(\cdot|\mu_k, \sigma_k^2)$, where μ_k and σ_k^2 are,

respectively, the mean and variance of x_k, the necessary transformation is standardization:

$$z_k = \frac{x_k - \mu_k}{\sigma_k}. \tag{17}$$

In other cases, more general techniques should be used, such as the Rosenblatt transformation (Rosenblatt 1952) and the Nataf transformation (Nataf 1962). To avoid introduction of additional notation, hereinafter, it is assumed without loss of generality that the vector x has been already transformed and it follows the standard multivariate Gaussian distribution,

$$\pi(x_1, \ldots, x_d) = \prod_{k=1}^{d} \phi(x_k), \tag{18}$$

where $\phi(\cdot)$ denotes the standard Gaussian PDF,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \tag{19}$$

The Subset Simulation Method

Unlike Direct Monte Carlo, where all computational resources are directly spent on sampling the input space $x^{(1)},\ldots, x^{(N)} \sim \pi(\cdot)$ and computing the values of the performance function $g(x^{(1)}),\ldots, g(x^{(N)})$, Subset Simulation first "probes" the input space χ by generating a relatively small number of i.i.d samples $x_0^{(1)}, \ldots, x_0^{(n)} \sim \pi(x)$, $n < N$, and computing the corresponding system responses $y_0^{(1)} = g(x_0^{(1)}), \ldots, y_0^{(n)} = g(x_0^{(n)})$. Here, the subscript 0 indicates the 0th stage of the algorithm. Since F is a rare event and n is relatively small, it is very likely that none of the samples $x_0^{(1)}, \ldots, x_0^{(n)}$ belongs to F, that is, $y_0^{(i)} < y^*$ for all $i = 1,\ldots,n$. Nevertheless, these Monte Carlo samples contain some useful information about the failure domain that can be utilized. To keep the notation simple, assume that $y_0^{(1)}, \ldots, y_0^{(n)}$ are arranged in the decreasing order, i.e., $y_0^{(1)} \geq \ldots \geq y_0^{(n)}$ (it is always possible to achieve this by renumbering $x_0^{(1)}, \ldots, x_0^{(n)}$ if needed). Then, $x_0^{(1)}$ and $x_0^{(n)}$ are, respectively, the closest to failure and the safest samples among $x_0^{(1)}, \ldots, x_0^{(n)}$, since $y_0^{(1)}$ and $y_0^{(n)}$ are the largest and the smallest responses. In general,
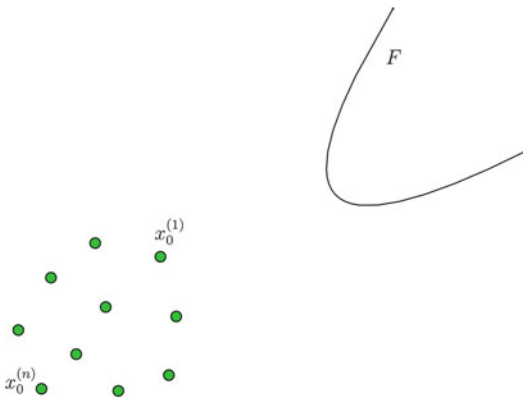
Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 1 Monte Carlo samples $x_0^{(1)}, \ldots, x_0^{(n)}$ and the failure domain F. $x_0^{(1)}$ and $x_0^{(n)}$ are, respectively, the closest to failure and the safest samples among $x_0^{(1)}, \ldots, x_0^{(n)}$

Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 2 The first intermediate failure domain F_1. In this schematic illustration, $n = 10$, $p = 0.2$, so that there are exactly $np = 2$ Monte Carlo samples in $F1$, $x_0^{(1)}, x_0^{(2)} \in F_1$

the smaller the i, the closer to failure the sample $x_0^{(i)}$ is. This is shown schematically in Fig. 1.

Let $p \in (0, 1)$ be any number such that np is integer. By analogy with Eq. 4, define the *first intermediate failure domain F_1* as follows:

$$F_1 = \left\{ x : g(x) > y_1^* \right\}, \qquad (20)$$

where

$$y_1^* = \frac{y_0^{(np)} + y_0^{(np+1)}}{2}. \qquad (21)$$

In other words, F_1 is the set of inputs that lead to the exceedance of the *relaxed threshold $y_1^* < y^*$*. Note that by construction, samples $x_0^{(1)}, \ldots, x_0^{(np)}$ belong to F_1, while $x_0^{(np + 1)}, \ldots, x_0^{(n)}$ do not. As a consequence, the Direct Monte Carlo estimate for the probability of F_1 which is based on samples $x_0^{(1)}, \ldots, x_0^{(n)}$ is automatically equal to p,

$$\mathbb{P}(F_1) \approx \frac{1}{n} \sum_{i=1}^{n} I_{F_1}\left(x_0^{(i)}\right) = p. \qquad (22)$$

The value $p = 0.1$ is often used in the literature, which makes F_1 a relatively frequent event. Figure 2 illustrates the definition of F_1.

The first intermediate failure domain F_1 can be viewed as a (very rough) conservative approximation to the target failure domain F. Since $F \subset F_1$, the failure probability p_F can be written as a product:

$$p_F = \mathbb{P}(F_1)\mathbb{P}(F|F_1), \qquad (23)$$

where $\mathbb{P}(F|F_1)$ is the conditional probability of F given F_1. Therefore, in view of Eq. 22, the problem of estimating p_F is reduced to estimating the conditional probability $\mathbb{P}(F|F_1)$.

In the next stage, instead of generating samples in the whole input space (like in DMC), the SS algorithm aims to populate F_1. Specifically, the goal is to generate samples $x_1^{(1)}, \ldots, x_1^{(n)}$ from the conditional distribution

$$\pi(x|F_1) = \frac{\pi(x)I_{F_1}(x)}{\mathbb{P}(F_1)} = \frac{I_{F_1}(x)}{\mathbb{P}(F_1)} \prod_{k=1}^{d} \phi(x_k). \quad (24)$$

First of all, note that samples $x_0^{(1)}, \ldots, x_0^{(np)}$ not only belong to F_1 but are also distributed according to $\pi(\cdot|F_1)$. To generate the remaining $(n - np)$ samples from $\pi(\cdot|F_1)$, which, in general, is not a trivial task, Subset Simulation uses the so-called Modified Metropolis algorithm (MMA). MMA belongs to the class of *Markov chain Monte Carlo* (MCMC) algorithms (Liu 2001;

Robert and Casella 2004), which are techniques for sampling from complex probability distributions that cannot be sampled directly, at least not efficiently. MMA is based on the original Metropolis algorithm (Metropolis et al. 1953) and specifically tailored for sampling from the conditional distributions of the form (24).

Modified Metropolis Algorithm

Let $x \sim \pi(\cdot|F_1)$ be a sample from the conditional distribution $\pi(\cdot|F_1)$. The Modified Metropolis algorithm generates another sample $\tilde{x}$ from $\pi(\cdot|F_1)$ as follows:

1. Generate a "candidate" sample ξ: For each coordinate $k = 1, \ldots, d$,
   (a) Sample $\eta_k \sim q_k(\cdot|x_k)$, where $q_k(\cdot|x_k)$, called the *proposal distribution*, is a univariate PDF for η_k centered at x_k with the symmetry property $q_k(\eta_k|x_k) = q_k(x_k|\eta_k)$. For example, the proposal distribution can be a Gaussian PDF with mean x_k and variance σ_k^2,

   $$q_k(\eta_k|x_k) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{(\eta_k - x_k)^2}{2\sigma_k^2}\right), \quad (25)$$

   or it can be a uniform distribution over $[x_k - \alpha, x_k + \alpha]$, for some $\alpha \geq 0$.
   (b) Compute the acceptance ratio

   $$r_k = \frac{\phi(\eta_k)}{\phi(x_k)}. \quad (26)$$

   (c) Define the kth coordinate of the candidate sample by accepting or rejecting η_k,

   $$\xi_k = \begin{cases} \eta_k, & \text{with probability } \min\{1, r_k\}, \\ x_k, & \text{with probability } 1 - \min\{1, r_k\}. \end{cases} \quad (27)$$

2. Accept or reject the candidate sample ξ by setting

   $$\tilde{x} = \begin{cases} \xi, & \text{if } \xi \in F_1, \\ x, & \text{if } \xi \notin F_1. \end{cases} \quad (28)$$

The Modified Metropolis algorithm is schematically illustrated in Fig. 3.
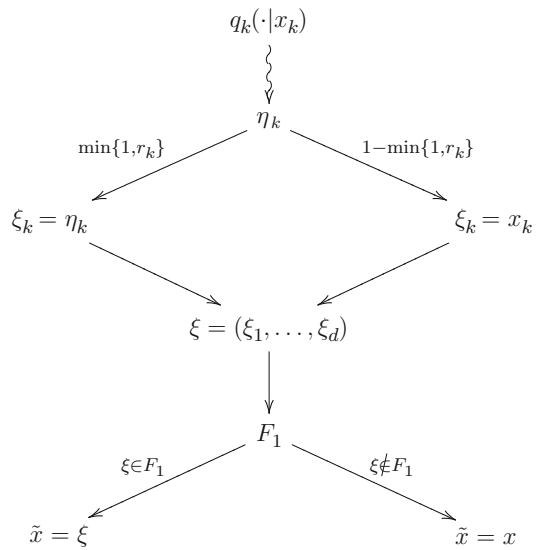
Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 3 Modified Metropolis algorithm

It can be shown that the sample $\tilde{x}$ generated by MMA is indeed distributed according to $\pi(\cdot|F_1)$. If the candidate sample ξ is rejected in Eq. 28, then $\tilde{x} = x \sim \pi(\cdot|F_1)$ and there is nothing to prove. Suppose now that ξ is accepted, $\tilde{x} = \xi$, so that the move from x to $\tilde{x}$ is a proper transition between two distinct points in F_1. Let $f(\cdot)$ denote the PDF of $\tilde{x}$ (the goal is to show that $f(\tilde{x}) = \pi(\tilde{x}|F_1)$). Then

$$f(\tilde{x}) = \int_{F_1} \pi(x|F_1)t(\tilde{x}|x)dx, \quad (29)$$

where $t(\tilde{x}|x)$ is the transition PDF from x to $\tilde{x} \neq x$. According to the first step of MMA, coordinates of $\tilde{x} = \xi$ are generated independently, and therefore $t(\tilde{x}|x)$ can be expressed as a product,

$$t(\tilde{x}|x) = \prod_{k=1}^{d} t_k(\tilde{x}_k|x_k), \quad (30)$$

where $t_k(\tilde{x}_k|x_k)$ is the transition PDF for the kth coordinate $\tilde{x}_k$. Combining Eqs. 24, 29, and 30 gives

$$f(\tilde{x}) = \int_{F_1} \frac{I_{F_1}(x)}{\mathbb{P}(F_1)} \prod_{k=1}^{d} \phi(x_k) \prod_{k=1}^{d} t_k(\tilde{x}_k|x_k) dx$$

$$= \frac{1}{\mathbb{P}(F_1)} \int_{F_1} \prod_{k=1}^{d} \phi(x_k) t_k(\tilde{x}_k|x_k) dx. \qquad (31)$$

The key to the proof of $f(\tilde{x}) = \pi(\tilde{x}|F_1)$ is to demonstrate that $\phi(x_k)$ and $t_k(\tilde{x}_k|x_k)$ satisfy the so-called detailed balance equation,

$$\phi(x_k)t_k(\tilde{x}_k|x_k) = \phi(\tilde{x}_k)t_k(x_k|\tilde{x}_k). \qquad (32)$$

If $\tilde{x}_k = x_k$, then Eq. 32 is trivial. Suppose that $\tilde{x}_k \neq x_k$, that is, $\tilde{x}_k = \xi_k = \eta_k$ in Eq. 27. The actual transition PDF $t_k(\tilde{x}_k|x_k)$ from x_k to $\tilde{x}_k \neq x_k$ differs from the proposal PDF $q_k(\tilde{x}_k|x_k)$ because the acceptance-rejection step in Eq. 27 is involved. To actually make the move from x_k to $\tilde{x}_k$, one needs not only to generate $\tilde{x}_k \sim q_k(\cdot|x_k)$ but also to accept it with probability $\min\left\{1, \frac{\phi(\tilde{x}_k)}{\phi(x_k)}\right\}$. Therefore,

$$t_k(\tilde{x}_k|x_k) = q_k(\tilde{x}_k|x_k)\min\left\{1, \frac{\phi(\tilde{x}_k)}{\phi(x_k)}\right\}, \quad \tilde{x}_k \neq x_k. \qquad (33)$$

Using Eq. 33, the symmetry property of the proposal PDF, $q_k(\tilde{x}_k|x_k) = q_k(x_k|\tilde{x}_k)$, and the identity $a \min\left\{1, \frac{b}{a}\right\} = b \min\left\{1, \frac{a}{b}\right\}$ for any $a, b > 0$,

$$\phi(x_k)t_k(\tilde{x}_k|x_k) = q_k(\tilde{x}_k|x_k)\phi(x_k)\min\left\{1, \frac{\phi(\tilde{x}_k)}{\phi(x_k)}\right\}$$

$$= q_k(x_k|\tilde{x}_k)\phi(\tilde{x}_k)\min\left\{1, \frac{\phi(x_k)}{\phi(\tilde{x}_k)}\right\}$$

$$= \phi(\tilde{x}_k)t_k(x_k|\tilde{x}_k), \qquad (34)$$

and the detailed balance in Eq. 32 is thus established. The rest is a straightforward calculation:

$$f(\tilde{x}) = \frac{1}{\mathbb{P}(F_1)} \int_{F_1} \prod_{k=1}^{d} \phi(\tilde{x}_k) t_k(x_k|\tilde{x}_k) dx$$

$$= \frac{1}{\mathbb{P}(F_1)} \prod_{k=1}^{d} \phi(\tilde{x}_k) \int_{F_1} t(x|\tilde{x}) dx$$

$$= \pi(\tilde{x}|F_1), \qquad (35)$$

since the transition PDF $t(x|\tilde{x})$ integrates to 1, and $I_{F_1}(\tilde{x}) = 1$.

Remark 2 A mathematically more rigorous proof of the Modified Metropolis algorithm is given in (Zuev et al. 2012).

Remark 3 It is worth mentioning that although the independence of input variables is crucial for the applicability of MMA, and thus for Subset Simulation, they need not be identically distributed. In other words, instead of Eq. 18, the joint PDF $\pi(\cdot)$ can have a more general form, $\pi(x) = \prod_{k=1}^{d} \pi_k(x_k)$, where $\pi^k(\cdot)$ is the marginal distributions of x_k which is not necessarily Gaussian. In this case, the expression for the acceptance ratio in Eq. 26 must be replaced by $r_k = \frac{\pi_k(\eta_k)}{\pi_k(x_k)}$.

Subset Simulation at Higher Conditional Levels

Given $x_0^{(1)}, \ldots, x_0^{(np)} \sim \pi(\cdot|F_1)$, it is clear now how to generate the remaining $(n - np)$ samples from $\pi(\cdot|F_1)$. Namely, starting from each $x_0^{(i)}$, $i = 1, \ldots, np$, the SS algorithm generates a sequence of $\left(\frac{1}{p} - 1\right)$ new MCMC samples $x_0^{(i)} = x_{0,0}^{(i)} \mapsto x_{0,1}^{(i)} \mapsto \ldots \mapsto x_{0,\frac{1}{p}-1}^{(i)}$ using the Modified Metropolis transition rule described above. Note that when $x_{0,j}^{(i)}$ is generated, the previous sample $x_{0,j-1}^{(i)}$ is used as an input for the transition rule. The sequence $x_{0,0}^{(i)}, x_{0,1}^{(i)}, \ldots, x_{0,\frac{1}{p}-1}^{(i)}$ is called a *Markov chain* with the *stationary distribution* $\pi(\cdot|F_1)$, and $x_{0,0}^{(i)} = x_0^{(i)}$ is often referred to as the "seed" of the Markov chain.

To simplify the notation, denote samples $\left\{x_{0,j}^{(i)}\right\}_{j=0,\ldots,\frac{1}{p}-1}^{i=1,\ldots,np}$ by $\{x_1^{(1)}, \ldots, x_1^{(n)}\}$. The subscript 1 indicates that the MCMC samples $x_1^{(1)}, \ldots, x_1^{(n)} \sim \pi(\cdot|F_1)$ are generated at the first conditional level of the SS algorithm. These conditional samples are schematically shown in Fig. 4. Also assume that the corresponding system responses $y_1^{(1)} = g(x_1^{(1)}), \ldots, y_1^{(n)} = g(x_1^{(n)})$ are arranged in the decreasing order, i.e., $y_1^{(1)} \geq \ldots \geq y_1^{(n)}$. If the failure event F is rare enough, that is, if p_F is sufficiently small, then it is very likely that none of the samples $x_1^{(1)}, \ldots, x_1^{(n)}$ belongs to F, i.e., $y_1^{(i)} < y^*$ for all $i = 1, \ldots, n$. Nevertheless, these MCMC samples can be used in the similar way the Monte Carlo samples $x_0^{(1)}, \ldots, x_0^{(n)}$ were used.
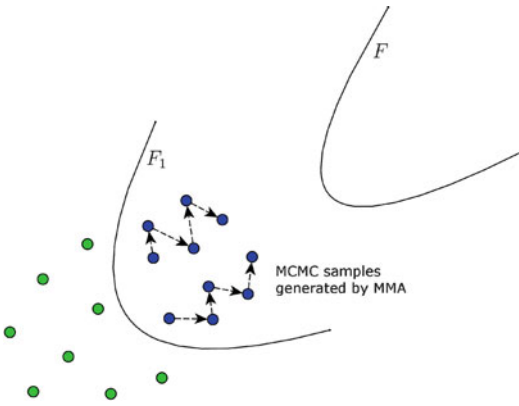
Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 4 MCMC samples generated by the Modified Metropolis algorithm at the first conditional level of Subset Simulation
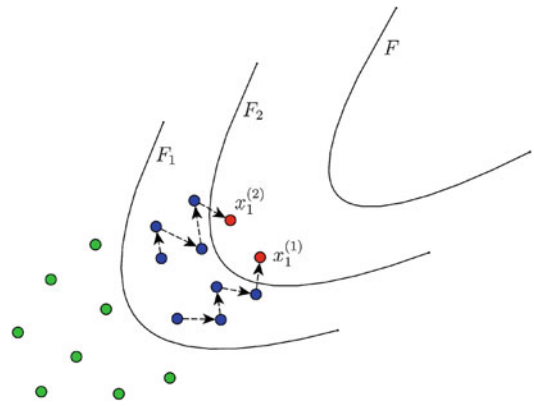
Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 5 The second intermediate failure domain $F2$. In this schematic illustration, $n = 10$, $p = 0.2$, so that there are exactly $np = 2$ MCMC samples in F_2, $x_1^{(1)}$, $x_1^{(2)} \in F_2$

By analogy with Eq. 20, define the *second intermediate failure domain F_2* as follows:

$$F_2 = \left\{ x : g(x) > y_2^* \right\}, \qquad (36)$$

where

$$y_2^* = \frac{y_1^{(np)} + y_1^{(np+1)}}{2}. \qquad (37)$$

Note that $y_2^* > y_1^*$ since $y_1^{(i)} > y_1^*$ for all $i = 1,\ldots,n$. This means that $F \subset F_2 \subset F_1$, and therefore, F_2 can be viewed as a conservative approximation to F which is still rough, yet more accurate than F_1. Figure 5 illustrates the definition of F_2. By construction, samples $x_1^{(1)}, \ldots, x_1^{(np)}$ belong to F_2, while $x_1^{(np+1)}, \ldots, x_1^{(n)}$ do not. As a result, the estimate for the conditional probability of F_2 given F_1 which is based on samples $x_1^{(1)}, \ldots, x_1^{(n)} \sim \pi(\cdot|F_1)$ is automatically equal to p,

$$\mathbb{P}(F_2|F_1) \approx \frac{1}{n} \sum_{i=1}^{n} I_{F_2}\left(x_1^{(i)}\right) = p. \qquad (38)$$

Since $F \subset F_2 \subset F_1$, the conditional probability $\mathbb{P}(F|F_1)$ that appears in Eq. 23 can be expressed as a product:

$$\mathbb{P}(F|F_1) = \mathbb{P}(F_2|F_1)\mathbb{P}(F|F_2). \qquad (39)$$

Combining Eqs. 23 and 39 gives the following expression for the failure probability:

$$p_F = \mathbb{P}(F_1)\mathbb{P}(F_2|F_1)\mathbb{P}(F|F_2). \qquad (40)$$

Thus, in view of Eqs. 22 and 38, the problem of estimating p_F is now reduced to estimating the conditional probability $\mathbb{P}(F|F_2)$.

In the next step, as one may have already guessed, the Subset Simulation algorithm populates F_2 by generating MCMC samples $x_2^{(1)}, \ldots, x_2^{(n)}$ from $\pi(\cdot|F_2)$ using the Modified Metropolis algorithm, defines the third intermediate failure domain $F_3 \subset F_2$ such that $\mathbb{P}(F_3|F_2) \approx \frac{1}{n}\sum_{i=1}^{n} I_{F_3}\left(x_2^{(i)}\right) = p$, and reduces the original problem of estimating the failure probability p_F to estimating the conditional probability $\mathbb{P}(F|F_3)$ by representing $p_F = \mathbb{P}(F_1)\mathbb{P}(F_2|F_1)\mathbb{P}(F_3|F_2)\mathbb{P}(F|F_3)$. The algorithm proceeds in this way until the target failure domain F has been sufficiently sampled so that the conditional probability $\mathbb{P}(F|F_L)$ can be accurately estimated by $\frac{1}{n}\sum_{i=1}^{n} I_F\left(x_L^{(i)}\right)$, where F_L is the Lth intermediate failure domain and $x_L^{(1)}, \ldots, x_L^{(n)} \sim \pi(\cdot|F_L)$ are the MCMC samples generated at the Lth conditional level. Subset Simulation can thus be viewed as a method that decomposes the rare failure event F into a sequence of

progressively "less-rare" nested events, $F \subset F_L \subset$ $\ldots \subset F_1$, where all intermediate failure events $F_1,\ldots,F_L$ are constructed adaptively by appropriately relaxing the value of the critical threshold $y_1^* < \ldots < y_L^* < y^*$.

Stopping Criterion

In what follows, the stopping criterion for Subset Simulation is described in detail. Let $n_F(l)$ denote the number of failure samples at the lth level, that is,

$$n_F(l) = \sum_{i=1}^{n} I_F\left(x_l^{(i)}\right), \qquad (41)$$

where $x_l^{(1)}, \ldots, x_l^{(n)} \sim \pi(\cdot|F_l)$. Since F is a rare event, it is very likely that $n_F(l) = 0$ for the first few conditional levels. As l gets larger, however, $n_F(l)$ starts increasing since Fl, which approximates F "from above," shrinks closer to F. In general, $n_F(l) \geq n_F(l-1)$, since $F \subset F_l \subset F_{l-1}$ and the np closest to F samples among $x_{l-1}^{(1)}, \ldots, x_{l-1}^{(n)}$ are present among $x_l^{(1)}, \ldots, x_l^{(n)}$. At conditional level l, the failure probability p_F is expressed as a product,

$$p_F = \mathbb{P}(F_1)\mathbb{P}(F_2|F_1)\ldots\mathbb{P}(F_l|F_{l-1})\mathbb{P}(F|F_l). \qquad (42)$$

Furthermore, the adaptive choice of intermediate critical thresholds $y_1^*, \ldots, y_l^*$ guarantees that the first l factors in Eq. 42 approximately equal to p, and, thus,

$$p_F \approx p^l \cdot \mathbb{P}(F|F_l). \qquad (43)$$

Since there are exactly $n_F(l)$ failure samples at the lth level, the estimate of the last conditional probability in Eq. 42 which is based on samples $x_l^{(1)}, \ldots, x_l^{(n)} \sim \pi(\cdot|F_l)$ is given by

$$\mathbb{P}(F|F_l) \approx \frac{1}{n}\sum_{i=1}^{n} I_F\left(x_l^{(i)}\right) = \frac{n_F(l)}{n}. \qquad (44)$$

If $n_F(l)$ is sufficiently large, i.e., the conditional event $(F|F_l)$ is not rare, then the estimate in Eq. 44 is fairly accurate. This leads to the following stopping criterion:

- If $\frac{n_F(l)}{n} \geq p$, i.e., there are at least np failure samples among $x_l^{(1)}, \ldots, x_l^{(n)}$, then Subset Simulation stops: the current conditional level l becomes the last level, $L = l$, and the failure probability estimate derived from Eqs. 43 and 44 is

$$p_F \approx \hat{p}_F^{\text{SS}} = p^L \frac{n_F(L)}{n}. \qquad (45)$$

- If $\frac{n_F(l)}{n} < p$, i.e., there are less than np failure samples among $x_l^{(1)}, \ldots, x_l^{(n)}$, then the algorithm proceeds by defining the next intermediate failure domain $F_{l+1} = \{x : g(x) > y_{l+1}^*\}$, where $y_{l+1}^* = (y_l^{(np)} + y_l^{(np+1)})/2$, and expressing $\mathbb{P}(F|F_l)$ as a product $\mathbb{P}(F|F_l) = \mathbb{P}(F_{l+1}|F_l)\mathbb{P}(F|F_{l+1}) \approx p \cdot \mathbb{P}(F|F_{l+1})$.

The described stopping criterion guarantees that the estimated values of all factors in the factorization $p_F = \mathbb{P}(F_1)\mathbb{P}(F_2|F_1)\ldots\mathbb{P}(F_L|F_{L-1})$ $\mathbb{P}(F|FL)$ are not smaller than p. If p is relatively large ($p = 0.1$ is often used in applications), then it is likely that the estimates $\mathbb{P}(F_1) \approx p$, $\mathbb{P}(F_2|F_1) \approx p,\ldots,\mathbb{P}(F_L|F_{L-1}) \approx p$, and $\mathbb{P}(F|F_L) \approx \frac{n_F(L)}{n}(\leq p)$ are accurate even when the sample size n is relatively small. As a result, the SS estimate in Eq. 45 is also accurate in this case. This provides an intuitive explanation as to why Subset Simulation is efficient in estimating small probabilities of rare events. For a detailed discussion of error estimation for the SS method, the reader is referred to Au and Wang (2014).

Implementation Details

In the rest of this section, the implementation details of Subset Simulation are discussed. The SS algorithm has two essential components that affect its efficiency: the parameter p and the set of univariate proposal PDFs $\{q_k\}$, $k = 1,\ldots,d$.

Level Probability

The parameter p, called the *level probability* in Au and Wang (2014) and the *conditional failure probability* in Zuev et al. (2012), governs how many intermediate failure domains F_l are needed

to reach the target failure domain F. As it follows from Eq. 45, a small value of p leads to a fewer total number of conditional levels L. But at the same time, it results in a large number of samples n needed at each conditional level l for accurate determination of F_l (i.e., determination of y_l^*) that satisfies $\frac{1}{n}\sum_{i=1}^{n} I_{F_l}\left(x_{l-1}^{(i)}\right) = p$. In the extreme case when $p \leq p_F$, no levels are needed, $L = 0$, and Subset Simulation reduces to the Direct Monte Carlo method. On the other hand, increasing the value of p will mean that fewer samples are needed at each conditional level, but it will increase the total number of levels L. The choice of the level probability p is thus a trade-off between the total number of level L and the number of samples n at each level. In the original paper (Au and Beck 2001a), it has been found that the value $p = 0.1$ yields good efficiency. The latter studies (Au and Wang 2014; Zuev et al. 2012), where the c.o.v. of the SS estimate $\hat{p}_F^{\text{SS}}$ has been analyzed, confirmed that $p = 0.1$ is a nearly optimal value of the level probability.

Proposal Distributions

The efficiency and accuracy of Subset Simulation also depends on the set of univariate proposal PDFs $\{q_k\}$, $k = 1,\ldots,d$, that are used within the Modified Metropolis algorithm for sampling from the conditional distributions $\pi(\cdot/F_l)$. To see this, note that in contract to the Monte Carlo samples $x_0^{(1)}, \ldots, x_0^{(n)} \sim \pi(\cdot)$ which are i.i.d., the MCMC samples $x_l^{(1)}, \ldots, x_l^{(n)} \sim \pi(\cdot|F_l)$ are *not independent* for $l \geq 1$, since the MMA transition rule uses $x_l^{(i)} \sim \pi(\cdot|F_l)$ to generate $x_l^{(i+1)} \sim \pi(\cdot|F_l)$. This means that although these MCMC samples can be used for statistical averaging as if they were i.i.d., the efficiency of the averaging is reduced if compared with the i.i.d. case (Doob 1953). Namely, the more correlated $x_l^{(1)}, \ldots, x_l^{(n)}$ are, the slower is the convergence of the estimate $P(F_{l+1}|F_l) \approx \frac{1}{n}\sum_{i=1}^{n} I_{F_{l+1}}\left(x_l^{(i)}\right)$, and, therefore, the less efficient it is. The correlation between samples $x_l^{(1)}, \ldots, x_l^{(n)}$ is due to proposal PDFs $\{q_k\}$, which govern the generation of the next sample $x_l^{(i+1)}$ from the current one $x_l^{(i)}$. Hence, the choice of $\{q_k\}$ is very important.

It was observed in Au and Beck (2001a) that the efficiency of MMA is not sensitive to the type of the proposal PDFs (Gaussian, uniform, etc.); however, it strongly depends on their *spread* (variance). Both small and large spreads tend to increase the correlation between successive samples. Large spreads may reduce the acceptance rate in Eq. 28, increasing the number of repeated MCMC samples. Small spreads, on the contrary, may lead to a reasonably high acceptance rate, but still produce very correlated samples due to their close proximity. As a rule of thumb, the spread of q_k, $k = 1,\ldots,d$, can be taken of the same order as the spread of the corresponding marginal PDF π_k (Au and Wang 2014). For example, if π is given by Eq. 18, so that all marginal PDFs are standard Gaussian, $\pi_k(x) = \phi(x)$, then all proposal PDFs can also be Gaussian with unit variance, $q_k(x/x_k) = \phi(x - x_k)$. This choice is found to give a balance between efficiency and robustness.

The spread of proposal PDFs can also be chosen adaptively. In Zuev et al. (2012), where the problem of optimal scaling for the Modified Metropolis algorithm was studied in more detail, the following nearly optimal scaling strategy was proposed: at each conditional level, select the spread such that the corresponding acceptance rate in Eq. 28 is between 30 % and 50 %. In general, finding the optimal spread of proposal distributions is problem specific and a highly nontrivial task not only for MMA but also for almost all MCMC algorithms.

Illustrative Examples

To illustrate Subset Simulation and to demonstrate its efficiency in estimating small probabilities of rare failure events, two examples are considered in this section. As it has been discussed in section "Engineering Reliability Problem," in reliability problems, the dimension d of the input space $\boldsymbol{\chi}$ is usually very large. In spite of this, for visualization and educational purposes, a linear reliability problem in two dimensions ($d = 2$) is first considered in section "Subset Simulation in 2D." A more realistic

high-dimensional example ($d = 10^3$) is considered in the subsequent section "Subset Simulation in High Dimensions."

Subset Simulation in 2D

Suppose that $d = 2$, i.e., the response variable y depends only on two input variables x_1 and x_2. Consider a linear performance function

$$g(x_1, x_2) = x_1 + x_2, \tag{46}$$

where x_1 and x_2 are independent standard Gaussian, $x_i \sim N(0, 1)$, $i = 1, 2$. The failure domain F is then a half-plane defined by

$$F = \{(x_1, x_2) : x_1 + x_2 > y^*\}. \tag{47}$$

In this example, the failure probability p_F can be calculated analytically. Indeed, since $x_1 + x_2 \sim N(0, 2)$ and, therefore, $\frac{x_1 + x_2}{\sqrt{2}} \sim \mathcal{N}(0, 1)$,

$$p_F = \mathbb{P}(x_1 + x_2 > y^*) = \mathbb{P}\left(\frac{x_1 + x_2}{\sqrt{2}} > \frac{y^*}{\sqrt{2}}\right)$$
$$= 1 - \Phi\left(\frac{y^*}{\sqrt{2}}\right), \tag{48}$$

where Φ is the standard Gaussian CDF. This expression for the failure probability can be used as a check on the SS estimate. Moreover, expressing y^* in terms of p_F,

$$y^* = \sqrt{2}\Phi^{-1}(1 - p_F), \tag{49}$$

allows to solve the inverse problem, namely, to formulate a linear reliability problem with a given value of the failure probability. Suppose that $p_F = 10^{-10}$ is the target value. Then the corresponding value of the critical threshold is $y^* \approx 9$.

Subset Simulation was used to estimate the failure probability of the rare event in Eq. 47 with $y^* = 9$. The parameters of the algorithm were chosen as follows: the level probability $p = 0.1$, the proposal PDFs $q_k(x/x_k) = \phi(x - x_k)$, and the sample size $n = 10^3$ per each level. This implementation of SS led to $L = 9$ conditional

levels, making the total number of generated samples $N = n + L(n - np) = 9.1 \times 10^3$. The obtained SS estimate is $\hat{p}_F^{SS} = 1.58 \times 10^{-10}$ which is quite close to the true value $p_F = 10^{-10}$. Note that, in this example, it is hopeless to obtain an accurate estimate by the Direct Monte Carlo method since the DMC estimate in Eq. 12 based on $N = 9.1 \times 10^3$ samples is effectively zero: the rare event F is too rare.

Figure 6 shows the samples generated by the SS method. The dashed lines represent the boundaries of intermediate failure domains F_l, $l = 1, \ldots, L = 9$. The solid line is the boundary of the target failure domain F. This illustrates how Subset Simulation pushes Monte Carlo samples (red) toward the failure region.

Subset Simulation in High Dimensions

It is straightforward to generalize the low-dimensional example considered in the previous section to high dimensions. Consider a linear performance function

$$g(x) = \sum_{i=1}^{d} x_i, \tag{50}$$

where $x_1, \ldots, x_d$ are i.i.d. standard Gaussian. The failure domain is then a half-space defined by

$$F = \left\{x : \sum_{i=1}^{d} x_i > y^*\right\}. \tag{51}$$

In this example, $d = 10^3$ is considered; hence the input space $\chi = \mathbb{R}^d$ is indeed high dimensional. As before, the failure probability can be calculated analytically:

$$p_F = \mathbb{P}\left(\sum_{i=1}^{d} x_i > y^*\right) = \mathbb{P}\left(\frac{\sum_{i=1}^{d} x_i}{\sqrt{d}} > \frac{y^*}{\sqrt{d}}\right)$$
$$= 1 - \Phi\left(\frac{y^*}{\sqrt{d}}\right). \tag{52}$$

This expression will be used as a check on the SS estimate.

First, consider the following range of values for the critical threshold, $y^* \in [0, 200]$.

Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 6 Samples generated by Subset Simulation: red samples are Monte Carlo samples generated at the 0th unconditional level, purple samples are MCMC sample generated at the 1st conditional level, etc. The dashed lines represent the boundaries of intermediate failure domains F_l, $l = 1,\ldots,L = 9$. The solid line is the boundary of the target failure domain F [Example 6.1]
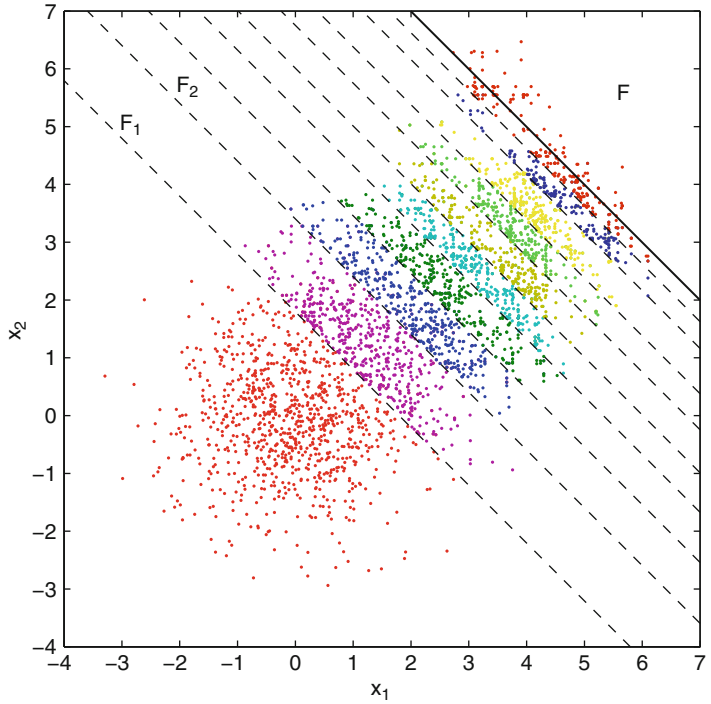
Figure 7 plots p_F versus y^*. The solid red curve corresponds to the sample mean of the SS estimates $\hat{p}_F^{SS}$ which is based on 100 independent runs of Subset Simulation. The two dashed red curves correspond to the sample mean $\pm$ one sample standard deviation. The SS parameters were set as follows: the level probability $p = 0.1$, the proposal PDFs $q_k(x/x_k) = \phi(x - x_k)$, and the sample size $n = 3 \times 10^3$ per each level. The solid blue curve (which almost coincides the solid red curve) corresponds to the true values of p_F computed from Eq. 52. The dark green curves correspond to Direct Monte Carlo: the solid curve is the sample mean (based on 100 independent runs) of the DMC estimates $\hat{p}_F^{DMC}$ in Eq. 12, and the two dashed curves are the sample mean $\pm$ one sample standard deviation. The total number of samples N used in DMC equals to the average (based on 100 runs) total number of samples used in SS. Finally, the dashed light green curves show the theoretical performance of Direct Monte Carlo, namely, they correspond to the true value of p_F (52) $\pm$ one theoretical standard deviation obtained from Eq. 14. The bottom panel of Fig. 7

shows the zoomed-in region that corresponds to the values $y^* \in [100, 160]$ of the critical threshold. Note that for relatively large values of the failure probability, $p_F < 10^{-3}$, both DMC and SS produce accurate estimates of p_F. For smaller values however, $pF < 10^{-5}$, the DMC estimate starts to degenerate, while SS still accurately estimates p_F. This can be seen especially well in the bottom panel of the figure.

The performances of Subset Simulation and Direct Monte Carlo can be also compared in terms of the coefficient of variation of the estimates $\hat{p}_F^{SS}$ and $\hat{p}_F^{DMC}$. This comparison is shown in Fig. 8. The red and dark green curves represent the sample c.o.v. for SS and DMC, respectively. The light green curve is the theoretical c.o.v. of $\hat{p}_F^{DMC}$ given by Eq. 15. When the critical threshold is relatively small $y^* < 60$, the performances of SS and DMC are comparable. As y^* gets large, the c.o.v. of $\hat{p}_F^{DMC}$ starts to grow much faster than that of $\hat{p}_F^{SS}$. In other words, SS starts to outperform DMC, and the larger the y^*, i.e., the more rare the failure event, the more significant the outperformance is.
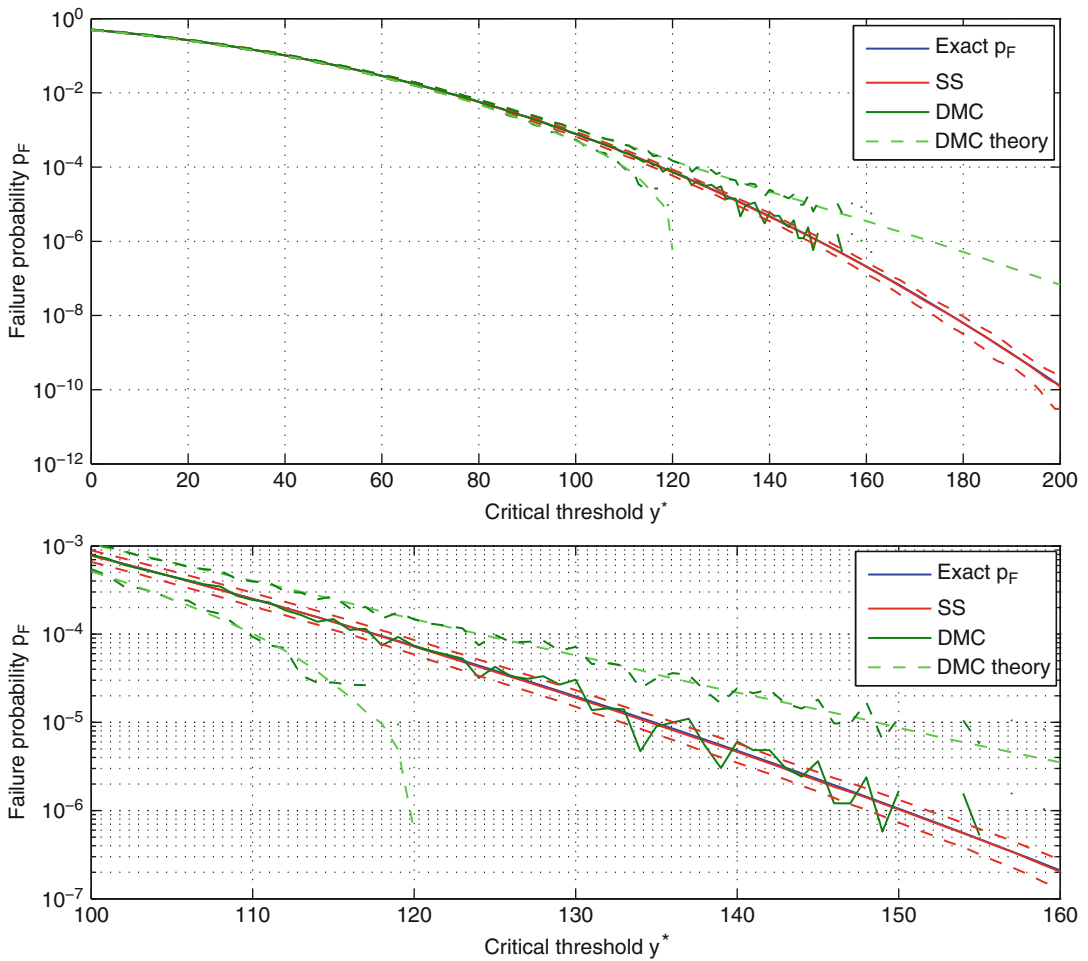
Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 7 Failure probability p_F versus the critical threshold y^* [Example 6.2]

The average total number of samples used in Subset Simulation versus the corresponding values of failure probability is shown in the top panel of Fig. 9. The staircase nature of the plot is due to the fact that every time p_F crosses the value p^k by decreasing from $p^k + \epsilon$ to $p^k - \epsilon$, an additional conditional level is required. In this example, $p = 0.1$ is used, that is why the jumps occur at $p_F = 10^{-k}$, $k = 1, 2,....$ The jumps are more pronounced for larger values of p_F, where the SS estimate is more accurate. For smaller values of p_F, where the SS estimate is less accurate, the jumps are more smoothed out by averaging over independent runs.

In Fig. 8, where the c.o.v's of SS and DMC are compared, the total numbers of samples (computational efforts) used in the two methods are the same. The natural question is then the following: by how much should the total number of samples N used in DMC be increased to achieve the same c.o.v as in SS (so that the green curve in Fig. 8 coincides with the red curve)? The answer is given in the bottom panel of Fig. 9. For example, if $p_F = 10^{-10}$, then $N = 1010$, while the computational effort of SS is less than 105 samples.

Simulation results presented in Figs. 7, 8, and 9 clearly indicate that (a) Subset Simulation
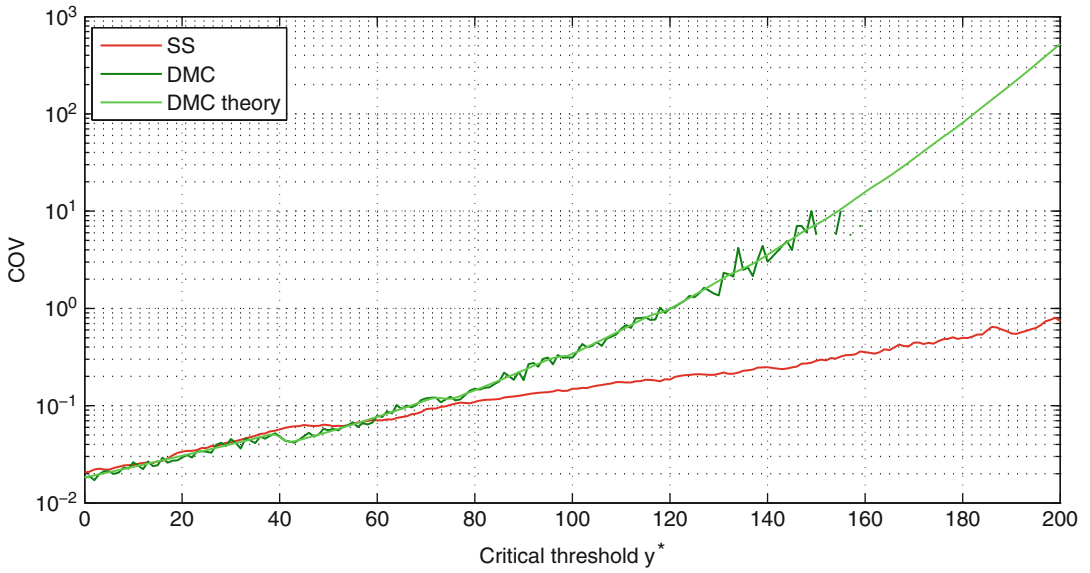
Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 8 C.o.v versus the critical threshold [Example 6.2]
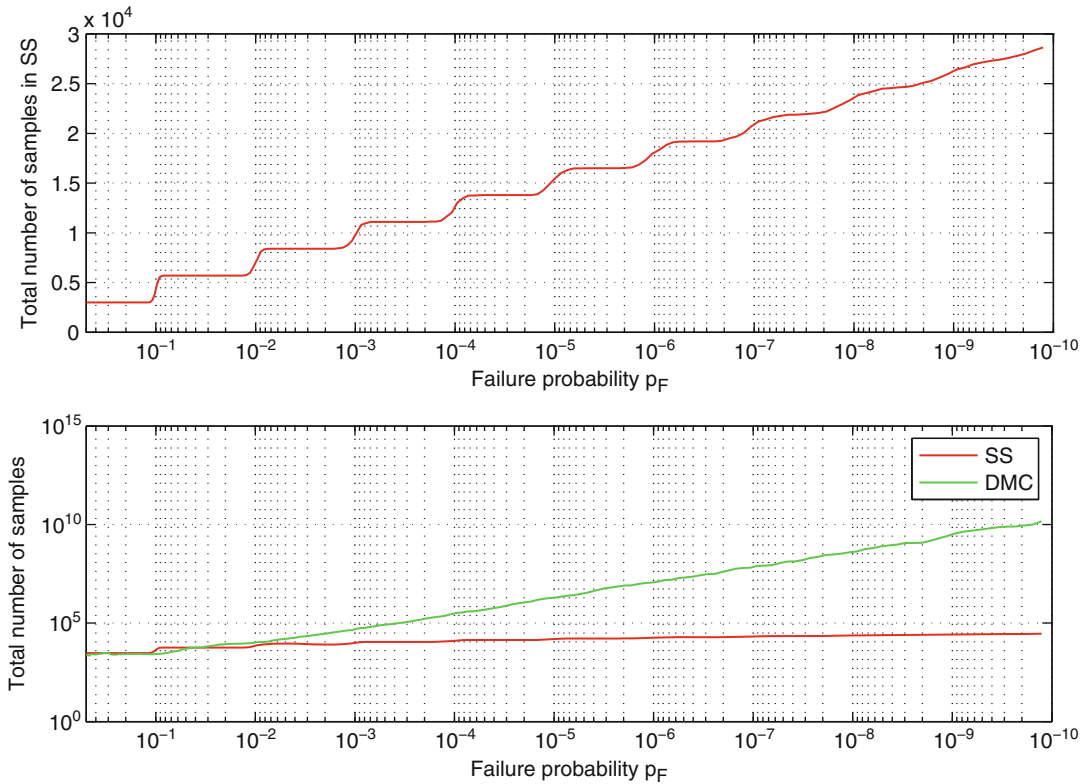
Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 9 Total number of samples versus the failure probability [Example 6.2]
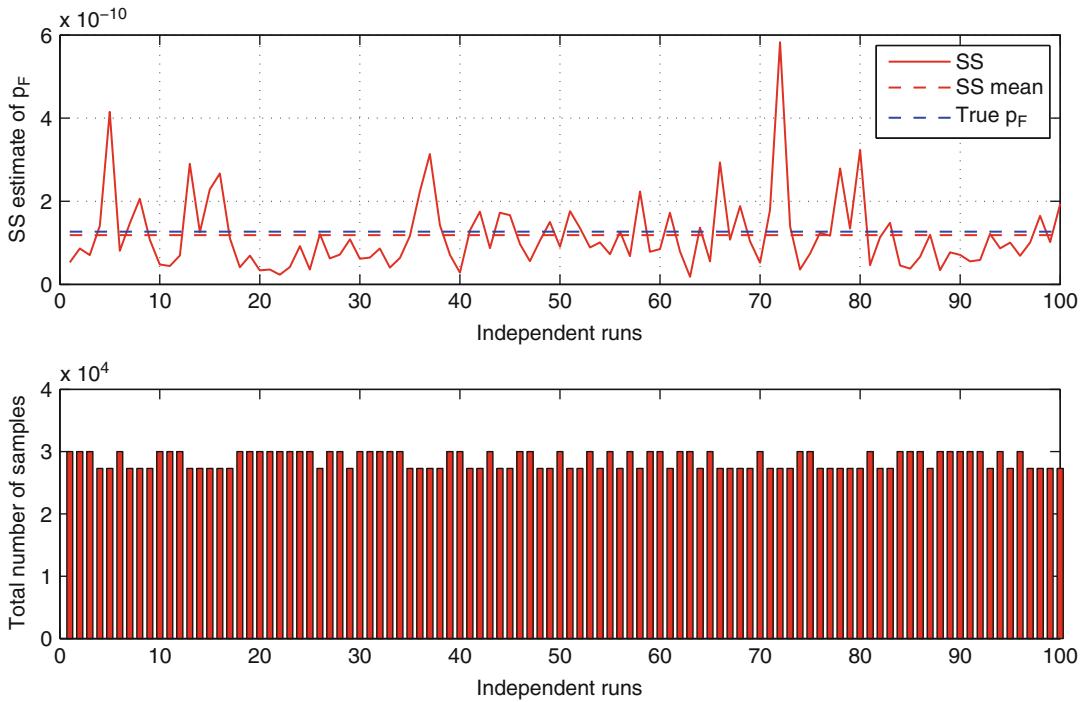
Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 10 Performance of Subset Simulation for 100 independent runs. The critical threshold is $y^* = 200$, and the corresponding true value of the failure probability is $p_F = 1.27 \times 10^{-10}$ [Example 6.2]
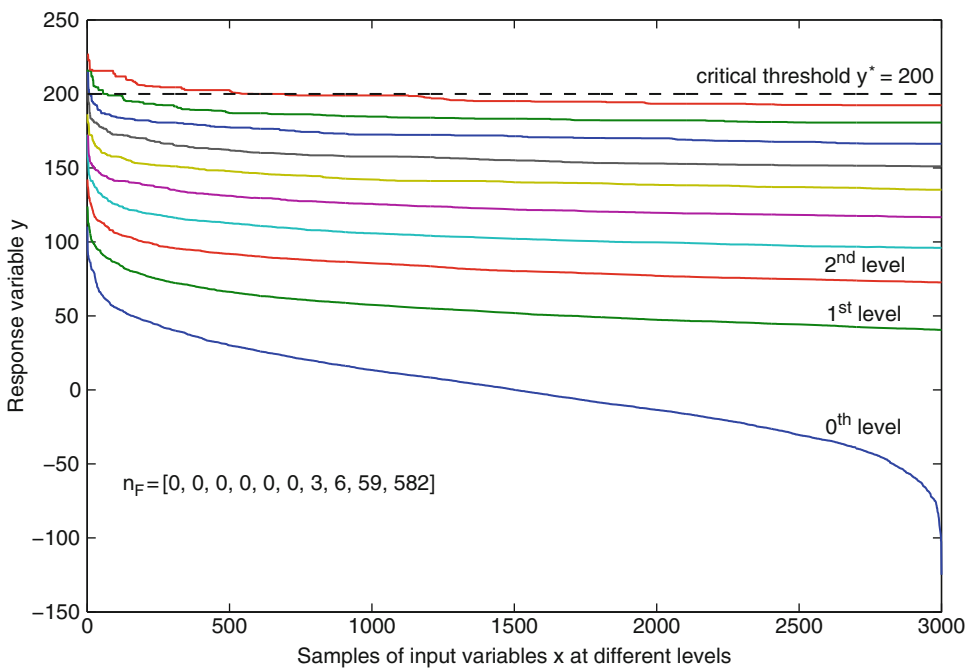
Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 11 System responses $y_1^{(1)} \geq \ldots \geq y_1^{(n)}$, $n = 3 \times 10^3$, for all levels, $l = 0, \ldots, L = 9$, for a fixed simulation run [Example 6.2]

S
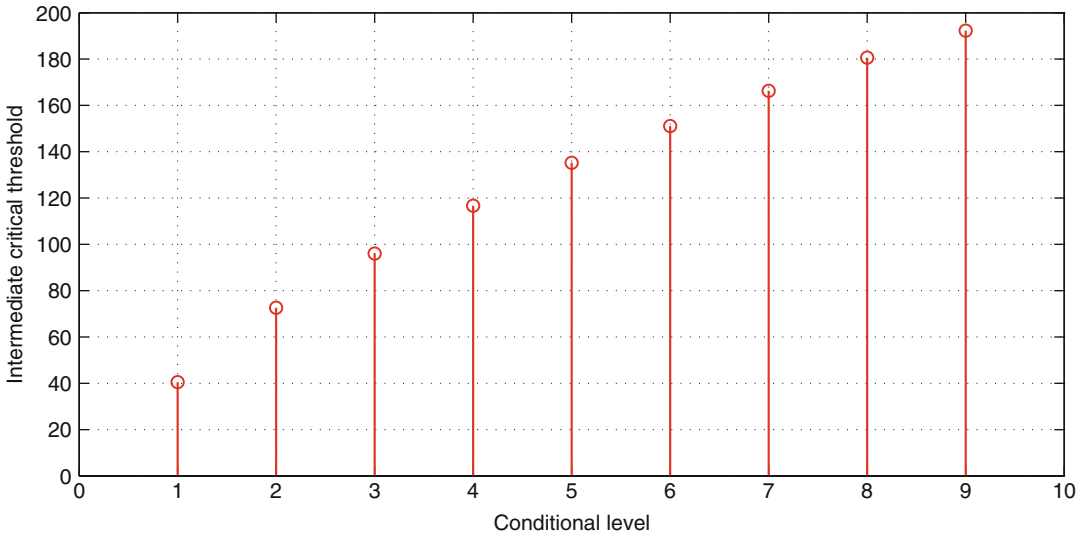
Subset Simulation Method for Rare Event Estimation: An Introduction, Fig. 12 Intermediate critical thresholds $y_1^{(*)}, \ldots, y_L^*, L = 9$, at different conditional levels in a fixed simulation run [Example 6.2]

produces a relatively accurate estimate of the failure probability and (b) Subset Simulation drastically outperforms Direct Monte Carlo when estimating probabilities of rare events.

Let us now focus on a specific value of the critical threshold, $y^* = 200$, which corresponds to a very rare failure event in Eq. 51 with probability $p_F = 1.27 \times 10^{-10}$. Figure 10 demonstrates the performance of Subset Simulation for 100 independent runs. The top panel shows the obtained SS estimate $\hat{p}_F^{SS}$ for each run. Although $\hat{p}_F^{SS}$ varies significantly (its c.o.v. is $\delta(\hat{p}_F^{SS}) = 0.74$), its mean value $\overline{\hat{p}_F^{SS}} = 1.18 \times 10^{-10}$ (dashed red line) is close to the true value of the failure probability (dashed blue line). The bottom panel shows the total number of samples used in SS in each run. It is needless to say that the DMC estimate based on $N \sim 3 \times 10^4$ samples would almost certainly be zero.

Figure 11 shows the system responses $y_l^{(1)} \geq \ldots \geq y_l^{(n)}$, n = 3×10^{-10} for all levels, $l = 0, \ldots, L = 9$, for a fixed simulation run. As expected, for the first few levels (six levels in this case), the number of failure samples $n_F(l)$, i.e., samples $x_l^{(i)}$ with $y_l^{(i)} = g(x_l^{(i)}) > y^*$, is zero. As Subset Simulation starts pushing the samples toward the failure domain, $n_F(l)$ starts increasing

with $n_F(6) = 3$, $n_F(7) = 6$, $n_F(8) = 59$, and, finally, $n_F(9) = 582$, after which the algorithm stopped since $n_F(9)/n = 0.194$ which is large than $p = 0.1$. Finally, Fig. 12 plots the intermediate (relaxed) critical thresholds $y_1^*, \ldots, y_L^*$ at different levels obtained in a fixed simulation run.

MATLAB Code

This section contains the MATLAB code for the examples considered in section "Illustrative Examples." For educational purposes, the code was written as readable as possible with numerous comments. As a result of this approach, the efficiency of the code was unavoidably scarified.

% Subset Simulation for Liner Reliability Problem	
% Performance function: g(x)=x1+...+xd	
% Input variables x1,...,xd are i.i.d. N (0,1)	
% Written by K.M. Zuev, Institute of Risk & Uncertainty, Uni of Liverpool	
clear;	
d=1000;	% dimension of the input space

(*continued*)

`YF=200;`	% critical threshold (failure <=> g(x)>YF)
`pF=1−normcdf (YF/sqrt(d));`	% true value of the failure probability
`n=3000;`	% number of samples per level
`p=0.1;`	% level probability
`nc=n*p;`	% number of Markov chains
`ns=(1−p)/p;`	% number of states in each chain
`L=0;`	% current (unconditional) level
`x=randn(d,n);`	% Monte Carlo samples
`nF=0;`	% number of failure samples
`for i=1:n`	
`y(i)=sum(x(:, i));`	% system response y=g(x)
`if y(i)>YF`	% y(i)>YF <=> x(:, i) is a failure sample
`nF=nF+1;`	
`end`	
`end`	
`while nF(L+1)/n<p`	% stopping criterion
`L=L+1;`	% next conditional lelvel is needed
`[y(L,:),ind] =sort(y (L,:),'descend');`	% renumbered responses
`x(:,:,L)=x(:,ind (:),L);`	% renumbered samples
`Y(L)=(y(L,nc)+y (L,nc+1))/2;`	% L^th intermediate threshold
`z(:,:,1)=x(:,1: nc,L);`	% Markov chain "seeds"

`% Modified Metropolis algorithm for sampling from pi(x / F L)`
`for j=1:nc`
`for m=1:ns`
`% Step 1:`
`for k=1:d`
`a=z(k,j,m)+randn; % Step 1(a)`
`r=min(1,normpdf(a)/normpdf(z(k,j, m))); % Step 1(b)`
`% Step 1(c):`
`if rand<r`	

(*continued*)

```
q(k)=a;
else
q(k)=z(k,j,m);
end
end
% Step 2:
if sum(q)>Y(L) % q belongs to F L
z(:,j,m+1)=q;
else
z(:,j,m+1)=z(:,j,m);
end
end
end
for j=1:nc
for m=1:ns+1
x(:,(j−1)*(ns+1)+m,L+1)=z(:,j,m); %
samples from pi(x / F_L)
end
end
clear z;
```

`nF(L+1)=0;`	
`for i=1:n`	
`y(L+1,i)=sum(x (:,i,L+1));`	% system response y=g(x)
`if y(L+1,i)>YF`	% then x(:,i,L+1) is a failure sample
`nF(L+1)=nF(L+1) +1;`	% number of failure samples at level L +1
`end`	
`end`	
`end`	
`pF SS=p^(L)*nF(L +1)/n;`	% SS estimate
`N=n+n*(1−p)*(L);`	% total number of samples

Summary

In this entry, a detailed exposition of Subset Simulation, an advanced stochastic simulation method for estimation of small probabilities of rare events, is provided at an introductory level. A simple step-by-step derivation of Subset Simulation is given, and important implementation details are discussed. The method is illustrated with a few intuitive examples.

After the original paper (Au and Beck 2001a) was published, various modifications of SS were proposed: SS with splitting (Ching et al. 2005a), hybrid SS (Ching et al. 2005b), and two-stage SS (Katafygiotis and Cheung 2005), to name but

S

a few. It is important to highlight, however, that none of these modifications offers a drastic improvement over the original algorithm. A Bayesian analog of SS was developed in Zuev et al. (2012). For further reading on Subset Simulation and its applications, a fundamental and very accessible monograph (Au and Wang 2014) is strongly recommended, where the method is presented from the CCDF (complementary cumulative distribution function) perspective and where the error estimation is discussed in detail.

Also, it is important to emphasize that Subset Simulation provides an efficient solution for general reliability problems without using any specific information about the dynamic system other than an input–output model. This independence of a system's inherent properties makes Subset Simulation potentially useful for applications in different areas of science and engineering.

As a final remark, it is a pleasure to thank Professor Siu-Kui Au whose comments on the first draft of the entry were very helpful; Professor James Beck, who generously shared his knowledge of and experience with Subset Simulation and made important comments on the pre-final draft of the entry; and Professor Francis Bonahon for his general support and for creating a nice atmosphere at the Department of Mathematics of the University of Southern California, where the author started this work.

Cross-References

▶ Reliability Analysis of Nonlinear Vibrating Systems-Spectral Approach
▶ Reliability Estimation and Analysis
▶ Response Variability and Reliability of Structures
▶ Seismic Reliability Assessment, Alternative Methods for
▶ Seismic Risk Assessment, Cascading Effects
▶ Site Response for Seismic Hazard Assessment
▶ Structural Reliability Estimation for Seismic Loading
▶ Structural Seismic Reliability Analysis

References

Au SK, Beck JL (2001a) Estimation of small failure probabilities in high dimensions by subset simulation. Probabilist Eng Mech 16(4):263–277

Au SK, Beck JL (2001b) First-excursion probabilities for linear systems by very efficient importance sampling. Probabilist Eng Mech 16(3):193–207

Au SK, Wang Y (2014) Engineering risk assessment and design with subset simulation. Wiley, Singapore, To appear

Bucher C (1990) A fast and efficient response surface approach for structural reliability problem. Struct Saf 7:57–66

Ching J, Au SK, Beck JL (2005a) Reliability estimation of dynamical systems subject to stochastic excitation using subset simulation with splitting. Comput Method Appl Mech Eng 194(12–16):1557–1579

Ching J, Beck JL, Au SK (2005b) Hybrid subset simulation method for reliability estimation of dynamical systems subject to stochastic excitation. Probabilist Eng Mech 20(3):199–214

Ditlevsen O, Madsen HO (1996) Structural reliability methods. Wiley, Chichester

Doob JL (1953) Stochastic processes. Wiley, New York

Engelund S, Rackwitz R (1993) A benchmark study on importance sampling techniques in structural reliability. Struct Saf 12(4):255–276

Faravelli L (1989) Response-surface approach for reliability analysis. J Eng Mech 115:2763–2781

Grooteman F (2008) Adaptive radial-based importance sampling method for structural reliability. Struct Saf 30(6):533–542

Hurtado JE (2004) Structural reliability. Statistical learning perspectives. Springer, Heidelberg

Hurtado JE, Alvarez DA (2003) A classification approach for reliability analysis with stochastic finite element modeling. J Struct Eng 129(8):1141–1149

Katafygiotis LS, Cheung SH (2005) A two-stage subset simulation-based approach for calculating the reliability of inelastic structural systems subjected to Gaussian random excitations. Comput Method Appl Mech Eng 194(12–16):1581–1595

Katafygiotis LS, Zuev KM (2007) Estimation of small failure probabilities in high dimensions by adaptive linked importance sampling. In: M. Papadrakakis, D.C. Charmpis, N.D. Lagaros, Y. Tsompanakis (Eds.), ECCOMAS thematic conference on computational methods in structural dynamics and earthquake engineering (COMPDYN), Rethymno, Crete, Greece, June 2007

Katafygiotis LS, Moan T, Cheung SH (2007) Auxiliary domain method for solving multi-objective dynamic reliability problems for nonlinear structures. Struct Eng Mech 25(3):347–363

Koutsourelakis PS, Pradlwarter HJ, Schuëller GI (2004) Reliability of structures in high dimensions, part I: algorithms and applications. Probabilist Eng Mech 19(4):409–417

Liu JS (2001) Monte Carlo strategies is scientific computing. Springer, New York

Madsen HO, Krenk S, Lind NC (2006) Methods of structural safety. Dover, Mineola

Melchers R (1999) Structural reliability analysis and prediction. Wiley, Chichester

Metropolis N (1987) The beginning of the Monte Carlo method. Los Alamos Science 15:125–130

Metropolis N, Ulam S (1949) The Monte Carlo method. J Am Stat Assoc 44:335–341

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21(6): 1087–1092

Nataf A (1962) Détermination des distributions de probabilité dont les marges sont donées. C R Acad Sci 225:42–43

Papadrakakis M, Papadopoulos V, Lagaros ND (1996) Structural reliability analysis of elastic–plastic structures using neural networks and Monte Carlo simulation. Comput Method Appl Mech Eng 136:145–163

Robert CP, Casella G (2004) Monte Carlo statistical methods. Springer, New York

Rosenblatt M (1952) Remarks on a multivariate transformation. Ann Math Stat 23:470–472

Schuëller GI, Bucher CG, Bourgund U, Ouypornprasert W (1989) On efficient computational schemes to calculate structural failure probabilities. Probabilist Eng Mech 4(1):10–18

Zuev KM, Katafygiotis LS (2011) Horseracing simulation algorithm for evaluation of small failure probabilities. Probabilist Eng Mech 26(2):157–164

Zuev KM, Beck JL, Au SK, Katafygiotis LS (2012) Bayesian postprocessor and other enhancements of subset simulation for estimating failure probabilities in high dimensions. Comput Struct 92–93: 283–296

Substructuring Methods for Finite Element Analysis

S. Abhinav, Debraj Ghosh and C. S. Manohar
Department of Civil Engineering, Indian Institute of Science, Bangalore, Karnataka, India

Synonyms

Component mode synthesis; Domain decomposition; Hybrid simulations

Introduction

The motivations for employing substructuring in finite element modeling vary from reduction of computational time, modal synthesis using substructure modes, combining experimental and numerical modeling approaches, equitable sharing of resources in parallel computing environment, and treatment of global/local non-linearities. The details of methods and tools accordingly also vary. An overview of related issues is presented in this entry.

Problems of computational structural mechanics of realistic systems involve inversion and eigenanalysis of large-size matrices and solutions of a large number of coupled ordinary differential equations or algebraic equations. These are computationally demanding tasks, and development of methods to reduce the computational efforts remains relevant notwithstanding advances in computational hardware. This is particularly true in problems of uncertainty quantification, reliability analysis, structural optimization, modeling of actively controlled systems, local/global sensitivity analysis, and problems of damage detection. Substructuring methods primarily serve to achieve reduction in computational effort in these types of problems. Here, the given structure is divided into a set of subsystems, each component is modeled separately, and the behavior of the built-up structure is inferred by synthesis of component behavior. The decomposition and synthesis steps are designed to achieve reduction in sizes of matrices to be handled and (or) reduction in time required to integrate equations of motion. In the context of computing using distributed memory multiprocessor computers, substructuring is used to distribute workload equally among all the processors. On a different note, the idea of substructuring is also attractive in experimental studies. Here again, the structure to be studied is divided into a set of subsystems, and one, more, or all of the subsystems can be studied experimentally in their uncoupled states, and a model for behavior of built-up structure is synthesized therefrom. In such studies, where both

computational and experimental studies are combined, the substructuring methods are termed as hybrid simulation methods. Furthermore, in treatment of transient dynamical problems, if the coupling between numerical substructures and experimental substructures is achieved in real time, then one gets real-time substructuring methods. Apart from computational advantages, the substructuring methods afford other benefits too: (a) structures made up of technologically diverse components having notably different dynamical characteristics can be separately studied and developed by different teams, (b) better insights can be gained on global behavior of built-up structures in terms of local behavior of components, and (c) a combination of computational and experimental tools can be brought to bear on study of large complex systems. The literature on substructuring in structural mechanics is vast, and comprehensive overviews can be found in the works of Hurty (1965), Craig (1995), Maia and de Silva (1997), Ewins (2000), Williams and Blakeborough (2001), de Klerk et al. (2008), Bursi and Wagg (2008), and Saouma and Sivaselvan (2008). This entry details a few select set of tools for substructuring and briefly touches upon a few other issues.

Problem Statement

The motivations for substructuring in finite element modeling vary and so do the details of the methods and criteria for assessing their success. The following is a list of questions which may be conceived in this context:

(a) Given modal characteristics (natural frequencies, mode shapes, damping ratios, and participation factors), set of frequency response functions, or structural matrices and forcing vectors, for a set of N_S substructures in their uncoupled state, how to synthesize the dynamical characteristics and response of the built-up system? How to deal with incompleteness in spatial, modal, and frequency domains? How to best configure the substructures?

(b) How to deal with presence of nonlinearities in one or more of the subsystems?

(c) Given an N degree of freedom dynamical system, how to partition the degrees of freedom into N_S subsets so that efficient schemes can be evolved to integrate the equation of motion of the built-up system?

(d) If some of the subsystems (linear or nonlinear) are studied experimentally and the remaining numerically, how to interface the subsystem responses? How to carry out this in real time for transient dynamic response analysis?

(e) Given the response of a built-up system, how to infer behavior of component subsystems? (This is known as the problem of inverse substructuring.)

(f) Given a computer with multiple processors (generally distributed memory), how to solve the governing equation by decomposing the domain followed by distributing the workload among the processors so that the execution cost reduces significantly? How to best maximize the usage of computational resources and use simultaneously all available processors equally well?

(g) How to divide a large structure into N_S subsets without a corresponding increase in the number of interface coordinates involved in the component mode synthesis step?

In the same vein, one could pose several other questions pertaining to uncertainty propagation, system identification, vibration energy flow modeling, and treatment of multiphysics problems (such as fluid–structure interactions, soil–structure interactions, primary–secondary structure interactions, etc.). The discussion in the following sections will focus on a few of these issues.

Fixed-Interface Modal Synthesis

The method is illustrated by considering the system shown in Fig. 1a (Maia and de Silva 1997). The given system is divided into two substructures labeled as A and B, and the two

Substructuring Methods for Finite Element Analysis, Fig. 1 (**a**) Built-up system with two substructures; (**b**) two substructures in uncoupled states
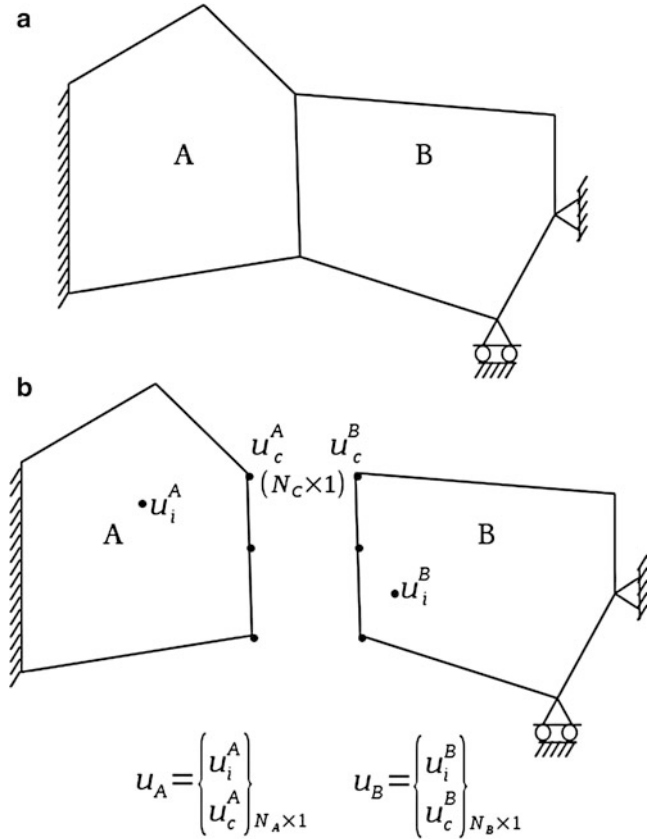
substructures in their uncoupled state are considered (Fig. 1b). The substructure degrees of freedom (dofs) are partitioned as shown into interior dofs (subscript i) and coupling dofs (subscript c). The equation for substructure A is written in the partitioned form as

$$\begin{bmatrix} M_{ii}^A & M_{ic}^A \\ M_{ci}^A & M_{cc}^A \end{bmatrix} \begin{Bmatrix} \ddot{u}_i^A \\ \ddot{u}_c^A \end{Bmatrix} + \begin{bmatrix} K_{ii}^A & K_{ic}^A \\ K_{ci}^A & K_{cc}^A \end{bmatrix} \begin{Bmatrix} u_i^A \\ u_c^A \end{Bmatrix} = \begin{Bmatrix} 0_i^A \\ f_c^A \end{Bmatrix}$$

$$(1)$$

Here $f_c{}^A(t)$ is the vector of interaction forces between A and B. Two separate analyses on this system are performed: first by assuming $u_c^A(t) = 0$ (fixed interface) and the second by taking $u_c^A(t) \neq 0$. In the first case, the governing equation is obtained as $M_{ii}^A \ddot{u}_i^A + K_{ii}^A u_i^A = 0_i^A$, and this equation is analyzed to determine the natural frequencies ω_r^A and modal vectors $\phi_r^A, r = 1, 2, \ldots, n_i^A$. The system response is

further represented using a k-term modal expansion as $u_{i(\text{Fix})}^A(t) = \left[\Phi_{ik_A}^A\right]\left\{p_{k_A}^A(t)\right\}$; the subscript k_A is used here to denote that the expansion has been truncated at k_A modes. In the second analysis, it is assumed that the interior and coupling dofs are related through conditions valid only under static conditions. This leads to the solution $u_{i(\text{Free})}^A(t) = -\left[K_{ii}^A\right]^{-1} K_{ic}^A u_c^A(t)$. The solution vector in Eq. 1 is now represented as

$$\begin{Bmatrix} u_i^A \\ u_c^A \end{Bmatrix} = \begin{bmatrix} \Phi_{ik}^A & -\left[K_{ii}^A\right]^{-1} K_{ic}^A \\ 0 & I \end{bmatrix} \begin{Bmatrix} p_{k_A}^A(t) \\ u_c^A(t) \end{Bmatrix}$$

$$= \Psi_k^A \begin{Bmatrix} p_{k_A}^A(t) \\ u_c^A(t) \end{Bmatrix} \qquad (2)$$

Here I is the identity matrix. Substituting this in Eq. 1, premultiplying by $\left[\Psi_{k_A}^A\right]^t$, and simplifying, one gets equation of the form

$$\begin{bmatrix} \tilde{M}^A_{kk} & \tilde{M}^A_{kc} \\ \tilde{M}^A_{ck} & \tilde{M}^A_{cc} \end{bmatrix} \begin{Bmatrix} \ddot{p}^A_{k_A}(t) \\ \ddot{u}^A_c(t) \end{Bmatrix} + \begin{bmatrix} \tilde{K}^A_{kk} & \tilde{K}^A_{kc} \\ \tilde{K}^A_{ck} & \tilde{K}^A_{cc} \end{bmatrix} \begin{Bmatrix} p^A_{k_A}(t) \\ u^A_c(t) \end{Bmatrix} = \begin{bmatrix} \Psi^A_{k_A} \end{bmatrix}^t \begin{Bmatrix} 0^A_i \\ f^A_c \end{Bmatrix}$$

with

$$\begin{bmatrix} \tilde{M}^A_{kk} & \tilde{M}^A_{kc} \\ \tilde{M}^A_{ck} & \tilde{M}^A_{cc} \end{bmatrix} = \begin{bmatrix} I & \left(\Phi^A_{ik}\right)^t \left(M^A_{ii}\alpha^A + M^A_{ic}\right) \\ \left(\alpha^A\right)^t M^A_{ii}\Phi^A_{ik} & \left(\alpha^A\right)^t \left(M^A_{ii}\alpha^A + M^A_{ic}\right) + \left(M^A_{ci}\alpha^A + M^A_{cc}\right) \end{bmatrix}$$

$$\begin{bmatrix} \tilde{K}^A_{kk} & \tilde{K}^A_{kc} \\ \tilde{K}^A_{ck} & \tilde{K}^A_{cc} \end{bmatrix} = \begin{bmatrix} \Lambda^A & \left(\Phi^A_{ik}\right)^t \left(K^A_{ii}\alpha^A + K^A_{ic}\right) \\ \left(\alpha^A\right)^t K^A_{ii}\Phi^A_{ik} & \left(\alpha^A\right)^t \left(K^A_{ii}\alpha^A + K^A_{ic}\right) + \left(K^A_{ci}\alpha^A + K^A_{cc}\right) \end{bmatrix}$$

$$\Lambda^A = \text{Diag}\left[\left(\omega^A_r\right)^2\right]; \alpha^A = -\left[K^A_{ii}\right]^{-1} K^A_{ic} \tag{3}$$

A similar analysis on substructure B leads to the equation

$$\begin{bmatrix} \tilde{M}^B_{kk} & \tilde{M}^B_{kc} \\ \tilde{M}^B_{ck} & \tilde{M}^B_{cc} \end{bmatrix} \begin{Bmatrix} \ddot{p}^B_{k_B}(t) \\ \ddot{u}^B_c(t) \end{Bmatrix} + \begin{bmatrix} \tilde{K}^B_{kk} & \tilde{K}^B_{kc} \\ \tilde{K}^B_{ck} & \tilde{K}^B_{cc} \end{bmatrix} \begin{Bmatrix} p^B_{k_B}(t) \\ u^B_c(t) \end{Bmatrix}$$
$$= \begin{bmatrix} \Psi^B_{k_B} \end{bmatrix}^t \begin{Bmatrix} 0^B_i \\ f^B_c \end{Bmatrix} \tag{4}$$

Returning to the built-up system in Fig. 1a and by imposing the conditions $u^A_c = u^B_c$ and $f^A_c + f^B_c = 0$ at the interfaces, one gets the governing equations for the built-up system as

$$\begin{bmatrix} I^A_{kk} & 0 & \tilde{M}^A_{kc} \\ 0 & I^B_{kk} & \tilde{M}^B_{kc} \\ \tilde{M}^A_{ck} & \tilde{M}^B_{ck} & \tilde{M}^A_{cc} + \tilde{M}^B_{cc} \end{bmatrix} \begin{Bmatrix} \ddot{p}^A_k(t) \\ \ddot{p}^B_k(t) \\ \ddot{u}_c(t) \end{Bmatrix}$$
$$+ \begin{bmatrix} \left(\omega^A_r\right)^2 & 0 & 0 \\ 0 & \left(\omega^B_r\right)^2 & 0 \\ 0 & 0 & \tilde{K}^A_{cc} + \tilde{K}^B_{cc} \end{bmatrix} \begin{Bmatrix} p^A_k(t) \\ p^B_k(t) \\ u_c(t) \end{Bmatrix} = 0 \tag{5}$$

The dynamic characteristics for the built-up system can now be determined using this model. The total number of dofs here is equal to the total number of substructure modes included in representing the fixed-interface responses for

systems A and B plus the number of interface dofs. This number is expected to be substantially less than the dof that one would get if the built-up system were to be analyzed without substructuring. This method of substructuring is also known as component mode synthesis and is a widely studied method in the existing literature. Various generalizations to include more than two substructures, presence of damping, substructure coupling through flexible elements, and response to external excitations have been explored.

Free-Interface Modal Synthesis

The fixed-interface method discussed in the preceding section requires the knowledge of subsystem structural matrices, and for this reason, they are not suited if one or more of the subsystems are studied experimentally (Maia and de Silva 1997). Also, creating fixed interfaces for substructures in experimental work is generally not feasible. An alternative formulation to overcome these difficulties is as follows. Consider the substructure Eq. 1 and determine the natural frequencies and mode shapes by solving the eigenvalue problem $K^A\phi^A = \omega^2_A M^A \phi^A$. This leads to the $N_A \times N_A$ matrices: $\Lambda^A = \text{diag}\left[\omega^2_{Ai}\right]$ and Φ^A such that $\left[\Phi^A\right]^t M^A \Phi^A = I$ and $\left[\Phi^A\right]^t K^A \Phi^A = \Lambda^A$. Now u^A is represented by a k_A term expansion $u^A(t) = \Phi^A_{k_A} p_{k_A}(t)$. The modal

matrix is further partitioned as $\Phi_{k_A}^A = \begin{bmatrix} \Phi_{k_A i}^A & \Phi_{k_A C}^A \end{bmatrix}^t$, and using Eq. 1, one gets

$$\mathrm{I}\ddot{p}_{k_A}(t) + \Lambda_{k_A}^A p_{k_A}(t) = \begin{bmatrix} \Phi_{k_A C}^A \end{bmatrix}^t f_c^A(t) \qquad (6)$$

A similar equation can also be obtained for subsystem B, and by combing these equations, one gets

$$[\mathrm{I}] \begin{Bmatrix} \ddot{p}_{k_A}(t) \\ \ddot{p}_{k_B}(t) \end{Bmatrix} + \begin{bmatrix} \Lambda_{k_A}^A & 0 \\ 0 & \Lambda_{k_B}^B \end{bmatrix} \begin{Bmatrix} p_{k_A}(t) \\ p_{k_B}(t) \end{Bmatrix}$$
$$= \begin{bmatrix} \begin{bmatrix} \Phi_{k_A C}^A \end{bmatrix}^t & 0 \\ 0 & \begin{bmatrix} \Phi_{k_B C}^B \end{bmatrix}^t \end{bmatrix} \begin{Bmatrix} f_c^A(t) \\ f_c^B(t) \end{Bmatrix} \qquad (7)$$

The condition $u_c^A(t) = u_c^B(t)$ at the interface is expressed as $\Phi_{k_A C}^A p_{k_A}(t) = \Phi_{k_B C}^B p_{k_B}(t)$. This equation is rewritten as

$$\begin{bmatrix} \Phi_{k_A C}^A & -\Phi_{k_B C}^B \end{bmatrix} \begin{Bmatrix} p_{k_A}(t) \\ p_{k_B}(t) \end{Bmatrix} = 0 \Rightarrow$$
$$Sp = 0 \text{ with } S = \begin{bmatrix} \Phi_{k_A C}^A & -\Phi_{k_B C}^B \end{bmatrix} \qquad (8)$$

Now, the matrix S is partitioned as $S = \begin{bmatrix} S_d & S_i \end{bmatrix}$ where S_d is a nonsingular square matrix and remaining part is S_i. This leads to

$$\begin{bmatrix} S_d & S_i \end{bmatrix} \begin{Bmatrix} p_d \\ p_i \end{Bmatrix} = 0 \Rightarrow p_d = -S_d^{-1} S_i p_i \text{ and}$$
$$\begin{Bmatrix} p_{k_A}(t) \\ p_{k_B}(t) \end{Bmatrix} = \begin{Bmatrix} p_d \\ p_i \end{Bmatrix} = \begin{bmatrix} -S_d^{-1} S_i \\ I \end{bmatrix} p_i = \Psi\{p_i\} \qquad (9)$$

Furthermore, by noting that $f_c^A(t) + f_c^B(t) = 0$, one gets

$$\mathrm{I} \begin{Bmatrix} \ddot{p}_{k_A}(t) \\ \ddot{p}_{k_B}(t) \end{Bmatrix} + \begin{bmatrix} \Lambda_{k_A}^A & 0 \\ 0 & \Lambda_{k_B}^B \end{bmatrix} \begin{Bmatrix} p_{k_A}(t) \\ p_{k_B}(t) \end{Bmatrix} = 0 \qquad (10)$$

By using Eq. 9 in the above equation and premultiplying by Ψ^t, one gets

$$M\ddot{q} + Kq = 0 \text{ with } M = \Psi^t \Psi; K = \Psi^t \begin{bmatrix} \Lambda_{k_A}^A & 0 \\ 0 & \Lambda_{k_B}^B \end{bmatrix} \Psi \qquad (11)$$

This equation can now be analyzed to deduce the built-up system natural frequencies and mode shapes. It is emphasized that the mass and stiffness matrices for the built-up system here are constructed in terms of the natural frequencies, and mode shapes of the subsystem in the uncoupled states and knowledge of substructure structural matrices are not needed in this formulation. This enables the introduction of experimentally studied substructures to be embedded into the modal synthesis of the built-up system.

Multilevel Substructuring

Consider an N-dof system being split into N_S substructures, such that dof of the largest substructure n satisfies the requirement $n << N$. If the value of N is large, the value of N_S also needs to be large in order to satisfy this requirement. This results in a large number of interface dofs appearing in the model thereby increasing the size of the synthesized system equation and consequent demands on computational efforts. The multilevel substructuring method addresses this issue by allowing for small substructures to be used without having to include a large number of interface dofs in the modal synthesis step.

The first step in multilevel substructuring is splitting the structure into multiple substructures, which form the first level of substructures. Then, each of these substructures is subdivided into multiple substructures, which form the second level. This process is continued until the substructures are as small as necessary. The substructures follow a tree topology, with the system to be analyzed at the top. At each level, each substructure has one parent one level above it and multiple-child substructures one level below. The equation of motion for each substructure (up to the penultimate level) can be

S

synthesized using its child substructures using component mode synthesis, and the interface elements belonging to those substructures alone appear in the synthesized system equation. Modal reduction is achieved in each of the substructures, and the interface dofs are retained. After the equations of motion of all substructures at a level are synthesized, the algorithm proceeds to the level above it. A fraction of the interface dofs at any level appears as interface dofs at the level above it, the rest turning into interior dofs, which can now be eliminated using modal reduction. Therefore, for the structure, the equation of motion can have, in the extreme case, only the interface dofs at the first level. This means that the method can be used to achieve a high degree of modal reduction. Additionally, at each level, the substructures can be analyzed in parallel, as in the case of single-level substructuring. At the higher levels, the number of substructures, and parallelizability, increases (see section "Domain Decomposition Techniques and Their Parallel Implementation" for issues related to parallelization).

The multilevel substructuring method has been used for linear and nonlinear structures with a large number of dofs (Papalukopoulos and Natsiavas 2007). Bennighof and Kaplan (1998) describe an adaptive multilevel substructuring procedure, where the contribution of each mode to the strain energy is used to decide whether or not to include the mode in the substructure analysis. This allows for a control on error due to modal reduction.

Nonlinear Substructures

Treatment of structural nonlinearities within the framework of ideas presented in the preceding sections offers several conceptual challenges. For problems involving linear substructures coupled through nonlinear elements, the fixed- and free-interface methods can be extended in a relatively easy manner. However, for problems involving globally distributed nonlinearities, the extensions

are not straightforward. An approach that employs fixed-interface component mode synthesis based on the idea of nonlinear normal modes has been developed by Apiwattanalunggarn et al. (2005). In a conservative structure vibrating in its nonlinear normal mode, the motion would be periodic and all points on the structure reach their respective extrema simultaneously and pass through zeros simultaneously (see Kerschen et al. 2009 for a detailed introduction). A few details of this approach are presented here.

Consider a nonlinear system divided into two nonlinear substructures A and B, as shown in Fig. 1a. The equation for substructure A is written in partitioned form as

$$\begin{bmatrix} M_{ii}^A & M_{ic}^A \\ M_{ci}^A & M_{cc}^A \end{bmatrix} \begin{Bmatrix} \ddot{u}_i^A \\ \ddot{u}_c^A \end{Bmatrix} + \begin{bmatrix} K_{ii}^A & K_{ic}^A \\ K_{ci}^A & K_{cc}^A \end{bmatrix} \begin{Bmatrix} u_i^A \\ u_c^A \end{Bmatrix} + \begin{Bmatrix} G_i^A(u_i^A, u_c^A) \\ G_c^A(u_i^A, u_c^A) \end{Bmatrix} = \begin{Bmatrix} 0_i^A \\ f_c^A \end{Bmatrix} \tag{12}$$

Here, $G_i^A(u_i^A, u_c^A)$ and $G_c^A(u_i^A, u_c^A)$ are the nonlinear forces associated with the interior and coupling dofs, respectively, and $f_c^A(t)$ is the vector of interaction forces between A and B. The procedure followed here has much in common with the fixed-interface method described earlier for linear substructures (section "Fixed-Interface Modal Synthesis"). In order to relate u_c^A to u_i^A, two separate analyses are done. First, fixed-interface linear modes are computed for A by assuming $u_c^A = 0$ and ignoring the nonlinear terms in Eq. 11. The governing equation is obtained as $M_{ii}^A \ddot{u}_{i(Fix)}^A + K_{ii}^A u_{i(Fix)}^A = 0$, which has natural frequencies ω_r^A and modal vectors $\phi_k^A, r = 1, 2, \ldots, n_i^A$. The system response represented by a modal expansion (truncated to k_A modes) is $u_{i(Fix)}^A = \left[\Phi_{i_{k_A}}^A \right] \left\{ p_{k_A}^A(t) \right\}$. Next, a relationship is obtained between u_c^A and u_i^A under static conditions, again ignoring nonlinear terms in Eq. 11. This leads to the relation $u_{i(Free)}^A(t) = - \left[K_{ii}^A \right]^{-1} K_{ic}^A u_c^A(t)$. The solution vector in Eq. 11 is represented as

$$\left\{\begin{array}{c} u_i^A \\ u_c^A \end{array}\right\} = \begin{bmatrix} \Phi_{i_{k_A}}^A & -\left[K_{ii}^A\right]^{-1} K_{ic}^A \\ 0 & I \end{bmatrix} \left\{\begin{array}{c} p_{k_A}^A(t) \\ u_c^A(t) \end{array}\right\} \tag{13}$$

Substituting Eq. 12 in Eq. 11 leads to

$$\begin{bmatrix} I & \tilde{M}_{ic}^A \\ \tilde{M}_{ci}^A & \tilde{M}_{cc}^A \end{bmatrix} \left\{\begin{array}{c} \ddot{p}_{k_A}^A \\ \ddot{u}_c^A \end{array}\right\} + \begin{bmatrix} \Lambda^A & 0 \\ \tilde{K}_{ci}^A & \tilde{K}_{cc}^A \end{bmatrix} \left\{\begin{array}{c} p_{k_A}^A \\ u_c^A \end{array}\right\}$$
$$+ \left\{\begin{array}{c} \tilde{G}_i^A\left(p_{k_A}^A, u_c^A\right) \\ \tilde{G}_c^A\left(p_{k_A}^A, u_c^A\right) \end{array}\right\} = \left\{\begin{array}{c} 0_i^A \\ f_c^A \end{array}\right\}$$
$$\tag{14}$$

Here, $\tilde{M}_{ic}^A, \Lambda^A, \tilde{M}_{ci}^A, \tilde{M}_{cc}^A, \tilde{K}_{ci}^A$, and $\tilde{K}_{cc}^A$ are given in Eq. 3, $\tilde{G}_i^A\left(p_{k_A}^A, u_c^A\right) = \left(\Phi_{i_{k_A}}^A\right)^T G_i^A\left(\Phi_{i_{k_A}}^A p_{k_A}^A, u_c^A\right)$, and $\tilde{G}_c^A\left(p_{k_A}^A, u_c^A\right) = G_c^A\left(\Phi_{i_{k_A}}^A p_{k_A}^A, u_c^A\right)$.

In Eq. 13, setting $u_c^A = 0$, one gets $\ddot{p}_{k_A}^A + \Lambda^A p_{k_A}^A + \tilde{G}_i^A\left(p_{k_A}^A, u_c^A\right) = 0_i^A$. Now, a nonlinear normal mode is constructed for this equation. The Mth coordinate of $p_{k_A}^A$, $p_{k_A,M}^A$, is taken as the master node, and all other dofs are related to it by the following constraints

$$p_{k_A,i}^A = X_i\left(p_{k_A,M}^A, \dot{p}_{k_A,M}^A\right), i = 1, 2, \ldots, k_A, i \neq M \tag{15}$$

$$\dot{p}_{k_A,i}^A = Y_i\left(p_{k_A,M}^A, \dot{p}_{k_A,M}^A\right), i = 1, 2, \ldots, k_A, i \neq M \tag{16}$$

These constraint relations are used to reduce A to a $N_C + 1$ degree of freedom system with the equation of motion

$$\begin{bmatrix} \hat{M}_{ii}^A & \hat{M}_{ic}^A \\ \hat{M}_{ci}^A & \hat{M}_{cc}^A \end{bmatrix} \left\{\begin{array}{c} \ddot{p}_{k_A,M}^A \\ \ddot{u}_c^A \end{array}\right\} + \begin{bmatrix} \hat{K}_{ii}^A & \hat{K}_{ic}^A \\ \hat{K}_{ci}^A & \hat{K}_{cc}^A \end{bmatrix} \left\{\begin{array}{c} p_{k_A,M}^A \\ u_c^A \end{array}\right\}$$
$$+ \left\{\begin{array}{c} \hat{G}_i^A\left(p_{k_A,M}^A, \dot{p}_{k_A,M}^A, u_c^A\right) \\ \hat{G}_c^A\left(p_{k_A,M}^A, \dot{p}_{k_A,M}^A, u_c^A\right) \end{array}\right\} = \left\{\begin{array}{c} 0 \\ f_c^A \end{array}\right\}$$
$$\tag{17}$$

Here, the terms $\hat{M}_{ii}^A, \hat{M}_{ic}^A, \hat{M}_{ci}^A, \hat{M}_{cc}^A, \hat{K}_{ii}^A, \hat{K}_{ic}^A, \hat{K}_{ci}^A, \hat{K}_{cc}^A$, $\hat{G}_i^A\left(p_{k_A,M}^A, \dot{p}_{k_A,M}^A, u_c^A\right)$, and $\hat{G}_c^A\left(p_{k_A,M}^A, \dot{p}_{k_A,M}^A, u_c^A\right)$ are obtained by substituting the constraints Eqs. 14 and 15 in Eq. 13. A similar procedure is followed for substructure B, and the equation of motion is obtained as

$$\begin{bmatrix} \hat{M}_{ii}^B & \hat{M}_{ic}^B \\ \hat{M}_{ci}^B & \hat{M}_{cc}^B \end{bmatrix} \left\{\begin{array}{c} \ddot{p}_{k_B,M}^B \\ \ddot{u}_c^B \end{array}\right\} + \begin{bmatrix} \hat{K}_{ii}^B & \hat{K}_{ic}^B \\ \hat{K}_{ci}^B & \hat{K}_{cc}^B \end{bmatrix} \left\{\begin{array}{c} p_{k_B,M}^B \\ u_c^B \end{array}\right\}$$
$$+ \left\{\begin{array}{c} \hat{G}_i^B\left(p_{k_B,M}^B, \dot{p}_{k_B,M}^B, u_c^B\right) \\ \hat{G}_c^B\left(p_{k_B,M}^B, \dot{p}_{k_B,M}^B, u_c^B\right) \end{array}\right\} = \left\{\begin{array}{c} 0 \\ f_c^B \end{array}\right\}$$
$$\tag{18}$$

Applying the constraints $u_c^A = u_c^B = u_c$ and $f_c^A + f_c^B = 0$, Eqs. 16 and 17 combine to form the following equation of motion for the built-up system:

S

$$\begin{bmatrix} \hat{M}_{ii}^A & 0 & \hat{M}_{ic}^A \\ 0 & \hat{M}_{ii}^B & \hat{M}_{ic}^B \\ \hat{M}_{ci}^A & \hat{M}_{ci}^B & \hat{M}_{cc}^A + \hat{M}_{cc}^B \end{bmatrix} \left\{\begin{array}{c} \ddot{p}_{k_A,M}^A(t) \\ \ddot{p}_{k_B,M}^B(t) \\ \ddot{u}_c(t) \end{array}\right\} + \begin{bmatrix} \hat{K}_{ii}^A & 0 & \hat{K}_{ic}^A \\ 0 & \hat{K}_{ii}^B & \hat{K}_{ic}^B \\ \hat{K}_{ci}^A & \hat{K}_{ci}^B & \hat{K}_{cc}^A + \hat{K}_{cc}^B \end{bmatrix} \left\{\begin{array}{c} p_{k_A,M}^A(t) \\ p_{k_B,M}^B(t) \\ u_c(t) \end{array}\right\}$$
$$+ \left\{\begin{array}{c} \hat{G}_i^A\left(p_{k_A,M}^A, \dot{p}_{k_A,M}^A, u_c^A\right) \\ \hat{G}_i^B\left(p_{k_B,M}^B, \dot{p}_{k_B,M}^B, u_c^B\right) \\ \hat{G}_c^A\left(p_{k_A,M}^A, \dot{p}_{k_A,M}^A, u_c^A\right) + \hat{G}_c^B\left(p_{k_B,M}^B, \dot{p}_{k_B,M}^B, u_c^B\right) \end{array}\right\} = 0$$
$$\tag{19}$$

The reduced system has $N_C + 2$ dofs. This procedure can be extended straightforwardly to more than two substructures.

Mesh Partitioning and Implicit–Explicit Schemes

Here time integration of equations of motion of nonlinear dynamical systems is considered. Methods to obtain numerical solutions can be either explicit or implicit in nature. In the context of linear dynamical systems, implicit methods of integration are preferred since they are unconditionally stable and permit the analyst to choose step size based on engineering judgment. On the other hand, explicit schemes for such systems are conditionally stable, and the step size is often controlled by the stability requirements. For nonlinear dynamical systems, implicit methods typically involve solution of nonlinear algebraic/transcendental equations at every time step. On the other hand, explicit schemes are non-iterative, and the computational effort needed at each time step is less than that for implicit schemes. Given this, in structural dynamic problems involving spatially localized nonlinearities, it becomes advantageous to bifurcate the degrees of freedom into those associated with linear and those with nonlinear parts of the system. Such a strategy has earlier been proposed by Hughes et al. (1979) and in more recent years has gained attention in the context of substructuring involving hybrid simulation (see section "Hybrid Simulations").

The essence of this strategy can be explained by considering an N degree of freedom dynamical system governed by the equation $M\ddot{U} + F(U,\dot{U}) = G(t), U(0) = U_0, \dot{U}(0) = \dot{U}_0$. The elements are bifurcated into an implicit and an explicit set, and the corresponding internal force vectors are $F^I(U,\dot{U})$ and $F^E(U,\dot{U})$, respectively. The equation of motion becomes $M\ddot{U} + F^I(U,\dot{U}) + F^E(U,\dot{U}) = G(t), U(0) = U_0, \dot{U}(0) = \dot{U}_0$. This equation is solved using a predictor–corrector scheme. The prediction step provides the initial estimates of the values of

displacement and velocity vectors at the $(n+1)^{th}$ time step, $n = 1, 2, \ldots$, as $\tilde{U}_{n+1}$ and $\dot{\tilde{U}}_{n+1}$, respectively. An implicit scheme is then used for the corrector step to obtain U_{n+1} and $\dot{U}_{n+1}$. The value of the explicit force vector used in the corrector step is $F^E\left(\tilde{U}_{n+1}, \dot{\tilde{U}}_{n+1}\right)$. If the implicit partition of the mesh is linear, $F^I(U,\dot{U})$ is of the form $C^I\dot{U} + K^I U$, where C^I and K^I are the implicit damping and mass matrices. The equation in the corrector step, $M\ddot{U}_{n+1} + C^I\dot{U}_{n+1} + K^I U_{n+1} = G(t_{n+1}) - F^E\left(\tilde{U}_{n+1}, \dot{\tilde{U}}_{n+1}\right)$, is also linear. Hence, a non-iterative implicit scheme can be used to obtain U_{n+1} and $\dot{U}_{n+1}$. In this procedure, although mesh partitioning is done, the implicit and explicit dofs are evaluated together. On the other hand, in problems of fluid–structure interaction, although the flow equation and the equation of motion of the structure are coupled, they are usually solved separately. In the strong coupling method (Ahn and Kallinderis 2006) based on predictor–corrector steps, an explicit scheme is used for the predictor step. The flow equation is then solved in order to provide the input for solving the equation of the structure. The corrector step involves solving the structure and flow equations in that sequence iteratively. Here, the time step used in the equations for the fluid and the structure has to be equal. If the structure is linear, implicit schemes can be used with a large time step. Both implicit and explicit schemes can be applied in solving the flow equations, with implicit schemes allowing for larger time steps, but requiring an iterative solution.

Domain Decomposition Techniques and Their Parallel Implementation

Present-day computer processors possess multicore architecture, and the usual sequential solvers are unable to use all the cores simultaneously. Therefore, the resource remains underutilized. Furthermore, with the advent of computer technology, parallel computers (often referred to as clusters) can be easily and economically built from off-the-shelf components.

Therefore, to take advantage of these inexpensive technologies, the solvers need to be parallelized.

As the name domain decomposition (DD) suggests, in these techniques the computational domain is decomposed into a set of subdomains, and a divide-and-conquer strategy is developed that permits using multiple processors more efficiently than the traditional single-domain approaches. For parallelizing a time-dependent (deterministic) problem, two options can be explored: (i) parallelizing in space and (ii) parallelizing in time. While there have been attempts on parallelizing in time for a few applications, it is generally very difficult due to the inherent sequential nature in the temporal domain. Whereas the spatial dependence is not sequential, although the subdomains are coupled. Therefore, to achieve parallelization in space, in a typical (iterative) DD method, (a) the spatial domain is divided into a number of subdomains, (b) each of the processors independently performs the computation related to one or more subdomains, and (c) at the end of each iteration, the processors communicate relevant information among them. As the iteration grows, the method converges toward achieving the global (considering the entire spatial domain) equilibrium and compatibility. A few dominant methods are component mode synthesis (CMS), finite element tearing and interconnecting (FETI) and its variants, and Schur complement (this is a direct method). One major use of DD methods is development of preconditioners for linear systems. When the condition number of the coefficient matrix A in a linear system $Ax = b$ becomes high, iterative techniques such as conjugate gradient becomes very slow. In this situation, a preconditioner matrix P is often used where PA has a low condition number, and finally $PAx = Pb$ is solved. DD is often used in developing P (Ghosh et al. 2009).

FETI methods are iterative methods where the subdomains communicate with each other through a set of Lagrange multipliers defined at subdomain boundaries. In a static problem, the equilibrium of each subdomain is satisfied at each iteration whereas the continuity of the displacement field is achieved at convergence.

The estimates from a DD method must be in good agreement with the exact result, that is, the results from a direct solver using a single domain. The success of a DD method is measured by its convergence rate toward the exact solution and by its scalability. In terms of scalability of an iterative method when the iteration count does not depend strongly on the problem size, then it is called numerical scalability. On the other hand, when an m-times bigger problem can be solved using m-times bigger computer (in terms of number of processors) in similar time, it is called parallel scalability. The efficiency of a DD method is often measured in terms of the ratio H/h where H denotes a characteristic size of a subdomain and h denotes element size.

A linear structural dynamics problem can be solved in two ways, either using a modal reduction or using a direct integration. For the modal approach, the domain decomposition can be used in the eigenvalue computation level, as described earlier in this entry (section "Fixed-Interface Modal Synthesis"). For direct integration, two separate approaches can be taken. First, casting the vibration problem in a single domain and adopting an implicit time integration scheme and then using the domain decomposition in the linear system solving level. To this end, the methods developed for solving elliptic equations can be directly used. For instance, when the Newmark-beta time integration is used, the system of linear algebraic equations can be solved using DD at every time step. The second option in the direct integration is to cast the structural dynamics problem in multiple domains itself and develop a time integrator. For this purpose, the compatibility and equilibrium among the subdomains must be ensured for all time instants or a selected set of time instants when different time steps are considered at different subdomains. In a FETI-type scheme, a Lagrange multiplier is used. Both these approaches are outlined below.

To explain the first option, consider a linear static problem (which essentially should be viewed as a system of linear algebraic equations)

$$Ku = f \qquad (20)$$

S

Now let the computational domain D be divided into N_s subdomains with the s-th subdomain denoted as $D^{(s)}$ and the boundary as $\partial D^{(s)}$. In the original formulation of FETI, these domains are completely unconnected; however, in a later development (FETI-DP), the corners of the subdomains are assumed to be connected. Following the FETI-DP formulation (Farhat et al. 2000), within the subdomain $D^{(s)}$, partition the displacement field $u^{(s)}$ as

$$u^{(s)} = \begin{bmatrix} u_{int}^{(s)} \\ u_{b_r}^{(s)} \\ u_{b_c}^{(s)} \end{bmatrix} = \begin{bmatrix} u_r^{(s)} \\ u_{b_c}^{(s)} \end{bmatrix} \text{ with } u_r^{(s)} = \begin{bmatrix} u_{int}^{(s)} \\ u_{b_r}^{(s)} \end{bmatrix}$$

(21)

where the subscript *int* means interior, b_c denotes the nodes at the corner, b_s denotes the nodes at the boundary but not on the corner, and r means the collection of *int* and b_r, that is, all the nodes except the corners – often referred to as *residual*. Accordingly, let the stiffness matrix and forcing vector be partitioned as

$$K^{(s)} = \begin{bmatrix} K_{rr}^{(s)} & K_{rc}^{(s)} \\ K_{cr}^{(s)} & K_{cc}^{(s)} \end{bmatrix}, f^{(s)} = \begin{bmatrix} f_r^{(s)} \\ f_{b_c}^{(s)} \end{bmatrix}$$

(22)

Let λ denote a vector-valued Lagrange multiplier defined globally over the subdomain interface dofs except the corner points, that is, on the dof denoted by b_r. Following these notations, the equilibrium for the interior and boundary (except the corner nodes) is written as

$$K_{rr}^{(s)} u_r^{(s)} + K_{rc}^{(s)} B_c^{(s)} u_c + B_r^{(s)T} \lambda = f_r^{(s)}; s = 1, 2, \ldots, N_s$$

(23)

and the equilibrium for the corner dofs is written as

$$\sum_{s=1}^{N_s} B_c^{(s)^T} K_{rc}^{(s)^T} u_r^{(s)} + \sum_{s=1}^{N_s} B_c^{(s)^T} K_{cc}^{(s)^T} B_c^{(s)} u_c$$
$$= \sum_{s=1}^{N_s} B_c^{(s)^T} f_{b_c}^{(s)}$$

(24)

where $B_c^{(s)}$ denotes a Boolean matrix mapping the global vector u_c containing all corner DOF to $u_{b_c}^{(s)}$ and $B_r^{(s)}$ denotes another Boolean matrix mapping the residual dof to $u_{b_r}^{(s)}$. The Boolean matrices are (usually rectangular) matrices with the entries as zeros or one. They are used to describe the binary relationship among the elements of two vectors. For instance, consider the Boolean relationship $B_c^{(s)} u_c = u_{b_c}^{(s)}$

where the Boolean matrix $B_c^{(s)}$ operates on the vector u_c enlisting all the corner DOFs and produces a vector $u_{b_c}^{(s)}$ of lower dimension that enlists only the corner DOFs in the subdomain $\mathcal{D}^{(s)}$. To explain it further, consider the ith row of the matrix $B_c^{(s)}$. The jth column in this row will be 1 if the jth element in the vector u_c corresponds to the ith element in the vector $u_{b_c}^{(s)}$, with all other entries being zero. The compatibility condition at the subdomain interface is stated as

$$u_{b_c}^{(s)} - u_{b_c}^{(q)} = 0 \text{ on } \partial D^{(s)} \cap \partial D^{(q)}$$

(25)

Upon a few algebraic operations, Eqs. 22, 23, and 24 are combined to a single system of linear algebraic equations with λ as the unknown vector. This resulting system is solved using a preconditioned conjugate gradient (PCG) method. The parallelization is achieved by distributing the subdomains among the processors and thereby distributing the operations involved in (Eqs. 22, 23, and 24).

To apply this method in a direct integration scheme, consider an implicit scheme such as Newmark-beta. Here the linear system of equations is needed to be solved at every time step

$$\hat{K} u_{t+\Delta t} = \hat{f}_{t+\Delta t}$$

(26)

where the coefficient matrix $\hat{K}$ and the vector $\hat{f}_{t+\Delta t}$ depend upon stiffness, mass and damping matrices, the parameters of the integration scheme, and the time step. In this case, the aforementioned DD solver can be used at every time step. Instead of an iterative method such as FETI, direct methods such as Schur complement can also be used which can be parallelized.

For explaining the Schur complement, consider the structure with two subdomains in Fig. 1a. Accordingly, the stiffness matrix, displacement vector, and the force vector are partitioned, and the static equilibrium condition is written as

$$
\begin{bmatrix} K_{ii}^A & 0 & K_{ic}^A \\ 0 & K_{ii}^B & K_{ic}^B \\ K_{ci}^A & K_{ci}^B & K_{cc}^A + K_{cc}^B \end{bmatrix} \begin{Bmatrix} u_i^A \\ u_i^B \\ u_c \end{Bmatrix} = \begin{Bmatrix} f_i^A \\ f_i^B \\ f_c \end{Bmatrix}
$$

(27)

To solve this equation, first the displacement at the boundary, u_c, is computed by solving

$$
\sum_{s=A,B} \left(K_{cc}^s - K_{ci}^s \left(K_{ii}^s \right)^{-1} K_{ic}^s \right) u_c
$$

$$
= \sum_{s=A,B} \left(f_c^s - K_{ci}^s \left(K_{ii}^s \right)^{-1} f_i^s \right)
$$

(28)

where the coefficient matrix is known as the Schur complement. Then the displacement at internal degrees of freedom is computed by solving

$$
K_{ii}^s u_i^s = f_i^s - K_{ic}^s u_c; s = A, B
$$

(29)

The computation in Eqs. 27 and 28 can be parallelized by performing the computation related to subdomains A and B in two different processors.

In the second approach (Prakash and Hjelmstad 2004), consider the equation of motion of a forced undamped motion as

$$
M\ddot{u}(t) + ku(t) = f(t)
$$

(30)

Now let the domain D be divided into N_s subdomains without the connectivity at the corner. Then the equation of motion for each subdomain can be written as

$$
M^{(s)}\ddot{u}^{(s)}(t) + K^{(s)}u^{(s)}(t) + B^{(s)^T}\lambda = f^{(s)}(t)
$$

(31)

where $B^{(s)}$ denotes a Boolean matrix mapping from the subdomain s to a global vector containing the dofs at all the inter-subdomain boundaries (interfaces) and λ is a Lagrange multiplier, this time with the units of velocity. Inter-subdomain compatibility is imposed in terms of velocity as

$$
\sum_{s=1}^{N_s} B^{(s)}\dot{u}^{(s)}(t) = 0
$$

(32)

Once again, using a Newmark-beta scheme, the discretized equation

$$
\mathcal{M}^{(s)}\ddot{\mathbf{u}}_{i+1}^{(s)}(t) + \mathcal{B}^{(s)}\lambda_{i+1}
$$

$$
= \mathbf{f}_{i+1}^{(s)}(t) - \mathcal{N}^{(s)}\mathbf{u}_i^{(s)}; s = 1, 2, \cdots, N_s
$$

(33)

where the index i denotes the time step,

$$
\mathbf{u}_i^{(s)} = \begin{Bmatrix} \ddot{u}_i^{(s)} \\ \dot{u}_i^{(s)} \\ u_i^{(s)} \end{Bmatrix}, \mathbf{f}_i^{(s)} = \begin{Bmatrix} f_i^{(s)} \\ 0 \\ 0 \end{Bmatrix}, \mathcal{B}^{(s)} = \begin{Bmatrix} B^{(s)^T} \\ 0 \\ 0 \end{Bmatrix},
$$

(34)

and the matrices $\mathcal{M}^{(s)}$ and $\mathcal{N}^{(s)}$ follow from the Newmark-beta scheme. Equations 32 and 31 are then solved to find the time history of response. Parallelization follows from the data independence of the subdomains in Eq. 32 by distributing the subdomains among the processors. Note that this method can be further refined to accommodate different time steps in different subdomains. In that case, the compatibility condition Eq. 32 is not enforced at every time step, but after a few time steps.

In all the DD methods described in this section, no reduction in the total number of degrees of freedom is made. Therefore, the error associated with static condensation of component mode synthesis where a set of degrees of freedom is eliminated from the calculation does not appear here. However, approximations arise from usage of iterative solvers. The main goal here is to reduce the total computational cost by distributing the total computational burden among a

S

number of processors. To this end, two important criteria must be followed for computational efficiency. First, the computational load must be equally distributed among the processors; this is called load balancing. Second, the inter-processor communication should be minimal, as communication is slower than computation within a processor. In a distributed memory system, the Message Passing Interface (MPI) is used for parallelization. Once a parallel algorithm is developed, often the data structure needs to be reworked. Then the computer program (written in, for instance, Fortran/C/C++) is augmented by a set of MPI commands to distribute and manage the computation among the processors (Karniadakis and Kirby 2003). The parallel computing would help not only in solving very large and complicated problems, but also in the context of reliability analysis where numerous repeated analyses are required.

Hybrid Simulations

The idea of substructuring also has applications in laboratory testing-based performance assessment of engineering systems for dynamic loads such as those included due to earthquakes. The test hardware here consists of computer-controlled servo-hydraulic actuators driving either a shake table or serving as loading devices in reaction-wall-based systems. In conventional test methods, these actuators are either in displacement or a force control and aim to apply prescribed time variations of dynamic loads on the test structure. This strategy however suffers from two drawbacks, namely, the need to geometrically scale and structure the study (dictated by limitations on payload capacity of the shake table and actuator force ratings) and the neglect of possible dynamic interaction between structural subsystems being tested (e.g., heavy machinery like turbines, rotors, and pumps) and the structure in which these systems are housed. Newer testing protocols aimed at overcoming these limitations which employ substructuring

schemes have been developed in recent years (Saouma and Sivaselvan 2008; Bursi and Wagg 2008; Williams and Blakeborough 2001). Two such strategies, namely, pseudodynamic testing and real-time substructuring, are discussed in the following:

The pseudodynamic test is carried out on a reaction-wall-based system using servo-hydraulic actuators under displacement control. The test structure under study is modeled as

$$
\begin{aligned}
M\ddot{X} + C\dot{X} &+ R[X(\tau), 0 \leq \tau \leq t] \\
&= -M\Gamma\ddot{x}_g(t), X(0) = X_0, \dot{X}(0) = \dot{X}_0 \quad (35)
\end{aligned}
$$

Here, the term R represents nonlinearities due to inelastic behavior of the structure, Γ is the influence matrix, and $\ddot{x}_g(t)$ is the earthquake-induced ground acceleration. The basic premise of this approach is that the inertial, viscous damping and external force characteristics can be numerically modeled while the inelastic nonlinear term R is obtained experimentally. This would mean that the terms $M\ddot{X}$, $C\dot{X}$, and $-M\Gamma\ddot{x}_g(t)$ constitute the numerical component of the test while R represents the experimental component. Furthermore, the time variable t is scaled to slow down such that the test involves the application of only static actions. A partial finite element model of the structure which characterizes the terms $M\ddot{X}$, $C\dot{X}$, and $-M\Gamma\ddot{x}_g(t)$ is first formulated, and it is embedded into the software that commands the servo-hydraulic actuators. The actuators are kept under displacement control. Beginning with an initial guess on stiffness characteristics, the finite element equations of motion are integrated to determine displacement vector $X(t)$ at $t = \Delta t$. These displacements are further applied on the test structure experimentally, and the reactions transferred to the walls are measured through load cells. These measured reactions serve to establish the term R in the governing equation of motion, which in turn is used to advance the integration steps. Thus the integration of equation of motion and measurement of nonlinear stiffness characteristics of

the structure go hand in hand leading to determination of structural response to a specified $\ddot{x}_g(t)$. The works of Takanashi and Nakashima (1987), Severn et al. (1989), Nakashima (2001), and Williams and Blakeborough (2001) provide comprehensive overviews on related topics. It is important to note that the earthquake-induced structural displacements are evaluated computationally and applied to the structure in a static manner. This test clearly does not permit experimental evaluation of time dependant nonlinear behavior, rate dependant stress–strain laws, and behavior of active elements, if any.

The real-time substructure test does not scale the time as is done in pseudodynamic testing. The test structure here is spatially divided into two parts: one which permits reliable numerical modeling and the other that requires experimental testing. The finite element model for the numerical substructure is embedded into the control software that controls the actuators, and the experimental substructure is placed on the test rig. Clearly the numerical and experimental models are coupled, and the integration of governing equations for the numerical model and the testing of experimental substructure take place hand in hand in real time. This requires online exchange of interfacial forces/displacements between the two substructures which in turn places demands on accurate sensing, fast data acquisition, and efficient numerical integration. Several issues related to time delays, measurement noise, and choice of integration schemes become critical in this context. Examples of recent studies include those by Sajeeb et al. (2009), Chen and Ricles (2010, 2012), and Gao et al. (2013).

approaches in an online or offline manner for steady-state/transient dynamics, and desire to equitably employ computational resources in a parallel computing environment. The challenges here pertain to choice of substructures, incompleteness in modal, spatial, and (or) frequency domains, dissipative mechanisms, local/global nonlinearities, and non-idealities in experimental work (such as measurement noise and time delays) and their role in selection of algorithmic parameters. The overview presented in this entry has endeavored to bring out currently available techniques and tools in the area of modeling.

Cross-References

▶ Classically and Nonclassically Damped Multidegree of Freedom (MDOF) Structural Systems, Dynamic Response Characterization of
▶ Modal Analysis

References

Ahn HT, Kallinderis Y (2006) Strongly coupled flow/structure interactions with a geometrically conservative ALE scheme on general hybrid meshes. J Comput Phys 219(2):671–696

Apiwattanalunggarn P, Shaw SW, Pierre C (2005) Component mode synthesis using nonlinear normal modes. Nonlinear Dyn 41(1–3):17–46

Bennighof JK, Kaplan MF (1998) Frequency window implementation of adaptive multi-level substructuring. J Vib Acoust 120(2):409–418

Bursi OS, Wagg D (eds) (2008) Modern testing techniques for structural systems: dynamics and control. Springer, New York

Chen C, Ricles J (2010) Tracking error-based servohydraulic actuator adaptive compensation for real-time hybrid simulation. J Struct Eng 136(4):432–440

Chen C, Ricles JM (2012) Large-scale real-time hybrid simulation involving multiple experimental substructures and adaptive actuator delay compensation. Earthq Eng Struct Dyn 41:549–569

Craig RR Jr (1995) Substructure methods in vibration. J Mech Des 117:207

Ewins DJ (2000) Modal testing: theory, practice and application. Research Studies Press, Baldock

Summary

A wide range of contexts in which substructuring is used in finite element modeling is reviewed. The motivations for substructuring include desire to reduce the model size, desire to combine experimental and numerical modeling

S

Farhat C, Lesoinne M, Pierson K (2000) A scalable dual–primal domain decomposition method. Numer Linear Algebra Appl 7:687–714

Gao X, Castaneda N, Dyke SJ (2013) Real time hybrid simulation: from dynamic system, motion control to experimental error. Earthq Eng Struct Dyn 42:815–832

Ghosh D, Avery P, Farhat C (2009) FETI-preconditioned conjugate gradient method for large-scale stochastic finite element problems. Int J Numer Methods Eng 80(6–7):914–931

Hughes TJ, Pister KS, Taylor RL (1979) Implicit-explicit finite elements in nonlinear transient analysis. Comput Methods Appl Mech Eng 17:159–182

Hurty WC (1965) Dynamic analysis of structural systems using component modes. AIAA J 3(4):678–685

Karniadakis GE, Kirby RM II (2003) Parallel scientific computing in C++ and MPI: a seamless approach to parallel algorithms and their implementation. Cambridge University Press, New York

Kerschen G, Peeters M, Golnival JC, Vakakis AF (2009) Nonlinear normal modes, part I: a useful framework for structural dynamicist. Mech Syst Signal Process 23:170–194

Klerk DD, Rixen DJ, Voormeeren SN (2008) General framework for dynamic substructuring: history, review and classification of techniques. AIAA J 46(5):1169–1181

Maia NMM, de Silva JMM (eds) (1997) Theoretical and experimental modal analysis. Research Studies Press, Taunton

Nakashima M (2001) Development, potential, and limitations of real–time online (pseudo–dynamic) testing. Philos Trans R Soc Lond Ser A Math Phys Eng Sci 359(1786):1851–1867

Papalukopoulos C, Natsiavas S (2007) Dynamics of large scale mechanical models using multilevel substructuring. J Comput Nonlinear Dyn 2(1):40–51

Prakash A, Hjelmstad KD (2004) A FETI-based multi-time-step coupling method for Newmark schemes in structural dynamics. Int J Numer Methods Eng 61:2183–2204

Sajeeb R, Manohar CS, Roy D (2009) A conditionally linearized Monte Carlo filter in nonlinear structural dynamics. Int J Nonlinear Mech 44:776–790

Saouma VE, Sivaselvan MV (eds) (2008) Hybrid simulation: theory, implementation and applications. CRC Press, London

Severn RT, Brownjohn JMW, Dumanoglu AA, Taylor CA (1989) A review of dynamic testing methods for civil engineering structures. In: Proceedings of the conference on civil engineering dynamics, University of Bristol, pp 1–24

Takanashi K, Nakashima M (1987) Japanese activities on on-line testing. J Eng Mech 113(7):1014–1032

Williams MS, Blakeborough A (2001) Laboratory testing of structures under dynamic loads: an introductory review. Philos Trans R Soc Lond Ser A Math Phys Eng Sci 359(1786):1651–1669

Sustained Earthquake Preparedness: Functional, Social, and Cultural Issues

Douglas Paton[1], Norio Okada[2], Julia Becker[3] and Li-ju Jang[4]
[1]School of Psychology, University of Tasmania, Launceston, TAS, Australia
[2]Graduate School of Science and Engineering, Kumamoto University, Kumamoto City, Japan
[3]GNS Science, Lower Hutt, New Zealand
[4]Department of Social Work, National Pingtung University of Science and Technology, Pingtung, Taiwan

Synonyms

Culture; Preparedness; Readiness; Risk communication

Introduction

When earthquakes occur, affected societies and their members suddenly find themselves having to deal with demands that differ considerably from anything they would encounter under normal conditions and in circumstances in which normal societal functions and resources are marked by their absence. The aftershock sequence that can accompany seismic events can prolong the period over which people have to deal with disruption. However, the degree of disruption and loss that people, communities, and societies experience is a function of the degree to which they have developed the knowledge, skills, and relationships required to anticipate, cope with, adapt to, and recover from earthquake consequences during both the initial event and the consequences they can encounter as they cycle through response and recovery processes with successive aftershocks. Furthermore, the fact that earthquakes occur without warning makes it imperative that people prepare prior to any event. Because it is impossible to predict when the next

earthquake will occur (it could be years or decades or longer into the future), preparedness strategies must ensure that once developed, preparedness is sustained by people and communities. The task of facilitating sustained preparedness occurs as part of a comprehensive risk management strategy.

Risk management offers ways in which societies and their members (individually and collectively) can make choices about mitigating risk and facilitating their preparedness to respond to seismic hazard events using a mix of mitigation and preparedness strategies. Mitigation and preparedness strategies play complementary roles in risk management.

Mitigation strategies describe ways societies can prepare themselves in advance of earthquakes occurring by undertaking activities designed to prevent or minimize the risk earthquakes pose to a society and its members. Mitigation encompasses, for example, land-use planning (e.g., precluding building in areas susceptible to liquefaction) and developing and implementing building codes and standards (e.g., specifying building codes to include, for instance, base isolation to increase building capacity to withstand ground shaking) and retrofitting existing buildings to facilitate their capacity to withstand the action of seismic hazards (at least up to a point). Building codes can reduce the adverse impacts of hazards and increase the likelihood of a building remaining available for people's use after the event. However, additional strategies may be required if it is possible to anticipate future earthquake events whose intensity or duration would exceed design parameters and retrofitting capability.

For example, a mitigation measure may function effectively when impacted by a 100-year event yet fail catastrophically if a 500-year (typically more intense) event is experienced. Thus, people could experience problems from events at the higher end of the spectrum of seismic intensities, magnitudes, and durations (e.g., continuing damage resulting from prolonged aftershocks such as what occurred in Christchurch) that could exceed the parameters of structural measures designed in part on

cost-benefit and political criteria (which reflect the level of risk mitigation a society is willing to pay for). This introduces into the risk management process a need for strategies designed to reduce the degree of loss and disruption experienced through encouraging personal and community preparedness.

The goal of preparedness strategies is to increase the likelihood of people and communities being in a position to be able to *respond* in planned and functional (resilient and adaptive) ways to the complex, challenging, and emergent challenges and demands that earthquakes create rather than having to *react* to them in ad hoc ways. For example, ensuring the physical integrity of the house and storing water and food helps people deal with the effects of ground shaking on the their home and on the loss of utilities. Taking steps to ensure the structural integrity of the home not only reduces the risk of injury and death to its inhabitants, but it also increases the likelihood of people having shelter during their recovery, reduces demands on societal resources for temporary accommodation, and increases the probability that people will remain in an area and be able to participate in social (mutual aid, social support), economic, and environmental recovery activities. The effectiveness of the mutual aid and social support people within communities can provide for each other and the quality of the working relationship between community members and civic agencies during and after earthquake events is a function of the degree to which community members have developed the knowledge, skills, and relationships with neighbors, community members, and civic agencies necessary to expedite local response and recovery initiatives before earthquakes occur (Paton and McClure 2013).

Preparedness programs aim to facilitate the proactive development of household, community, and societal capabilities (e.g., develop emergency plans and resources, capacity for self-reliance, ability to work with others to confront local problems, etc.) in ways that increase people's ability to anticipate what they may have to contend with and develop their capacity to cope with, adapt to, and recover from the physical, personal, and social consequences that

earthquakes create (Paton and McClure 2013). Achieving this goal falls within the remit of comprehensive risk management (i.e., where responsibility is shared between community members and agencies in ways that increase the likelihood of their respective activities playing complementary roles) in ways that devolve (partial) responsibility for personal and local risk management to individuals and communities. Involving people in risk management is pursued using risk communication and community outreach strategies.

In order to design and implement risk communication and community outreach strategies, it is important to know what to communicate and engage people about. The content of risk communication tends to reflect the fact that, fundamentally, risk is a product of the likelihood of earthquakes occurring and the consequences that arise when they do occur. Communicating about both, and particularly about consequences, is essential. Communicating information about the earthquake likelihood is not a good predictor of preparedness (McClure et al. 1999; Mileti and Darlington 1995). Preparedness is more likely when people believe they are likely to suffer negative consequences from an earthquake if they do not prepare (Palm et al. 1990). This suggests that risk communication strategies profit from including information about earthquake consequences and encourage citizens to personalize this risk in ways that focus on how they can prepare in ways that can prevent or minimize adverse earthquake consequences (Paton and McClure 2013). To do so requires several pieces of information.

The first concerns identifying what people need to know and be able to do (individually and collectively) to cope with and adapt to the earthquake hazard consequences that emerge and evolve over the course of the response, recovery, and rebuilding phases of earthquake disasters. That is, identifying earthquake consequences and the physical and social demands they create for people and identifying the strategies required to prepare for these consequences and demands (i.e., what being comprehensively prepared looks like). It is also important to identify why some people prepare and others less so or not at all.

Knowledge of the latter informs how risk communication and community outreach strategies can be developed to facilitate sustained, comprehensive preparedness. Finally, because earthquakes occur in culturally diverse locations, a third issue is ascertaining whether or not preparedness theories and practices are applicable across cultures. The first issue to be tackled is identifying what comprehensive preparedness is.

Comprehensive Earthquake Preparedness

The goal of preparedness strategies is to facilitate people's ability to deal with the full spectrum of demands that unfold over prolonged periods of response and recovery. People need to be able to deal with direct (e.g., impact of ground shaking, liquefaction, and aftershocks on buildings and infrastructure) and secondary (e.g., loss of lifelines like water, power, and sewerage services and the consequent need for people to be self-reliant and able to continue to function as well as possible in the absence of normal services) hazard consequences. As physical and social rebuilding occurs, people also need to be prepared to contend with challenges emanating from recovery processes (e.g., dealing with government agencies, insurance companies, builders, etc.) that may persist for weeks, months, or years. Furthermore, they may, as was evident in Christchurch, New Zealand, following the 2011 earthquake, have to cycle through response and recovery several times as they deal with the implications of aftershock sequences that prolonged people's experience of the physical and social consequences of earthquake activity (Paton et al. 2014). Recognition of the diverse issues people could encounter led to the development of functional typologies of earthquake preparedness that maps the content of preparedness programs onto the demands and challenges people have to contend with over time.

Functional Typologies

Russell et al. (1995) developed a preparedness typology that comprised three factors: structural,

survival, and planning preparedness. Structural actions encompass activities that secure the house (e.g., secure house to foundations) and its contents (e.g., securing water heaters and tall furniture) to prevent contents from injuring inhabitants (e.g., from ground shaking). Survival actions facilitate people's capacity for self-reliance during periods of disruption (e.g., ensuring a supply of water/dehydrated or canned food for several days, having a radio with spare batteries, etc.). Finally, planning includes, for example, developing household hazard plans and attending meetings to learn about earthquakes and how to deal with their consequences. The latter introduces a social dimension into how preparedness is conceptualized. A subsequent factor analytic study (Lindell et al. 2009) proposed that preparedness comprise direct action (e.g., learn how to shut off utilities, have a 4-day supply of canned food, strap heavy objects, etc.) and capacity building (e.g., join an earthquake-related organization, attend meetings about earthquake hazards) factors. These studies identify a need for preparedness strategies to facilitate the development of structural, survival, planning, and capacity building (social) activities. The importance of including all these aspects within a typology of comprehensive earthquake preparedness has been reinforced by studies of what people affected by earthquakes identified as being required to increase their ability to deal with the consequences of earthquakes.

For example, in Christchurch, NZ, residents affected by the 2011 earthquake reinforced the importance of developing (pre-event) structural and survival preparedness and household emergency planning (Paton et al. 2014). Christchurch respondents reported how their lack of survival and planning preparedness made coping with the loss of essential societal utilities (e.g., loss of water, power, and sewerage services) more challenging than it would have been had they prepared before the earthquake struck. Respondents also called for preparedness programs to address dealing with loss of or disruption to people's livelihoods (both directly from damage to place of work and indirectly from being relocated or injured), developing psychological preparedness,

adapting to changes in living conditions and loss of and disruption to family and social relationships, and accommodating the fact that the period of loss and disruption people have to confront extends to months and years. A need for effort to be directed to developing social relationships and social competencies within a preparedness strategy was identified by those affected by the 2011 earthquake.

Respondents suggested that preparedness programs should develop the ability of neighborhood and community group members to collaborate in ways that facilitate their ability to effectively confront local physical and social demands (e.g., removing rubble, providing mutual support, setting up community meeting places, taking care of those with special needs, organizing local efforts to repair homes, identifying and meeting local needs) when responding under conditions in which normal societal resources and functions are absent. Respondents also called for effort to be directed to preparedness programs including content to develop, for example, community leadership, social inclusion, prioritizing problems, problem-solving and decision-making skills. The latter, in conjunction with developing community capacity to represent diverse community and local needs to agencies and government will facilitate the ability of communities to proactively secure the resources they need to take responsibility for their own recovery (Paton et al. 2014).

Taken together, this discussion highlights a need for preparedness to encompass structural, survival, planning, and community capacity capabilities. The Christchurch work introduced a need to complement these functional categories with those addressing livelihood and psychological and community-societal relationship (e.g., between communities and governments, NGOs, business, etc.) preparedness. These are summarized in Fig. 1 (right-hand column). While the Christchurch research discussed what people realized they needed during a disaster, the focus of preparedness research is on developing these capacities and relationships prior to the occurrence of a hazard event. A need to research the latter reflects the finding that comprehensive preparedness is the exception rather than the rule.

S
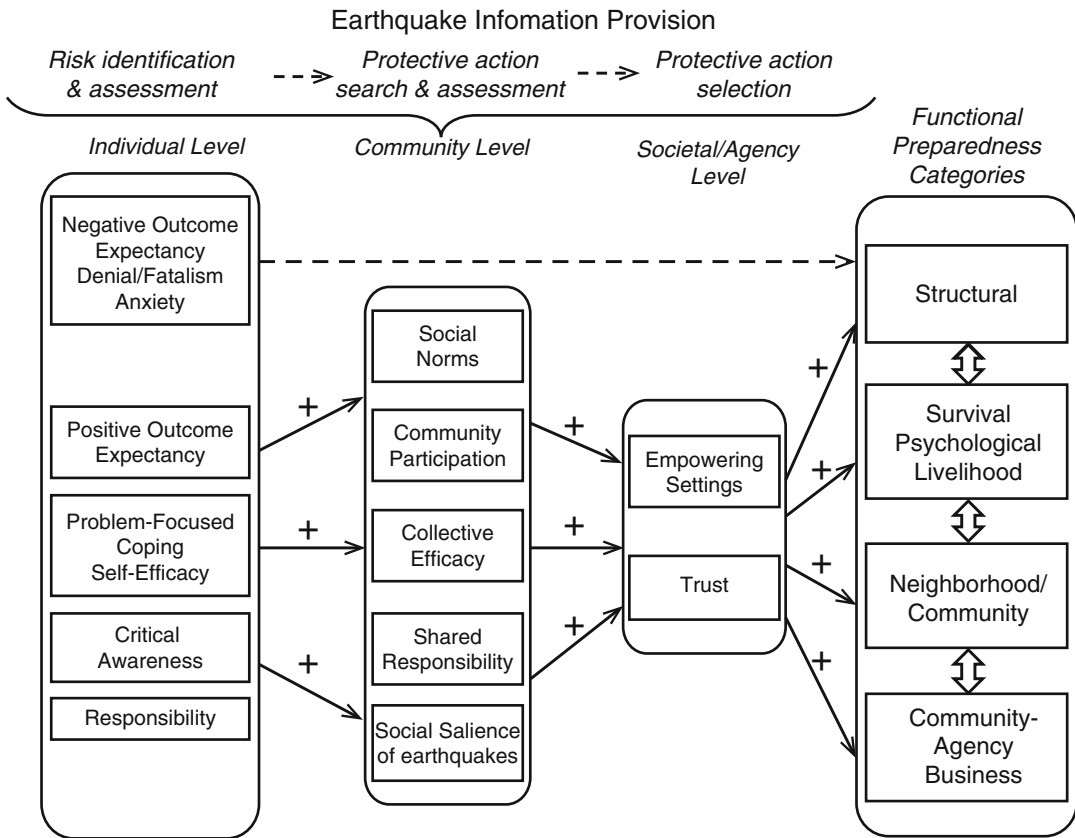
Sustained Earthquake Preparedness: Functional, Social, and Cultural Issues, Fig. 1 Summary of the relationship between protective action decision-making; personal-, social-, and agency-level predictors; and the functional categories of earthquake preparedness

Despite the considerable time, effort, and expenditure invested in earthquake public education programs, levels of preparedness for earthquakes typically remain low (Lechliter and Willis 1996; Lindell et al. 2009; Paton et al. 2014; Paton and McClure 2013). Finding considerable diversity in existing levels of preparedness introduces the second issue discussed in this chapter; the need to account for why some people prepare comprehensively, some much less so, and others not at all.

Accounting for Differences in Preparedness

When making decisions about uncertain, threatening events, several dispositional or individual-level predictors of preparedness have been identified. Some, such as fatalism, reduce the likelihood of preparing. Others influence how people make judgments about who is responsible for their preparedness. Some increase the likelihood of people preparing.

People who are fatalistic about hazard activity are unlikely to prepare (Turner et al. 1986). Another dispositional characteristic that reduces preparedness is denial (Crozier et al. 2006). If people believe that they have no control over an earthquake and/or its activity, they can attempt to cope with this by denying the seriousness of the risk. Household and personal preparedness is also less likely if people believe that risk management agencies are responsible for all aspects of public safety and so transfer responsibility for community preparedness from themselves to agencies (Paton and McClure 2013). Certain characteristics of people's social relationships influence preparedness.

If people believe that their significant others (parents, spouses, friends, peer group, etc.) hold favorable attitudes toward preparedness or that performance of a specific behavior is likely to be interpreted favorably by significant others, they are more likely to perform these actions (Smith and Terry 2003). Prevailing social norms regarding preparedness expectations of significant others and personal motivation to act in ways consistent with these expectations increases the likelihood of preparing (McIvor and Paton 2007). The work on individual predictors has been further developed in several theories of earthquake preparedness.

Person-relative-to-event (PrE) theory (Duval and Mulilis 1995) describes how the degree to which people perceive themselves at risk from an earthquake interacts with person (self-efficacy, outcome efficacy) and event (severity of event and probability of occurrence of the event) variables to predict preparedness. The PrE model argues that someone who appraises their resources as sufficient in both quality and quantity *relative* to the perceived magnitude of a particular threat will engage in more problem-focused coping activities than one who appraises their personal resources as insufficient relative to the seriousness of the threat. This work introduces a need to include perceived vulnerability, severity, response efficacy (outcome expectancy), problem-focused coping, and anxiety (fear-arousing persuasive communication) as predictors of earthquake preparedness (see Fig. 1).

Building on ideas advanced in the PrE theory, Paton et al. (2005) developed a theory that identified how critical awareness or the relative social salience (i.e., the extent to which people think and talk about earthquakes and earthquake preparedness with others in their community) of earthquakes was a key motivator of preparedness. Paton et al.'s (2005) critical awareness theory demonstrated that the anxiety, outcome expectancy (response efficacy), self-efficacy, and problem-focused/action coping mediated the relationship between critical awareness and preparedness and so identified a need for these variables to be included as predictors of earthquake

preparedness (see Fig. 1). Other theories have focused on how people's information search and evaluation influences the degree to which they take protective action.

The Protective Action Decision Model (PADM) (e.g., Lindell and Hwang 2008) argues that people's decisions about earthquake preparedness reflects the judgments they make about three factors: the threat posed by an earthquake, identifying the actions available and capable of protecting them against earthquake consequences, and identifying the sources from whom information about hazard and protective measures can be obtained. The PADM proposes that this involves people proceeding through several stages of information seeking and evaluation: detection/warning, psychological preparation, logistical preparation, and protective action selection and implementation (see Fig. 1). The information required for people to act at each stage can be sourced from the physical environment and/or people's social context (e.g., other people in their community, emergency management, media, etc.).

Lindell's work discussed the fact that people do not take information from expert sources at face value and introduced a need to consider both information about earthquake risk and preparedness and the quality of the (communication) relationship between the sources and recipients of earthquake risk information. An important contextual element of this relationship is that it occurs under conditions of uncertainty. People may have no or limited experience of earthquakes, and the infrequent nature of earthquake activity makes it impossible for them to independently test the validity of the information about risk and preparedness. Under these circumstances, an important aspect of the relationship between source of information and its recipients is the degree to which recipients trust sources of information. Levels of risk acceptance and people's willingness to take responsibility for their own safety is increased, and decisions to take steps to actively manage their risk more likely, if people believe that their relationship with risk management agencies is fair and empowering (e.g., agencies are perceived as trustworthy, as

acting in the interest of community members) (e.g., Paton 2008).

Paton (2008) developed a community engagement theory that captured key elements of these relationships, empowerment and trust, and the personal beliefs and social competencies that influenced preparedness decision-making under uncertainty. This theory has demonstrated an ability to predict earthquake preparedness (Paton 2013).

Paton (2008) argued that if people hold negative outcome expectancy beliefs (belief that earthquakes are uncontrollable and their consequences are too catastrophic for personal action to make any difference to peoples' safety), their likelihood of preparing is reduced. In contrast, holding positive outcome expectancy beliefs (beliefs that personal actions can enhance personal safety and/or mitigate hazard consequences) motivates people to start to prepare. However, the infrequent and complex nature of earthquakes means that believing that preparing can be effective does not necessarily equate with knowing what to do or how to prepare. To reduce uncertainty and guide action, people look to other community members and to expert sources (Paton 2008). Paton's community engagement theory introduced a need to include outcome expectancy, community participation and collective efficacy (which serve to socially construct risk beliefs and identify how risk can be reduced though social action by articulating community members preparedness plans), empowerment, and trust as predictors of earthquake preparedness (see Fig. 1).

Figure 1 summarizes the several variables that were implicated as predictors in the theories discussed above. It also includes how the PADM suggestions about information search and action might relate to these other variables. Collectively, this depiction could serve as a framework for future research into earthquake preparedness. The theories introduced above have empirically demonstrated their ability to account for differences in levels of earthquake preparedness. This confers upon them a capacity to inform the development of risk communication and community outreach strategies (see below)

by identifying how information content and personal and social competencies and relationships complement one another in ways that facilitate the ability of people to make earthquake preparedness decisions under conditions of uncertainty. However, given the fact that earthquakes affect people in countries characterized by considerable social and cultural diversity, it cannot be automatically assumed that they will retain their predictive utility if applied in different cultural contexts. Developing a comprehensive understanding of earthquake preparedness requires considering this cultural dimension.

Cross-Cultural Issues in Earthquake Preparedness

One reason for exercising caution in this regard derives from the fact that the theories introduced above were developed and tested in countries that are culturally highly individualistic. The variables they tested reflected this. In contrast, Asian countries, where a disproportionate number of significant and damaging earthquakes occur, tend to fall at the collectivistic end of the cultural spectrum in which social processes play relatively more important roles in people's decisions and actions (Triandis 1995). This raises issues regarding the applicability of Western-developed theories in Asian cultures that typically lie at the collectivistic end of these cultural dimensions. Examining whether theory can be applied across the individualistic-collectivistic spectrum is important (Paton et al. 2013).

A preliminary investigation of cultural equivalence was undertaken by examining whether Paton's (2008) community engagement theory (see above) of preparedness could account for differences in earthquake preparedness in countries (New Zealand, Japan, and Taiwan) that differ with regard to their relative positions on the individualistic-collectivistic cultural spectrum (see Fig. 2). This theory was selected for this initial comparison because its inclusion of both individual-level (outcome expectancy) and community-/social-level process (community participation and collective efficacy) variables provides a useful starting point for studying cultures that
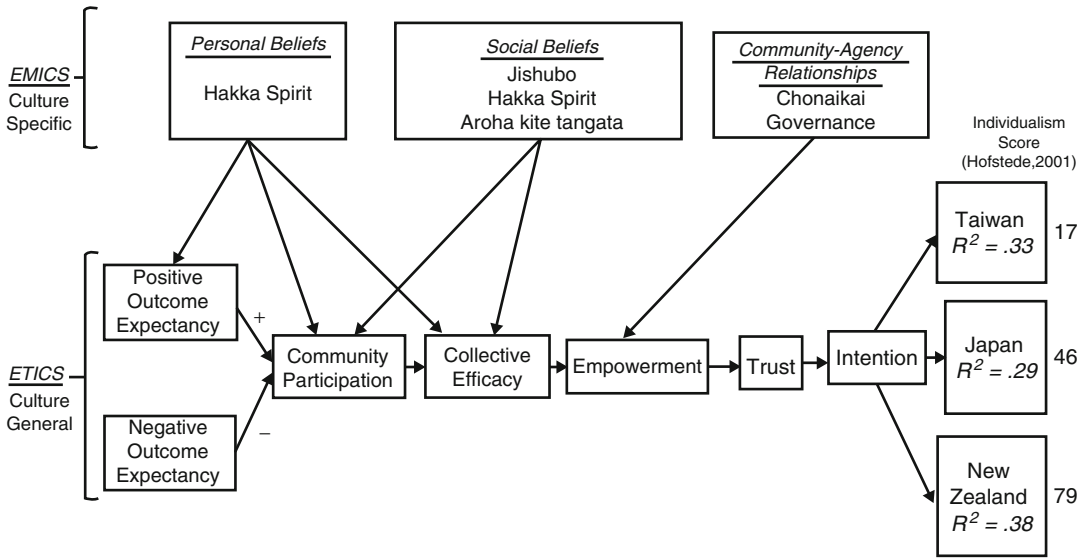
Sustained Earthquake Preparedness: Functional, Social, and Cultural Issues, Fig. 2 Summary of cross-cultural studies of earthquake preparedness and the relative contribution of culture-general (etic) and culture-specific (emic) processes to earthquake preparedness

differ with regard to their implicit emphasis on individual (individualistic cultures) and social (collectivistic) characteristics. For example, the relative position of a country on the I-C dimension could influence the comparative importance of person-level versus group-level factors on preparedness decisions (Paton et al. 2013). Empirical analysis of the ability of this theory to predict earthquake preparedness in New Zealand, Japan, and Taiwan (Fig. 2) supports the view that Western theory can be used across cultural borders (Paton et al. 2013). In Fig. 2, the R^2 value refers to how well the model can explain differences in people's intention to prepare for earthquakes (e.g., in Taiwan, the model explained 33% (or .33) of these differences). Models that predict over 28% are regarded as offering a good explanation (Sheeran 2002).

Evidence for cross-cultural equivalence should not, however, be taken to imply that cultural factors can be ignored. Cultural studies distinguish between etic (culturally universal) and emic (culture-specific) processes. While the cross-cultural analyses discussed above identified how the same process (Fig. 2) could operate across different countries (i.e., it describes an etic process), this does not mean

that the process is enacted in the same way in different cultures. For example, community participation was important in each country. However, the origins of participatory activities and how they are sustained and enacted to influence people's risk beliefs and preparedness options differ from country to country. Thus, emic processes (i.e., how and why participation occurs in a specific culture) capable of affecting earthquake preparedness need to be identified. Doing so can provide valuable insights into the theoretical and practical aspects of earthquake preparedness (see below). This is illustrated here with reference to New Zealand Māori and to how Jishubo (in Japan) and the Hakka Spirit (in Taiwan) describe emics that affect earthquake preparedness.

The importance of accommodating emic characteristics was evident in New Zealand. Analysis of unique cultural characteristics capable of facilitating social preparedness was evident in Māori populations. The Māori cultural value of "aroha ki te tangata" (love to all people) represents an emic characteristic (Fig. 2) capable of informing understanding of how community participation influences preparedness in this population (Paton et al. 2014).

Jishu-bosai-soshiki (Jishubo) are neighborhood-based "autonomous organizations for disaster reduction" unique to Japan (Bajek et al. 2008). Jishubo is a resource available to local government that can be used to organize earthquake preparedness events within communities (e.g., drills, workshops) and facilitate local response and rescue activities (e.g., rescue residents, provide initial first aid, supply food and water to survivors) should an earthquake occur. Jishubo are organized and operate through Chonaikai ("community councils"), a traditional Japanese neighborhood governance system that includes within its functions the maintenance of safe and secure communities through implementing various community activities (e.g., disaster risk reduction and crime prevention). This example illustrates how culture-specific (emic) characteristics can influence how community (Jishubo) and community-agency (Chonaikai) relationships (i.e., they represent unique ways of enabling community participation and empowerment) inform earthquake preparedness in ways unique to Japanese society (Bajek et al. 2008) (see Fig. 2). Another example of how culture-specific social mechanisms can influence earthquake preparedness is found in Taiwan.

Jang and LaMendola (2006) discussed how the Hakka Spirit, a cultural characteristic of the Hakka people residing in Tung Shih in Taiwan (affected by the 921 earthquake that hit Taiwan in 1999) influences earthquake preparedness. The Hakka Spirit describes an emic process, a unique set of cultural beliefs and practices that encompasses qualities such as frugality, diligence, self-reliance, responsibility, and persistence that combine to predispose people to prepare for adversity. Jang and LaMendola discussed how farming practices that evolved to limit the impact of typhoons on people's livelihoods generated enduring outcome expectancy beliefs (regarding the effectiveness of collective actions in response to hazard consequences). These were linked to increasing the likelihood that people would act to secure their homes from earthquake consequences. Culturally specific reciprocal support and collaborative problem-solving practices, developed over time as part of an adaptive response to meeting the challenges imposed by agricultural life in remote mountain communities, represent the source of the community participation and collective efficacy described in the above analyses (Fig. 2).

Practical Implications

Being able to identify (Fig. 1) factors that facilitate and hinder earthquake preparedness can inform the development of practical risk communication and community outreach strategies. For example, fatalistic attitudes can be reduced by asking people what can be done to help specific vulnerable groups, such as people living in unsafe buildings or young children in schools (McClure et al. 2001; Turner et al. 1986). When people deliberated about preparedness in vulnerable groups, they became less fatalistic and were more likely to perceive preparedness actions as more manageable. Denial and negative outcome expectancy can be countered by providing information that focuses people's attention specifically on how preparedness actions can mitigate the consequences of hazard activity and increase the degree of control people perceive themselves having over earthquake consequences (Lehman and Taylor 1987) and by sourcing information about preparedness and its effectiveness from comparable (to the target one) communities (Paton and McClure 2013).

Positive outcome expectancy beliefs can be developed by presenting people with a small number of items initially, starting with relatively easily adopted items and introducing progressively more complex tasks over time (Lindell and Perry 2000). By presenting specific explanations about how preparedness actions reduce risk and mitigate hazard consequences progressively over time, sustained adoption is more likely (Paton and McClure 2013). Social and relationship predictors (e.g., community participation, collective efficacy, social norms, etc.) can be enhanced by integrating community development and risk management strategies and by ensuring that community outreach adopted empowerment and community engagement principles (Paton and McClure 2013). This raises the possibility that one reason for the success of

Chonaikai is by integrating its disaster risk reduction activities with other community activities.

The work discussed above introduced how earthquake preparedness theory can possess etic quality that allows it to be used in a broad range of cultural contexts. Demonstrating that occidental theories retain their predictive utility in oriental contexts has several theoretical and practical benefits. For example, it would provide a common theoretical basis for collaborative learning and research across national borders. Demonstrating a degree of cross-cultural theoretical equivalence provides risk management agencies in different countries with access to a wider range of risk management options (which are implemented by accommodating local emic processes). It would make information about how to facilitate preparedness available to countries that lack the resources to undertake this work themselves and inform the development of humanitarian aid planning knowing that they have a robust framework that could be used to assist earthquake recovery irrespective of the country they find themselves being deployed at short notice. It remains important, however, to accommodate emic, culture-specific processes in the development of comprehensive earthquake preparedness theory and practice (e.g., by mapping local cultural practices (emic) onto this etic framework). Researching emic cultural influences on how earthquake preparedness and recovery processes are developed, sustained, and enacted also opens up new research and outreach opportunities for developing progressively comprehensive theories of earthquake preparedness.

Summary

The three countries that were specifically discussed in this chapter, New Zealand, Japan, and Taiwan, along with many others, all sit on the circum-Pacific seismic belt (the Pacific Ring of Fire). Approximately 90 % of the world's earthquakes and 81 % of its largest earthquakes occur along the Ring of Fire. Earthquakes will remain an inevitable aspect of the natural history of these countries. However, people can prepare in ways that can reduce their risk of experiencing adverse earthquake consequences and recover more promptly and effectively should an earthquake occur. This chapter first identified a need for preparedness to cover structural, planning, psychological, livelihood, community, and community-agency relationship categories (Fig. 1). Despite the evident advantages of being prepared, the fact that some people prepare comprehensively and others less so or not at all identifies a need for robust theories of hazard preparedness to inform DRR practice. This chapter also recognized the need for preparedness theory and strategy to accommodate the social and cultural diversity inherent in countries whose citizens must live with seismic risk. By demonstrating how theory can predict earthquake preparedness in culturally diverse countries, the chapter illustrated the potential for collaborative learning, research, and practice across national borders. However, it is important to ensure that national work accommodates the unique cultural resources present in each country that can be used to facilitate sustained earthquake preparedness.

Cross-References

▶ Building Disaster Resiliency Through Disaster Risk Management Master Planning
▶ Community Recovery Following Earthquake Disasters
▶ Earthquake Disaster Recovery: Leadership and Governance
▶ Interim Housing Provision Following Earthquake Disaster
▶ Resilience to Earthquake Disasters
▶ Site Response for Seismic Hazard Assessment
▶ Social Media Benefits and Risks in Earthquake Events

References

Bajek R, Matsuda Y, Okada N (2008) Japan's Jishu-bosai-soshiki community activities: analysis of its role in participatory community disaster risk management. Nat Hazards 44:281–292

Crozier M, McClure J, Vercoe J, Wilson M (2006) The effects of land zoning information on judgments about earthquake damage. Area 38:143–152

Duval TS, Mulilis J-P (1995) A person-relative-to-event (PrE) approach to negative threat appeals and earthquake preparedness: a field study. J Appl Soc Psychol 29:495–516

Jang L, LaMendola W (2006) The Hakka spirit as a predictor of resilience. In: Paton D, Johnston D (eds) Disaster resilience: an integrated approach. Charles C. Thomas, Springfield, pp 174–189

Lechliter GJ, Willis FN (1996) Living with earthquakes: beliefs and information. Psychol Rec 46:391–396

Lehman D, Taylor SE (1987) Date with an earthquake: coping with a probable, unpredictable disaster. Pers Soc Psychol Bull 13:546–555

Lindell MK, Hwang SN (2008) Households' perceived personal risk and responses in a multi-hazard environment. Risk Anal 28:539–556

Lindell MK, Perry RW (2000) Household adjustment to earthquake hazard: a review of research. Environ Behav 32:461–501

Lindell MK, Arlikatti S, Prater CS (2009) Why people do what they do to protect against earthquakes risk: perceptions of hazard adjustment attributes. Risk Anal 29:1072–1088

McClure J, Walkey F, Allen M (1999) When earthquake damage is seen as preventable: attributions, locus of control and attitudes to risk. Appl Psychol 48:239–256

McClure J, Allen MW, Walkey F (2001) Countering fatalism: causal information in news reports affects judgements about earthquake damage. Basic Appl Soc Psychol 23:109–121

McIvor D, Paton D (2007) Preparing for natural hazards: normative and attitudinal influences. Disaster Prev Manag 16:79–88

Mileti DS, Darlington JD (1995) Societal response to revised earthquake probabilities in the San Francisco Bay area. Int J Mass Emerg Disasters 13:119–145

Palm R, Hodgson M, Blanchard RD, Lyons D (1990) Earthquake insurance in California. Westview Press, Boulder

Paton D (2008) Risk communication and natural hazard mitigation: how trust influences its effectiveness. Int J Glob Environ Issues 8:2–16

Paton D (2013) Disaster resilient communities: developing and testing an all-hazards theory. J Integr Disaster Risk Manag 3:1–17

Paton D, McClure J (2013) Preparing for disaster: building household and community capacity. Charles C. Thomas, Springfield

Paton D, Smith L, Johnston D (2005) When good intentions turn bad: promoting natural hazard preparedness. Aust J Emerg Manag 20:25–30

Paton D, Okada N, Sagala S (2013) Understanding preparedness for natural hazards: a cross cultural comparison. J Integr Disaster Risk Manag 3:18–35

Paton D, Johnston D, Mamula-Seadon L, Kenney CM (2014) Recovery and development: perspectives from New Zealand and Australia. In: Kapucu N, Liou KT (eds) Disaster & development: examining global issues and cases. Springer, New York

Russell LA, Goltz JD, Bourque LB (1995) Preparedness and hazard mitigation actions before and after two earthquakes. Environ Behav 27:744–770

Sheeran P (2002) Intention-behaviour relations: A conceptual and empirical review. In: Stroebe W, Hewstone M (eds) European review of social psychology. Wiley, Chichester, pp 1–36

Smith JR, Terry DJ (2003) Attitude-behaviour consistency: The role of group norms, attitude accessibility, and mode of behavioural decision-making. European J of Soc Psychol 33:591–608

Triandis HC (1995) Individualism and collectivism. Westview, Boulder

Turner RH, Nigg JM, Paz DH (1986) Waiting for disaster: earthquake watch in California. University of California Press, Los Angeles

Symmetric Triaxial Seismometers

Bruce Townsend
Nanometrics, Inc., Kanata, Ontario, Canada

Synonyms

Galperin configuration seismometer; Homogeneous triaxial seismometer

Introduction

Of the class of seismic instruments measuring ground motion known as triaxial seismometers, most provide three signal outputs that represent mutually orthogonal motions in the East, North, and Vertical (or X, Y, and Z) directions (see entry "▶ Principles of Broadband Seismometry"). Of these, some are designed with three independent internal sensors that are respectively sensitive to motions in the three XYZ directions, and so directly measure the single vertical and two horizontal motion degrees of freedom. Some triaxial seismometers use a different configuration known as a Galperin arrangement (Galperin 1955), also known as a "symmetric triaxial" or "homogeneous" design (Melton and Kirkpatrick 1970).

In the Galperin configuration, the sensing elements are also arranged to be mutually orthogonal, but instead of one axis being vertical, all three are inclined upwards from the horizontal at precisely the same angle, as if they were aligned with the edges of a cube balanced on a corner. The operational principles of a design based on the Galperin configuration have many similarities as well as some important differences relative to one based on the conventional XYZ arrangement. The Galperin configuration presents some significant benefits for the users, owners, and manufacturers of seismometers that use this topology, as well as implying some trade-offs users should be aware of.

Benefits of a symmetric triaxial seismometer include being able to more easily distinguish external noise sources from internal ones, that mass centering does not compromise mutual orthogonality of the three axes, assurance of well-matched responses of the three outputs, and ability to achieve higher performance in smaller packages. A trade-off made by the Galperin approach is that unlike conventional XYZ seismometers that can suffer the failure of one axis element while continuing to provide valid output signals for the remaining two directions, all elements in a symmetric triaxial system must function for any of its XYZ outputs to be valid.

The first modern broadband seismometer to be based on a Galperin configuration was the Streckeisen STS-2, an observatory-grade vault instrument introduced in 1990. Nanometrics introduced a symmetric triaxial seismometer with a 240 s lower corner period, the Trillium 240, in 2004. There are now a wide variety of symmetric triaxial seismometers available including models designed for borehole, posthole, ocean bottom, and vault installation and ranging from ultracompact to the large form factors.

## Principles of Operation

Inertial Sensing Principles
Measuring motion must always be done relative to some frame of reference. Seismometers by definition use an inertial reference; that is, motion is sensed relative to a proof mass that is suspended in such a way that the proof mass tends to remain stationary while the seismometer that is coupled to the earth (or a structure) moves relative to it. The sensing element includes some means of measuring the motion of the sensing frame to its suspended proof mass. The proof mass in most broadband seismometers is a boom hinged at one end and balanced near a null point by a mainspring. A capacitive displacement transducer is often used to sense boom deflection.

A broadband seismometer achieves its wide response and exceptional linearity by means of a force-feedback electronic control loop that includes a forcing coil constantly working to keep the mass at dead center. The electric current required to counteract any deflection of the proof mass is proportional to the acceleration of the frame. In a sense, the forcing coil must at each instant apply exactly the same acceleration to the proof mass as the seismometer is experiencing for the proof mass to "keep up with" (stay stationary relative to) the seismometer. The seismometer output is derived from a signal within the control circuit that is proportional to the time integral of acceleration, providing a signal representative of velocity.

Broadband Axis Essential Elements
A typical arrangement for a broadband seismometer axis is to have a boom hinged at one end with two capacitor plates and a wire-wound coil mounted on it to serve as the displacement transducer and forcing coil, respectively. The mounting frame has one fixed capacitor plate arranged to sit between the two on the boom, and a permanent magnet and yoke fitting into and around the boom's forcing coil. This arrangement is mounted within the pressure vessel of the seismometer in one attitude for the vertical axis (the boom and capacitor plates horizontal) but in the upright attitude for the horizontal axes (the boom and capacitor plates vertical with the hinge at the bottom). Usually, the electronic feedback circuit together with the power conditioning, output signals, controllers, and related circuitry is on printed circuit boards arranged above the three sensing elements.

S
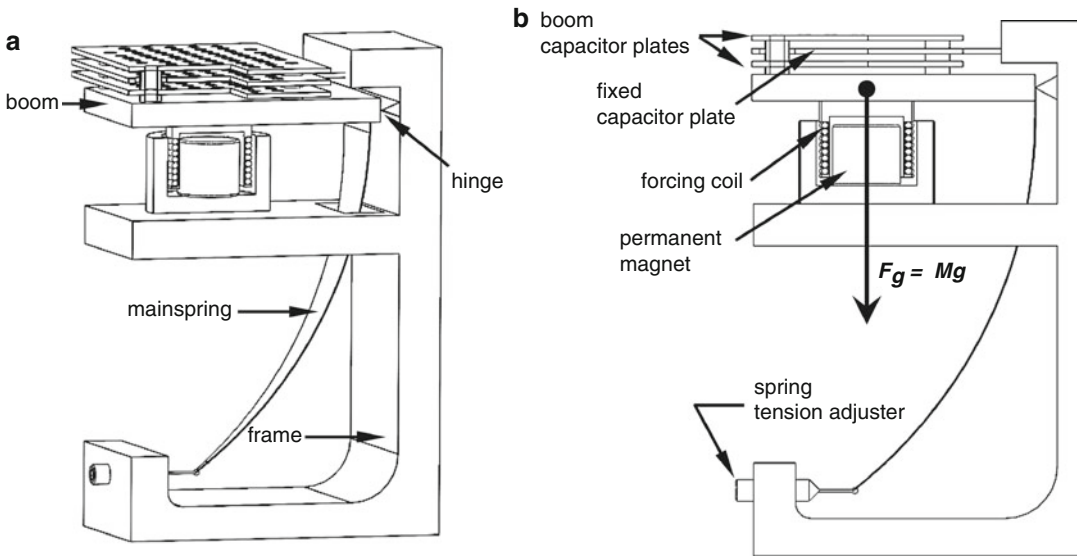
Symmetric Triaxial Seismometers, Fig. 1 Force-feedback seismometer vertical axis construction. (**a**) Perspective view. (**b**) Side view (© Nanometrics)

Figure 1 shows conceptual mechanical diagrams of a typical broadband axis construction configured to sense vertical motion; both perspective and side views are shown. This and subsequent diagrams are highly simplified for the purpose of illustrating solely the essential functional elements. The hinged boom with its attached forcing coil (shown in cutaway) and two outer capacitor plates have a combined proof mass M and together with the mainspring make a mechanical oscillator. The electric forcing coil attached to the movable boom acts around a permanent magnet fixed to the frame and is driven by force-feedback control loop electronics to maintain the proof mass at its measurement null point, which is where the fixed center capacitor plate is equally centered between the two outer capacitor plates. The boom is pulled down by gravity with force F_g that must be counterbalanced by the mainspring applying torque at the hinge point.

Figure 2 shows the essential construction of a horizontal axis. The principles of the horizontal and vertical designs are evidently the same, but with the horizontal axis oriented so as to be sensitive to sideway acceleration and insensitive to vertical motion. A significant difference between
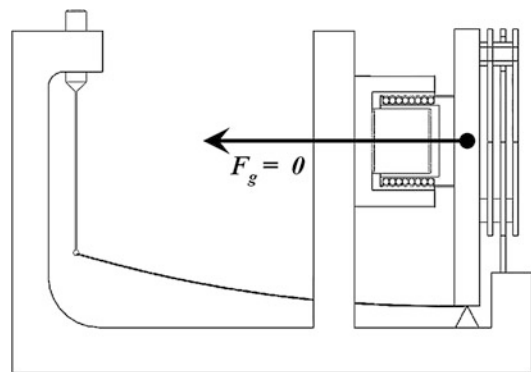
Symmetric Triaxial Seismometers, Fig. 2 Force-feedback seismometer horizontal axis construction (© Nanometrics)

the two designs is that the horizontal axis mainspring only supplies a restoring force when the boom deflects from its rest position and does not counterbalance any gravitational force ($F_g = 0$), whereas the vertical axis mainspring both supplies a restoring force and counterbalances the weight of the boom against gravity ($F_g = \mathbf{Mg}$).

See entry "▶ Principles of Broadband Seismometry" for more information on the operational principles.

The XYZ Sensor Configuration

The conventional vertical/horizontal or XYZ arrangement within a triaxial seismometer has one sensing element arranged so that it is sensitive to vertical (Z) motion and two sensing elements that are sensitive to horizontal motions in the X and Y directions. This corresponds to the output signals preferred for most seismology purposes; one generally wants to record the vertical, East and North components. However, the design of the vertical sensing element is unavoidably distinct from that of the two horizontal elements. The constant gravitational acceleration of the Earth that acts on (in fact, defines) the up direction and is absent in the horizontal directions means the vertical and horizontal elements must be of different designs.

The essential differences forced on the design by the presence or absence of the constant $g_o \cong$ **9.81 m/s^2** acceleration of gravity are the spring suspending the boom and the mass centering mechanism if there is one. Other design differences may also arise from practical design considerations such as the physical orientation of the vertical axis sensor being different with respect to the seismometer case, common electronics circuit boards, connectors, and the like. The vertical axis will have a hinged boom lying in the horizontal plane needing a strong spring to suspend it against gravity, and will tend to be tall and narrow as in Figure 1. A horizontal axis, having an inverted pendulum oriented in the vertical plane, will tend to be wider and shorter as in Figure 2. This creates challenges for the designer when attempting to fit one vertical and two horizontal axis elements in a single enclosure and can result in larger enclosures and/or further asymmetries between the horizontal and vertical designs to optimize physical fit.

The primary function of the mainspring that suspends the proof mass is to apply a restoring force to the mass in the direction of its center position, so that when the mass is deflected by some acceleration acting on it, the spring acts in the opposite direction to restore the position of the mass. The spring constant (the force the spring applies per unit of deflection) is generally as weak as can be practically achieved, so as to ensure the mass may freely move relative to the frame in which it is suspended. However, the mainspring of a vertical sensor must also balance the suspended proof mass against the acceleration of gravity, whereas the mainspring of a horizontal sensor does not. For example, a proof mass of $M = \mathbf{200}$ **g** in a vertical axis of the type shown in Fig. 1 experiences a downwards force due to gravity of $F_g = \mathbf{Mg} = \mathbf{1.96}$ N that applies a torque at the hinge point, and the spring must apply the equivalent counteracting torque to keep the mass balanced at the center position.

Adjusting the sensing axis so that the boom is positioned at the center position of the displacement transducer is known as "mass centering." This is a mechanical operation which is done for some seismometers by the operator turning an adjustment screw, and for others by microprocessor-controlled motors within the seismometer making the adjustments automatically, initiated by an external command or signal. On a horizontal axis, this is often done by tilting the axis (e.g., Guralp CMG-3T, Streckeisen STS-1) one way or the other within a range of a few degrees, which allows gravity to act sideways on the axis in proportion to the sine of the tilt angle (which for small angles is directly proportional to the tilt angle). If the mass is decentered either because the spring is applying a decentering force or because the seismometer is tilted off level, applying a small amount of gravity to one side or the other of the axis by tilting it in the opposite direction deflects the mass back to its center position. This method is depicted in Fig. 3, showing a horizontal axis on its internal mounting plate in an off-level situation. In Fig. 3a, gravity has pulled the boom off-center and its two capacitor plates are not centered about the middle capacitor plate that is fixed to the frame. In Fig. 3b, the axis has been tilted mechanically to re-center the boom mass. Note that the direction of sensitivity changes as the axis is tilted.

Tilting a vertical axis to center the mass is not practical because the effect of gravity on the axis is proportional to the cosine of the tilt angle, which for small angles has almost no effect: a vertical axis would have to be tilted by 10.7°
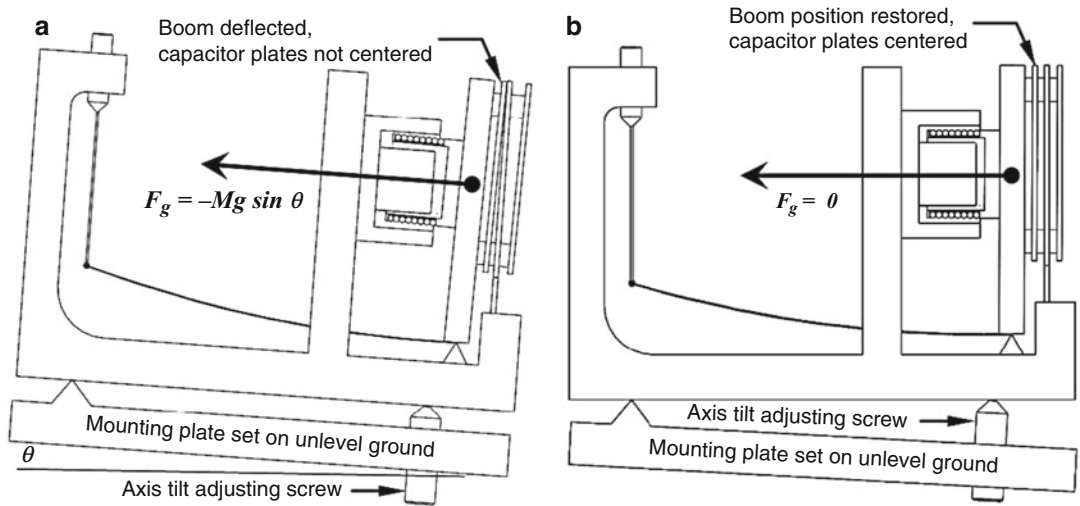
Symmetric Triaxial Seismometers, Fig. 3 Mass centering a horizontal axis by tilting. (**a**) Tilted axis on unlevel ground. (**b**) Axis internally tilted to re-center mass (© Nanometrics)

Symmetric Triaxial Seismometers, Fig. 4 Two methods for mass centering a vertical axis (© Nanometrics)
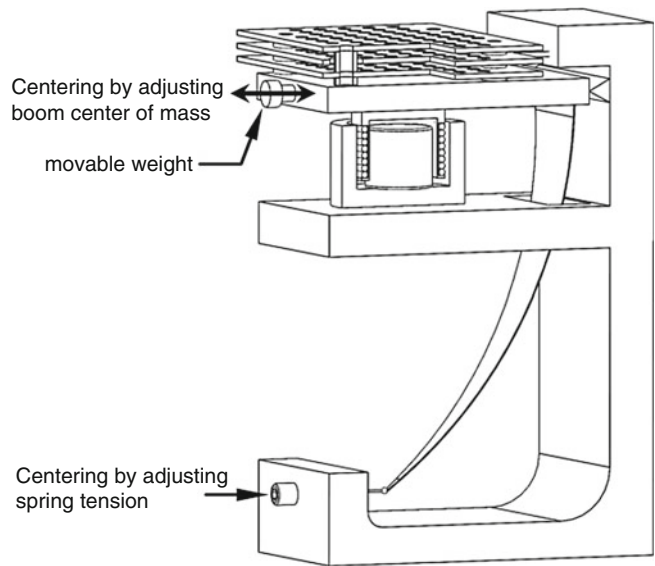
to counteract the same deflection as for a 1° tilt of the horizontal axis element. Instead, the mass centering for a vertical axis is done in one of two ways: by adjusting the tension of the main spring to change the force applied to the boom (e.g., Guralp CMG-3T) or by changing the position of the center of gravity of the hinged proof mass by moving an adjustable slug horizontally along the mass relative to the hinge point (Streckeisen STS-1). Figure 4 shows how both

methods work in principle. In practice the mechanisms are of course more elaborate and are often motorized. Note that, unlike the method of tilting the axis, the direction of sensitivity does not change as the spring tension is changed or the center of mass of the boom is adjusted.

A non-Galperin three-component broadband seismometer includes two horizontal axes and one vertical axis, usually mounted within a common pressure vessel. Because axis tilting
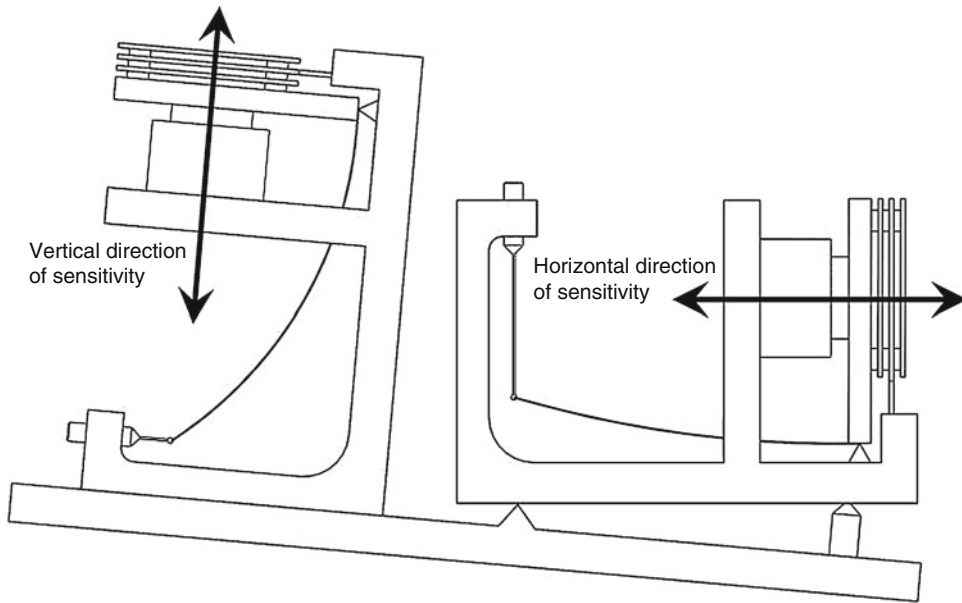
Symmetric Triaxial Seismometers, Fig. 5 Vertical and horizontal component non-orthogonality (© Nanometrics)

is used for mass centering the horizontal elements but not the vertical element, the mutual orthogonality of the XYZ components is compromised by the degree of tilt applied. This phenomenon is illustrated in Fig. 5, which shows one of the two horizontal axes and the vertical axis side by side on the same tilted base, the two axis elements having been centered by different methods resulting in non-orthogonal directions of sensitivity.

Practical design considerations also tend to encourage differences in implementation between the vertical and horizontal sensors. Asymmetries between the vertical and horizontal sensor elements arise for several practical reasons, despite the principles of operation being much the same. Besides the mainspring designs that must be different (one counters gravity, the other does not) and the significantly different mass centering mechanisms, the physical orientation and aspect height/width ratio of the two types of axis are asymmetrical. As the two horizontal axes and one vertical axis mount side by side on a common base plate (or are stacked vertically for a borehole configuration), such practical details, such as mounting features, wiring harness, and circuit board connectorization,

fit, and placement within the enclosure and other considerations generally lead to substantive design differences between the vertical and horizontal elements.

The Galperin Sensor Configuration

An alternative configuration to the conventional XYZ horizontal/vertical arrangement was proposed by Galperin, in which the three sensing elements of a triaxial seismometer are arranged in such a way that permits them to be identical in every respect. This is achieved by rotating the orthogonal sensing axes from the XYZ frame of reference to an orientation where each of the three (still mutually orthogonal) axes is inclined up from the horizontal plane at the same angle.

Figure 6 illustrates a symmetric triaxial arrangement. The directions of sensitivity of the new U, V, and W axes are shown overlaid on the XYW coordinate system. In this orientation, the edges (or directions of sensitivity, or axes) are often given the designations U, V, and W (Wielandt 2002), creating a new "UVW" coordinate system. The projections downwards of the three UVW axes onto the horizontal plane are lines radiating from the center equally spaced 120° apart. This arrangement is called symmetrical

Fig. 6 Isometric view of
UVW and XYZ coordinate
systems (© Nanometrics)
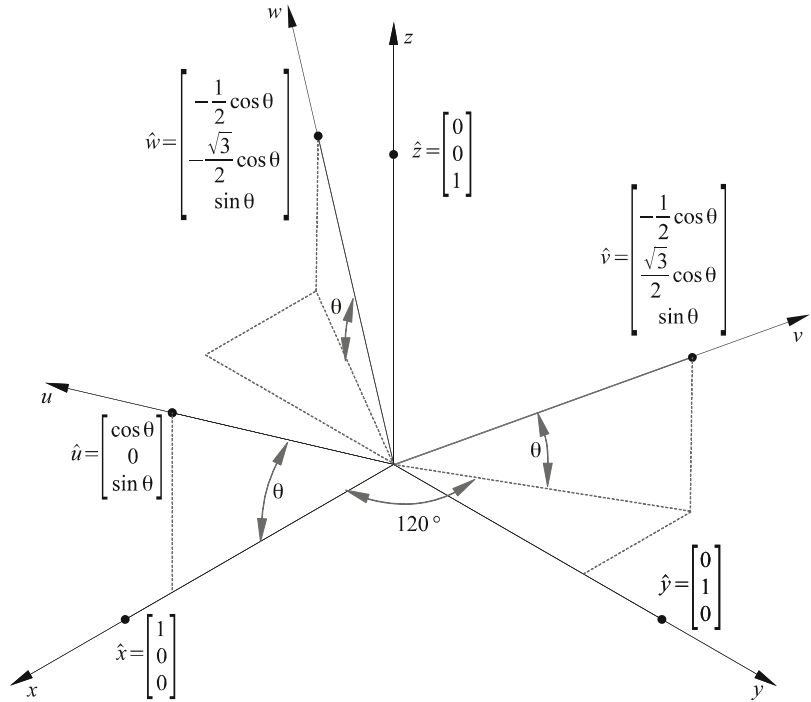
because each axis "sees" the same proportion
of gravitational acceleration and this allows
the axes to be constructed identically. It is possible
to choose any upwards inclination angle θ
for the three axes for this arrangement to be sym-
metrical (such as 45°), but setting
$\theta = \tan^{-1}\left(\frac{1}{\sqrt{2}}\right) \cong 35.26°$ also makes the UVW
axes mutually orthogonal.

The UVW system is then a simple rotation
from the XYZ. This can be visualized by thinking
of the XYZ directions of sensitivity as the three
edges of a cube radiating from one of its eight
corners, where the cube is resting on a horizontal
surface. The X and Y directions are along two
perpendicular edges of the cube's bottom face,
and the Z direction is the vertical edge of the cube
that joins the same corner where the X and
Y edges meet. Now visualize the cube being tilted
upwards with only the common corner resting on
the flat surface, and balance it so that the topmost
corner of the cube is directly above its bottom
corner. The edges that meet at the bottom corner
are still of course mutually orthogonal but now
form the same angle upwards with respect to the
flat surface base.

The trigonometry of this "balanced cube"
arrangement results in the angle formed between
any of the UVW axes and the horizontal plane
being $\theta = \tan^{-1}\left(\frac{1}{\sqrt{2}}\right) \cong 35.26°$. The comple-
mentary angle with respect to the vertical is
then $\theta_0 = 90° - \theta \cong 54.74°$.

Figure 6 employs the convention used in
Nanometrics Trillium seismometers whereby
the projection of U in the XY plane points in the
X direction, and the UVW system is right-handed
like the XYZ coordinate system (Nanometrics
Inc. 2013). The convention for Streckeisen seis-
mometers is that the projection of U in the XY
plane is antiparallel to X and that the UVW coor-
dinate system is left-handed (Streckeisen and
Messgeräte 1995).

Note that the term "symmetric triaxial" is not
in itself sufficiently specific to fully define the
Galperin configuration, as any triaxial orientation
with horizontal projections spaced 120° apart and
where each axis of sensitivity forms the same
angle with respect to vertical (say, e.g., 45°
instead of 54.74°) is a symmetric arrangement.
Each axis still experiences gravity equally and
they are arranged symmetrically with respect to
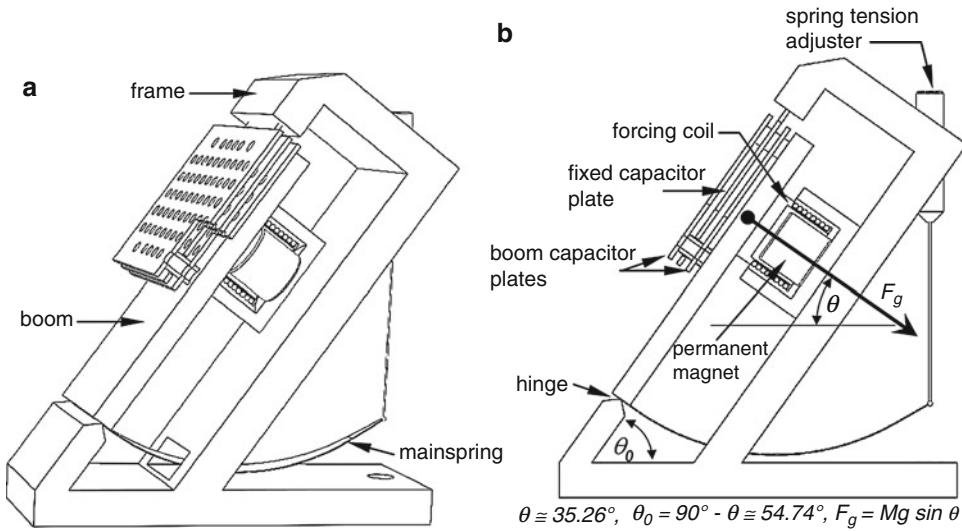
Symmetric Triaxial Seismometers, Fig. 7 Oblique axis construction suitable for a symmetric triaxial seismometer.
(**a**) Perspective view. (**b**) Side view (© Nanometrics)

each other, but they no longer meet at right angles as the edges of a cube. The Galperin topology requires the three axes to be mutually orthogonal as well as having the same angle with respect to the vertical direction, which then fully constrains their mutual orientation. Nevertheless, the "symmetric triaxial" designation is usually assumed to refer to the more specific Galperin topology.

Figure 7 shows conceptual mechanical diagrams of the structure of an axis suitable for a symmetric triaxial seismometer, showing the boom set at an oblique angle, but otherwise based on the same principle as a vertical axis, with the mainspring suspending the weight of the boom against gravity ($F_g = \text{Mg} \sin \theta$). For the specific case of a Galperin axis set at a $\theta \cong 35.26°$ angle, $\sin \theta = \frac{1}{\sqrt{3}}$. Because each axis experiences the same static acceleration due to gravity, $\frac{1}{\sqrt{3}} g_0$, the mainspring counterbalancing gravity is identical on all three axes. Because the direction of sensitivity of each axis is inclined by the same angle $\theta_0 = 90 - \theta$ with respect to the vertical, the physical design of the axis can be made identical for all three UVW component directions.

The directions of sensitivity with respect to the X or East direction are of course different for each axis – they are equally spaced 120° apart

as shown in Fig. 6 – but that is achieved by arranging the three identical sensing elements next to each other on the same horizontal base and pointing them in different but equally spaced directions as shown in Fig. 8. This is typically how vault seismometers are configured internally. Seismometers for borehole or posthole installations usually have the three axis elements stacked in a vertical column, one above the other pointing in different directions, but the principle is the same. The diagrams in Fig. 8 illustrate that there is more than one suitable way to dispose the three Galperin axis elements within a seismometer. Note that the relative directions of X, Y, Z, U, V, and W are the same in both arrangements; the three axis elements have just been translated to different positions on the base. The leftmost arrangement of axis elements is employed in the Nanometrics Trillium 120P vault instrument, for example, while the Streckeisen STS-2 vault instrument employs an arrangement similar to the one shown on the right.

Mass centering of a Galperin-type axis may be achieved by any of the methods used with horizontal or vertical axes: tilting the axis, adjusting the mainspring, or manipulating the center of gravity of the boom. Examples of each method
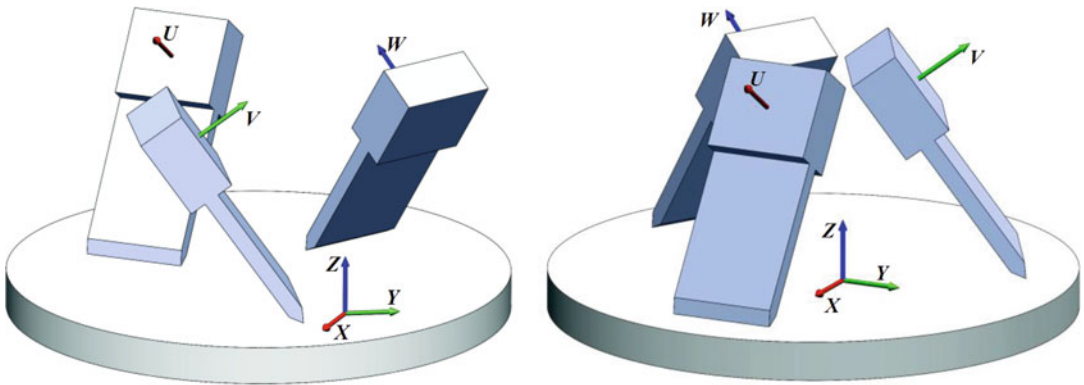
Symmetric Triaxial Seismometers, Fig. 8 Two Galperin-type axis arrangements for a vault instrument (© Nanometrics)

used on Galperin-type instruments are the Streckeisen STS-2 vault instrument that moves a slug on the boom, the various models of Nanometrics Trillium observatory-grade seismometers that adjust the tension on the mainspring, and the Geotech KS54000 borehole seismometer that tilts each axis in its vertical stack.

Mutually orthogonality UVW (and therefore XYZ) directions of sensitivity are maintained because the masses of all three axes are centered using the same technique, taking advantage of the symmetry of the design. Unlike in Fig. 5 that shows how a vertical and horizontal axis can become mutually non-orthogonal, the Galperin designs do not adjust one axis using a different technique than the others. For example, the direction of sensitivity of the Galperin axis depicted in Fig. 7 is fixed by its geometry to be at 90° to the face of the boom. Adjusting the tension of the spring to re-center a deflected mass does not alter the direction of sensitivity, and because the three axes are permanently fixed to a common mounting plate as in Fig. 8, the three UVW directions remain mutually orthogonal.

The commonly desired XYZ horizontal/vertical signals are readily derived from the UVW signals using a vector algebraic transformation:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 & -1 & -1 \\ 0 & \sqrt{3} & -\sqrt{3} \\ \sqrt{2} & \sqrt{2} & \sqrt{2} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

The inverse transformation derives UVW from XYZ:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 & 0 & \sqrt{2} \\ -1 & \sqrt{3} & \sqrt{2} \\ -1 & -\sqrt{3} & \sqrt{2} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The transformation to the conventional XYZ frame of reference may be done at any point, such as by seismic analysis processing software just before the signals are plotted or used as input to seismic analysis algorithms that assume the XYZ reference frame. However, as a convenience to the operators of seismic signal data centers and the users of the seismic data, most Galperin-type seismometers effect the signal transforms within the instrument by summing the UVW signals in the required proportions using a precision analog mixing circuit. Some seismometers (e.g., Nanometrics Trillium models) allow the operator to optionally configure the instrument to output raw UVW instead of the mixed XYZ if desired, though this is primarily used for instrument troubleshooting rather than for routine recording of seismic data.

Benefits and Drawbacks of the Galperin Configuration

The choice between a Galperin and a conventional configuration is usually regarded as a secondary consideration in specifying an

instrument, as the internal topology of the seismometer does not of itself dictate the performance or reliability of an instrument. However, while either configuration is capable of achieving reliable high performance, the symmetric triaxial configuration does offer some significant benefits to users. The two topologies have distinct characteristics that are useful to consider when selecting, installing, operating, or troubleshooting a broadband seismometer.

Commonly Cited Trade-Offs

Response Matching

Because the designs of a horizontal axis and a vertical axis differ significantly, it requires special effort in design and diligence in manufacture to match the transfer functions of the two different constructions, if they are to provide the same response to ground motion in terms of passband sensitivity, lower corner frequency/phase response, and the upper corner response. Indeed, there is no constraint other than designer or user preference that requires the response of the vertical to match the response of the horizontal component; the two could be markedly different with the manufacturer providing the specifications for each component by giving, for each axis, a frequency/phase response plot or a set of poles and zeros and passband sensitivity. In contrast, because the Galperin configuration requires its cardinal XYZ outputs to be derived as weighted sums of all the UVW components, the design and the manufacturing must ensure that all three components are very closely matched; otherwise the transfer functions of the XYZ outputs become complicated. The symmetrical nature of the Galperin topology helps produce matched responses because the three axes are constructed identically. Manufacturers of Galperin seismometers supplement this advantage with precision machining, trimming of electronics values, and other calibration measures to precisely match axis responses. The precision matching of U, V, and W components in a Galperin design provides assurance to the user that the X, Y, and Z transfer functions are likely to be consistent and the same as any of the U, V, or W transfer functions.

Response matching is easily tested on a Galperin instrument by providing a vertical signal on a high-quality vertical shake table or by injecting the same calibration signal into all three UVW components simultaneously (which is equivalent to providing a calibration input in the vertical direction). The output should be pure vertical; the extent to which signal is present on the X and Y outputs is a measure of response difference between the UVW elements. By sweeping the frequency of the input and recording the vertical and horizontal outputs, the quality of the response matching can be measured across the frequency band.

In contrast, subjecting a conventional XYZ design seismometer to a pure vertical stimulus tests only the vertical axis and does not provide any insight into how well the responses of the three axis components are matched.

Manufacturing Trade-Offs

Some aspects of the design and manufacture of Galperin instruments are simplified or facilitated, while others are made more complex. Obvious advantages to the manufacturer derive from having only one electromechanical axis design within a seismometer. That means there are fewer unique part types to manage, only one style of axis to build and test, and the identical axes can be readily swapped for troubleshooting or repair purposes. On the other hand, a symmetric triaxial seismometer requires a UVW-to-XYZ analog electronic mixing circuit that is not needed in an XYZ design, adding somewhat to the cost of manufacture.

Because the vertical output signal of a symmetric triaxial design is an equally weighted sum of the UVW components, a clean vertical signal provides complete assurance that all three components are well behaved. As the horizontal signals at most surface sites are usually noisier than the vertical due to tilt effects, especially at long periods, it is often feasible to test and qualify the vertical channel but more difficult to do so for the horizontals. This permits the manufacturers of Galperin designs to reliably test the performance of all three sensor elements by qualifying the vertical signal, providing

assurance to the users that all three outputs meet the published specifications. In contrast, the X and Y components of a conventional design cannot be assessed by examining the vertical channel signal.

Reliability Trade-Offs

An often-cited drawback relating to the Galperin configuration is that all three UVW components must be functional and well behaved for any of the XYZ outputs to work (Graizer 2009). It is held that this contrasts with a conventional seismometer where any one of the three components could fail without compromising the other two. This can be seen as a corollary of the fact that a good vertical output from a Galperin configuration provides assurance that all three sensor elements are good.

The spontaneous catastrophic failure of just an individual axis element within an operating seismometer is relatively uncommon. It is more likely for catastrophic failure to occur or be evident at installation time (in which case most operators would choose not to deploy even if there was a good vertical still operating) or for a failure to occur in a way that affects the entire system, such as an electronics fault.

Many seismometers with Galperin configurations provide the ability to remotely choose between UVW and XYZ outputs. In this case, if a failed axis is detected and a site visit is not feasible, high-quality biaxial data can still be recorded after the flip of a switch.

It is possible for an element to develop noise spontaneously, and in this case the conventional design has the advantage: a Galperin design would mix the noisy channel into its vertical output (where it is most easily detected), whereas the conventional XYZ design would only manifest the noise on the output associated with one degraded axis.

A manufacturer can minimize the likelihood of excessive channel noise by ensuring the seismometer self-noise is tested before leaving the factory and by designing the instrument for long-term operation to ensure performance is maintained through the instrument's operating life.

Other Benefits of a Symmetric Triaxial Design

Using the UVW Orientations to Discriminate Noise Sources

The symmetric triaxial configuration makes it possible to distinguish phenomena that occur within one axis element from those that occur independently of one axis element. The operator can use this to advantage, helping diagnose and resolve problems associated with the environment, installation, ancillary equipment, or the seismometer itself. This is discussed in greater detail in the section "Using the Galperin Topology to Help Discriminate Noise Sources" below.

Assuring the Mutual Orthogonality of Component Signals

The three identical and symmetrically arranged sensor elements of the Galperin configuration are fixed to be mutually orthogonal during manufacture. A mass centering operation, whether moving a slug to shift the boom center of gravity or adjusting mainspring tension to return the boom to center, does not alter the mutual orthogonality of the directions of sensitivity. In the first case the sensitivity direction of each axis is shifted by the same amount and so they remain mutually orthogonal. In the latter case, the tension of the main spring does not alter direction of sensitivity and so orthogonality is also preserved. However, a conventional XYZ design tilts the horizontal axes but adjusts the mainspring tension for the vertical, causing the horizontal directions of sensitivity to change relative to the Z axis. This causes the horizontal components to shift relative to the vertical by an angle equivalent to the tilt of the seismometer housing itself. Effectively, the vertical of an XYZ design remains aligned to the housing while the horizontals remain perpendicular to gravity.

Using the Galperin Topology to Help Discriminate Noise Sources

Using a seismometer that has a Galperin design can facilitate identifying noise sources by showing whether they are specific to one axis element

or not. This section discusses noise sources and some methods for troubleshooting noise artifacts that use the characteristics of the symmetric tri-axial configuration to advantage.

Noise Sources

A noise artifact is usually apparent only once it is in the seismic record. This record is furthermore the sum of all effects applied by the entire system including the environment, site, seismometer mount, the seismometer itself, cables and connectors, the digitizer, and signal transmission and post-processing. [In this context the "seismometer mount" is a general term that refers to the structure the seismometer is mechanically coupled to. That may be a concrete pier, a gimbal mount such as for an ocean bottom system, surrounding granular substrate as in the case of a buried instrument, a metal bracket affixed to a building structure, and others.] All these components of the end-to-end system must be considered when investigating noise or other signal impairments.

Noise may be generated by processes within the seismometer or by external processes acting on the seismometer. Some noise is expected or unavoidable, such as the manufacturer's characterized self-noise, or may be indicative of instrument defects such as persistent spurious transient noise event ("pops") or installation deficiencies such as thermally driven noise or the seismometer shifting or tilting. It is first useful to review some typical sources of internal and external noise.

The sources of noise (defined as any type of additive unwanted signal) in a seismic record include inherent instrument noise (self-noise), environmental effects, excessive instrument noise, installation-driven noise, and ancillary equipment noise. A seismometer's self-noise is usually specified by the manufacturer, typically as a power spectral density (PSD) plot, and this establishes a performance baseline for an installation. Good installation practice will include mitigating sources of environmental noise as much as is practical. Excessive instrument noise is defined as noise originating within the seismometer in addition to its published self-noise characteristic and may represent variability in the manufacturing, defects, or failures. Installation-driven noise is defined as a non-seismic signal that originates from poor or defective installation practices. Lastly, noise may also be added by downstream ancillary equipment such as the digitizer.

Environmental and Installation-Driven Sources of Noise

Seismometers respond in varying degrees to many stimuli besides seismic motion, and while seismometer designers try to maximize the immunity or insensitivity of the instrument to everything except translation ground motion, this is difficult to achieve. The installation must also be designed to maximize the effectiveness of the seismometer. An inadequate installation of even the highest performance seismometers can yield very poor results, showing high noise levels and susceptibility to non-seismic environmental inputs. There is therefore significant benefit in tools and techniques that help diagnose installation deficiencies and distinguish them from instrument problems.

Environmental influences on a seismometer that may cause unwanted output signal (noise) include temperature changes, seismic wave-induced tilt, Earth's magnetic field or fields generated from local equipment, wind-induced ground motion, atmospheric pressure changes, electrical interference induced on the output signal, and other effects. Many environmental influences can be mitigated, such as by providing a temperature stable vault and locating the installation far from vertical structures such as trees that conduct wind noise into the ground. Some are more difficult or even impossible to eliminate, such as apparent vertical accelerations induced by changes in atmospheric pressure, because the apparent acceleration due to gravitation cannot be distinguished from other kinds of acceleration and is relatively insensitive to the depth of the site (Zürn and Wielandt 2007).

Installation-driven noise sources are those that induce motion on the seismometer due to defects in the installation itself. A common example is poor mechanical coupling of the instrument to the ground, which may allow the seismometer to

move relative to the structure being monitored. This can be due to poor mounting, inadequate substrate, poor coupling to the ground, improperly locked mounting feet, or cables applying strain on the seismometer. Such problems most commonly produce tilt noise, but vertical noise is also possible, such as for a seismometer "bouncing" up and down on adjustable feet that have not been tightly locked. Another example is convection-driven air currents within a vault inducing low-frequency horizontal tilt noise. Such noise is produced by slow-moving air currents acting on the seismometer to expand or shrink the sides of its case unevenly and thus tilting it.

A seismometer exhibiting excessive response to environmental stimuli may be indicative of an instrument defect. For example, a leak in the pressure vessel that allows atmospheric air pressure changes to pump air in or out of the seismometer will manifest as a non-seismic vertical signal as the buoyancy of the proof masses responds to internal air density changes. This is most commonly seen in instruments that have access ports that an operator may open, such as to insert a screwdriver for manual mass centering, and that may not be properly closed or where the seals may have degraded.

Sources of Internally Generated Seismometer Noise

All seismometers generate noise internally, that is, they would produce a signal even if the ground was not moving at all. This inherent noise is characteristic of the design and is usually specified in some detail by the manufacturer, although in the case of passive seismometers, the manufacturer will instead usually provide a few parameters from which the self-noise can be estimated. Noise is inherent in electronics circuits due to a multitude of phenomena in conductors and semiconductors, such as thermal noise and flicker noise. The suspended mass and spring has inherent noise due to the Brownian motion of air molecules buffeting the mass, air damping the boom motion, as well as other energy-loss mechanisms. The summation of these effects produces a stationary noise spectrum that is commonly

referred to as the seismometer's "self-noise," often represented as a PSD graph and commonly shown together with the New Low Noise Model (NLNM) and New High Noise Model (NHNM) (Peterson 1993). Reputable manufacturers provide this information for each model and produce instruments with performance consistent with their published specifications.

Excessive instrument noise is internally generated noise that exceeds the self-noise specification of the seismometer. There are three broad classes of excessive internal noise:

1. Stationary broadband noise that significantly exceeds the instrument's published self-noise specification
2. Nonstationary transient noise events, such as pops or spikes
3. Stationary narrowband noise, such as tones or oscillations

A common technique for measuring noise is frequency domain analysis, such as plotting the acceleration power spectral density (PSD) of the seismic data over some time period. While tones are readily distinguishable from broadband noise, nonstationary events such as pops have a broadband spectral characteristic (typically proportional to $1/f^2$) that can mislead the troubleshooter. However, small pops may not be readily identified by examining the time-domain time-series data as they may have amplitudes too small to be distinguished from the background seismic activity.

The Galperin Topology and Noise Source Identification

Being able to demonstrate that a noise phenomenon is peculiar to one particular axis element and not the other two, or conversely that a noise is registered by more than one axis, can be a powerful technique to help isolate and identify noise sources and causes. The Galperin topology facilitates this because its sensor reference coordinate system (UVW) is rotated from the vertical/horizontal (XYZ) coordinate system, providing a powerful way to distinguish instrument noise

Symmetric Triaxial Seismometers, Table 1 Channel characteristics of some possible noise sources (© Nanometrics)

Channels showing non-seismic noise	Potential noise sources						
	Axis mechanics or electronics	Tilt motion	XY oriented seismometer mount (e.g.,gimbal)	UVW-XYZ conversion electronics, cable, digitizer	Varying temperature or pressure leaks	Varying magnetic fields	Power supply, electrical/magnetic interference
U, V or W only	Probable						
X or Y only		Possible	Possible	Possible			
Z only				Possible	Possible		
X and Y in some proportion but not Z		Probable					
Equally on all X/Y/Z channels				Possible			Possible
Unequally on X/Y/Z channels						Possible	Possible

sources from environmental or installation phenomena. Many external noise phenomena have a characteristic direction – often horizontal or vertical – and in any case are unlikely to align with one of the U, V, or W directions. Likewise, noise that is associated with a U, V, or W direction provides a clear indication that the noise source is internal and associated with that axis.

Table 1 provides general guidelines for narrowing down the sources of an observed noise signal and provides a starting point for troubleshooting. This presumes the operator is able to perform the essential basics of signal analysis: conversion from XYZ to UVW domains, plotting time series and frequency domain plots such as a PSD, removing instrument frequency response, and the like. Table 1 also presumes the seismometer has a Galperin topology.

Once a noise artifact is recognized and classified (e.g., stationary broadband, nonstationary transient, or stationary narrowband) and characterized as much as possible, the next step is to determine whether it consistently manifests on a specific X, Y, Z, U, V or W channel or a particularly suggestive combination of these channels.

This is equivalent to asking which direction the signal represents. Signals specific to X, Y, or Z suggest East/West, North/South, and up/down motions or phenomena that mimic these. Signals specific to U, V, or W suggest motions in those axis directions or, more likely, phenomena within the axis mimicking this direction. A signal that manifests on X and Y but not (or very little) on Z represents motion in a horizontal direction that is not East/West or North/South, or a phenomena such as tilt that mimics that. A signal that is equal on X, Y, and Z is more likely to be a common-mode electronics noise problem than a real motion in the direction represented.

Because it is unusual to have real seismic motion act exactly along any of these X, Y, Z, U, V, or W directions, signal artifacts with these characteristic directions are more likely to be of non-seismic origin. An exception could be a site expected to have motions exactly along a specific direction, such as a seismometer located near the top of a swaying structure, oriented so that its X channel (say) is aligned with the direction of sway.

Having narrowed the potential sources, it is usually possible to make further deductions from

S

the character of the signal artifact observed. The next sections discuss typical phenomena and their classifications according to possible sources listed in Table 1.

Axis Mechanics or Electronics: UVW Channel Noise

Noise artifacts apparent on one U, V, or W channel and not on the others point squarely to a noise source associated with that axis. Of course this can only be recognized by transforming the seismic data into the UVW domain to observe that a noise previously apparent on all XYZ channels is suddenly associated with just a single one of the Galperin axis elements. Noise that is vertical (Z only) or horizontal (some combination of X and Y) cannot have a source within a Galperin-type axis.

A noise event commonly attributed to highly sensitive broadband seismometers is that of pops, otherwise known as spikes or steps. These are sudden sharp excursions in the seismic waveform that appear as if there had been a sudden step in acceleration. This can be the result of spontaneous movements at a microscopic level within the axis mechanism or sudden shifting of the seismometer itself due to installation or mounting problems. If a pop is evident only on one U, V, or W axis, it is due to spontaneous movement within the axis, such as some mechanical stress or strain suddenly relaxing. Inadequate design, assembly flaws, component defects, or rough handling can give rise to excessive rate of pops in an axis. One indication of excessive axis pops is when the PSD is plotted for the U, V, and W signals and there is a well-defined $1/f^2$ spectrum in the lowest frequency band of one of the axes. It is well known that seismometers are more prone to pop noise when first manufactured, installed in a new location, or acclimating to a new environment as stresses are relieved.

Excessive broadband axis noise (specific to a U, V, or W axis) can be due to electronics noise, for example, excessive current noise due to components not meeting specification.

Excessive axis-specific narrow band noise is less common, but can occur if there is excessive loop gain or insufficient phase margin in the control loop of one axis. This would generally suggest a manufacturing or design defect and could also be the consequence of component failure or rough handling.

A signal impairment specific to one U, V, or W axis such as zero signal or railed signal suggests an axis failure, such as failure to center, or a failure of the electronic circuitry associated with that axis.

Likewise, a non-seismic noise or other impairment (such as zero signal or railed signal) that is not limited to a U, V, or W channel but manifests as a horizontal or vertical signal cannot be an axis problem but must have some other origin, as indicated in Table 1. The following sections discuss these other potential noise origins.

Tilt Effects: XY Channel Noise

The most common sources of noise are installation deficiencies and environmental stimuli, often where inadequate installation fails to counter environmental drivers such as temperature changes or air currents. Some of these effects are quite subtle but can create significant noise artifacts. It is quite common to see horizontal noise levels for a seismic station be significantly greater than the vertical noise floor (Hanka 2002). While horizontal noise could be ascribed to noisier horizontal components in a conventional XYZ design, with a Galperin topology with identical UVW elements, noisier horizontal data can definitively be attributed to a process external to the seismometer.

Because tilt mixes gravitational acceleration into the horizontal channels in direct proportion to the tilt, and because gravity is significant compared with the small accelerations due to background seismicity, even sudden tilts of less than a millionth of a degree will be registered by broadband seismometers and appear as noise artifacts. The effect is far less significant for vertical because gravity mixes to the vertical in proportion to the cosine of the tilt angle, which is a near-zero effect for small angles.

Seismic waves traveling on the surface of the Earth that produce vertical translation motion (like a cork bobbing on the waves of an ocean)

also cause tilt: one can visualize a raft on the same ocean that tilts one way as it rises up one side of the wave and then tilts the other way as it falls. The rise/fall is vertical motion the seismometer will faithfully record, but the tilt induces a small varying sideway component of gravity producing a noise signal on the horizontal channels in addition to the real horizontal motion. Because tilt motion is inherent in seismic motion at the Earth's surface, horizontal tilt noise cannot always be mitigated. However, excessive tilt noise and in particular sudden discrete tilt events are usually indicative of installation problems that should be addressed.

Noise artifacts that appear on X and/or Y but not Z are horizontal and often represent the seismometer tilting in some manner. There are many potential causes and sources of tilt, and so it is useful to further characterize the noise. Sudden isolated transitions are suggestive of rapid discrete tilt events, oscillations can indicate rocking behavior, and broadband horizontal noise may indicate continual random tilting. Steps that occur with a consistent polarity suggest a series of tilts in the same direction, as may happen if the seismometer is slumping in discrete steps to one side. Steps that alternate polarity may indicate a back-and-forth tipping action. The direction of the tilt is indicated by the proportion of X-to-Y noise, and some combination of X and Y (with very little Z) noise is a further confirmation of probable tilt action.

Atmospheric pressure changes acting on a sealed vessel such as a seismometer will deform the vessel to some extent, and because tilts much less than 10^{-6} of a degree can generate noise, even miniscule deformations can matter. Seismometer designers generally take great care to ensure those deformations do not act to tilt the internal sensing elements; otherwise horizontal noise driven by weather changes would be inherent in the instrument (Widmer-Schnidrig and Kurrle 2006). In a well-designed instrument, pressure-driven tilt should not be evident. It is also difficult for a manufacturing or component defect to give rise to pressure-tilt sensitivity; the design either has sufficient pressure immunity or it does not. The inherent symmetry of a Galperin-type topology also provides more opportunity to the designer to optimize pressure vessels with highly symmetric internal mounting arrangements that attenuate pressure vessel deformation effects.

XY-Oriented Seismometer Mount: X or Y Noise

For Galperin-type seismometers, tilt that is aligned strictly with either the X (East) or Y (North) direction is of special interest: while it is possible for tilt to coincide with the X or Y, there may be a specific reason for tilt to occur in a well-defined and aligned orientation. For example, an unlocked foot that is positioned in line with East or North could be suspected, or the seismometer may be mounted on some structure that may tend to tilt in a particular direction. With a conventional XYZ seismometer, X- or Y-aligned noise could either be related to the internal X or Y axes or be due to external tilting. A symmetric triaxial construction experiencing such noise clearly points to an external cause, and when the tilt direction is highly aligned with X or Y, or in some other specific direction in line with a particular mounting feature such as a foot, it provides a useful indication of the likely trouble spot.

A special instance of this is when a Galperin-type seismometer is mounted within a moveable platform, such as a leveling gimbal for ocean bottom deployment. If tilt events are clear in the data and strictly oriented in line with the potential movement of the platform, it is suggestive of gimbal platform movement.

Output Stage Electronics, Cable, and Digitizer: X, Y, Z, or XYZ Noise

Single X, Y, or Z Channel Noise
The only electronics internal to a typical symmetric triaxial seismometer which are specific to a single X, Y, or Z output channel are the UVW-to-XYZ coordinate-conversion circuits that combine the UVW signals to produce XYZ outputs, the signal output drivers, internal cabling, and external connector. It is rare to see noise problems originate from these areas, but it

is possible for the failure of an electronic component to cause a failure of one output channel (or one side of a differential output channel that then would manifest itself as a half-amplitude signal on one channel).

Because the signal is in the XYZ domain from the seismometer onwards, the cable and digitizer must be suspected when troubleshooting. Because the output stage electronics are relatively simple and usually reliable, it is more common for the cause of single-channel or common-mode XYZ failures or noise to be associated with the downstream digitizer or the cable connecting the seismometer to the digitizer.

It is useful to determine if noise apparent on a single output channel is predominantly or exclusively evident on that channel. If there is no attenuated version of that noise on the other channels, it is likely to represent a failure or deficiency of hardware associated with that channel, such as a connector pin or digitizer channel problem. Noise appearing predominantly on one output channel but also on the others in an attenuated form points to an effect predominantly aligned with that channel but not exclusive to it, such as tilt along the X direction.

Predominant But Not Exclusive Z Channel Noise

A special case of single-channel noise is Z channel noise. That is because the Z (vertical) channel is produced by an equally weighted sum of the UVW axis signals, and so any phenomena that affects all axis elements equally will manifest predominately on the Z channel. Therefore, Z channel noise could originate from within the seismometer due to a common-mode effect acting on all axis elements or from an external fault on the Z channel in the cable or digitizer. One can usually determine which by the nature of the noise or impairment and also by determining if the noise or impairment is exclusively or just predominately on the Z channel. A common-mode effect acting on all axis elements will predominately manifest on the Z channel, but usually there are slight differences in sensitivity to these effects from axis to axis, so some attenuated version of that noise on the X and

Y channels may be present. If the impairment is external, such as a noisy digitizer Z channel, it is much less likely to leak into the X and Y channels.

Common-Mode XYZ Channel Noise

A noise source or impairment that affects all output channels more or less equally (known as a common-mode characteristic) could be associated with the seismometer electronics that serve common functions such as power conversion and control circuits, but is more likely an external common-mode effect, such as noise induced on all input channels to the digitizer by electrical interference. Common-mode XYZ cannot be problems with any or all of the internal Galperin axis elements, because these would manifest in the U, V, W, or Z directions, respectively, never in the direction corresponding to equal XYZ signals.

Temperature Changes and Pressure Leaks: Z Channel Noise

The most common sources of predominately Z channel broadband noise are environmental effects acting on the seismometer, such as varying temperature or pressure leaks. Seismometers respond to temperature changes in part because thermoelastic effects act on the spring that suspends the proof mass against gravity, causing apparent changes in acceleration. These are slow effects and therefore affect low-frequency noise performance. The three axis elements will usually have a similar but not exactly equal sensitivity to temperature changes, manifesting as predominately Z channel noise with much smaller effects on X and Y.

A leak in the pressure vessel of the seismometer would allow air to pump in or out driven by atmospheric pressure changes, and this would cause the proof masses to rise or fall buoyed by changing air density inside the vessel. This creates a very pronounced noise that is strongly vertical. Pressure leaks mostly arise in seismometers that provide service access ports such as for manual mass centering that may not have been properly closed or where the seals have deteriorated.

Magnetic Field Effects: Unequal XYZ Noise

Broadband seismometers are designed to be as insensitive as possible to changes in external magnetic fields, but some sensitivity is almost inevitable (Forbriger et al. 2010). Springs designed to be relatively insensitive to temperature changes are often fabricated from magnetic materials and so have small but measurable forces applied by the Earth's magnetic field. These can manifest as low-amplitude noise at long periods. Each axis will be sensitive to the direction as well as the strength of the geomagnetic field. Because the direction of the Earth's magnetic field differs depending on location, the effect of changing magnetic fields is not equal on all axis elements, nor are the relative proportions among axis elements consistent or predictable. These are typically small effects and usually only of concern for very high-quality low-noise sites with very low-noise seismometers. Mitigation strategies can include augmenting magnetic shielding.

Electronic Interference: Equal or Unequal XYZ Noise

Noise that appears on all channels may arise from interference of nearby electrical equipment. A common-mode noise appearing equally (or nearly equally) on all channels could indicate a conducted or radiated noise source from a power source such as a battery charger cycling, a nearby motor, or similar problem. The interference is unlikely to be affecting the seismometer itself, as noise induced into the three Galperin axis elements manifests as predominately vertical (Z) channel signals. Induced noise affecting all channels is likely to suggest a cable or digitizer vulnerability or issue of some sort. Mitigation typically includes determining the source if possible, improving shielding or grounding, and changing cables, digitizers, or power sources if faults are suspected.

Summary

A seismometer that is designed using a Galperin arrangement has three identical sensing elements with their directions of sensitivity arranged to be mutually orthogonal, and inclined up from the horizontal plane at an angle of $\theta_0 \cong 35.26°$. This is a subclass of the more general "symmetric triaxial" configuration in which three axis elements are identical and symmetrically arranged with respect to each other. The signals from the three sensors, designated U, V, and W, are usually remixed by analog circuitry within the seismometer to provide a vertical and two horizontal outputs equivalent to a conventional XYZ seismometer. The benefits of the symmetric triaxial approach pertain to site and instrument troubleshooting, channel response matching, ease of manufacturing and testing, reliability, and channel orthogonality. The most frequently cited drawback is that all three axis elements must be functional for any of the three XYZ seismic signal outputs to be valid. The high reliability of the best seismometer designs largely mitigates this concern, and it is furthermore avoided by choosing a seismometer which allows the user to remotely select UVW signals to be output in place of the mixed XYZ in the event of a single axis failure. The second cited drawback inherent in the symmetric triaxial configuration is the requirement for an internal analog electronic mixing circuit to convert UVW to XYZ, although in a well-designed seismometer, this is invisible to the user.

A primary benefit of the Galperin approach is that, because the UVW coordinate systems are different from the XYZ horizontal/vertical reference system, noise sources and faults within an axis are distinguishable from problems that manifest in primarily vertical or horizontal directions. A corollary to this is that a properly functioning vertical channel on a Galperin instrument is a reliable indication that all three UVW elements are good, since the vertical is an equal sum of the three Galperin axes. Response matching is necessary for Galperin instruments to enable proper remixing to XYZ outputs, providing assurance of consistent response independent of the direction of motion. Galperin designs also generally assure rigid orthogonality of the XYZ outputs, whereas the tilting of the horizontal elements in the mass centering operation of conventional designs can produce deviations in X and Y orientation relative to the fixed vertical axis. Manufacturing benefits include having one

S

rather than two axis designs to build and improved ability to test and troubleshoot.

A review of noise sources and characteristics and how their directions align relative to the separate Galperin and horizontal/vertical coordinate systems demonstrates the powerful utility of using a symmetric triaxial configuration to help discriminate and diagnose potential noise problems with the site, installation, or seismometer.

In summary, the Galperin geometry has proven to offer significant benefits to seismometer users and manufacturers and through use and experience has proven to have significant additional benefits in terms of remotely troubleshooting station noise.

References

Forbriger T, Widmer-Schnidrig R, Wielandt E, Hayman M, Ackerley N (2010) Magnetic field background variations can limit the resolution of seismic broad-band sensors. Geophys J Int 183(1):303–312

Galperin EI (1955) Azimuthal method of seismic observations (in Russian). Gostoptechizdat 80

Graizer V (2009) The response to complex ground motions of seismometers with galperin sensor configuration. Bull Seismol Soc Am 99(2B):1366–1377

Hanka W (2002) Parameters which influence the very long-period performance of a seismological station: examples from the GEOFON Network, Section 7.4.4. In: Bormann P (ed) New manual of seismological observatory practice, vol 1. GeoForschungsZentrum, Potsdam, pp 64–74

Melton BS, Kirkpatrick BM (1970) The symmetric triaxial seismometer – its design for application to long-period seismometry. Bull Seismol Soc Am 60(3):717–739

Nanometrics Inc. (2013) Trillium posthole user guide. (17217R5). Nanometrics, Inc., Kanata

Peterson J (1993) Observations and modeling of seismic background noise. Open-file report 93–322, U. S. Geological Survey, Albuquerque

Streckeisen G, Messgeräte AG (1995) Portable very-broad-band tri-axial seismometer STS-2 manual, 50

Widmer-Schnidrig R, Kurrle D (2006) Evaluation of installation methods for Streckeisen STS-2 seismometers. http://www.geophys.uni-stuttgart.de/~widmer/ge2.pdf. Retrieved 25 Oct 2013

Wielandt E (2002) Seismic sensors and their calibration, Chapter 5. In: Bormann P (ed) New manual of seismological observatory practice, vol 1. GeoForschungsZentrum, Potsdam

Zürn W, Wielandt E (2007) On the minimum of vertical seismic noise near 3 mHz. Geophys J Int 168:647–658

System and Damage Identification of Civil Structures

Babak Moaveni[1] and Joel P. Conte[2]
[1]Department of Civil and Environmental Engineering, Tufts University, Medford, MA, USA
[2]Department of Structural Engineering, University of California at San Diego, CA, USA

Synonyms

Damage identification; Nondestructive evaluation; Structural damage; Structural health monitoring; System identification

Introduction

Definitions

Structural damage	Changes in a structural system that adversely affect its performance
System identification	Process of extracting the properties (often dynamic) of a system from measured output or input-output data
Damage identification	Process of detecting, localizing, and quantifying damage in a system
Structural health monitoring	Process of global damage identification in a structural system
Nondestructive evaluation	Process of local damage identification in a structural system

Major structural failures in recent years have brought the public's attention to the urgent need for improved infrastructure monitoring and maintenance. In their latest report card for America's infrastructure (American Society of Civil Engineers 2009), the American Society of Civil

Engineers (ASCE) described an infrastructure that is poorly maintained, unable to meet current and future demands, and, in some cases, unsafe. Expanding and improving structural health monitoring [C1] (SHM) for damage assessment and maintenance is essential for establishing sustainable and resilient civil infrastructure systems and ensuring they can meet the needs of future users. SHM refers to the process of damage identification in a structural system. Damage identification can be performed at four levels: (1) detection, i.e., estimating the existence of damage; (2) localization; (3) quantification; and (4) prognosis, i.e., predicting the remaining useful life of the structure. System identification refers to the process of extracting a system's dynamic properties from the measured data. System identification can be performed to extract damage-sensitive features of a system from its measured data to be used for the estimation of damage (Aktan et al. 1997). It can also be used to "realize" (i.e., identify) a mathematical dynamic model that best predicts the measured output data from the measured input excitation or the statistical correlation structure of the measured output data in the case of output-only measurements. The realized model can then be used to predict the structural response to different deterministic input excitations or the response correlation matrix to different stochastic excitations. Recent advances in sensing and computational power provide an opportunity for improved system and damage identification of large-scale and complex civil structures. The ASCE technical report (Catbas et al. 2013) provides a comprehensive review of the current state-of-the-art approaches, methods, and technologies for effective practices of system and damage identification methods on civil structures. While many researchers have successfully applied various identification approaches to numerical and/or small-scale lab models of civil structures, the literature lacks successful applications to real-world structures in the real loading environment. A recent special issue of the ASCE Journal of Structural Engineering (Moaveni et al. 2013a) was devoted to help bridge this gap.

System Identification

As previously mentioned, system identification refers to the process of estimating the dynamic properties of a structural system from its measured data. System identification methods can be categorized into input–output versus output-only, parametric versus nonparametric, time-domain versus frequency-domain, or linear versus nonlinear methods. Input–output system identification methods also referred to as experimental modal analysis [C2] (EMA) methods extract the dynamic system properties based on measurement data of both the dynamic response and the input excitation. On the other hand, output-only methods or operational modal analysis (OMA) methods [C3] are used when only the measured (ambient) response of a structure is available. OMA methods are suitable when the input excitation is not measureable and can be assumed as a broadband random signal such as wind loads on a building or bridge or vehicular traffic on a bridge. These methods are very useful in situations involving large-scale civil structures which are difficult to excite experimentally.

A system identification method is considered parametric if a mathematical dynamic model (often formulated in state-space) is "realized" in a first step and the dynamic properties of the system estimated from the realized model in the second step. Nonparametric system identification methods directly estimate the dynamic parameters of a system from transformation of data, e.g., Fourier transform or power-spectral density estimation. Time-domain identification methods estimate the dynamic parameters of a system by directly using the measured response time histories, while frequency-domain methods use the Fourier transformation or power-spectral density estimation of the measured time histories. There is also a class of time-frequency methods such as the short-time Fourier transform and the wavelet transform. These methods are commonly used for identification of time-varying systems in which the dynamic properties are time-variant. Linear system identification methods are based on the assumption that the system behaves linearly and

S

the identified dynamic properties correspond to an underlying linear system. Nonlinear methods [C4] assume specific types of nonlinearities in the system and its dynamic response and provide parameter estimates of the considered nonlinear model class. Nonlinear system identification is a challenging task due to the complex nature of nonlinearities encountered in civil structures and is the subject of ongoing research by many investigators.

A large number of system identification methods have been used successfully for identification of civil structures. Some of the more commonly used system identification methods in the literature include the peak picking method (both input–output and output-only, nonparametric, frequency domain, linear); the frequency-domain decomposition and enhanced frequency-domain decomposition methods (output-only, nonparametric, frequency domain, linear); the class of prediction error methods such as the auto-regressive (AR), auto-regressive with moving average (ARMA), auto-regressive with exogenous input (ARX), ARMAX, ARIMA, CARIMA, output error, and Box-Jenkins methods (input–output except AR, parametric, time domain, linear); subspace-based methods such as the eigensystem realization algorithm (ERA), the natural excitation technique combined with ERA (NExT-ERA), stochastic subspace identification (SSI), and deterministic–stochastic subspace identification (DSI) (output-only except DSI, parametric, time domain, linear); time-frequency methods such as the short-time Fourier transform, the wavelet transform, Hilbert-Huang transform, and proper orthogonal decomposition in combination with Hilbert transform or the wavelet transform (output-only, nonparametric, time frequency domain, nonlinear); and time-varying prediction error methods such as TV-ARX and TV-ARMA (input–output, parametric, time domain, nonlinear).

Damage Identification

Damage identification methods for structural systems are generally categorized into two groups:

(1) global methods (also referred to as SHM methods) and (2) local methods (also referred to as nondestructive evaluation or NDE methods). The global methods are usually applied for macroscopic damage identification of large-scale structures when there is no a priori knowledge about the location of damage and the instrumentation of the monitored structure is relatively scarce. In this case, damage is generally detected as the equivalent or effective loss of stiffness in large segments of structure (i.e., substructures). The local methods are applied when the monitoring region is smaller in size (due to, e.g., knowledge about the potential damage location(s)) and more dense (spatially) sensor data are available. NDE methods are commonly applied for monitoring aerospace structures, pipelines, and train tracks among others. These methods are capable of providing more accurate detection, localization, and quantification of damage; however, they are not in general suitable for application to larger structural systems such as building and bridges. Among NDE methods, acoustic emission, infrared thermography, and ultrasound methods can be named. Among global damage identification methods, vibration-based methods [C5] have received increased attention in the civil engineering research community in the last decade. The basic concept behind vibration-based damage identification is that the dynamic parameters of a structure are functions of its physical properties (mass, damping, and stiffness). Therefore, in concept, changes in these physical properties due to structural damage are detectable through changes in the identified dynamic parameters. One of the most common classes of methods is finite element (FE) model updating (Friswell and Mottershead 1995) which is able to detect, locate, and quantify damage. This method consists of updating the physical parameters of a FE model of the structure by minimizing an objective function that expresses the discrepancy between FE-predicted and experimentally measured response or dynamic features extracted from the response that are sensitive to damage. Dynamic features that are most commonly used for damage identification are modal parameters (especially natural frequencies and mode shapes) and the

System and Damage Identification of Civil Structures, Fig. 1 Dowling hall footbridge (*left*) and placement of concrete blocks on the bridge (*right*)

process of extracting modal parameters from vibration data is referred to as modal identification or system identification.

Damage identification methods can also be categorized into model-based methods and model-free or data-driven methods. In model-based damage identification methods, a physics-based model of the structure such as a FE model is used in the process of damage identification. These methods account for the known information about the geometry and material properties of the structure and are usually capable of detecting, localizing, and quantifying damage (level 3). Data-driven methods directly make use of the measured data on a structure and are usually capable of detection and sometimes localization of damage (level 1 or level 2). Some of the damage identification methods in the literature include methods based on changes in modal parameters such as natural frequencies and mode shapes; methods based on changes in features derived from modal parameters such as curvature mode shapes, modal strain energy, and modal flexibility; methods based on signal processing such as principal component analysis, Kalman filtering, and Hilbert–Huang transformation; and finally methods based on FE model updating.

Sample Case Studies

In this section, applications of system and damage identification methods to four large-scale civil structures are reviewed. Note that this is

not a comprehensive list of successful applications by any means and only includes a few of the authors' past work based on which some of the conclusions are drawn.

Dowling Hall Footbridge

Dowling Hall Footbridge, a full-scale footbridge located on the Tufts University campus as shown in Fig. 1, is equipped with a continuous monitoring system that measures continuously ambient vibration and temperature of the bridge. Modal parameters of the footbridge are identified every hour using an automated SSI method (Moser and Moaveni 2011; Moaveni and Behmanesh 2012). Correlation of the identified natural frequencies with ambient temperatures was investigated and different models were proposed to capture the relationship. Effects of physical damage were physically simulated on this test-bed structure by loading a small segment of the footbridge deck with concrete blocks (Fig. 1, right) so as to locally modify the bridge inertia properties since no damage could be inflicted to the bridge to modify its stiffness properties. Deterministic and probabilistic FE model updating approaches were implemented for damage identification of the bridge (Behmanesh and Moaveni 2014). Both methods could successfully estimate the location and extent of damage (here change in mass). However, the probabilistic (Bayesian) method could also provide the level of confidence in the damage identification results by estimating the posterior probability distributions of the FE model updating parameters.

Lessons Learned: The maximum a posteriori (MAP) estimates of the updating model parameters match the values of the physically simulated damage (i.e., local change in mass) and are in good agreement with the optimum parameter values obtained from the deterministic FE model updating approach. Effects of the number of data sets used in the damage identification process (i.e., "value" of added data) were investigated by using different subsets of available data and it was found that the estimation uncertainty of the updating model parameters is significantly reduced by adding more data sets to the likelihood function, which yields more accurate model updating results. However, the reduction in estimation uncertainty becomes progressively less significant as the number of data sets keeps increasing. In the application of deterministic FE model updating, addition of more data sets does not necessarily reduce the estimation uncertainty of the model updating results. Thus, probabilistic FE model updating approaches based on multiple sets of measurement data are strongly recommended for structural damage identification purposes as they can provide a measure of confidence to the estimated damage values.

Seven-Story Shear Wall Building Section

A full-scale seven-story reinforced concrete (RC) shear wall building section, shown in Fig. 2, was tested on the UCSD-NEES shake table in the period October 2005–January 2006. The shake table tests were designed to damage the building progressively through several historical earthquake ground motions reproduced on the shake table. At several levels of damage, ambient vibration tests and low-amplitude white-noise base excitation tests were applied to the building which responded as a quasi-linear system with dynamic parameters evolving as a function of structural damage. Six different state-of-the-art system identification methods including three output-only and three input–output methods were used to estimate the modal parameters (natural frequencies, damping ratios, and mode shapes) of the building in its undamaged (baseline) and various damage states

System and Damage Identification of Civil Structures, Fig. 2 Seven-story shear wall test structure

(Moaveni et al. 2011). Deterministic and probabilistic finite element (FE) model updating strategies were applied for vibration-based damage identification of the test specimen (Moaveni et al. 2010; Simeon et al. 2013). Three damage identification cases are considered based on modal parameters identified using ambient vibration, 0.03 and 0.05 g white-noise base excitation test data, respectively. The damage identification results obtained for these three cases do not match exactly, but they are consistent with the actual damage observed in the building which shows a concentration of damage at the bottom two stories of the web wall. The ambient vibration data satisfy better the assumption of system linearity and therefore are more appropriate as input for linear FE model updating.

Lessons Learned: From the results of the system identification study, it was observed that the identified natural frequencies of the three longitudinal

(in the direction of shaking) vibration modes decrease with increasing level of damage, while the corresponding identified modal damping ratios do not follow a clear trend as a function of structural damage. The (effective) natural frequencies of the three longitudinal vibration modes identified based on higher amplitude structural response data are in general lower than their counterparts identified based on low-amplitude response data at all damage states considered, especially for the first longitudinal vibration mode. This is most likely due to the fact that the test structure was nonlinear (even at the relatively low levels of excitation considered in the system identification study) with effective modal parameters depending strongly on the amplitude of the excitation and therefore of the structural response. In general, higher modal damping ratios were identified for the three longitudinal vibration modes during the higher amplitude base excitation tests. This is due to the fact that the additional hysteretic energy dissipation (due to inelastic action of the material) at higher level of response nonlinearity is identified as equivalent viscous damping by the linear system identification methods used. It should be emphasized that these "inflated" identified equivalent viscous damping ratios are not to be used to represent the viscous damping component in a nonlinear FE model of a structure that explicitly accounts for the nonlinear material behavior. The damage identification results obtained using a linear FE model updating approach are sensitive to the amplitude of the white-noise base excitation to which the structure is assumed to respond quasi-linearly. With increasing level of base excitation and structural damage, the level of nonlinearity in the structural response increases. Therefore, the assumption that the structure behaves as a quasi-linear dynamic system is violated and a linear dynamic model is not strictly able to represent well the behavior of a damaged structure. The spatial distribution (i.e., relative amplitudes) of the identified damage factors however is not sensitive to the amplitude of base excitation and was found in good agreement with the

damage observed or inferred (from strain sensors) in the building specimen. The Bayesian FE model updating approach also succeeded in identifying the damage in the test structure and the results obtained were found in good agreement with their deterministic counterparts. The probabilistic model updating method also provided an estimate for the uncertainties in the identified damage values. It was found that the data used contained little information about the state of the top stories of the building, as shown by the fact that the uncertainty of the parameters representing this portion of the structure could not be reduced through Bayesian processing of the observed data.

Three-Story RC Frame with Masonry Infilled Walls

A deterministic linear FE model updating strategy was applied for vibration-based damage identification of a 2/3-scale, three-story, two-bay, infilled RC frame, tested on the UCSD-NEES outdoor shake table (Moaveni et al. 2013b). Figure 3 shows the test structure on the shake table. FE model updating was first used to calibrate the FE model at the undamaged state which served as the reference/baseline state and then applied for damage identification at a number of damage states of the structure. These damage states correspond to states of increasing damage of the physical specimen which was subjected to a sequence of earthquake excitations of increasing intensity. The damage identification results indicate that the severity of structural damage increases as the structure is exposed to stronger earthquake excitations. The damage identification method correctly identifies the spatial distribution of damage in the structure, with the most severe damage at the bottom story and the least damage at the top story. The method also captures the fact that the extent of damage is more significant in the infill walls than in the columns. The analytical modal parameters obtained from the updated FE models are in good agreement with their experimentally identified counterparts, an indication of the accuracy of the updated FE models.

System and Damage Identification of Civil Structures, Fig. 3 Front view (*left*) and side view of the three-story reinforced concrete frame with masonry infilled walls (*right*)

System and Damage Identification of Civil Structures, Fig. 4 Three-story precast concrete parking garage test structure

Lessons Learned: Comparison of the damage identification results and the seismic shake table test results shows that the level of damage identified may not reflect the loss of the structural strength, since the loss of stiffness (commonly defined as structural damage) is not well correlated to the loss of strength. This motivates the need for new research based on nonlinear FE model updating based on calibrated, mechanics-based, nonlinear FE models of structural systems able to capture stiffness degradation and strength deterioration phenomena.

Three-Story Precast Concrete Parking Garage
This study performed damage assessment through system identification of a three-story half-scale precast concrete building resembling a parking garage tested under earthquake type loading on the NEES-UCSD Shake Table in 2008 (Belleri et al. 2014). Figure 4 shows the parking garage test structure on the NEES-UCSD shake table. The system identification was performed using the DSI method. The effective modal parameters of the structure at different damage states show significant changes after

each of the high intensity earthquake tests. In general, the identified natural frequencies decrease and the identified damping ratios increase as the structure is exposed to base excitations of increasing intensity. Changes in the identified modal parameters are correlated with the observed damage in the structure. The analysis of the identified mode shapes allow to point to the location of damage. Specifically, localized changes in the first and second spatial derivatives of mode shapes, modal rotations, and curvatures indicate local losses of shear and bending stiffness, respectively. Local significant changes in modal rotations and curvatures allowed localizing both visually detected damage in the third floor mid-span joint and visually undetected damage in the second floor mid-span joint, thus providing a useful tool for precast floor damage detection.

Lessons Learned: The accuracy and spatial resolution of the damage identification results depend significantly on the accuracy and completeness of the identified modal parameters. The estimation variability (or uncertainty) of the modal parameters can be influenced by several factors such as the number and types of sensors, measurement noise, length of the data time windows used for system identification, and the system identification method used, in addition to changes in the environmental conditions such as the ambient air temperature and relative humidity during data collection. The variability in the identified modal parameters due to non-damage related factors needs to be smaller than the changes in these parameters due to damage or compensated for in order to identify the actual damage in the structure using vibration measurements.

Summary

This paper provides an overview of the concepts of system identification and damage identification of civil structures. First, the need for system and damage identification of civil structures is underlined and the current gap in the knowledge is discussed. System identification and damage identification processes are defined and some of the popular methods for application to civil structures are mentioned. Finally, a few successful applications to large-scale civil structures are provided as illustrations and conclusions are drawn from these studies.

Cross-References

▶ Modal Analysis
▶ Nonlinear System Identification: Particle-Based Methods
▶ Operational Modal Analysis in Civil Engineering: An Overview
▶ Vibration-Based Damage Identification: The Z24 Bridge Benchmark

References

Aktan A, Farhey D, Helmicki A, Brown D, Hunt V, Lee K, Levi A (1997) Structural identification for condition assessment: experimental arts. J Struct Eng ASCE 123(12):1674–1684

American Society of Civil Engineers (ASCE) (2009) Report card for America's infrastructure. http://www.asce.org/reportcard/2009

Behmanesh I, Moaveni, B (2014) Probabilistic identification of simulated damage on the Dowling Hall Footbridge through Bayesian FE model updating. Structural Control and Health Monitoring, in press

Belleri A, Moaveni B, Restrepo JI (2014) Damage assessment through structural identification of a large-scale three-story precast concrete structure. Earthq Eng Struct Dyn 43(1):61–76

Catbas FN, Kijewski-Correa TL, Aktan AE (eds) (2013) Structural identification of constructed systems: approaches, methods and technologies for effective practice of St-Id. Technical report, ASCE, Structural Engineering Institute, Reston

Friswell MI, Mottershead JE (1995) FE model updating in structural dynamics. Klumer, Boston

Moaveni B, Behmanesh I (2012) Effects of changing ambient temperature on finite element model updating of the Dowling Hall Footbridge. Eng Struct 43:58–68

Moaveni B, He X, Conte JP, Restrepo JI (2010) Damage identification study of a seven-story full-scale building slice tested on the UCSD-NEES shake table. Struct Saf 32(5):347–356

S

Moaveni B, He X, Conte JP, Restrepo JI, Panagiotou M (2011) System identification study of a seven-story full-scale building slice tested on the UCSD-NEES shake table. J Struct Eng ASCE 137(6):705–717

Moaveni B, Hurlebaus S, Moon F (2013a) Editorial of special issue on real-world application of structural identification and health monitoring methodologies. J Struct Eng ASCE 139:1637–1638

Moaveni B, Stavridis A, Lombaert G, Conte JP, Shing PB (2013b) Finite element model updating for assessment of progressive damage in a three-story infilled RC frame. J Struct Eng ASCE 139:1665–1674

Moser P, Moaveni B (2011) Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge. Mech Syst Signal Process 25(7): 2336–2357

Simeon E, Moaveni B, Conte JP, Lombaert G (2013) Uncertainty quantification in the assessment of progressive damage in a seven-story full-scale building slice. J Eng Mech ASCE 139(12):1818–1830