


Assessment of Frequency versus Time Domain enhanced technique for response-only modal dynamic identification under seismic excitation

Fabio Pioldi¹ · Egidio Rizzi¹ 

Received: 14 April 2017 / Accepted: 1 November 2017 / Published online: 11 November 2017
© Springer Science+Business Media B.V., part of Springer Nature 2017

Abstract In the present study, response-only system identification is explored towards estimating modal dynamic properties of buildings under earthquake excitation. Seismic structural response signals are adopted with two different implemented Operational Modal Analysis (OMA) techniques, namely a *refined Frequency Domain Decomposition* algorithm and an *improved data-driven Stochastic Subspace Identification (SSI-DATA)* procedure, working in the Frequency Domain and in the Time Domain, respectively. These algorithms are specifically conceived to operate with seismic structural responses and at simultaneous heavy damping (in terms of identification challenge), towards achieving consistent estimations of natural frequencies, mode shapes and modal damping ratios. Classical OMA assumptions shall not contemplate short-duration, non-stationary earthquake-induced response signals. Nevertheless, the present enhanced output-only algorithms allow for estimating all strong ground motion modal parameters. First, a linear three-storey frame structure under ten selected earthquake base-excitations is considered, in order to assess the two OMA techniques at *synthetic* seismic response input. Comprehensive results are presented, by comparing the two developed methods, between achieved modal estimates and sought target values. Second, an existing instrumented building (CESMD database) is analyzed, by adopting *real* seismic response signals, reaching again effective modal estimates and corroborating the previous necessary condition analysis. According to this investigation, best up-to-date, re-interpreted, output-only techniques may effectively be used for potential Structural Health Monitoring purposes in the Earthquake Engineering range.

Keywords Operational Modal Analysis (OMA) · refined Frequency Domain Decomposition (rFDD) · improved Data-Driven Stochastic Subspace Identification (SSI-DATA) · synthetic and real earthquake structural response input · heavy damping

✉ Egidio Rizzi
egidio.rizzi@unibg.it

¹ Dipartimento di Ingegneria e Scienze Applicate, Scuola di Ingegneria (Dalmine), Università degli Studi di Bergamo, viale G. Marconi 5, 24044 Dalmine, BG, Italy

1 Introduction

The dynamic description of mechanical and civil engineering systems constitutes a fundamental target towards the characterization of structural modal properties. To this purpose, not only ambient or experimentally controlled vibrations, but also earthquake-induced excitations may be adopted (Ntotsios et al. 2009; Sevim et al. 2016). In Earthquake Engineering, specifically developed identification procedures may be adopted in order to detect structural modifications and to achieve reliable strong ground motion modal parameter estimates.

In the Earthquake Engineering field, most of known identification procedures pertain directly to Experimental Modal Analysis (EMA), where both input and output signals need to be available for achieving an appropriate operation of the estimation procedures (Heylen et al. 2006; Steiger et al. 2016). During the last decades, Operational Modal Analysis (OMA) has been increasingly adopted in this field (Rainieri and Fabbrocino 2014; Bindi et al. 2015). With OMA procedures, only structural responses (input signals for the identification technique) need to be known, which makes it particularly suitable for treating ambient vibrations, but in principle also for processing earthquake excitations.

Generally, for OMA applications, the unknown excitation input acting on the structure shall be considered to be similar, in main characteristics, to that of a (stationary) white noise signal, as it may be recorded under typical ambient or operational excitations. In the field of OMA algorithms, the adoption of input channels coming from (short-duration, non-stationary) earthquake-induced structural response signals has been considered quite a few times in the dedicated literature, either in the Time or in the Frequency Domain. On that, a brief summary on recent literature contributions is reported below.

By adopting earthquake-induced structural response signals, most of current OMA techniques refer to *Time Domain* methods. In Pridham and Wilson (2004), a Stochastic Subspace Identification (SSI) approach was combined to an Expectation Maximization method in order to identify shear-type frames under seismic base excitation. In Lin et al. (2005), an Ibrahim time domain method with modified random decrement was adopted, by relying only on a few floor acceleration seismic responses. In Kun et al. (2009), Taylor's first-order approximations for the identification equations were adopted and solved with damped iterative Least-Squares (LS). In Ghahari et al. (2013), spatial time-frequency distributions were used for blind identification of strong motion modal parameters. In Gouache et al. (2013), OMA under harmonic transient input was attempted through a phase analysis. In Pioldi and Rizzi (2016a, b), a Full Dynamic Compound Inverse Method (FDCIM) was developed, to simultaneously identify output-only modal parameters, element-level structural features and earthquake input excitation, through a two-stage iterative identification algorithm.

Conversely, by adopting OMA techniques at seismic input in the *Frequency Domain*, only a few notable exceptions have been investigated in the field. In Ventura et al. (2005), a real street overpass was studied through ambient vibrations and earthquake ground motions, by a commercial software version of Frequency Domain Decomposition (FDD) (Brincker et al. 2001) (and of SSI, Peeters and De Roeck 1999). In Mahmoudabadi et al. (2007), a method based on iterative Least-Squares was proposed to identify classically damped linear systems. In Michel et al. (2010), the seismic response of a building to weak earthquakes was studied through a FDD algorithm.

By making reference to such output-only identification techniques, earthquake-induced structural responses are adopted here as input channels for a *refined Frequency Domain Decomposition (rFDD)* algorithm (Brincker et al. 2001) (Frequency Domain) and an

improved Data-Driven Stochastic Subspace Identification (SSI-DATA) technique (Peeters and De Roeck 1999) (Time Domain). Both methods were implemented autonomously within MATLAB, for rFDD in Pioldi et al. (2015a, b), Pioldi and Rizzi (2017) and for SSI-DATA starting from Pansieri (2016) and as definitely presented here. The traditional versions of these algorithms rely on the typical assumption of white noise input, which no longer holds with seismic response input. On the contrary, the present methods have been specifically developed to deal with earthquake-induced structural responses and simultaneous heavy damping (in terms of identification challenge, i.e. with modal damping ratios larger than only a few percents and up to 10% or higher).

In the present investigation, the autonomously-implemented rFDD and SSI-DATA techniques are separately adopted first to identify the modal properties of a reference linear heavy-damped 3-dof shear-type frame under ten different selected strong ground motions. This work shall constitute a first basis for a comparison between the two OMA identification methods within the Earthquake Engineering range, in order to inspect their positive and negative aspects, concerning their reliability of correctly identifying strong ground motion modal parameters, within the linear range of seismic response, as recorded by *synthetic* response signals. From the performed analyses, the achieved estimates are compared among them and also with the known target values computed before stage, in order to extract general and specific considerations on the efficiency and consistency of both OMA algorithms and to investigate and compare their effectiveness in identifying all current strong ground motion modal parameters. Second, a real structural case is considered, based on a single earthquake record and attached *real* seismic response signals. The two identification techniques are further employed, again with very consistent results, which further corroborates and completes the previous necessary condition analysis with synthetic seismic response signals.

Presentation is structured as follows. In Sects. 2 and 3, necessary main theoretical backgrounds (Sects. 2.1, 3.1) and enhancements (Sects. 2.2, 3.2) of the developed rFDD and SSI-DATA algorithms are outlined, respectively. In Sect. 4 the selected earthquake dataset and the adopted numerical model for the first synthetic analyses are presented (Sect. 4.1), jointly with the results separately achieved by the two OMA identification methods for a reference three-storey frame (Sect. 4.2). Then, real recorded seismic response signals for an existing instrumented building are finally processed by the two identification techniques and presented in Sect. 5. In the end, salient conclusions on the whole investigation are gathered in closing Sect. 6.

2 Fundamentals of the present refined Frequency Domain Decomposition algorithm

2.1 Classical FDD theoretical background

Classical FDD theory is based on a general input/output expression as a function of frequency ω for a MDoF system (Bendat and Piersol 1986; Brincker et al. 2001):

$$\mathbf{G}_{yy}(\omega) = \overline{\mathbf{H}}(\omega)\mathbf{G}_{xx}(\omega)\mathbf{H}^T(\omega) \quad (1)$$

where $\mathbf{G}_{xx}(\omega) \in \mathcal{R}^{r \times r}$ is the input Power Spectral Density (PSD) matrix (excitations), r is the number of input channels; $\mathbf{G}_{yy}(\omega) \in \mathcal{R}^{m \times m}$ is the output PSD matrix (responses), m is

the number of output response signals; $\mathbf{H}(\omega) \in \mathcal{R}^{m \times r}$ is the Frequency Response Function (FRF) matrix; overbar denotes complex conjugate and apex symbol T transpose.

FRF $\mathbf{H}(\omega)$ may also be written in pole/residue form as Heylen et al. (2006):

$$\mathbf{H}(\omega) = \sum_{k=1}^n \frac{\mathbf{R}_k}{i\omega - \lambda_k} + \frac{\overline{\mathbf{R}}_k}{i\omega - \bar{\lambda}_k} \tag{2}$$

where n is the number of modes, λ_k and $\bar{\lambda}_k$ are the k th poles (in complex conjugate pairs) of the FRF function (Heylen et al. 2006), and $\mathbf{R}_k = \phi_k \mathbf{\Gamma}_k^T \in \mathcal{R}^{m \times r}$ is the residue matrix (Brincker et al. 2001; Reynders 2012), obtained by the product between mode shape vector $\phi_k \in \mathcal{R}^{m \times 1}$ and modal participation factor vector $\mathbf{\Gamma}_k \in \mathcal{R}^{r \times 1}$.

When all output measurements are taken as input references (i.e. when $m = r$), $\mathbf{H}(\omega)$ becomes a square matrix. So, by defining the residue matrix of PSD output $\mathbf{A}_k \in \mathcal{R}^{m \times m}$, which corresponds to k th pole λ_k , as (Brincker et al. 2001):

$$\mathbf{A}_k = \sum_{s=1}^n \left(\frac{\mathbf{R}_s}{-\lambda_k - \lambda_s} + \frac{\overline{\mathbf{R}}_s}{-\lambda_k - \bar{\lambda}_s} \right) \mathbf{G}_{xx} \mathbf{R}_k^T \tag{3}$$

output PSD matrix $\mathbf{G}_{yy}(\omega)$ of Eq. (1), after substituting Eqs. (2) and (3), and by applying the Heaviside partial fraction expansion theorem, may be reduced to the following final pole/residue form (Brincker et al. 2001; Wang et al. 2005):

$$\mathbf{G}_{yy}(\omega) \simeq \sum_{k=1}^n \frac{\mathbf{A}_k}{i\omega - \lambda_k} + \frac{\mathbf{A}_k^H}{-i\omega - \bar{\lambda}_k} + \frac{\overline{\mathbf{A}}_k}{i\omega - \bar{\lambda}_k} + \frac{\mathbf{A}_k^T}{-i\omega - \lambda_k} \tag{4}$$

where Hermitian apex symbol H denotes complex conjugate and transpose. Under the assumption of *stationary white noise input* (Bendat and Piersol 1986), PSD matrix $\mathbf{G}_{xx}(\omega)$ degenerates to a real-valued non-negative single scalar constant \mathbf{G}_{xx} .

When the structure is *lightly damped* (small modal damping ratios $\zeta_k \ll 1$), the pole in the vicinity of k th modal frequency ω_k can be expressed in approximate form as (Brincker et al. 2001):

$$\mathbf{A}_k \simeq \frac{\overline{\mathbf{R}}_k \mathbf{G}_{xx} \mathbf{R}_k^T}{2\zeta_k \omega_k} = \frac{\bar{\phi}_k \mathbf{\Gamma}_k^H \mathbf{G}_{xx} \mathbf{\Gamma}_k^T \phi_k}{2\zeta_k \omega_k} = d_k \bar{\phi}_k \phi_k^T \tag{5}$$

where index k spans modes $k = 1, \dots, n$, ζ_k is the modal damping ratio and term d_k can be proven to be a real scalar. So, by substituting Eq. (5) into Eq. (4), in the narrow band with spectrum lines in the vicinity of a modal frequency, Eq. (4) can be simplified to:

$$\mathbf{G}_{yy}(\omega) \simeq \sum_{k=1}^n \frac{d_k \bar{\phi}_k \phi_k^T}{i\omega - \lambda_k} + \frac{d_k \bar{\phi}_k \phi_k^T}{-i\omega - \bar{\lambda}_k} = \bar{\Phi} \left\{ \text{diag} \left[\text{Re} \left(\frac{2d_k}{i\omega - \lambda_k} \right) \right] \right\} \Phi^T \tag{6}$$

where Φ is the eigenvector matrix, gathering all n eigenvectors ϕ_i as columns.

Therefore, the first step of classical FDD methods is the estimation of the PSD matrix of system responses $\mathbf{G}_{yy}(\omega)$ in previous Eq. (6) by time correlation first, in the Time Domain, and then Fourier transform to the Frequency Domain. Then, its transpose shall be decomposed by performing a Singular Value Decomposition (SVD) at each discrete frequency line $\omega = \omega_i$:

$$\mathbf{G}_{yy}^T(\omega = \omega_i) \simeq \mathbf{\Phi}_i \left\{ \text{diag} \left[\text{Re} \left(\frac{2d_j}{i\omega_i - \lambda_j} \right) \right] \right\} \mathbf{\Phi}_i^H = \mathbf{U}_i \mathbf{S}_i \mathbf{U}_i^H \tag{7}$$

where $\mathbf{\Phi}_i$ is the i th eigenvector matrix, gathering all m eigenvectors ϕ_{ij} as columns, and index $j = 1, \dots, m$, being m the number output response signals, i.e. the input channels for the algorithm. In parallel, \mathbf{U}_i is a unitary complex matrix holding singular vectors \mathbf{u}_{ij} and \mathbf{S}_i is a real diagonal matrix holding Singular Values (SV) s_{ij} .

Starting from the SVD in Eq. (7), the identification of mode q can be made around a modal peak in the frequency domain, which can be located by an appropriate peak-picking procedure on the SV representations. Then, the response PSD matrix of the q th mode, in correspondence of identified damped modal frequency ω_q , can be approximated as (Brincker et al. 2001):

$$\mathbf{G}_{yy}^T(\omega_i = \omega_q) \simeq \phi_q \left\{ \text{diag} \left[\text{Re} \left(\frac{2d_q}{i\omega_q - \lambda_q} \right) \right] \right\} \phi_q^H = s_1 \mathbf{u}_{q1} \mathbf{u}_{q1}^H \tag{8}$$

where first singular vector \mathbf{u}_{q1} at q th resonance frequency ω_q leads to an estimate of related mode shape vector $\hat{\phi}_q = \mathbf{u}_{q1}$. Associated Singular Value s_1 is the Auto-PSD function of the corresponding SDoF system, which may be detected by comparing the identified mode shape ϕ_q with the surrounding singular vectors around the peak. For this purpose, the Modal Assurance Criterion (MAC) index may be classically used (Brincker et al. 2001):

$$\text{MAC}(\phi_q, \phi_s) = \frac{|\phi_q^H \phi_s|^2}{|\phi_q^H \phi_q| |\phi_s^H \phi_s|} \tag{9}$$

If MAC index is 1, the two compared vectors are considered to be identical; if 0, clearly distinct (orthogonal).

Then, for the modal damping ratio identification it is possible to operate as follows. The Inverse Discrete Fourier Transform (IDFT) (Time Domain) of the located q th Auto-PSD function (Frequency Domain) allows to obtain an estimate of the SDoF Auto-Correlation Function (ACF) related to the located resonance peak. In this process, remaining parts of the Auto-PSD function are simply reset to zero (Brincker et al. 2001). All ACF extrema (i.e. peaks and valleys), which represent the free amplitude decay of a damped SDoF system may be detected by peak-picking of peaks and valleys within an appropriate time window. The logarithmic decrement, which is classically defined as $\delta_j = (2/j) \ln(r_0/|r_j|)$, can be estimated by a linear regression on δj and $2 \ln(|r_j|)$, where $j = 1, 2, \dots$ is an integer index counter of the j th ACF extreme and r_0, r_j are the initial and the j th extreme value of the ACF, respectively. Then, modal damping ratio ζ_q can be typically estimated as:

$$\zeta_q = \frac{\delta_q}{\sqrt{4\pi^2 + \delta_q^2}} \tag{10}$$

Finally, knowing the estimated modal damping ratio, the undamped natural frequency can be obtained from the estimated damped modal frequency by dividing it by factor $\sqrt{1 - \zeta_k^2}$.

2.2 Main enhancements of the present rFDD algorithm

Main assumptions of classical FDD methods consist of (stationary) white noise input, light damping (modal damping ratios in the order of 1%) and geometrically-orthogonal mode shapes of close modes (Brincker et al. 2001).

The present rFDD method, whose original theoretical background has been reported in Pioldi et al. (2015a, b), conceptually derives from classical FDD methods (Brincker et al. 2001), but has been specifically developed to deal with earthquake-induced structural response signals and concurrent heavy damping (in terms of FDD identification challenge, i.e. for realistic modal damping ratios up to 10%).

Pioldi et al. (2015a, b) have discussed the theoretical validity and efficacy of the present rFDD technique, through the use of synthetic seismic response signals in the linear range, possibly affected by simulated noise of different levels. Trials with real earthquake responses and damage scenarios in the non-linear range have been effectively performed as well in Pioldi et al. (2017). In Pioldi and Rizzi (2017), further rFDD computational strategies have been introduced, by adopting excitation data from the complete FEMA P695 earthquake database, towards achieving an extensive validation in the Earthquake Engineering range. In Pioldi et al. (2016), the rFDD technique has been also applied to frames under Soil-Structure Interaction (SSI) effects, towards obtaining the identification of flexible- and fixed-base modal parameters from earthquake-induced structural response signals.

For the sake of completeness, the following computational steps (and references quoted therein) summarize the main workflow of the present rFDD algorithm:

- Suitably-developed filtering applied to the structural response input signals (earthquake-induced structural responses) before starting the modal identification process (Pioldi et al. 2015a).
- Coupling of rFDD to a time-frequency Gabor Wavelet Transform (GWT), towards achieving a correct evaluation of the time-frequency features of the signals and a best setup for rFDD identification (Pioldi and Rizzi 2017).
- Processing of the auto- and cross-correlation matrix entries, by aiming at obtaining clearer and well-defined SVs out of seismic response signals (Pioldi et al. 2015b).
- Integrated PSD matrix computation, implementing simultaneously both Wiener-Khinchin and Welch's modified periodogram methods (Pioldi et al. 2015a, b). The Wiener-Khinchin's approach works especially well with short signals, allowing for a clearer peak detection, not only on the first SV curve, but also on the subsequent ones. Welch's method, instead, implements averaging and windowing before the frequency-domain convolution, allowing to achieve better mode shapes, despite for the not so good separation of the signals in the modal space. Then, the integrated PSD matrix computation aims at extracting better modal estimates, by taking simultaneous advantage of both PSD evaluation methods.
- Iterative loop and optimization algorithm towards achieving effective modal damping ratio estimates, especially under heavy-damping identification conditions (Pioldi et al. 2015a).
- Coupled Chebyshev Type II bandpass filters computational procedure, aiming at enhancing the SDOF spectral bells towards estimation improvement, when challenging seismic input and heavy-damping conditions apply (Pioldi and Rizzi 2017).

- Estimation of modal parameters by operating on different SVs and on their composition, to detect each SV contribution and to reconstruct the original SDoF spectral bells (Pioldi et al. 2015a).
- Inner procedure for frequency resolution enhancement, without the need of higher frequency sampling, as first outlined in Pioldi et al. (2015a).
- Combined use of different MAC indexes towards modal validation purposes (Pioldi et al. 2015a, 2017). After a preliminary “peak-picking” (Pioldi et al. 2015a), the use of Modal Assurance Criterion (MAC) and Modal Phase Collinearity (MPC) indexes (Pioldi et al. 2015b) becomes necessary to discern “spurious peaks” from true (physical) modal ones. In particular, these indexes have been used to discard spurious peaks that exhibit a complex-number character (i.e. displaying modal deflection phases that significantly deviate from 0 or π).

Consistently, the rFDD results reported later in Sect. 4 demonstrate the robustness of the developed rFDD algorithm in returning reliable modal parameter estimates at seismic response input and concurrent heavy damping. These are going to be compared to those independently achieved by a separate SSI-DATA implementation, as introduced in the next section.

3 Fundamentals of the present improved Data-Driven Stochastic Subspace Identification algorithm

3.1 Classical SSI theoretical background

A classical system of m second-order differential equations of motion of a dynamical linear structural system may be written as (*spatial model*):

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \bar{\mathbf{B}}\mathbf{f}'(t) \tag{11}$$

where \mathbf{M}, \mathbf{C} and $\mathbf{K} \in \mathcal{R}^{m \times m}$ are the mass, damping and stiffness matrices, matrix $\bar{\mathbf{B}} \in \mathcal{R}^{m \times m}$ defines the location of the input channels, $\mathbf{f}'(t) \in \mathcal{R}^{m \times 1}$ is the input force vector and $\ddot{\mathbf{u}}(t), \dot{\mathbf{u}}(t)$ and $\mathbf{u}(t) \in \mathcal{R}^{m \times 1}$ are the vectors of total acceleration, velocity and displacement structural responses.

By switching to *State-Space* form, the m second-order differential equations of motion in Eq. (11) can be rewritten into $2m$ first-order differential equations as (Van Overschee and De Moor 1996):

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}_c\mathbf{x}(t) + \mathbf{B}_c\mathbf{f}'(t) \\ \mathbf{y}(t) = \mathbf{C}_c\mathbf{x}(t) + \mathbf{D}_c\mathbf{f}'(t) \end{cases} \tag{12}$$

being the first equation the *state equation*, in terms of *state vector* of responses $\mathbf{x}(t) \in \mathcal{R}^{2m \times 1}$, $\mathbf{x}(t) = \{\mathbf{u}(t) \dot{\mathbf{u}}(t)\}^T$ and its derivative $\dot{\mathbf{x}}(t) = \{\dot{\mathbf{u}}(t) \ddot{\mathbf{u}}(t)\}^T$, and the second equation the *observer equation*, in terms of *observer vector* of responses $\mathbf{y}(t) \in \mathcal{R}^{m \times 1}$ (either displacements, velocities and/or, typically, accelerations). From Eq. (12) *state matrix* $\mathbf{A}_c \in \mathcal{R}^{2m \times 2m}$, *input matrix* $\mathbf{B}_c \in \mathcal{R}^{2m \times m}$, *output matrix* $\mathbf{C}_c \in \mathcal{R}^{m \times 2m}$ and *feed-through matrix* $\mathbf{D}_c \in \mathcal{R}^{m \times m}$ are defined as follows, where subscript c denotes continuous time (Van Overschee and De Moor 1996):

$$\begin{aligned}
 \mathbf{A}_c &= \begin{bmatrix} \mathbf{0}_{n \times n} & \mathbf{I}_{n \times n} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix}, & \mathbf{B}_c &= \begin{bmatrix} \mathbf{0}_{n \times n} \\ \mathbf{M}^{-1}\bar{\mathbf{B}} \end{bmatrix} \\
 \mathbf{C}_c &= [-\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C}], & \mathbf{D}_c &= \mathbf{M}^{-1}\bar{\mathbf{B}}
 \end{aligned}
 \tag{13}$$

Here, the structures of matrices \mathbf{C}_c and \mathbf{D}_c refer to the adoption of *total acceleration responses* for the observer vector, namely $\mathbf{y}(t) = \ddot{\mathbf{u}}(t)$.

In the OMA context, structures are excited by unmeasurable and spatially distributed input excitations; this means that the information from excitation $\mathbf{f}'(t)$ is not available. Also, experimental tests yield measurements taken at discrete time instants, while Eqs. (11)–(13) are actually expressed in continuous time. For a given sampling time interval Δt , continuous-time equations can be discretized and solved at discrete time instants $t_k = k\Delta t, k = 1, \dots, N$, with N being the total number of sampling points of the signal.

By taking into account the k th time instant and assuming unknown/unmeasured input (treated as white noise), classical SSI-DATA theory takes as a typical starting point of the identification process the following OMA Stochastic State-Space model in discrete-time notation (Van Overschee and De Moor 1996):

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{w}_k \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \end{cases}
 \tag{14}$$

being $\mathbf{x}_k = \{\mathbf{u}_k \ \mathbf{u}_{k+1}\}^T$ the *state vector* of responses and its derivative $\mathbf{x}_{k+1} = \{\mathbf{u}_{k+1} \ \mathbf{u}_{k+2}\}^T$, and \mathbf{y}_k the *observer vector* of responses (either displacements, velocities and/or, typically, accelerations, as considered here). Notation $\mathbf{u}_k, \mathbf{u}_{k+1}$ and \mathbf{u}_{k+2} refers to the discrete-time counterparts of continuous-time vectors of displacement $\mathbf{u}(t)$, velocity $\dot{\mathbf{u}}(t)$ and acceleration $\ddot{\mathbf{u}}(t)$ responses, respectively. There, also \mathbf{A} and \mathbf{C} are the discrete-time counterparts of \mathbf{A}_c and \mathbf{C}_c matrices, while vectors \mathbf{w}_k and $\mathbf{v}_k \in \mathcal{R}^{2m \times 1}$ are zero mean, stationary white noise stochastic processes, representing process noise and measurement noise, respectively (Van Overschee and De Moor 1996). These \mathbf{w}_k and \mathbf{v}_k stochastic processes become necessary and shall be included, in order to describe real measurement data, which are also driven by uncertainty and noise.

Then, the salient mathematical and computational steps described in the following summarize the main workflow of the present SSI-DATA algorithm.

The first step of classical SSI-DATA identification algorithms is the computation of the so-called *block Hankel matrix* $\mathbf{H}_{0|2i-1} \in \mathcal{R}^{2mi \times j}$ of responses, which is directly calculated from the measurement data (Van Overschee and De Moor 1996), i.e. system structural responses \mathbf{y}_k (in the present case, total accelerations), as:

$$\mathbf{H}_{0|2i-1} = \frac{1}{\sqrt{j}} \begin{bmatrix} \mathbf{y}_0 & \cdots & \mathbf{y}_{j-1} \\ \vdots & \ddots & \vdots \\ \mathbf{y}_{i-1} & \cdots & \mathbf{y}_{i+j-2} \\ \mathbf{y}_i & \cdots & \mathbf{y}_{i+j-1} \\ \vdots & \ddots & \vdots \\ \mathbf{y}_{2i-1} & \cdots & \mathbf{y}_{2i+j-2} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_{0|i-1} \\ \mathbf{Y}_{i|2i-1} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_p \\ \mathbf{Y}_f \end{bmatrix}
 \tag{15}$$

where the two partition sub-matrices refer to past $\mathbf{Y}_p = \mathbf{H}_{0|i-1}$ and future $\mathbf{Y}_f = \mathbf{H}_{i|2i-1}$ output channel matrices, where subscripts on the lefthand and righthand sides of delimiter | denote the first and the last element of the first column, respectively, of block Hankel

matrix $\mathbf{H}_{0|2i-1}$. So, matrices \mathbf{Y}_p and \mathbf{Y}_f are defined by splitting matrix $\mathbf{H}_{0|2i-1}$ in two equal parts of i block rows, where number of block rows i shall be determined in agreement with condition $m \cdot i \geq n$ (Peeters 2000), being m the number of output (acquisition) channels and n the so-called *system order* (i.e. the dimension of square matrix \mathbf{A} in the identification process, i.e. the rank of diagonal matrix Σ_1 defined below). Number of columns j of block Hankel matrix $\mathbf{H}_{0|2i-1}$ is usually taken as $j = N - 2i + 1$, which implies that all recorded data samples are used (Van Overschee and De Moor 1996). Thus, all following quantities showing subscript i refer to the assumed number of block rows. By observing Eq. (15), it is clear that the block Hankel matrix consists of the repetition of the same element in each anti-diagonal term.

The second computational step is based on the calculation of *projection matrix* $\mathbf{P}_i \in \mathcal{R}^{mi \times mi}$, i.e. the orthogonal projection of the row space of future output channels \mathbf{Y}_f into the row space of past output channels \mathbf{Y}_p , which can be expressed as (Van Overschee and De Moor 1996):

$$\mathbf{P}_i = \mathbf{Y}_f / \mathbf{Y}_p = \mathbf{Y}_f \mathbf{Y}_p^T (\mathbf{Y}_p \mathbf{Y}_p^T)^\dagger \mathbf{Y}_p = \mathbf{O}_i \hat{\mathbf{S}}_i \tag{16}$$

where symbol \dagger indicates Moore-Penrose pseudo-inverse, whilst the factorization of projection matrix \mathbf{P}_i into the product of *observability matrix* $\mathbf{O}_i \in \mathcal{R}^{mi \times n}$ and *Kalman filter state sequence* $\hat{\mathbf{S}}_i \in \mathcal{R}^{n \times mi}$ defines a main theorem of SSI-DATA (Van Overschee and De Moor 1996; Rainieri and Fabbrocino 2014), where n is the selected system order, as detailed in the following.

Then, through the application of specific weighting matrices \mathbf{W}_1 and \mathbf{W}_2 to projection matrix \mathbf{P}_i , a SVD may be derived, by holding non-zero singular values only (Van Overschee and De Moor 1996; Rainieri and Fabbrocino 2014), as:

$$\mathbf{W}_1 \mathbf{P}_i \mathbf{W}_2 = [\mathbf{U}_1 \mathbf{U}_2] \begin{bmatrix} \Sigma_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} = \mathbf{U}_1 \Sigma_1 \mathbf{V}_1^T \tag{17}$$

where \mathbf{U}_k and $\mathbf{V}_k \in \mathcal{R}^{mi \times n}$, $k = 1, 2$ are the singular vector matrices, and $\Sigma_1 \in \mathcal{R}^{n \times n}$ is the diagonal matrix holding the non-zero singular values (which allows to estimate the rank of matrix \mathbf{P}_i). The selection of the dimension (n) of square diagonal matrix Σ_1 fixes system order n of the State-Space model, which is adopted for the subsequent computational steps.

As concerning weighting matrices $\mathbf{W}_1 \in \mathcal{R}^{mi \times mi}$ and $\mathbf{W}_2 \in \mathcal{R}^{mi \times mi}$, they may be defined according to different weighting proposals (Van Overschee and De Moor 1996), namely Principal Component (PC), Unweighted Principal Component (UPC) and Canonical Variate Analysis (CVA). Accordingly, the corresponding weighting matrices can be defined as follows:

- for the Principal Component (PC) weighting:

$$\mathbf{W}_1 = \mathbf{I}, \quad \mathbf{W}_2 = \mathbf{Y}_p^T \left(\frac{1}{j} \mathbf{Y}_p \mathbf{Y}_p^T \right)^{-\frac{1}{2}} \mathbf{Y}_p; \tag{18}$$

- for the Unweighted Principal Component (UPC) weighting:

$$\mathbf{W}_1 = \mathbf{I}, \quad \mathbf{W}_2 = \mathbf{I}; \tag{19}$$

- for the Canonical Variate Analysis weighting (CVA) (of a main use in the following):

$$\mathbf{W}_1 = \left(\frac{1}{j} \mathbf{Y}_f \mathbf{Y}_f^T \right)^{-\frac{1}{2}}, \quad \mathbf{W}_2 = \mathbf{I}; \tag{20}$$

where j is again the number of columns of block Hankel matrix $\mathbf{H}_{0|2i-1}$ (taken as $j = N - 2i + 1$ in the present case, i.e. all recorded data samples are used).

Then, starting from Eq. (17), observability matrix \mathbf{O}_i and Kalman filter state sequence $\hat{\mathbf{S}}_i$ may be computed as:

$$\mathbf{O}_i = \mathbf{U}_1 \Sigma_1^{1/2} \mathbf{T}, \quad \hat{\mathbf{S}}_i = \mathbf{O}_i^\dagger \mathbf{P}_i \tag{21}$$

where $\mathbf{T} \in \mathcal{R}^{n \times n}$ is a further possible weighting matrix, generally taken as an identity matrix, $\mathbf{T} = \mathbf{I}$ (as done here). By taking into account Eqs. (16) and (21), Kalman filter state sequence $\hat{\mathbf{S}}_{i+1}$ and output sequence $\mathbf{Y}_{i|i}$ may be calculated as shown in Van Overschee and De Moor (1996). Especially, $\mathbf{Y}_{i|i}$ comes directly from a different partition of block Hankel matrix $\mathbf{H}_{0|2i-1}$:

$$\mathbf{H}_{0|2i-1} = \begin{bmatrix} \mathbf{Y}_{0|i-1} \\ \mathbf{Y}_{i|i} \\ \mathbf{Y}_{i+1|2i-1} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_{0|i} \\ \mathbf{Y}_{i+1|2i-1} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_p^+ \\ \mathbf{Y}_f^- \end{bmatrix} \tag{22}$$

where superscript symbols $+$ and $-$ stay for addition and for subtraction of one block row to the original \mathbf{Y}_p and \mathbf{Y}_f matrices. Then, from projection matrix \mathbf{P}_{i-1} it is possible to obtain:

$$\mathbf{P}_{i-1} = \mathbf{Y}_f^- / \mathbf{Y}_p^+ = \mathbf{Y}_f^- \mathbf{Y}_p^{+T} (\mathbf{Y}_p^+ \mathbf{Y}_p^{+T})^\dagger \mathbf{Y}_p^+ = \mathbf{O}_{i-1} \hat{\mathbf{S}}_{i+1} \tag{23}$$

and Kalman filter state sequence $\hat{\mathbf{S}}_{i+1}$ can be determined as:

$$\hat{\mathbf{S}}_{i+1} = \mathbf{O}_{i-1}^\dagger \mathbf{P}_{i-1} \tag{24}$$

At this stage, discrete-time State-Space matrices \mathbf{A} and \mathbf{C} may be computed through an asymptotically-unbiased least squares estimate as (Van Overschee and De Moor 1996):

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{S}}_{i+1} \\ \mathbf{Y}_{i|i} \end{bmatrix} \hat{\mathbf{S}}_i^\dagger. \tag{25}$$

Then, the eigenvalue decomposition of discrete-time state matrix $\mathbf{A} = \boldsymbol{\Psi} \boldsymbol{\mathcal{M}} \boldsymbol{\Psi}^{-1}$ allows for the estimation of the modal parameters, through matrices $\boldsymbol{\Psi}$ and $\boldsymbol{\mathcal{M}}$, holding discrete-time eigenvectors $\boldsymbol{\psi}_r$ and eigenvalues μ_r , respectively. Afterwards, discrete-time eigenvalues μ_r are converted to continuous-time eigenvalues λ_r , as $\lambda_r = \ln(\mu_r) / \Delta t$ (Rainieri and Fabbrocino 2014), so that the so-called *system poles* are obtained. Finally, the r th mode shape, natural frequency, damped modal frequency and modal damping ratio estimates may be computed as (Rainieri and Fabbrocino 2014):

$$\boldsymbol{\phi}_r = \mathbf{C} \boldsymbol{\psi}_r; \quad f_r = \frac{|\lambda_r|}{2\pi}, \quad f_{r,d} = \frac{\text{Im}(\lambda_r)}{2\pi}, \quad \zeta_r = -\frac{\text{Re}(\lambda_r)}{|\lambda_r|}. \tag{26}$$

3.2 Main enhancements of the present improved SSI-DATA algorithm

Main assumptions of classical SSI methods consist of (stationary) white noise input and adequately long structural response signals, to achieve a suitable *stabilization* of the estimated poles. Light damping (modal damping ratios in the order of $1 \div 2\%$) leads to better estimates, too, since it may reduce the occurrence of noise poles or of mathematical poles (e.g. false stable poles characterized by a positive real part and a negative damping ratio) (Van Overschee and De Moor 1996; Rainieri and Fabbrocino 2014).

The present SSI-DATA algorithm, whose original theoretical background comes from the general formulation in Van Overschee and De Moor (1996), as outlined in Sect. 3.1 above, may be intended as a first implementation attempt, for SSI algorithms, to deal with earthquake-induced structural response signals, at concurrent heavy structural damping in terms of modal identification challenge.

Therefore, the following fundamental items summarize the main steps and issues related to the present SSI-DATA implementation:

- A first issue is to appropriately define weighting matrices \mathbf{W}_1 and \mathbf{W}_2 . After first extensive simulations performed in Pansieri (2016), under white noise input or seismic excitation, it was shown that the Canonical Variate Analysis weighting (CVA) (Van Overschee and De Moor 1996), with weighting matrices \mathbf{W}_1 and \mathbf{W}_2 as given in Eq. (20), turns out to be the most stable and performing weighting option towards achieving reliable estimates at seismic response input and concurrent heavy damping (as demonstrated later in Sects. 4, 5). This type of weighting, as opposed to widely-used Principal Component (PC) [Eq. (18)] and Unweighted Principal Component (UPC) [Eq. (19)] weightings, returns even less noise or mathematical poles and looks mostly able to separate true physical modes from possible spurious earthquake harmonics.
- For the correct selection of system order n and for the determination of the stable poles (i.e. the poles where frequency, mode shape and modal damping ratio estimates show to be stable and not deriving from noise or mathematical poles), a *stabilization diagram* may be constructed from the SSI-DATA identification outcomes (Cara et al. 2013). It displays the poles that are obtained according to different considered system orders, as a function of the estimated frequency lines. The Singular Value (SV) curves extracted from SVD of SSI-DATA output spectral matrix $\mathbf{G}_{yy}(\omega)$ may be reported too, within the same stabilization diagram. This matrix may be typically calculated from the estimated SSI model as outlined in Peeters (2000), by adopting the estimated *next state-output covariance matrix* and the *output covariance matrix* (Rainieri and Fabbrocino 2014). As an original alternative in the present work, a novel integrated arrangement considers instead the use of SV curves, which are computed from the SVD of direct output spectral matrix $\mathbf{G}_{yy}(\omega)$. This is calculated through a routine of the previous rFDD algorithm, by adopting Welch's modified periodogram (Pioldi et al. 2017). Such proposed *integration of SSI and FDD information* demonstrates to provide a reliable tool to support the individuation of the stable SSI poles within the stabilization diagram, especially when dealing with earthquake-induced structural response signals and at concurrent heavy damping, as handled here.
- Anyway, the most severe issue in the present SSI identification keeps lying in the fact that seismic response signals are characterized by rather short durations (specifically with respect to ambient vibration recordings). This directly affects the achievable estimates, since the poles intrinsically display a harder stabilization. So, the maximum

system order employed in the analysis shall be incremented, jointly with a careful setting of the adopted number of block rows in the Hankel matrix. This has been specifically pursued in the present implementation (see also preliminary investigation results in Pansieri 2016). In this way, better natural frequency and mode shape estimates may be achieved. As regarding to the modal damping ratios, they appear to be the most challenging parameters to be detected by the present SSI identification, especially in relation to the cited very short durations of the seismic response signals.

4 Numerical attempts and identification outcomes

4.1 Adopted earthquake dataset and numerical models

The current output-only algorithms require as input channels the recorded structural responses of the considered building. In the present work, these responses are first obtained from *synthetic* signals calculated from a heavy-damped *linear three-storey frame*. Out of several preliminary MDOF trials in Pansieri (2016), this 3-DoF case has been selected since it already constitutes a good and challenging synthetic sample for the present identification methodologies within the seismic engineering range, providing a consistent estimation of all the modal characteristics of the three modes and allowing for a compact presentation of all the achieved results. Also, a coherence subsists with the forthcoming analysis of a real MDOF case that will be exposed in subsequent Sect. 5, which will target and report as well the successful identification of the first three modes of vibration of an existing monitored building.

In the identification analyses, total storey accelerations are considered as input data for the identification algorithms. These numerical response recordings are generated prior to the modal identification by taking as base acceleration an earthquake signal out of a set of *ten selected seismic ground motions* (see Table 1).

The ten adopted earthquake records have been chosen as representative ones from a variety of available seismic excitations, displaying rather different peculiar characteristics, i.e. time-frequency spectra [see e.g. further information reported in Pioldi et al. (2015b)], duration, sampling, epicentral distance, magnitude M and PGA. Also, they have been

Table 1 Main features of the adopted set of ten selected earthquake base excitations

Earthquake	Date	Station	Dur. (s)	f_s (Hz)	M	Comp.	PGA (g)
(AQ) L'Aquila	06/04/2009	AQV	100	200	5.8	WE	0.659
(CH) Maule	27/02/2010	Angle S/N 760	180	100	8.8	WE	0.697
(EC) El Centro	18/05/1940	0117	40	100	7.1	NS	0.312
(IV) Imperial Valley	15/10/1979	01260	58	100	6.4	NS	0.331
(KO) Kobe	17/01/1995	Nishi Akashi	41	100	6.9	WE	0.510
(LP) Loma Prieta	17/10/1989	47459	40	50	7.0	WE	0.359
(NO) Northridge	17/02/1994	24436	60	50	6.7	WE	1.778
(NZ) Christchurch	03/09/2010	163541	82	50	7.1	NS	0.752
(TA) Tabas	16/09/1978	70 Boshrooyeh	43	50	7.3	WE	0.929
(TO) Tohoku	11/03/2011	MYG004	300	100	9.0	NS	2.699

specifically selected as potential challenging instances for the present rFDD and SSI-DATA output-only identification purposes, given their strong non-stationary nature. Seismic signals have been imputed as is, as base excitation, and then considered as unknown in the subsequent output-only identification process.

The simulated structural responses are calculated by direct time integration of the equations of motion, via Newmark's (average acceleration) method. The use of synthetic signals shall fulfil a first necessary condition for the algorithms' effectiveness, since modal parameters are determined via direct modal analysis before identification and adopted as known targets for validation purposes.

The frame structure that has been adopted for initial verification is a reference three-storey shear-type frame, subjected to each single base excitation instance from the above adopted set of ten strong ground motions. This reference 3-DoF case has been characterized by a modal damping ratio $\zeta_k = 7\%$, for all the structural modes, a rather high value for the present OMA identification purposes, especially within the seismic engineering scenario. Structural and modal dynamic characteristics of the adopted three-storey frame (Pioldi et al. 2015a) are reported in Table 2.

4.2 Results from synthetic output-only modal dynamic identification

By taking as base excitation the single instances from the set of ten selected earthquake recordings presented in Sect. 4.1 (Table 1), separate dynamic identification analyses have been performed here with the present rFDD and SSI-DATA algorithms, in order to identify all strong ground motion modal parameters.

Both rFDD and SSI-DATA identification methods adopt constant time series lengths of 400 s and a 0.0025 Hz frequency resolution, by applying the method outlined in Pioldi et al. (2015a), which allows to increase the frequency resolution of the recordings, despite for the shortness of the seismic histories.

As concerning *rFDD*, the following parameters have been adopted for the performed analysis:

- Butterworth low-pass filtering, order 8, cut-off frequency 15 Hz, applied to the earthquake-induced structural response signals (input channels for the rFDD algorithm);

Table 2 Properties of the analyzed reference three-storey frame (Pioldi et al. 2015a)

	1	2	3
<i>Floor</i>			
Stiffness $\times 10^3$ (kN/m)	202.83	202.83	202.83
Mass $\times 10^3$ (kg)	144	144	144
<i>Mode</i>			
Natural frequency (Hz)	2.658	7.448	10.76
Mode shape [1]	$\begin{Bmatrix} 0.328 \\ 0.591 \\ 0.737 \end{Bmatrix}$	$\begin{Bmatrix} -0.737 \\ -0.328 \\ 0.591 \end{Bmatrix}$	$\begin{Bmatrix} 0.591 \\ -0.737 \\ 0.328 \end{Bmatrix}$
Assumed modal damping ratio	7%	7%	7%

- Different decimation settings, as a function of the frequency sampling of the recordings, i.e. no decimation (50 Hz signals), decimation of order 2 (100 Hz signals) and of order 5 (200 Hz signals);
- Integrated PSD matrix computation through both Welch’s Modified Periodogram, generally set with 1024-points Hanning smoothing windows and 66.7% overlapping (2048-points Hanning smoothing windows when no decimation is applied), and Wiener-Khinchin’s method, set with a de-trended biased correlation matrix (see Pioldi et al. 2015b for more details).

Then, as concerning SSI-DATA, the performed analysis adopted the following settings:

- Butterworth low-pass filtering, order 8, cut-off frequency 25 Hz, no decimation of the signals;
- Block Hankel matrix parameters: number of block rows set to $i = 50$ (in general, adopted as variable in the range $30 \leq i \leq 80$), number of columns $j = N - 2i + 1$, for all the analyzed cases;
- Stabilization diagram parameters: maximum order $n = 150$ (in general, adopted as $n = m \cdot i$, as a function of number of response channels m and number of block rows i), detection of stable poles with tolerance levels set to $\Delta f_k = |(f_k - f_{k+1})|/f_k < 0.01$, $\Delta \zeta_k = |(\zeta_k - \zeta_{k+1})|/\zeta_k < 0.075$ and $\Delta MAC_k = 1 - MAC(\phi_k, \phi_{k+1}) < 0.02$ for frequencies, modal damping ratios and MACs, respectively, being k the current model order. Notice that higher tolerance value 7.5% on the damping ratios has been set based on the experience that has been gained by running the various cases, pointing out (expected) higher dispersion on the damping ratio estimations, basically as the lowest that has anyway allowed to produce the scores in the diagrams that will be reported in the paper;
- Stabilization diagram SVs computed through SVD of the PSD matrix as estimated via Welch’s Modified Periodogram, set with 2048-points Hanning smoothing windows and 66.7% overlapping.

A sample of the outcomes from the two rFDD and SSI-DATA algorithms, operating here at seismic response input and concurrent heavy damping, is represented in Fig. 1. In this representation, the Singular Value Decomposition with peak-picking of the modes of vibration (left) and the enhanced *Global stabilization diagram* with the first rFDD Singular Value, representing rFDD and SSI-DATA all together (right), respectively, are depicted.

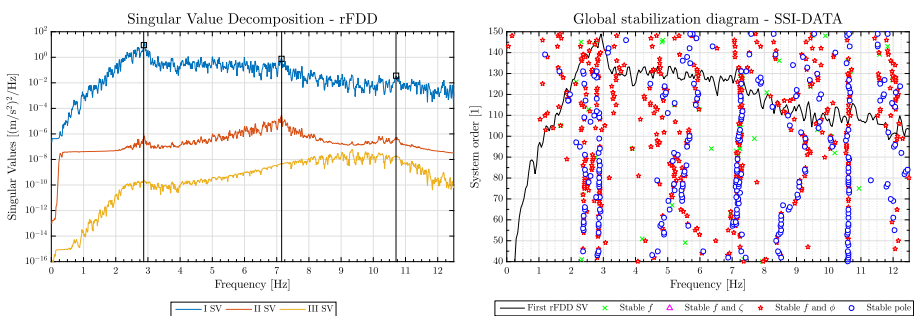


Fig. 1 Singular Value Decomposition and peak-picking of the modes of vibration (rFDD) and enhanced Global Stabilization Diagram with 1st rFDD SV (SSI-DATA), three-storey frame, synthetic response signals, rFDD (left) and SSI-DATA (right) algorithms, earthquake of Maule (CH) (Table 1)

The Global stabilization diagram displays the stable poles, at increasing system order n , concerning frequency, damping ratio and mode shape. Stable poles are selected as poles with a stable frequency, damping ratio and mode shape, between two consecutive system orders (they are marked with blue circles in Fig. 1).

The case of Maule (CH) excitation has been selected here because it constitutes a non-stationary case characterized by several earthquake harmonics in the structural frequency range of interest, which makes modal identification rather challenging. Anyway, similar features have been displayed also by the other SVs curves and stabilization diagrams obtained from the other seismic excitations.

As concerning rFDD outcomes, the three modal peaks are detectable from the frequency lines on the first SV. Also, modal peaks are repeated on the remaining SVs, which is a clear index of existence of the related mode of vibration. Anyway, MAC and MPC indexes may be simultaneously adopted towards modal peak selection and validation, as extensively outlined in Pioldi et al. (2015b, 2017).

Then, regarding SSI-DATA, the achieved enhanced Global Stabilization Diagram, coupled with the representation of the 1st rFDD SV (see Sect. 3.2), shows three clear lines of stabilization of the poles. These lines may be discerned from the remaining poles, which are spurious poles coming from the earthquake harmonics or from numerical or mathematical poles. For example, the left line of stabilization for the first modal frequency contains several noise and mathematical poles, with poor consistency also by adopting different and combined MAC indexes, while the right one clearly defines better poles, and constitutes the line that is chosen for the identification purposes. Then, the use of the rFDD SV considerably helps in the selection of the correct stable poles, since it provides a better indication with respect to classically adopted PSDs within the stabilization diagrams. For example, in the right plot in Fig. 1 there appears a bifurcation of stable pole lines for the first mode of vibration, which can be anyway discerned by reading as well the underlying SV representation.

Thus, by the strategies and settings reported above, complete synthetic output-only analyses with the two OMA methods have been performed. A synopsis from all the achieved results is reported in Fig. 2, where the estimates in terms of absolute deviations of rFDD and SSI-DATA identified natural frequencies and modal damping ratios and achieved MAC indexes for the estimated mode shapes are depicted. Estimates are reported in terms of absolute deviations of estimated natural frequencies ($\Delta f = |(f_{est} - f_{targ})/f_{targ}|$) and modal damping ratios ($\Delta \zeta = |(\zeta_{est} - \zeta_{targ})/\zeta_{targ}|$) from the target parameters, and of achieved MAC indexes for the estimated mode shapes on the target ones ($MAC = |\phi_{est}^H \phi_{targ}|^2 / (|\phi_{est}^H \phi_{est}| |\phi_{targ}^H \phi_{targ}|)$).

The rFDD estimated frequencies show deviations that are always below 5%, except for the last modes of the NO, LP and TA earthquake cases (Table 1), where deviations increase up to 9%. SSI-DATA estimated frequencies show more scattered deviations, with values up to 14%. The estimated modal damping ratios display very low deviations, at around 10%, for the rFDD algorithm. This is not true for the SSI-DATA algorithm, where modal damping ratios display discrepancies raising up to 70%. By this method, the order of magnitude is more or less caught, but deviations look much higher and often rather unacceptable in engineering terms. However, it should be recalled that really tough heavy-damping identification conditions have been considered here. Generally, the frequency and damping estimates from the rFDD cases show to be much closer to the target values than from the SSI ones. In this sense, *Frequency Domain rFDD appears superior to Time Domain SSI-DATA*, within the considered seismic engineering scenario at simultaneous

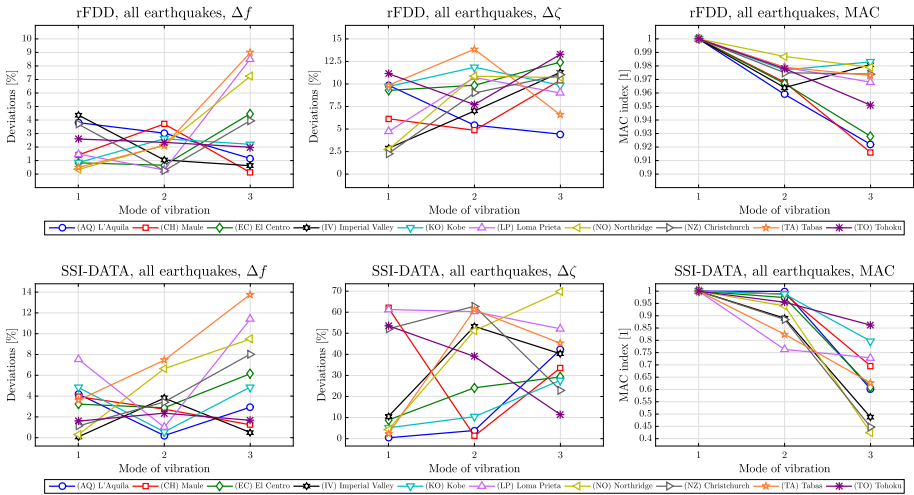


Fig. 2 Deviations of estimated natural frequencies (first column), modal damping ratios (second column) and MAC indexes (third column), three-storey frame, synthetic response signals, rFDD (first row) and SSI-DATA (second row) algorithms, complete considered earthquake dataset

heavy damping and according to the present implementations and their achieved level of refinement.

MAC values are always higher than 0.91 for the rFDD instances, for all the modes. With SSI-DATA, MAC indexes perform slightly less well. For the first two modes, MAC values are always higher than 0.75, with acceptable values in engineering terms. The third modes, instead, display some problems, especially with the IV, NO and NZ cases (Table 1), which return quite unreliable mode shapes. Thus, also in terms of mode shape estimates, rFDD performs better than SSI-DATA, in the present seismic and heavy damping context.

Then, global results on the achieved modal estimates are further summarized in Fig. 3, where the absolute deviations of estimated natural frequencies and modal damping ratios, and the MAC indexes are represented, in terms of suitably-designed dispersion diagrams (Pioldi and Rizzi 2017). The synthetic estimates for the adopted three-storey frame have been condensed all together, by displaying the minimum, the mean and the maximum (absolute) deviations, in blue, black and red coloured lines and markers, respectively. Then, the normalized truncated Gaussian Probability Density Functions (PDF) related to the dispersion of the estimates have been depicted for each mode, jointly with an indication of standard deviation σ of the estimated values. In the present case, frequency and modal damping ratio deviations shall turn out strictly positive (since absolute percentage deviations are adopted), while MAC indexes shall vary between 0 and 1. By taking into account such boundaries, truncated Gaussians are fitted on the achieved estimates, for each examined case.

These truncated Gaussians represent the probability of appearance of a certain deviation, as associated to each estimate, between the minimum and the maximum value, and are centred on the mean value. As it is possible to be appreciated, the maximum deviations are always on the Gaussian tails, while the minimum deviations lay in the Gaussian center. This confirms the goodness of the achieved modal identification estimates, especially for the rFDD outcomes.

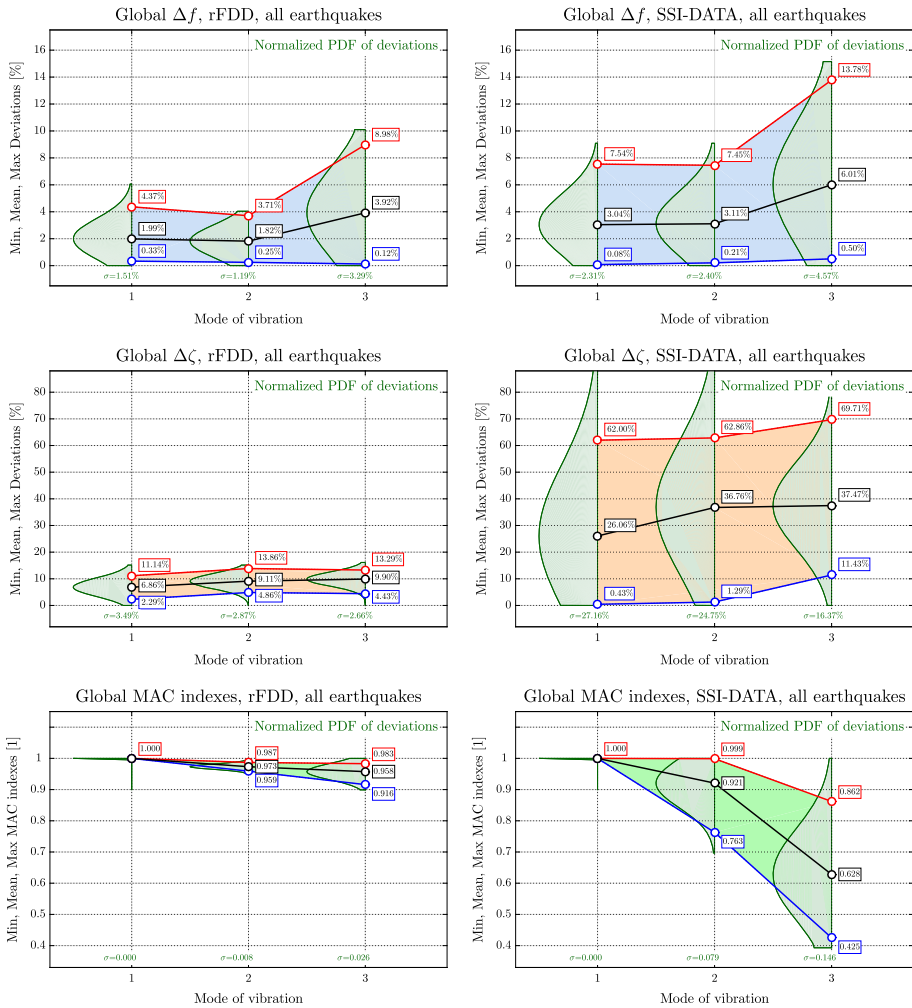


Fig. 3 Dispersion diagrams of deviations of estimated natural frequencies (first row), modal damping ratios (second row) and MAC indexes (third row), three-storey frame, synthetic response signals, rFDD (first column) and SSI-DATA (second column) algorithms, complete considered earthquake dataset. Minimum, mean and maximum values, and standard deviations are indicated

Finally, the achieved results are also reported in statistical form in Fig. 4, through appropriate *boxplots*. In these representations, each boxplot relates to the natural frequencies, modal damping ratios and MAC values estimated from the performed analyses. In each boxplot the inner rectangular box represents the central 50% of the identified parameters, while the centred line indicates their median. Then, the right and left boundary segments depict the 25% and 75% quantiles of the related statistical distributions. Finally, the vertical through-plot dashed green lines mark the known targeted modal parameters for the identification procedure.

Presented boxplots confirm again the goodness of the achieved results, especially as concerning the rFDD outcomes. Natural frequency estimates show to substantially catch the expected target values. Also MAC indexes reveal very good mode shape estimates, also

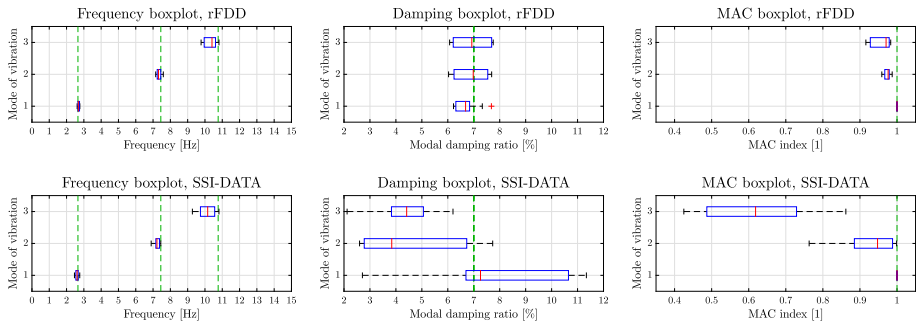


Fig. 4 Boxplot diagrams for estimated natural frequencies, modal damping ratios and MAC values, three-storey frame, synthetic response signals, rFDD (first row) and SSI-DATA (second row) algorithms, complete considered earthquake dataset

for the last modes of vibration, where SSI-DATA returns less accurate results, though still acceptable in engineering terms. Finally, modal damping ratios display very good estimates for the rFDD outcomes, by showing a rather contained dispersion, while for the SSI results some troubles appear, especially on the second and third modes of vibration.

The effect of artificial noise of a controlled amount added to the synthetic response signals does not impede the output-only identification analysis, up to noise levels that could be associated to those typical of practical instrumentations, as extensively demonstrated in Pioldi et al. (2015a, b). This is also testified by the effective processing of real signals, likely endowed with a certain amount of inherent noise, which is presented next, for final validation (see also Pioldi et al. 2017).

5 rFDD and improved SSI-DATA analyses with real earthquake-induced structural response signals

Next to the numerical analysis earlier reported in Sect. 4, where synthetic response signals have been adopted first, as a necessary validation condition, the present rFDD and improved SSI-DATA OMA approaches are now applied to *real* earthquake-induced structural response recordings. The selected building is the San Bruno six-storey office building (in short SBOB), California (Fig. 5). Data are taken from the Center of Engineering Strong Motion Data (CESMD) online database (CESMD Database 2016), and represent one of the likely few well-documented and available cases, already studied in the dedicated EMA literature (Marshall et al. 1992; Celebi 1996) (notice that here OMA is instead attempted).

This six-storey building is constituted by RC moment resisting frames, with individual spread footings, located in the city of San Bruno, in the metropolitan area of San Francisco. The design was made in 1978 and displays a plan of 60.96 m \times 27.43 m, and a height of 23.77 m. More information on the building, its history and characteristics may be found in Marshall et al. (1992) and Celebi (1996).

Thanks to the California Strong Motion Instrumentation Program, the building was instrumented in 1985 (CSMIP Station n. 58490). The recording system consisted of 13 accelerometers, on four levels of the building: four channels were devoted for the NS direction, eight channels were devised to the WE direction and one channel was setup for

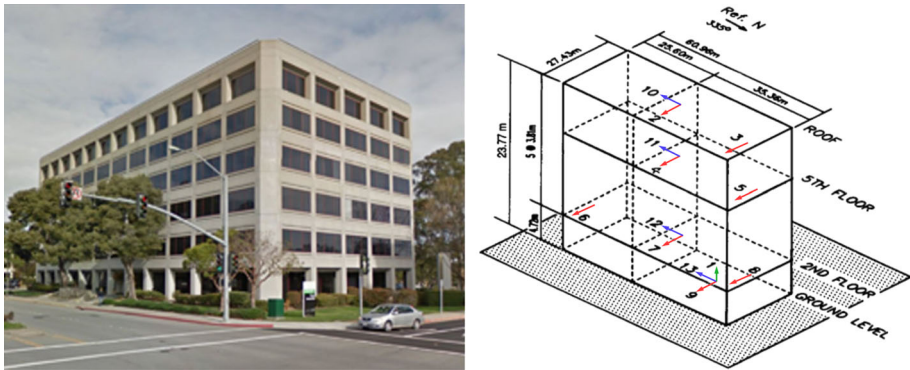


Fig. 5 External view of the San Bruno six-storey office building (SBOB), California, and three-dimensional sensor layout (adapted from Marshall et al. 1992). Input data are taken from the CESMD database (storey acceleration responses)

the UP direction. Fig. 5 briefly represents the overall building dimensions and the instrumentation layout.

In this work, the adopted SBOB seismic response data refer to the local earthquake excitation of Loma Prieta (1989), already considered in the previous synthetic analysis (see Table 1), now with main local characteristics as recorded at the building site and reported in Table 3. Recorded data belong to the total accelerations of the seven WE channels and of the three NS channels (the channels at the ground floor have not been considered, both for the OMA purposes of the present analysis and for their low signal-to-noise ratio).

The developed rFDD and SSI-DATA algorithms have been adopted to analyse the real earthquake-induced structural response data coming from the local Loma Prieta earthquake excitation. Sample of the outcomes from the two identification techniques are represented in Fig. 6. In this figure, the Singular Value Decomposition with modal peak-picking and the enhanced Global Stabilization Diagram with the first rFDD Singular Value, for the NS response component, are depicted. Despite for the real seismic responses, the *first three modes* of vibration can still be detected from the graphs.

Then, main results, in terms of estimated natural frequencies and modal damping ratios, are reported in Table 4. The analyses and the resulting modal parameter estimates have been subdivided into the two spatial components of the excitation, namely the NS and the WE responses. The first three modes of vibration can be identified as well, and deviations between the two methods are reported, as $\Delta f = [(f_{SSI} - f_{rFDD})/f_{rFDD}]%$ and $\Delta \zeta = [(\zeta_{SSI} - \zeta_{rFDD})/\zeta_{rFDD}]%$. These deviations show to be very limited concerning the natural frequencies, with a maximum of 8.14% on the third natural frequency of the NS case. Then, also the modal damping ratios reveal rather reasonable deviations, with a maximum of about 30% for the first mode of the NS cases. The only exceptions are related to the second modes of vibration, where deviations increase, for both NS and WE

Table 3 Main local characteristics and properties of the Loma Prieta earthquake (CESMD database). PGA and PSA values are referred to as the maximum between NS and WE components

Adopted earthquake	Date	Dur. (s)	f_s (Hz)	M	Dist. (km)	PGA (g)	PSA (g)
Loma Prieta (LP)	17/10/1989	60	50	7.0	81.0	0.14	0.32

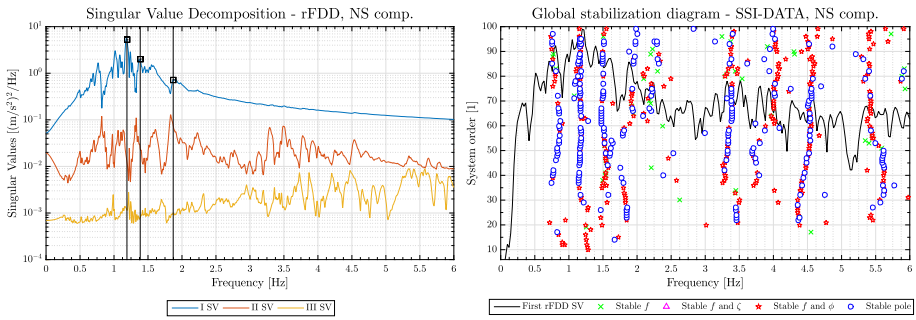


Fig. 6 Singular Value Decomposition and peak-picking of the modes of vibration (rFDD) and enhanced Global Stabilization Diagram with 1st rFDD SV (SSI-DATA), San Bruno six-storey office building, real response signals, rFDD (left) and SSI-DATA (right) algorithms, earthquake of Loma Prieta (LP), NS component (Table 3)

Table 4 Identified natural frequencies f_i , modal damping ratios ζ_i and related deviations, computed between rFDD and SSI-DATA estimates, San Bruno six-storey office building, real response signals, earthquake of Loma Prieta (LP), NS and WE components (Table 3)

	Mode	Loma Prieta (NS)			Loma Prieta (WE)		
		rFDD	SSI-DATA	Δ rFDD-SSI	rFDD	SSI-DATA	Δ rFDD-SSI
f_i (Hz)	1	1.185	1.187	0.133%	0.9957	0.9872	0.851%
	2	1.386	1.348	2.73%	1.286	1.267	1.52%
	3	1.873	1.720	8.14%	2.302	2.231	3.11%
ζ_i (%)	1	6.38%	8.37%	31.2%	3.99%	3.51%	– 12.2%
	2	2.54%	7.83%	208%	4.96%	11.2%	125%
	3	4.55%	5.50%	20.9%	6.06%	5.00%	– 17.5%

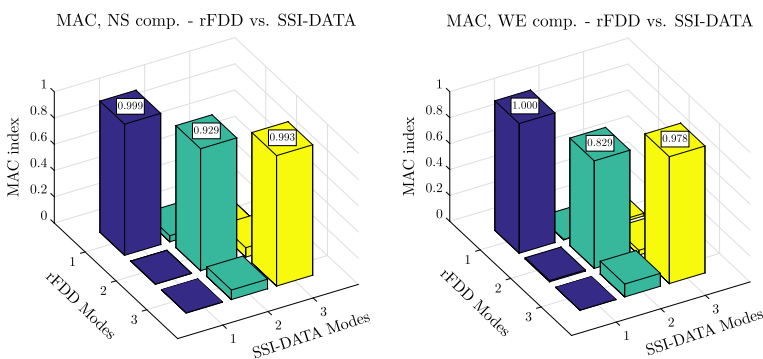


Fig. 7 MAC indexes for the estimated mode shapes, computed between rFDD and SSI-DATA eigenvector estimates, San Bruno six-storey office building, real response signals, earthquake of Loma Prieta (LP), NS and WE components (Table 3)

components. In these cases, SSI-DATA shows much higher modal damping ratios than for rFDD; this is probably due to the presence of *two very close modes* of vibration, i.e. the first and the second, which may negatively affects the damping estimates.

Mode shapes, instead, are reported in terms of MAC indexes in Fig. 7, i.e. by adopting $MAC(\phi_{rFDD}, \phi_{SSI}) = |\phi_{rFDD}^H \phi_{SSI}|^2 / (|\phi_{rFDD}^H \phi_{rFDD}| |\phi_{SSI}^H \phi_{SSI}|)$, as calculated between the rFDD and the SSI-DATA outcomes. Here, 3D MAC barplots are made by combining the rFDD and the SSI mode shapes, for each mode of vibration. On the diagonal terms, where the same modes are combined, MAC values turn out very close to one, as expected (since the first rFDD mode shape is combined with the first SSI mode shapes, and so on), while off-diagonal terms shall be close to zero, as effectively detected (since different modes are combined to each other, resulting to be orthogonal among them).

Notice that the direct comparison with the two rFDD and SSI-DATA methods (with respect to natural frequencies, modal damping ratios and mode shapes), after the earlier numerical analysis with synthetic response signals made in Sect. 4, helps with the validation of the achieved strong ground motion modal parameters. In fact, by adopting the two validated method, it is possible to achieve a set of rFDD and SSI-DATA modal parameters, which can be “self-compared” in order to reach final identification outcomes.

Finally, the barplots proposed in Fig. 8 represent the achieved natural frequencies and modal damping ratios, computed with the rFDD and SSI-DATA methods for the LP earthquake responses, NS and WE components. In particular, jointly with the *estimated* parameters, also the *target* ones are proposed, towards further comparison purposes. Target modal parameters are taken from the work of Marshall et al. (1992), Celebi (1996). For the NS case, the first three natural frequencies and the first modal damping ratio are available, while for the WE case the first two natural frequencies and the first modal damping ratio are provided. These target parameters are marked in red in Fig. 8, and deviations are calculated on them as $\Delta f = [(f_{ID} - f_{TARGET})/f_{TARGET}] \%$ for natural frequencies and $\Delta \zeta = [(\zeta_{ID} - \zeta_{TARGET})/\zeta_{TARGET}] \%$ for modal damping ratios. The very limited deviations between the estimated and the target values confirm once again the reliability of the proposed rFDD and SSI-DATA techniques, also in detecting strong ground motion modal parameters with real earthquake-induced structural response signals.

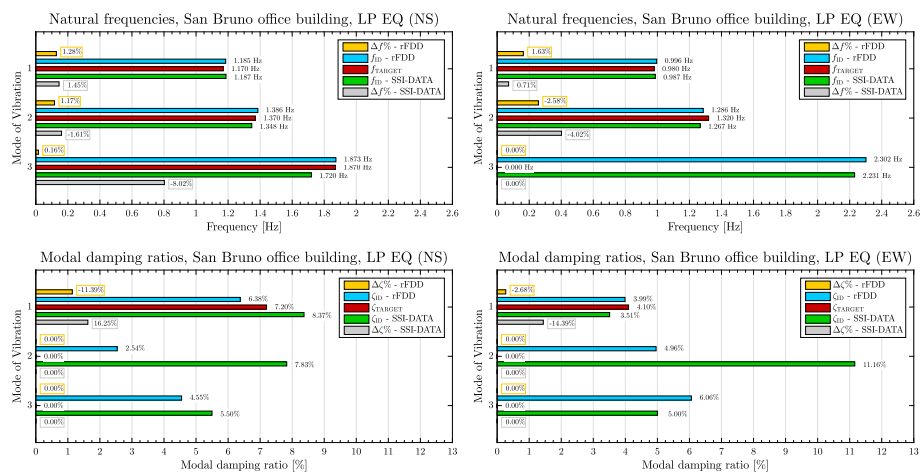


Fig. 8 Natural frequency and modal damping ratio barplots for rFDD and SSI-DATA, San Bruno six-storey office building, real response signals, earthquake of Loma Prieta (LP), NS and WE components (Table 3). Percentage deviations from the target values (where available) are indicated

6 Conclusions

In this work, two different output-only identification algorithms have been adopted to achieve strong ground motion modal parameter estimations. Simultaneous heavy damping, in terms of identification challenge, has been considered all together. In this framework, both *synthetic* (Sect. 4) and *real* (Sect. 5) seismic response signals have been effectively analyzed.

Consistently, two enhanced OMA methods have been originally developed and implemented within MATLAB, by referring either to the Frequency Domain (*refined Frequency Domain Decomposition, rFDD*) or to the Time Domain (*improved Data-Driven Stochastic Subspace Identification, SSI-DATA*). Starting from classical implementations (Sects. 2.1, 3.1), the present algorithms have been specifically implemented to deal with seismic vibration and simultaneous heavy damping. Specifically, a series of further peculiar strategies, described in Sects. 2.2 and 3.2, have been successfully implemented to handle such a challenging Earthquake Engineering identification context.

By adopting the two output-only methods, consistent estimates have been achieved from both *synthetic* (reference linear three-storey shear-type frame) and *real* (six-storey instrumented RC building in California) seismic response signals. The present rFDD and SSI-DATA methodologies show to be rather robust in terms of global modal identification capability, within the considered seismic engineering range. In these terms, readable estimates of all the modal parameters have been achieved.

The obtained results from the adoption of synthetic earthquake-induced response signals have been demonstrated and validated by the consistent outcomes from the use of real seismic response recordings. With the present enhanced output-only implementations (to be used separately or all together), the challenging modal identification within the Earthquake Engineering range may effectively be performed. As a specific identification outcome from the analyzed cases, *rFDD overall looks superior to SSI-DATA*, at this particular stage of development and implementation. In such a sense, better modal estimates may be achieved, especially for the modal damping ratios.

Future developments shall concern additional refinements of both methods, especially by focusing on the improvement of the present SSI-DATA implementation, and more extensive analyses and case studies.

Funding This study was funded by public research support from “*Fondi di Ricerca d’Ateneo ex 60%*” and a ministerial doctoral grant and funds at the ISA Doctoral School, University of Bergamo, Department of Engineering and Applied Sciences (Dalmine).

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

- Bendat J, Piersol A (1986) Random data, analysis and measurements procedures. Wiley, New York
- Bindi D, Petrovic B, Karapetrou S, Manakou M, Boxberger T, Raptakis D, Ptilakis KD, Parolai S (2015) Seismic response of an 8-story RC-building from ambient vibration analysis. Bull Earthq Eng 13:2095–2120
- Brincker R, Zhang L, Andersen P (2001) Modal identification of output-only systems using Frequency Domain Decomposition. Smart Mater Struct 10(3):441–445
- Cara FJ, Juan J, Alarcon E, Reynders E, De Roeck G (2013) Modal contribution and state space order selection in Operational Modal Analysis. Mech Syst Signal Process 38(2):276–298

- Celebi M (1996) Comparison of damping in buildings under low-amplitude and strong motions. *J Wind Eng Ind Aerodyn* 59(2):309–323
- CESMD Database (2016) Center for Engineering Strong Motion Data (CESMD), a cooperative database effort from U.S. Geological Survey (USGS), California Geological Survey (CGS) and Advanced National Seismic System (ANSS). <http://www.strongmotioncenter.org>. Accessed 30 June 2016
- Ghahari SF, Abazarsa F, Ghannad MA, Tacioglu E (2013) Response-only modal identification of structures using strong motion data. *Earthq Eng Struct Dyn* 42(11):1221–1242
- Gouache T, Morlier J, Michon G, Coulange B (2013), Operational Modal Analysis with non stationary inputs. In: Proceedings of the 5th international modal analysis conference (IOMAC-2013), vol 1, 13–15 May 2013, Guimaraes, Portugal, pp 1527–1534
- Heylen W, Lammens S, Sas P (2006) Modal analysis theory and testing. Katholieke Universteit Leuven, Departement Werktuigkunde, Leuven, Belgium
- Kun Z, Law S, Zhongdong D (2009) Condition assessment of structures under unknown support excitation. *Earthq Eng Eng Vib* 8(1):103–114
- Lin C, Hong L, Ueng Y, Wu K, Wang C (2005) Parametric identification of asymmetric buildings from earthquake response records. *Smart Mater Struct* 14(4):850–861
- Mahmoudabadi M, Ghafory-Ashiani M, Hosseini M (2007) Identification of modal parameters of non-classically damped linear structures under multi-component earthquake loading. *Earthq Eng Struct Dyn* 36(6):765–782
- Marshall RD, Phan LT, Celebi M (1992) Measurement of structural response characteristics of full-scale buildings: comparison of results from strong-motion and ambient vibration records. US Department of Commerce—National Institute of Standards and Technology, Report No. NISTIR 4884 ([online](#))
- Michel C, Guéguen P, El Arem S, Mazars J, Panagiotis K (2010) Full-scale dynamic response of an RC building under weak seismic motions using earthquake recordings, ambient vibrations and modelling. *Earthq Eng Struct Dyn* 39(4):419–441
- Ntotsios E, Karakostas C, Lekidis V, Panetsos P, Nikolaou I, Papadimitriou C, Salonikos T (2009) Structural identification of Egnatia Odos bridges based on ambient and earthquake induced vibrations. *Bull Earthq Eng* 7:485–501
- Pansieri S (2016) Time domain output-only modal dynamic Stochastic Subspace Identification algorithm for seismic applications, M.Sc. Thesis in Building Engineering, Advisor E. Rizzi, Co-advisor F. Pioldi, Università degli Studi di Bergamo, Scuola di Ingegneria, 29 March 2016, p 203
- Peeters B (2000) System identification and damage detection in civil engineering. Ph.D. Thesis, Katholieke Universiteit Leuven, Departement Burgerlijke Bouwkunde, Leuven, Belgium
- Peeters B, De Roeck G (1999) Reference-based stochastic subspace identification for output-only modal analysis. *Mech Syst Signal Process* 13(6):855–878
- Pioldi F, Rizzi E (2016a) A Full Dynamic Compound Inverse Method for output-only element-level system identification and input estimation from earthquake response signals. *Comput Mech* 58(2):307–327. <https://doi.org/10.1007/s00466-016-1292-0>
- Pioldi F, Rizzi E (2016b) Full Dynamic Compound Inverse method: extension to general and Rayleigh damping. *Comput Mech* 59(4):539–553. <https://doi.org/10.1007/s00466-016-1347-2>
- Pioldi F, Rizzi E (2017) A refined Frequency Domain Decomposition tool for structural modal monitoring in earthquake engineering. *Earthq Eng Eng Vib* 16(3):627–648. <https://doi.org/10.1007/s11803-017-0394-9>
- Pioldi F, Ferrari R, Rizzi E (2015a) Output-only modal dynamic identification of frames by a refined FDD algorithm at seismic input and high damping. *Mech Syst Signal Process* 68–69(February 2016):265–291. <https://doi.org/10.1016/j.ymssp.2015.07.004>
- Pioldi F, Ferrari R, Rizzi E (2015b) Earthquake structural modal estimates of multi-storey frames by a refined FDD algorithm. *J Vib Control* 23(13):2037–2063. <https://doi.org/10.1177/1077546315608557>
- Pioldi F, Salvi J, Rizzi E (2016) Refined FDD modal dynamic identification from earthquake responses with soil-structure interaction. *Int J Mech Sci* 127(July 2017):47–61. <https://doi.org/10.1016/j.ijmecsci.2016.10.032>
- Pioldi F, Ferrari R, Rizzi E (2017) Seismic FDD modal identification and monitoring of building properties from real strong-motion structural response signals. *Struct Control Health Monit* e1982:1–20. <https://doi.org/10.1002/stc.1982>
- Pridham B, Wilson J (2004) Identification of base-excited structures using output-only parameter estimation. *Earthq Eng Struct Dyn* 33(1):133–155
- Rainieri C, Fabbrocino G (2014) Operational Modal Analysis of civil engineering structures. An introduction and guide for applications. Springer, Berlin
- Reynders E (2012) System identification methods for (Operational) Modal Analysis: review and comparison. *Arch Comput Methods Eng* 19(1):51–124

- Sevim B, Atamturktur S, Altunisik AC, Bayraktar A (2016) Ambient vibration testing and seismic behavior of historical arch bridges under near and far fault ground motions. *Bull Earthq Eng* 14:241–259
- Steiger R, Feltrin G, Weber F, Nerbano S, Motavalli M (2016) Experimental modal analysis of a multi-storey light-frame timber building. *Bull Earthq Eng*. <https://doi.org/10.1007/s10518-015-9828-9>
- Van Overschee P, De Moor B (1996) Subspace identification for linear systems—theory, implementation, applications. Kluwer, London
- Ventura CE, Brincker R, Andersen P (2005) Dynamic properties of the Painter Street Overpass at different levels of vibration. In: Proceedings of the 6th international conference on structural dynamics (EURODYN), vol 1. Paris, France, pp 167–172
- Wang T, Zhang L, Tamura Y (2005) An operational modal analysis method in frequency and spatial domain. *Earthq Eng Eng Vib* 4(2):295–300