

Empirical scoring of ground motion prediction equations for probabilistic seismic hazard analysis in Italy including site effects

S. Barani¹ · D. Albarello² · D. Spallarossa¹ · M. Massa³

Received: 18 February 2016 / Accepted: 25 October 2016 / Published online: 2 November 2016
© Springer Science+Business Media Dordrecht 2016

Abstract In this study, we apply an empirical scoring method to evaluate the feasibility of probabilistic seismic hazard analyses at regional scale in Italy accounting for site amplification, which is taken into account through the application of a set of ground motion prediction equations (GMPEs) defined for specific ground types. Precisely, this method calculates the agreement (in terms of likelihood) between the hazard results computed using a specific hazard model and the number of ground-motion exceedances at a set of reference sites. Such a procedure is applied to quantify the likelihood of the outcomes of different hazard models, each based on a specific GMPE, with respect to the observations at 56 accelerometric sites operating in Italy for at least 25 years. Indirectly, this allows evaluating the influence of the selected GMPE in providing reliable hazard estimates. Seven possible GMPEs, applicable in active shallow crustal regions like Italy, have been examined taking into account the correlation among the hazard estimates provided by each computational model at the reference sites. Our results indicate that, although using GMPEs for pre-defined soil categories provides only a broad assessment of the hazard (since it ignores specific site response), large-scale hazard maps of Italy that are compatible with observations can be provided when suitable GMPEs are considered. A restricted number of GMPEs was found appropriate to this scope.

Keywords Probabilistic seismic hazard analysis · Ground motion prediction equation · Site effects · Statistical scoring

Electronic supplementary material The online version of this article (doi:10.1007/s10518-016-0040-3) contains supplementary material, which is available to authorized users.

✉ S. Barani
barani@dipteris.unige.it

¹ Dipartimento di Scienze della Terra dell’Ambiente e della Vita, Università degli Studi di Genova, Genova, Italy

² Dipartimento di Scienze Fisiche della Terra e dell’Ambiente, Università degli Studi di Siena, Siena, Italy

³ Istituto Nazionale di Geofisica e Vulcanologia, Milano, Italy

1 Introduction

Probabilistic seismic hazard (PSH) estimates represent a basic element for ruling effective earthquake-resistant design. Hazard maps are commonly developed at national scale and then implemented in seismic codes [e.g., “NEHRP Recommended Seismic Provisions” (Building Seismic Safety Council 2009), “Norme Tecniche per le Costruzioni” (Ministero delle Infrastrutture e dei Trasporti 2008)]. Due to their importance, testing effectiveness of such estimates is becoming increasingly important and several procedures have been proposed [Albarello and D’Amico (2015) and references therein]. The interest of the scientific community to this problem is also enhanced by the coexistence of several alternative seismic hazard maps provided by different computational models equally reliable (at least in principle). Therefore, selecting best performing maps for seismic regulations becomes a mandatory task for scientists and practitioners.

In principle, scoring hazard estimates can be performed by comparing them against observations available at accelerometric sites operating for quite long time intervals (e.g., Albarello and D’Amico 2008; Albarello et al. 2015a; Barani et al. 2016). A critical aspect of this kind of analysis is that, in conventional PSH models for national (or regional) assessments, the effect of local soil deposits is generally neglected. Hence, hazard maps for rock conditions define only a basic level for the definition of the expected ground motion at a site. On the other hand, observations available for scoring are generally collected at sites where the seismic bedrock is overlain by soil deposits and this makes any comparison between hazard estimates and observations quite problematic. To overcome this problem, one may correct observations to obtain the hypothetical ground motion on reference rock (e.g., $V_{s,30} > 800$ m/s) through the application of site (de-) amplification factors derived from seismic codes (Albarello and Peruzza 2016) or via deconvolution of the ground motion time histories recorded at each reference accelerometric station with the site transfer function (Castellaro and Albarello 2016). Alternatively, site effects may be directly included into the PSH models.

The scopes of the S2-2012 and S2-2014 DPC-INGV Projects are both probabilistic soil hazard assessments in Italy (Task 4 Working Group 2013; Barani and Spallarossa 2015a, b) and the evaluation of the consistency of the assessed hazard with observations by means of empirical scoring procedures and statistical testing methods (Albarello et al. 2013, 2015b). Besides the deliverables of the S2 DPC-INGV Projects (cited above), a comparison of probabilistic methods that incorporate site effects into ground motion hazard calculations can be found in the article of Barani and Spallarossa (2016). In principle, all PSHA methods that account for site-specific characteristics in a rigorous manner imply the determination of the ground response for a variety of target soil models representative of the actual local site conditions (e.g., Bazzurro and Cornell 2004a; Barani et al. 2013, 2014a, b). Therefore, the application of methods for site-specific PSHA to large areas requires an extensive knowledge of regional geology both to define representative soil models for the numerical ground response analyses and to associate the results from these analyses to areas with similar depositional history and properties. This kind of approach was applied by Pelli et al. (2006) and Haase et al. (2011) to compute the hazard in relatively small areas in northern Italy and Indiana (USA), respectively. Similarly, it could be applied in regions of the world where extensive seismic microzonation studies have been carried out, leading to large-scale evaluations of seismic amplification effects. However, the application of site-specific methods appears unfeasible (or, at least, unsuitable) for large-scale PSHAs. Attempts of large-scale hazard mapping inclusive of site

effects are those by Petersen et al. (1997), Romeo et al. (2000), Cramer (2006), and Kalkan et al. (2010). All these studies incorporate site effects by simply implementing ground motion prediction equations (GMPEs) defined for specific soil classes (few classes are generally considered), which in most cases correspond to those provided by national building codes. Basically, this approach assumes that the soil conditions at each node of the calculation grid resemble those at the stations in the database considered for the development of the GMPEs selected for the PSHA. It appears clear that this approach ignores detailed site-specific information and, therefore, produces only an approximate assessment of the local hazard. Nevertheless, it could be useful to verify the actual feasibility of this approach, at least to provide first-order hazard estimates to be used for regional-scale risk assessments.

In line with the scopes of the S2 DPC-INGV Projects, this study presents an attempt to evaluate the feasibility of the latter approach by evaluating the agreement (in terms of likelihood) between the outcomes from different hazard models, each one based on a specific GMPE, and the observations at a set of reference accelerometric sites. The effectiveness of different hazard models corresponding to seven alternative GMPEs suitable for PSHA in Italy (Table 1) has been examined via a scoring test. Indirectly, this analysis allows evaluating the performance of the selected GMPEs. As known, GMPEs play a crucial role on the hazard results, particularly due to the impact of the aleatory variability in the ground motion prediction (e.g., Strasser et al. 2009; Barani et al. 2015). Hence, recent research related to PSHA has led to the release of an increasing number of models, with the result of improving the accuracy of the predictions but with no tangible reduction of the aleatory variability in ground motions (Barani et al. 2016).

The analysis presented here is similar, to some extent, to the one proposed in Barani et al. (2016). First, a PSHA is carried out by performing different computational runs sharing the same input models and parameters values (e.g., geometry of earthquake source, seismicity rates, b -value, M_{\max} value) except for the GMPEs, which are applied one at a time separately. Then, scores are assigned to each hazard model by applying a probabilistic test, which calculates the likelihood of the outcomes from each model in relation to available observations at a set of 56 recording sites operating in Italy for at least 25 years (Table 2; Fig. 1). The analysis has been limited to these sites as they present a V_S profile down to 30 m. Differently from Barani et al. (2016), the scoring test adopted here is carried out taking into account the correlation among the hazard estimates at the reference recording sites (Albarello and Peruzza 2016). This approach, repeated for each one of the GMPEs considered, allows us to compare the performance of each model.

2 Empirical scoring procedure

The approach applied here is the same as the one presented in the article of Albarello and Peruzza (2016). It consists of comparing the outcomes of a PSH model against observations at a set of S recording sites, each one in operation for a Δt_s period of time. To this end, the time span covered by both should be the same. Commonly, the outcome of a PSHA at a given site is expressed in terms of ground motion value y^* with a certain probability P (e.g., 10%) of exceedance in a specified time period $\Delta t \neq \Delta t_s$ (e.g., $\Delta t = 50$ years). Assuming that the seismic process underlying the hazard assessment is Poissonian (as in this study), the mean annual rate of exceeding y^* is given by (e.g., Kramer 1996):

Table 1 List of GMPEs considered in the present study

GMPE	No. of events	No. of recordings	Magnitude type and range	Distance type and range (km)	Ground motion parameters	Definition of H component	Definition of rock site (m/s)	Region	References
ITA08	107	561	M_w 4.0–6.9	$R_{epi} \leq 100$	PGA, PGV, SA up to 2 s	ENV_{xy}	$V_{S,30} > 800$	Italy	Bindi et al. (2010)
AB10	131	532	M_w 5.0–7.6	$R_{jb} \leq 100$	PGA, PGV, PSA up to 3 s	GM_{xy}	$V_{S,30} > 750$	Europe	Akkar and Bommer (2010)
ITA10	218	1213	M_w 4.0–6.9	$R_{jb} \leq 200$	PGA, PGV, SA up to 2 s	GM_{xy}	$V_{S,30} > 800$	Italy	Bindi et al. (2011)
AKK14	221	1041	M_w 4.0–7.6	$R_{hyp} \leq 200$	PGA, PGV, PSA up to 4 s	GM_{xy}	$V_{S,30} > 750$	Europe	Akkar et al. (2014a, b)
BND14	365	2126	M_w 4.0–7.6	$R_{hyp} \leq 300$	PGA, PGV, PSA up to 3 s	GM_{xy}	$V_{S,30} > 800$	Europe	Bindi et al. (2014a, b)
BOR14	350	15,000	M_w 3.0–8.5	$R_{jb} \leq 400$	PGA, PGV, PSA up to 10 s	RotD50	$V_{S,30} > 760$	Global	Boore et al. (2014)
CZZ14	98	1880	M_w 4.5–7.9	$R_{rup} \leq 150$	PGA, PGV, DRS up to 10 s	GM_{xy}	$V_{S,30} > 800$	Global	Cauzzi et al. (2015)

The “distance type” field indicates the distance metrics used in the computations

R_{epi} : epicentral distance, R_{jb} : shortest distance from a site to the surface projection of the rupture surface (Joyner and Boore, 1981), R_{hyp} : hypocentral distance, R_{rup} : closest distance from a site to the ruptured fault, ENV_{xy} : larger horizontal component of the ground motion, GM_{xy} : geometric mean of the ground motion relative to the two horizontal components, RotD50: median single-component horizontal ground-motion across all non-redundant azimuths

SA spectral acceleration, PSA pseudo-spectral acceleration, PSV pseudo-spectral velocity, DRS displacement response spectrum

Table 2 List of accelerometric sites considered in the scoring test

Site	Station code	T_s (years)	Site category						
			ITA 08	AB 10	ITA 10	AKK 14 ^b	BND14 ^a	CZZ14 ^a	BOR 14 ^b
Alfonsine	ALF	38	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Ariano Irpino	ARI	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Asiago	ASG	37	R	R	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Auletta	ALT	39	R	R	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Bagnoli Irpino	BGI	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Barcis	BRC	41	R	R	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Benevento	BNV	36	ShA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Bersezio	BRZ	38	R	R	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Bevagna	BVG	39	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Bovino	BVN	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Brasimone	BRM	38	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Brienza	BRN	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Cagli	CGL	35	R	R	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Calitri	CLT	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Cascia	CSC	39	ShA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Castel Viscardo	CSD	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Castelfranco 5	CST	39	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Conegliano 5	CNG	41	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Cortina d'Ampezzo	CRD	39	R	R	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Demonte	DMN	38	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Fornovo	FRN	38	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Gildone	GLD	37	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Gioia Sannitica	GSN	32	ShA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Grumento Nova	GRM	24	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Lauria	LRS	39	R	R	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Maiano	MAI	36	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Malcesine	MLC	41	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Matelica	MTL	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Mazara del Vallo	MZR	37	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Meldola	MLD	39	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Mercato San Severino	MRT	38	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Mirandola	MRN	38	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Monselice	MNS	41	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Naso	NAS	37	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Nicosia	NCS	37	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Nizza Monferrato	NZZ	38	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$

Table 2 continued

Site	Station code	T_s (years)	Site category						
			ITA 08	AB 10	ITA 10	AKK 14 ^b	BND14 ^a	CZZ14 ^a	BOR 14 ^b
Nocera Umbra	NCR	39	ShA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Novellara	NVL	38	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Peglio	PGL	25	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Pennabilli	PNN	31	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Pinerolo	PNR	38	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Rincine	RNC	38	R	R	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Roccamonfina	RCC	39	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
S. Giorgio la Molara	SGR	39	ShA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
S. Sofia	STS	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
San Severo	SSV	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Sannicandro Garganico	SNN	38	R	R	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Senigallia	SNG	38	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Sirolo	SRL	39	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Sortino	SRT	37	R	R	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Sturmo	STR	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Tolmezzo 1	TLM1	40	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Torre del Greco	TDG	34	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Tortona	TRT	38	ShA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Tricarico	TRR	39	DA	SA	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$
Villa San Giovanni-1	VLS2	39	DA	SS	EC8	$V_{S,30}$	EC8, $V_{S,30}$	EC8, $V_{S,30}$	$V_{S,30}$

Ground types are assigned following the classification scheme proposed by the Eurocode 8—EC8 (Comité Européen de Normalisation—CEN 2004) or, for those GMPEs using an alternative definition, the scheme adopted by the GMPE authors. Up-to-date $V_{S,30}$ values and site classes can be found in Felicetta et al. (2016) and ITACA Working Group (2016)

R rock, *DA* deep alluvium (i.e., thicker than 20 m), *ShA* shallow alluvium (i.e., thinner than 20 m), *SA* stiff alluvium (i.e., $V_{S,30} > 750$ m/s), *SS* soft soil (i.e., $V_{S,30} < 360$ m/s)

^a BND14 and CZZ14 allow for soil effects by means of site factors based on EC8 ground types or, alternatively, through a $V_{S,30}$ -dependent site coefficient

^b AKK14 and BOR14 allow for soil effects by means of a $V_{S,30}$ -dependent coefficient

$$\lambda_{y^*} = - \frac{\ln(1 - P[Y > y^*])}{\Delta t} \tag{1}$$

The reciprocal of λ_{y^*} is the mean return period (MRP, hereinafter).

Hence, given the *i*-th PSH model (H_i) and the resulting ground motion value y^* associated with a given MRP at the *s*-th site, the probability $P_{s,i}$ of exceeding y^* during an exposure time Δt_s is:

$$P_{s,i} = P_{s,i}[Y > y^*] = 1 - e^{-\lambda_{y^*} \Delta t_s} \tag{2}$$

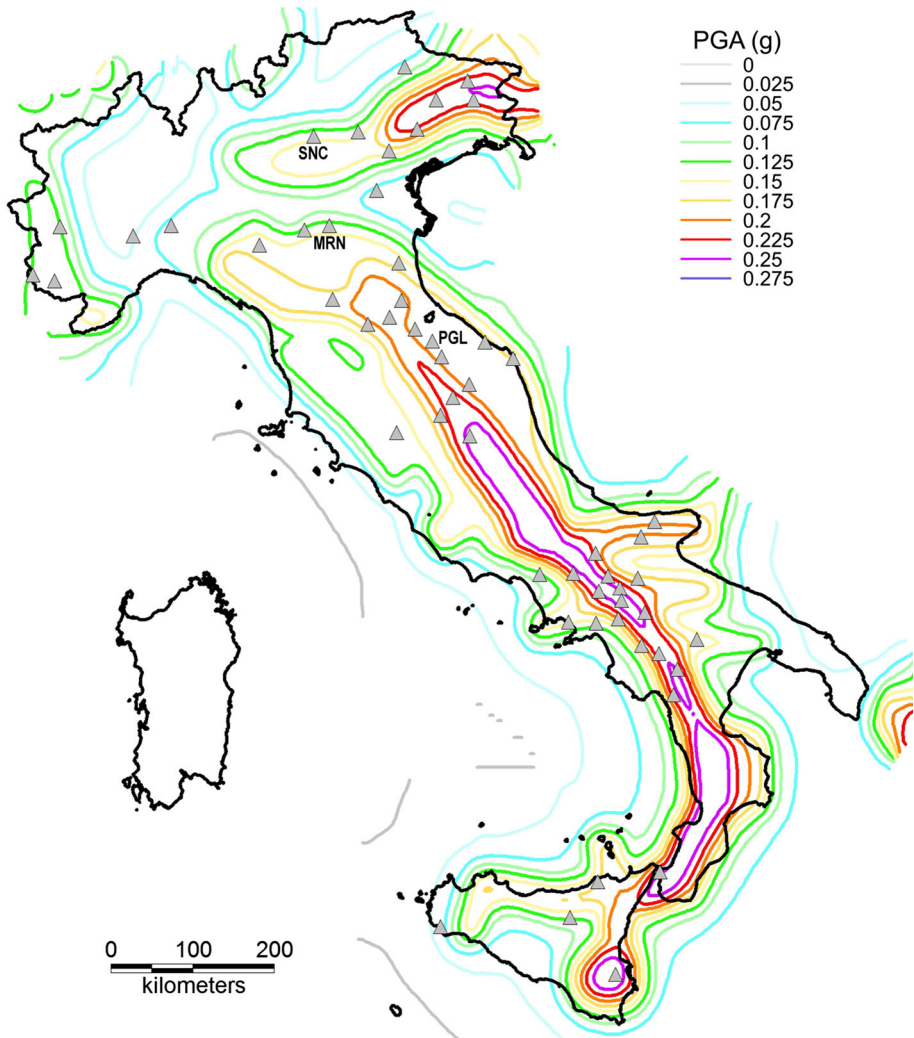


Fig. 1 Distributions of the accelerometric sites considered in the scoring test. *Contour lines* of the Italian PGA hazard map for a mean return period of 475 years (MPS Working Group 2004) are also displayed. *MRN* (Mirandola), *SNC* (Soncino), and *PGL* (Peglio) indicate the sites where we calculated the uniform hazard spectra shown in Fig. 6. *SNC* is not included in the list of stations used for the scoring test

Since each PSH model provides different values of y^* , each one corresponding to a particular MRP, models can be scored by considering each realization as an independent forecast. Since lower exceedance probabilities correspond to longer MRPs and to higher y^* values, different scores can be obtained for different points along a hazard curve.

Given a model H_i , scores are computed by comparing the number N_i^* of sites where the threshold y^* is exceeded (i.e., the observed ground motion $g_s \geq y^*$), which can be simply determined by counting, with the number $\mu(N_i)$ of exceedances expected on the basis of that hazard model. If S is large enough, according to the central limit theorem, numerical simulations show that N_i is a normal random variate with mean $\mu(N_i)$ and standard deviation $\sigma(N_i)$ (Albarelli and Peruzza 2016). Specifically:

$$\mu(N_i) = \sum_{s=1}^S P_{s,i} \tag{3}$$

and

$$\sigma(N_i) = c_i \sqrt{\sum_{s=1}^S P_{s,i}(1 - P_{s,i})} \tag{4}$$

where c_i , which is defined in the next section, is a correction coefficient that accounts for the mutual correlation of the hazard estimates at the S sites; c_i is equal to 1 if model outcomes are mutually uncorrelated at all S sites and increases as the correlation among the hazard estimates increases.

Hence, the new random variate Z_i

$$Z_i = \frac{N_i^* - \mu(N_i)}{\sigma(N_i)} \tag{5}$$

follows the standardized normal distribution. This implies that the likelihood L_i of observing at least N_i^* ground motion exceedances given the i -th PSH model is:

$$L_i = L(N_i^* | H_i) = L(|Z_i|) = \frac{4}{\sqrt{\pi}} \int_{|Z_i|}^{\infty} e^{-t^2} dt \tag{6}$$

The likelihood L_i can be interpreted as the degree of belief in the hypothesis that the observed number of exceedances N_i^* is an outcome of the seismicity process described by the i -th model.

It is worth noting that the hypothesis about the earthquake occurrence process assumed in the hazard assessment does not necessarily apply to observations. The general structure of the test (Eqs. 5, 6) still holds in the case that a “non-Poissonian” earthquake occurrence model is adopted to compute the expected number of sites where exceedances occur.

To evaluate the degree of belief in a hazard model (i.e., the confidence in the model H_i given the observation of N_i^* exceedances), $Q(H_i | N_i^*)$, additional restrictive hypotheses are required. In the assumption that the set of n_H hazard models is exhaustive and that at least one of these models is representative of the actual seismicogenic process one can obtain, by the Bayes theorem:

$$Q_i = Q(H_i | N_i^*) = \frac{p(H_i)}{\sum_{j=1}^{n_H} p(H_j) L(N_i^* | H_j)} L(N_i^* | H_i) \tag{7}$$

where $p(H_i)$ is the ex-ante degree of belief in the i -th model (see Albarello and D’Amico 2015). The likelihood L_i and the confidence Q_i are inherently different. The former quantifies the “absolute” performances of the i -th model irrespective of the other models. Very low values of L_i indicate that model outcomes are apparently incompatible with the observations since the probability of observing a similar outcome is low. On the other hand, Q_i provides the relative “score” of the i -th model conditioned to the performances of the other competing models, assuming that at least one of them is representative of the underlying natural process. Thus, the likelihood L_i allows testing the feasibility of calculating Q_i in the sense that, if none of the considered models provide results compatible with

observations (e.g., $L_i < 5\%$ for all n_H models), Eq. 7 cannot be applied (for a detailed discussion, see Albarello and D’Amico 2015).

Equation 7 allows estimating the values of Q_i for one specific realization of the model H_i (e.g., a spectral acceleration value corresponding to a specific MRP). This implies that the observation N_i^* (i.e., the number of observed exceedances) relative to the i -th model depends on both the number of MRPs (n_{MRP}) and response periods (n_T). In principle, we can consider $M = n_{MRP} \cdot n_T$ combinations of MRPs and T values. For the m -th combination, we get the observation $N_{i,m}^*$. In the case of n_H hazard models, we get $M \cdot n_H$ values of Q_i where m and i are in the ranges $[1, M]$ and $[1, n_H]$, respectively. By iteratively applying Eq. 7, one can obtain the overall degree of belief associated with the i -th model (hereinafter called “score”) given that $N_{i,M}^*$ exceedances occur, $Q_i = Q(H_i|N_{i,M}^*)$:

$$\left\{ \begin{aligned} Q_i &= Q(H_i|N_{i,1}^*) = \frac{p(H_i)}{\sum_{j=1}^{n_H} p(H_j)L(N_{j,1}^*|H_j)} L(N_{i,1}^*|H_i) & m = 1 \\ Q_i &= Q(H_i|N_{i,m}^*) = \frac{Q(H_i|N_{i,m-1}^*)}{\sum_{j=1}^{n_H} Q(H_j|N_{j,m-1}^*)L(N_{j,m}^*|H_j)} L(N_{i,m}^*|H_i) & m = 2, 3, \dots, M \end{aligned} \right. \tag{8}$$

In this study, the values of Q_i are computed assuming that $p(H_i) = 1/n_H$.

3 Numerical simulations to estimate c_i values

In order to score the PSH models, the relevant c_i values must be determined in advance. These values depend on the relative positions of the S sites, of the seismogenic sources, and on the GMPEs considered in the i -th computational model. In the case of the models considered here, the position of the accelerometric sites, as well as that of the seismogenic sources, is kept fixed. Thus, differences among the c_i values are expected to depend on GMPEs.

For each GMPE considered, the value of c is computed following a procedure similar to the one proposed by Albarello and Peruzza (2016). In that study, the authors generate a specified number of virtual seismic catalogues by randomly picking a number of epicenters from a given earthquake catalogue. A random magnitude is then assigned to each event by assuming a specified recurrence relationship for the whole Italy. It is clear that this procedure does not account for the role of seismogenic areas (where the generation of earthquakes is assumed as equiprobable) in the PSH model.

To overcome this limitation, a different approach has been considered here. This approach is essentially a Monte Carlo simulation-based PSH assessment similar to the one proposed by Musson (2009). For each source area of the seismogenic zonation used in the PSHA of Italy (MPS Working Group 2004), we generate K ($=1000$) earthquake catalogues of Δt_k years ($=1000$ years). In each source area, the generation of earthquakes above a threshold magnitude m_{min} is assumed to be a random Poissonian process with a mean annual rate $v(m_{min})$. A truncated Gutenberg–Richter distribution with parameter b is assumed to characterize magnitudes in the range $[m_{min}, M_{max}]$. The values of $v(m_{min})$, b , m_{min} , and M_{max} for each source area are provided in Barani et al. (2009). To generate the

seismic catalogues for each single seismogenic area, we first compute incremental annual activity rates, $n(m)$, for each magnitude bin in the interval $m_{\min} - M_{\max}$ [bins with amplitude equal to 0.23 magnitude units are assumed according to the MPS Working Group (2004)]. For each bin, we randomly generate a number of events proportional to the activity rate associated with that bin (i.e., corresponding to the nearest integer of the product $n(m) \cdot \Delta t_k$). A random magnitude value m , compatible with that magnitude bin, is then attributed to each virtual event. The epicenter location of each event is determined by discretizing each seismogenic area into cells with spacing of 0.01° in both latitude and longitude: each earthquake is then attributed to the barycenter of a cell selected randomly (with uniform distribution) from the pool of cells falling in that area. Similarly, a random depth is assigned to each epicenter assuming that the seismogenic thickness of each source zone is drawn by a normal density with mean value and standard deviation equal to those assumed by the MPS Working Group (2004). The mechanism of faulting assigned to each event is the same as that of the pertinent source zone.

Once all K virtual catalogues are generated, a number of events compatible with a period of time $\overline{\Delta t_s}$ of interest (in this study, we assume $\overline{\Delta t_s} = 39$ years, which nearly corresponds to the average lifetime of the 56 reference recording stations since their installation) is randomly selected from each of them (this implies that K random extractions are performed). More precisely, for each source zone, $v(m_{\min}) \cdot \overline{\Delta t_s}$ events are selected from each one of the K virtual data sets. Alternatively, for each source area, one may select $v(m_{\min}) \cdot \overline{\Delta t_s}$ events from a single data set including $K \cdot \Delta t_k$ years of virtual data and repeats the random extraction K times. For each event extracted from the k -th catalogue, we determine the ground motion level at the s -th site (for the computation of c , a rock condition (i.e., $V_{S,30} > 800$ m/s) is assumed for all S sites) by randomly picking a ground motion value from the probability distribution $P[Y > y^* | m, r]$ associated with a specified GMPE (ground motion distributions are truncated at $\varepsilon = 3$ standard deviations), where Y indicates the median value of the ground motion predicted given m and r , and y^* is a certain ground motion threshold given in input. Note that r indicates either the epicentral distance or the hypocentral distance or the rupture distance, or the Joyner and Boore (1981) distance according to the GMPE used (see Table 1).

From here on, the procedure to the computation of c is analogous to that described in the article of Albarello and Peruzza (2016). Specifically, the value of c for the i -th model is computed as:

$$c_i = \frac{\sigma_{corr}}{\sigma_{uncorr}} = \frac{\sqrt{\sum_{n_S=0}^S [n_S - n_S p_i(n_S)]^2 p_i(n_S)}}{\sqrt{\sum_{s=1}^S P_{s,i}(1 - P_{s,i})}} \tag{9}$$

where σ_{corr} and σ_{uncorr} have the meaning of $\sigma(N_i)$ (Eq. 4) in the case of mutual correlation among the hazard estimates at the S benchmark sites and when no correlation exists (i.e., $c_i = 1$), n_S indicates the number of sites where the threshold y^* is jointly exceeded $D_{k,i}(n_S)$ times, $p_i(n_S)$ indicates the probability associated with n_S

$$p_i(n_S) = \frac{1}{K} \sum_{k=1}^K D_{k,i}(n_S) \tag{10}$$

and $P_{s,i}$ indicates again the probability of exceeding y^* at the s -th site

$$P_{s,i} = \frac{1}{K} \sum_{k=1}^K E_{s,k} \tag{11}$$

where $E_{s,k} = 1$ when y^* is exceeded (at least once) during the k -th run, otherwise $E_{s,k} = 0$.

In this study, the procedure is repeated for different ground motion thresholds y^* (i.e., c is computed for each ground motion value y^* specified in input) of the following ground motion parameters: peak ground acceleration (PGA), 0.15, 1, and 2 s 5%-damped spectral acceleration [$S_a(T)$]. Note that, for each value y^* , we compute the average frequency (probability) of exceedance during $\overline{\Delta t}_s$ years over the S sites. Assuming a Poisson process, the average annual rate of exceeding y^* can be therefore determined. Thus, one can compute the value of c corresponding to a particular mean return period (MRP).

As shown in Fig. 2, c varies not only as a function of the GMPE but also as a function of the ground motion level y^* and spectral period. In particular, c tends to increase with increasing spectral period. At short periods [i.e., PGA and $S_a(0.15\text{ s})$], c varies around unity for both very low and high ground motion levels while assumes the greatest values for moderate ground motions [i.e., $0.04\text{ g} < \text{PGA} < 0.15\text{ g}$ and $0.1\text{ g} < S_a(0.15\text{ s}) < 0.3\text{ g}$]. At longer periods [i.e., $S_a(1\text{ s})$ and $S_a(2\text{ s})$], c is greater for moderate to low acceleration values [i.e., $0.03\text{ g} < S_a(1\text{ s}) < 0.08\text{ g}$ and $S_a(2\text{ s}) < 0.03\text{ g}$]. Among all

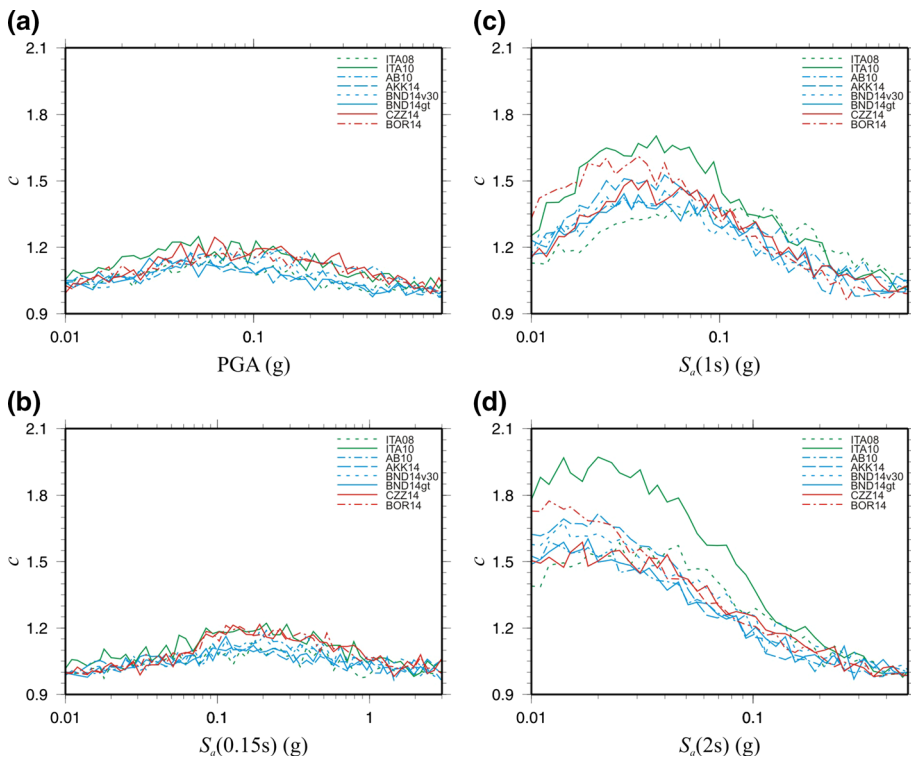


Fig. 2 Variation of the parameter c as a function of PGA (a), $S_a(0.15\text{ s})$ (b), $S_a(1\text{ s})$ (c), and $S_a(2\text{ s})$ (d). Note that distinction is made between BND14v30 and BND14gt but not in the case of CZZ14. For CZZ14, distinction is not necessary as, contrary to BND14, the regression coefficients related to the magnitude, distance, and style of faulting functions are independent of the site amplification term

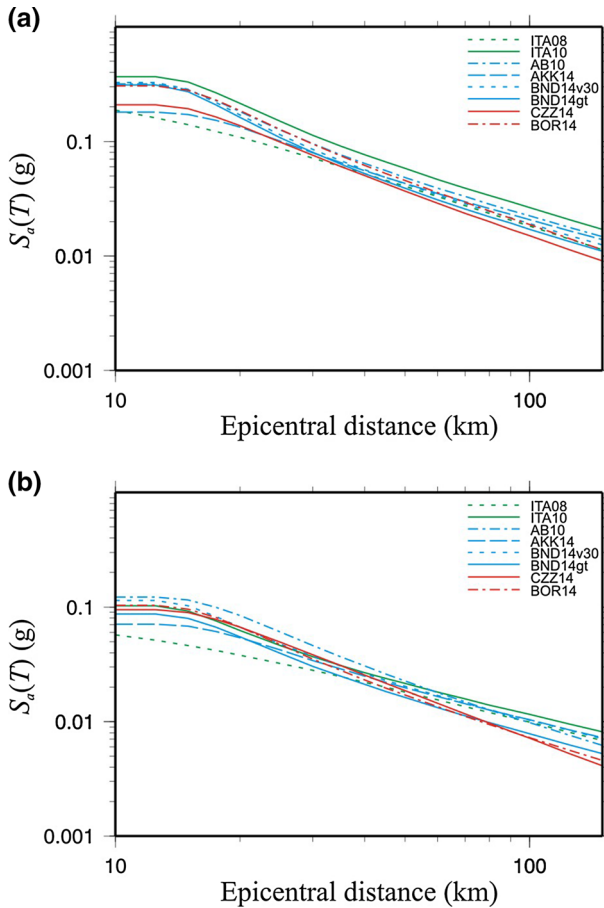


Fig. 3 Scaling of ground-motion with distance computed for the GMPEs listed in Table 1 considering an earthquake with $M_W = 6.5$ and $V_{S,30} = 800$ m/s. **a** $S_d(1$ s) and **b** $S_d(2$ s)

GMPEs considered, ITA10 is the one providing the highest correlation levels, particularly at longer spectral periods. As shown in Fig. 3, this behavior may be attributed to the slower attenuation of the ground motions provided by this GMPE for larger magnitude events (i.e., approaching the upper limit of applicability, $M_W = 6.9$) which, as known, tend to control the long-period hazard (e.g., Barani et al. 2009).

The values of c_i employed in the scoring test are provided as electronic supplementary material.

4 Scoring test

4.1 A note on seismic hazard computation

In order to assess the effectiveness of the hazard models based on the GMPEs in Table 1 in providing reliable (i.e., compatible with observations) hazard estimates, we carried out

different computational runs by varying one GMPE at a time while keeping constant the remaining components of the PSH model. In each run, the ground motion hazard is computed at each site according to the site classifications presented in Table 2. The reference seismic hazard model considers the same set of input models and parameters adopted by Barani et al. (2009) for the disaggregation of the Italian ground motion hazard maps. In order to avoid conversions between ground motion values, hazard estimates are computed either in terms of ENV_{xy} or GM_{xy} depending on the GMPE (see Table 1 for acronym definitions). Only BOR14 was corrected to predict GM_{xy} (instead of RotD50) ground motions by applying the empirical conversion models of Barani et al. (2015) in conjunction with those of Beyer and Bommer (2006). According to Bommer et al. (2005), the aleatory variability related to the conversion of the median ground motion was carried across into the aleatory variability of each GMPE. Concerning the source-to-site distance, each GMPE is applied according to the relevant distance metrics (see Table 1). Note that CZZ14 and BND14 model site amplification either as a continuous function of $V_{S,30}$ or by means of ground categories. Both options are considered (the suffixes “v30” and “gt” are used after the GMPE acronym) in order to explore their effects on the hazard.

4.2 Results of the scoring test

In order to compare the outcomes of each PSH model with observations, a set of S accelerometric sites is considered, each operating for a Δt_s time interval (Table 2). As stated previously, the set of benchmark sites includes stations that have been operating in Italy for at least 25 years. Specifically, the largest PGA and $S_a(T)$ values (either in terms ENV_{xy} or GM_{xy} depending on the GMPE) have been considered for each station (Luzi et al. 2008; Pacor et al. 2011).

Figure 4 summarizes the results of the scoring test showing the overall likelihood Q_i associated with the hazard models corresponding to the GMPEs in Table 1 for separate ground shaking parameters [PGA, $S_a(0.15\text{ s})$, $S_a(1\text{ s})$, and $S_a(2\text{ s})$]. To this end, the values of L_i obtained through Eq. 6 for different MRPSs (30, 50, 72, 101, 140, 201, 475, 975, and 2475 years) are used with Eq. 8. The values of L_i are provided as electronic supplementary material. Note that, for an MRP of 30 years and $T = 0.15\text{ s}$, L_i is found to be always lower than a 5% significance level. This behavior may be indicative that, for this spectral period, one component of the PSH model (e.g., recurrence model, GMPE) makes the hazard estimates for very short MRPs not compatible with the observations. Hence, this MRP is excluded from the computation of the scores presented in Fig. 4 for $T = 0.15\text{ s}$.

Among the PSH models considered, those implementing ITA10 and BND14 (particularly in the “gt” form) are clearly the best performing. The model based on BND14 gets higher scores at 0.15 and 2 s. On the other hand, the model implementing ITA10 appears more effective to forecast the PGA and the spectral acceleration corresponding to a 1 s period. This model exhibits a good performance also for a 2 s response. For this spectral period and for $T = 1\text{ s}$, the model based on AB10 performs very similarly to that based on ITA10. The remaining models get lower scores and, consequently, appear less suitable for PSHAs that include site effects. It is worth noting that, although AKK14 and BND14 are both based on the same strong motion archive (Akkar et al. 2013), the former always provides lower scores. Although the original archive is the same, the final regression data sets are not. In addition, also the functional forms are different. Both these factors contribute to differences in the predicted median ground motions and, more generally, in the ground motion probability distributions. In particular, major differences concern the

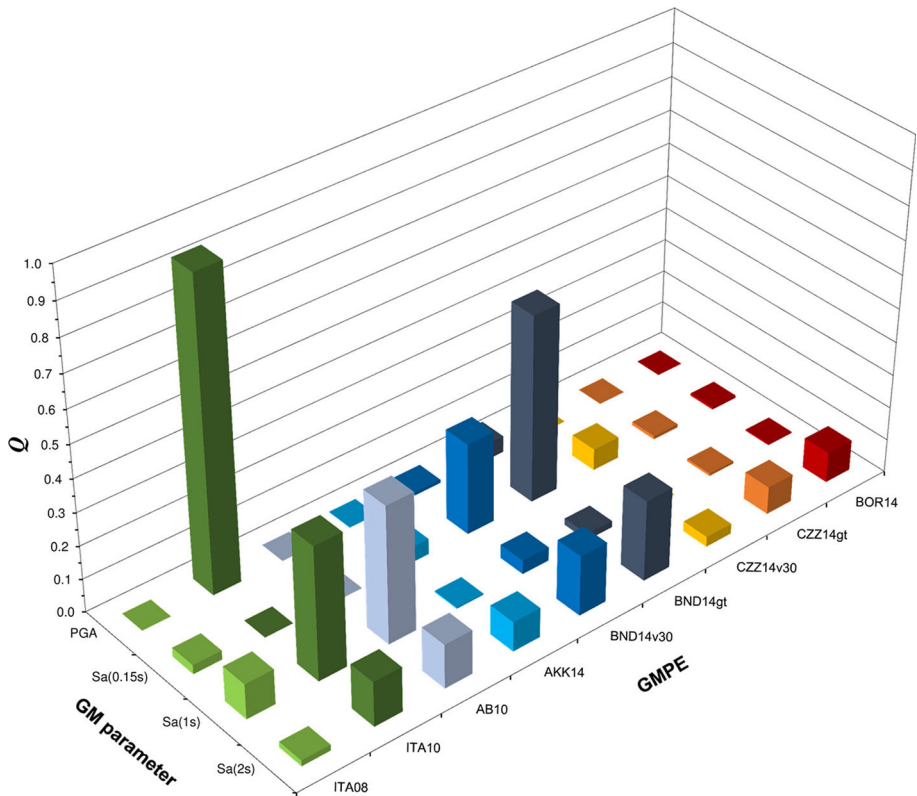


Fig. 4 Q -scores (see Eq. 8) associated with the PGA-, $S_a(0.15\text{ s})$ -, $S_a(1\text{ s})$ -, and $S_a(2\text{ s})$ -hazard models corresponding to the GMPEs in Table 1. GMPEs developed for the same region (or by the same research group) are displayed using different tones of the *same color* (e.g., Italian GMPEs are in *green*)

attenuation with distance (particularly at longer response periods and increasing magnitudes) and site conditions. Other minor differences are related to magnitude scaling for M_W greater than 6.75 (for details, see Douglas et al. 2014).

Combining the likelihood values computed for all the MRPs and ground motion parameters considered (again by applying Eq. 8) leads to the diagram in Fig. 5, which summarizes the overall performance of all models tested. BND14 is definitely the dominating (i.e., best performing) model. Note that the hazard model based on ITA10 is significantly penalized by the low score at 0.15 s (see Fig. 4), which, in turn, is due to the very low likelihood value determined for an MRP of 2475 years (see the L_i values in the electronic supplementary material). Repeating the scoring test without taking the hazard results for that MRP into account, the model incorporating ITA10 gets indeed the largest overall score. Comparing the scores of BND14v30 with those associated with BND14gt and those of CZZ14v30 with those corresponding to CZZ14gt indicates that modeling site effects by means of a $V_{S,30}$ -dependent soil coefficient instead of using ground types does not lead to improved performances. Using ground types provides slightly higher scores and, therefore, appears more appropriate for large-scale hazard evaluations.

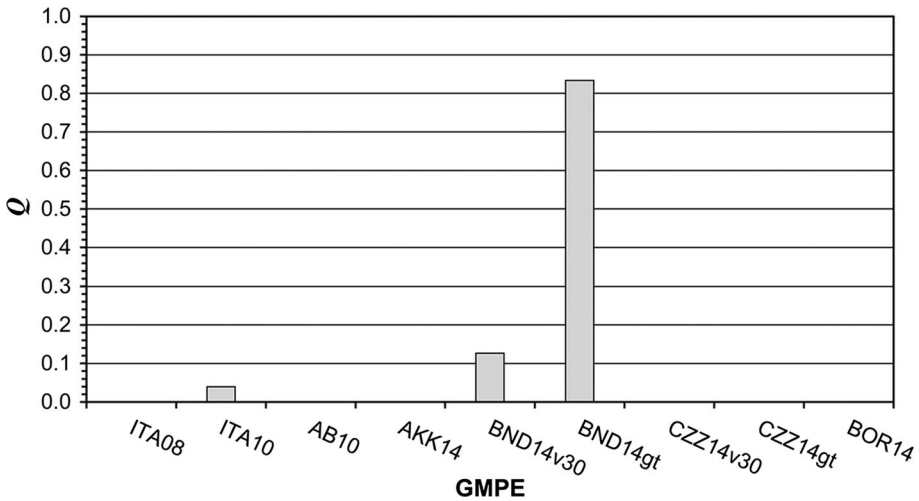


Fig. 5 Overall scores associated with the hazard models considered in Fig. 4

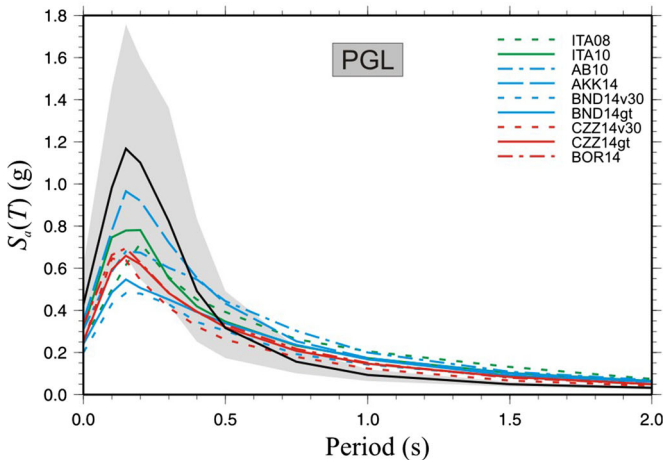
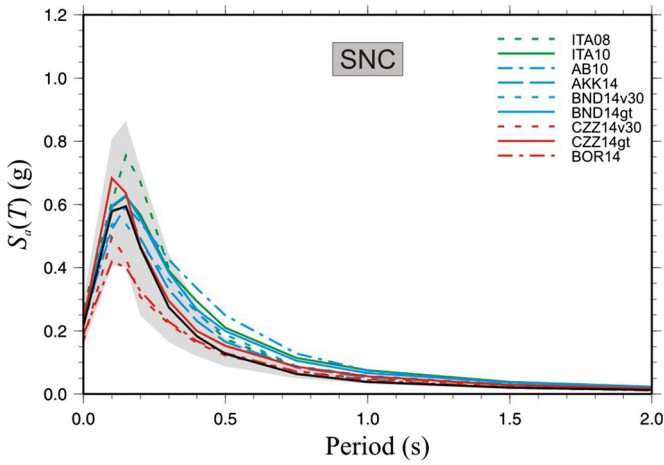
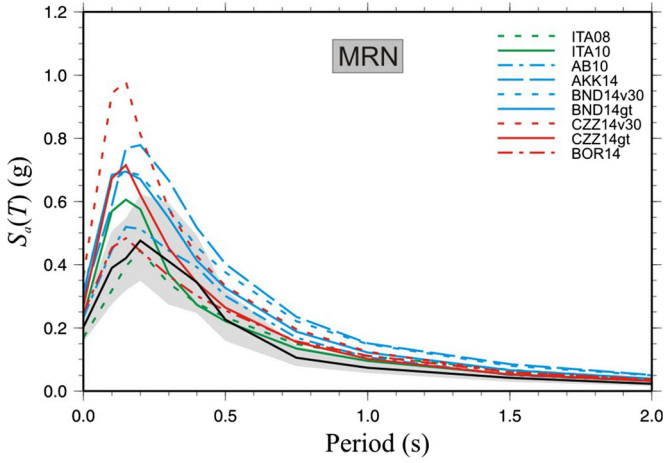
5 Drawbacks of using GMPEs for generic site conditions in PSHA

Previous results have shown that, although GMPEs provide only a generic evaluation of site amplification effects, reliable large-scale hazard maps can be computed when suitable models are selected. However, some spatial variations in the results presented above may exist. Moreover, given the approximate nature of the site amplification term included in the GMPEs, some overestimation or underestimation of the hazard may occur in particular areas due to variations in regional geology. Due to the limited sample of benchmark sites, we cannot verify this statistically. Thus, it may be useful to compare the hazard estimates determined through the application of GMPEs for generic site conditions with those resulting from a more rigorous site-specific hazard approach. The comparison is presented here for three sites representative of different soil conditions and hazard levels in Italy (see Fig. 1). Specifically, we have considered the same three sites examined in the article of Barani and Spallarossa (2016): Mirandola (MRN), Soncino (SNC) (this site is not included in the list of stations in Table 2), and Peglio (PGL). MRN and SNC are both located in the Po Plain (northern Italy). However, they present different soil conditions. MRN presents a thick layer of approximately 100 m of soil deposits (mainly sand and clay) above the bedrock (i.e., rock/soil with average shear wave velocity greater than or equal to 800 m/s) while SNC only 17 m (mainly gravel). The third site, PGL, is located in central Italy, in an area characterized by a higher hazard level. Here, the bedrock is overlaid by approximately 40 m of soil and soft rock (clays, clayey marls and gypsum) with shear wave velocity lower than 800 m/s. Although the comparison is provided for three sites only, it may serve to give the reader a clearer idea of the strengths and weaknesses of the method for soil hazard assessment examined in this work.

Figure 6 compares the soil uniform hazard spectra (UHSs) for an MRP of 475 years resulting from the application of the GMPEs in Table 1 with the average UHS determined through the application of the multi-SAF (SAF is for Soil Amplification Function) approach for site-specific PSHA proposed by Barani and Spallarossa (2016). This approach

Fig. 6 Soil UHSs for an MRP of 475 years resulting from the application of the GMPEs in Table 1 for the sites of Mirandola (MRN), Soncino (SNC), and Peglio (PGL). The *black curves* and the *shaded areas* indicate the average site-specific UHSs and the uncertainty bands between the 2nd and 98th percentile UHSs determined by applying the multi-SAF approach proposed by Barani and Spallarossa (2016)

is based on the convolution method described in the article of Bazzurro and Cornell (2004b). Specifically, for each site, 200 numerical ground response analyses, corresponding to 200 randomizations of the reference soil model, are performed (see Barani and Spallarossa 2016). The results from each analysis are then used to define a regression model for soil amplification (termed SAF) for each spectral period of interest. The average multi-SAF UHS is then computed by assuming that each one of the seven GMPEs corresponds to a branch of a logic tree. In this exercise, GMPEs are each assigned equal weights. For a given spectral period, each rock hazard curve from the logic tree is convolved with each one of the 200 SAFs. Globally, 1400 site-specific hazard curves are computed for each spectral period of interest for each site. From this set of hazard curves, we calculate the mean hazard (i.e., the mean annual rate of exceedance of specified ground motion intensities) and the hazard corresponding to different percentiles. Note that the hazard curves are computed by removing the ergodic assumption in the ground motion variability (e.g., Anderson and Brune 1999; Al Atik et al. 2010; Rodriguez-Marek et al. 2011, 2014), thus to avoid double counting of the uncertainty related to site amplification, which is already considered in the numerical ground response analysis via Monte Carlo simulations. This was achieved here by roughly reducing the standard deviations of the logarithmic ground motion by an amount of 15% (Luzi et al. 2014; Rodriguez-Marek et al. 2011). The average multi-SAF UHSs are used as terms of comparison to analyze the effect of each GMPE on the hazard. Precisely, we calculate the difference (Δ) between the spectral acceleration values displayed by each UHS in Fig. 6 and those computed for the same MRP using the multi-SAF approach. The values of Δ derived for the PGA, $S_a(0.15\text{ s})$, $S_a(1\text{ s})$, and $S_a(2\text{ s})$ hazard are shown in Figs. 7, 8, 9 for MRN, SNC, and PGL, respectively. At MRN, where thick soil deposits are present and moderate soil nonlinearity has been observed from the numerical simulations performed by Barani and Spallarossa (2016), all GMPEs with the exception of ITA08 at shorter periods lead to hazard estimates that are greater than those resulting from the multi-SAF approach, which takes into account the nonlinear behavior of soils. Among the GMPEs considered, only AKK14 and BOR14 include a nonlinear site amplification term. The latter model, along with ITA08, ITA10, and AB10, provides a better agreement with the multi-SAF hazard estimates, particularly at low periods (Fig. 7a, b). At this site, using $V_{S,30}$ to model site effects instead of ground types does not seem beneficial. The values of Δ associated with BND14gt and CZZ14gt are always lower than those obtained by applying BND14v30 and CZZv30. For the site of Soncino (Fig. 8), the reader may observe a general better agreement (with respect to MRN) between the hazard estimates relative to each single GMPE and those obtained with the multi-SAF method. Except for a period of 0.15 s (Fig. 8b), which is close to the site resonance (Barani and Spallarossa 2016), differences Δ in the hazard estimates are always negligible. At this period, except for ITA08, CZZ14v30, and BOR14, all other GMPEs lead to hazard results compatible with those obtained with the multi-SAF method. Concerning PGL, which is the site (among those considered) with the highest hazard, Fig. 9 clearly shows that, at shorter periods (Fig. 9a, b), all GMPEs underestimate the multi-SAF hazard by significant amounts. This is attributable to the ineffectiveness of GMPEs in properly capturing the actual site amplification, which at this site is concentrated between 3 Hz and 10 Hz (Barani and Spallarossa 2016). The opposite occurs when longer periods



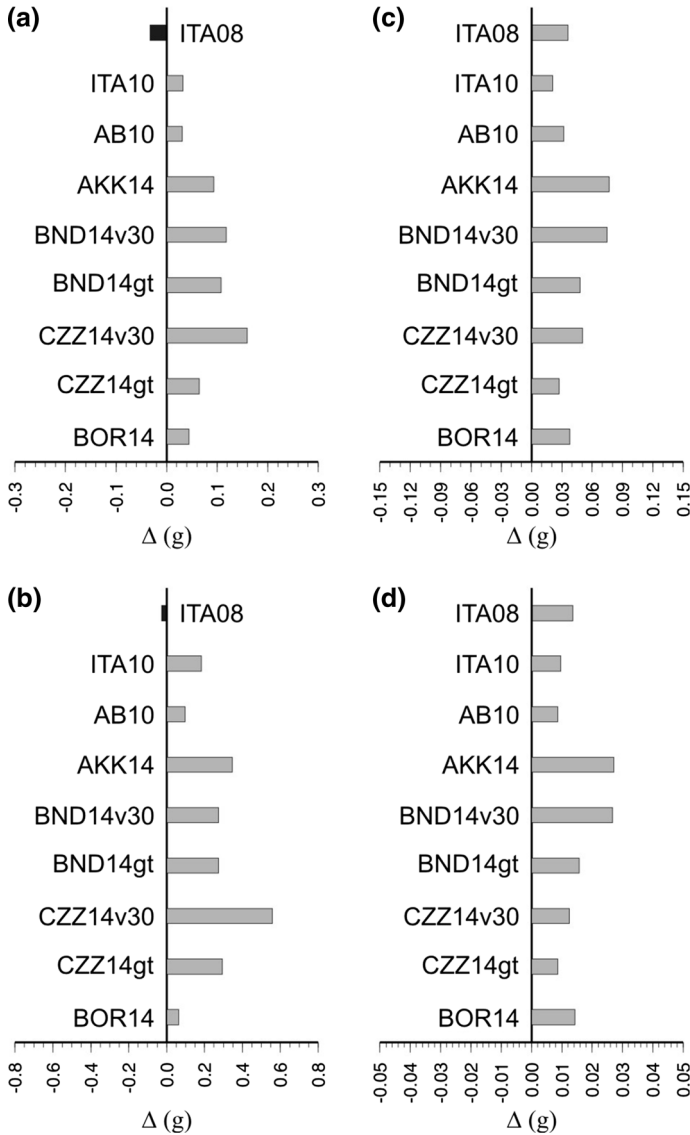


Fig. 7 Difference (Δ) between the PGA (a), 0.15 s (b), 1 s (c), and 2 s (d) spectral acceleration hazard values for an MRP of 475 years computed by using the GMPEs in Table 1 and the hazard values obtained from the multi-SAF approach for the site of Mirandola (MRN). Underestimations with respect to the multi-SAF estimates are indicated by black bars. Bars in gray indicate overestimation

are considered (Fig. 9c, d). For 1 s and 2 s response, indeed, all GMPEs lead to hazard values that are greater than those provided by the multi-SAF method. At these periods, the lower values of Δ are provided by CZZ14v30. This finding, along with observation that BND14v30 has led to smaller Δ values compared to BND14gt, indicates that, at least for this site, incorporating $V_{S,30}$ in the site amplification term is helpful in improving hazard estimates at medium-to-long spectral periods.

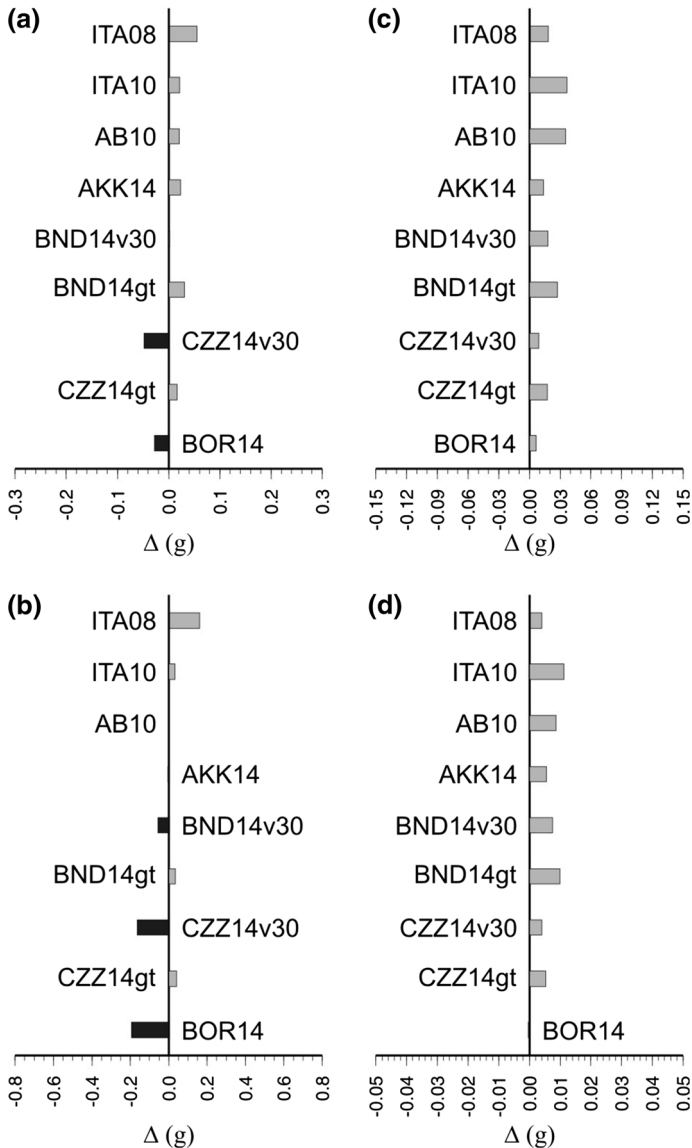


Fig. 8 Same as Fig. 7 but for the site of Soncino (SNC)

6 Discussion and conclusions

Based on our results, reliable large-scale hazard mapping of Italy inclusive of site effects can be achieved through a conventional PSHA based on a restricted number of GMPEs. In particular, if the same input assumptions considered in this work are adopted, two GMPEs, BND14gt and ITA10, appears appropriate. Their relative scores might serve as ‘weights’ expressing the degree of belief in the selected models which, as a first step, were assumed equiprobable. Note that in the application presented here, we have not taken into account

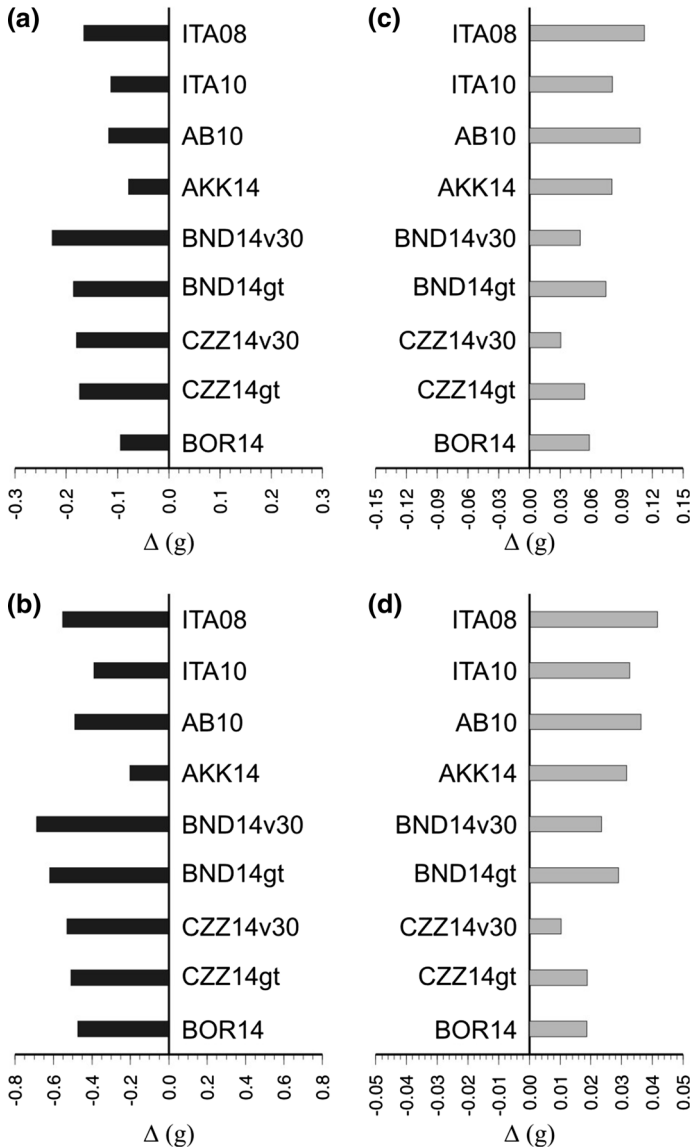


Fig. 9 Same as Fig. 7 but for the site of Peglio (PGL)

the epistemic uncertainties related, for instance, to earthquake sources, activity rates, b -values, and maximum magnitude. We have simply varied one GMPE at a time while keeping constant a particular set of input assumptions, which, in our case, is that providing hazard results closer to the median values obtained using the entire logic tree adopted for the PSHA of Italy (Barani et al. 2009). In the case of logic tree applications, a multi-factorial approach (e.g., Rabinowitz and Steinberg 1991; Barani et al. 2007) is advisable, thus to account for all epistemic uncertainties into the annual rates of exceedance associated with each branch to be examined. Actually, indeed, the results of a logic-tree-based

PSHA are influenced by all assumptions considered in the logic tree. As a consequence, the results of the scoring test are dependent on the entire hazard model adopted. In other words, the score associated with a particular assumption of the PSH model is influenced by the remaining logic-tree branches. The advantage of a multi-factorial approach is that it allows one to evaluate the effect on the hazard introduced by the uncertainty of a single assumption by simultaneously varying all other input parameters and models. In such a way, all epistemic uncertainties are carried across into the annual rate of exceedance of a certain level of ground motion (or alternatively, into the ground motion that is exceeded with a certain probability) associated with a particular branch of the logic tree. In this perspective, the empirical scoring test presented here can be profitably applied to evaluate the likelihood of the various components of a PSH model. This can be helpful to prune unlikely branches from “leafy” logic trees.

Although the use of GMPEs for generic site conditions is found to be particularly appropriate for large scale soil hazard applications, some caution is required in handling the assessed hazard. At specific sites or in particular areas, such an approach could lead to hazard estimates that may be either over-conservative (e.g., in the case of non-linear soil response) or under-conservative with respect to those determined by the application of more rigorous site-specific PSH approaches, particularly in the frequency range around the site fundamental frequency. This should be carefully considered by the final users of large-scale soil hazard maps, especially when handling such estimates in specific regions characterized by particular geological conditions and/or for the design of important facilities for which, however, site-specific approaches are strongly recommended (e.g., Abrahamson et al. 2004; Rodriguez-Marek et al. 2014).

Acknowledgements We are thankful to the guest editor Francesca Pacor, Gabriele Ameri, and Marco Pagani for their thorough review and precious suggestions, which brought significant improvement to the work. This study has benefited from funding of the Projects S2-2012 and S2-2014, provided by the Italian Presidenza del Consiglio dei Ministri—Dipartimento della Protezione Civile (DPC) (<http://sites.google.com/site/ingvdpc2014progettos2/>). This paper does not necessarily represent DPC official opinion and policies.

References

- Abrahamson NA, Coppersmith KJ, Koller M, Roth P, Sprecher C, Toro GR, Youngs R (2004) Probabilistic seismic hazard analysis for Swiss nuclear power plant sites (PEGASOS Project). Final Report, Volume 1. <http://www.swissnuclear.ch/upload/cms/user/PEGASOSProjectReportVolume1-new.pdf>. Accessed 15 April 2016
- Akkar S, Bommer JJ (2010) Empirical equations for the prediction of PGA, PGV and spectral accelerations in Europe, the Mediterranean region and the Middle East. *Seismol Res Lett* 81:195–206
- Akkar S, Sandikkaya MA, Bommer JJ (2014a) Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. *Bull Earthq Eng* 12:359–387
- Akkar S, Sandikkaya MA, Bommer JJ (2014b) Erratum to: Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. *Bull Earthq Eng* 12:389–390
- Akkar S, Sandikkaya MA, Şenyurt M, Azari Sisi A, Ay BO, Traversa P, Douglas J, Cotton F, Luzi L, Hernandez B, Godey S (2013) Reference database for seismic ground-motion in Europe (RESORCE). *Bull Earthq Eng* 12:311–339
- Al Atik L, Abrahamson N, Bommer JJ, Scherbaum F, Cotton F, Kuehn N (2010) The variability of ground-motion prediction models and its components. *Seismol Res Lett* 81:794–801
- Albarelo D, D’Amico V (2008) Testing probabilistic seismic hazard estimates by comparison with observations: an example in Italy. *Geophys J Int* 175:1088–1094
- Albarelo D, D’Amico V (2015) Scoring and testing procedures devoted to probabilistic seismic hazard assessment. *Surv Geophys* 36:269–293

- Albarelo D, Peruzza L (2016) Accounting for spatial correlation in the empirical scoring of probabilistic seismic hazard estimates. *Bull Earthq Eng*. doi:10.1007/s10518-016-9961-0
- Albarelo D, Peruzza L, D'Amico V (2013) D6.1—Report on model validation procedures (Technical report). Internal report of the S2-2012 Project. <https://sites.google.com/site/ingvdpc2012/progettos2/deliverables/d6>. Last access July 2016
- Albarelo D, Peruzza L, D'Amico V (2015a) A scoring test on probabilistic seismic hazard estimates in Italy. *Nat Hazards Earth Syst Sci* 15:171–187
- Albarelo D, Peruzza L, Goretti A (2015b) D6.1—Validation of PSHA: methodological updates and new results (Technical report). Internal report of the S2-2014 Project. <https://sites.google.com/site/ingvdpc2014/progettos2/deliverables>. Last Accessed July 2016
- Anderson JG, Brune J (1999) Probabilistic seismic hazard analysis without the ergodic assumption. *Seismol Res Lett* 70:19–28
- Barani S, Spallarossa D (2015a) D4.1—Site-specific PSHA for previously selected sites (Technical report). Internal report of the S2-2014 Project. <https://sites.google.com/site/ingvdpc2014progettos2/deliverables>. Last access July 2016
- Barani S, Spallarossa D (2015b) D4.2—Site-specific PSHA for newly selected sites, guidelines (Technical report). Internal report of the S2-2014 Project. <https://sites.google.com/site/ingvdpc2014progettos2/deliverables>. Last access July 2016
- Barani S, Spallarossa D (2016) Soil amplification in probabilistic ground motion hazard analysis. *Bull Earthq Eng*. doi:10.1007/s10518-016-9971-y
- Barani S, Spallarossa D, Bazzurro P, Eva C (2007) Sensitivity analysis of seismic hazard for Western Liguria (North Western Italy): a first attempt towards the understanding and quantification of hazard uncertainty. *Tectonophysics* 435:13–35
- Barani S, Spallarossa D, Bazzurro P (2009) Disaggregation of probabilistic ground-motion hazard in Italy. *Bull Seismol Soc Am* 99:2638–2661
- Barani S, De Ferrari R, Ferretti G (2013) Influence of soil modeling uncertainties on site response. *Earthq Spectra* 29:705–732
- Barani S, Spallarossa D, Bazzurro P, Pelli F (2014a) The multiple facets of probabilistic seismic hazard analysis: a review of probabilistic approaches to the assessment of different hazards caused by earthquakes. *Boll Geofis Teor Appl* 55:17–40
- Barani S, Massa M, Lovati S, Spallarossa D (2014b) Effects of surface topography on ground shaking prediction: implications for seismic hazard analysis and recommendations for seismic design. *Geophys J Int* 197:1551–1565
- Barani S, Albarelo D, Spallarossa D, Massa M (2015) On the influence of horizontal ground shaking definition on probabilistic seismic hazard analysis. *Bull Seismol Soc Am* 105:2704–2712
- Barani S, Albarelo D, Massa M, Spallarossa D (2016) Influence of twenty years of research on ground motion prediction equations on probabilistic seismic hazard in Italy. *Bull Seismol Soc Am* 107
- Bazzurro P, Cornell CA (2004a) Ground-motion amplification in nonlinear soil sites with uncertain properties. *Bull Seismol Soc Am* 94:2090–2109
- Bazzurro P, Cornell CA (2004b) Nonlinear soil-site effects in probabilistic seismic-hazard analysis. *Bull Seismol Soc Am* 94:2110–2123
- Beyer K, Bommer JJ (2006) Relationships between median values and between aleatory variabilities for different definitions of the horizontal component of motion. *Bull Seismol Soc Am* 96:1512–1522
- Bindi D, Luzi L, Massa M, Pacor F (2010) Horizontal and vertical ground motion prediction equations derived from the Italian Accelerometric Archive (ITACA). *Bull Earthq Eng* 8:1209–1230
- Bindi D, Pacor F, Luzi L, Puglia R, Massa M, Ameri G, Paolucci R (2011) Ground motion prediction equations derived from the Italian strong motion database. *Bull Earthq Eng* 9:1899–1920
- Bindi D, Massa M, Luzi L, Ameri G, Pacor F, Puglia R, Augliera P (2014a) Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset. *Bull Earthq Eng* 12:391–430
- Bindi D, Massa M, Luzi L, Ameri G, Pacor F, Puglia R, Augliera P (2014b) Erratum to: Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset. *Bull Earthq Eng* 12:431–448
- Bommer JJ, Scherbaum F, Bungum H, Cotton F, Sabetta F, Abrahamson NA (2005) On the use of logic trees for ground-motion prediction equations in seismic hazard analysis. *Bull Seismol Soc Am* 95:377–389
- Boore DM, Stewart JP, Seyhan E, Atkinson GM (2014) NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. *Earthquake Spectra* 30:1057–1085
- Building Seismic Safety Council (2009) NEHRP recommended provisions for seismic regulations for new buildings and other structures. Federal Emergency Management Agency, Report No. P-750, Washington, DC, USA

- Castellaro S, Albarello D (2016) Reconstructing seismic ground motion at reference site conditions: the case of accelerometric records of the Italian National Accelerometric Network (RAN). *Bull Earthq Eng*. doi:[10.1007/s10518-016-0032-3](https://doi.org/10.1007/s10518-016-0032-3)
- Cauzzi C, Faccioli E, Vanini M, Bianchini A (2015) Updated predictive equations for broadband (0.01–10 s) horizontal response spectra and peak ground motions, based on a global dataset of digital acceleration records. *Bull Earthq Eng* 13:1587–1612
- Comitè Européen de Normalisation (2004) Eurocode 8: Design of structures for earthquake resistance, Part 1: General rules, seismic actions and rules for buildings. Belgium, Brussels
- Cramer C (2006) Quantifying uncertainty in site amplification modeling and its effects on site-specific seismic-hazard estimation in the upper Mississippi embayment and adjacent areas. *Bull Seismol Soc Am* 96:2008–2020
- Douglas J, Akkar S, Ameri G, Bard P-Y, Bindi D, Bommer JJ, Bora SS, Cotton F, Derras B, Hermkes M, Kuehn NM, Luzi L, Massa M, Pacor F, Riggelsen C, Sandikkaya MA, Scherbaum F, Stafford PJ, Traversa P (2014) Comparisons among the five ground-motion models developed using RESORCE for the prediction of response spectral accelerations due to earthquakes in Europe and the Middle East. *Bull Earthq Eng* 12:341–358
- Felicetta C, D'Amico M, Lanzano G, Puglia R, Russo E, Luzi L (2016) Site characterization of Italian accelerometric stations. *Bull Earthq Eng*. doi:[10.1007/s10518-016-9942-3](https://doi.org/10.1007/s10518-016-9942-3)
- Haase JS, Choi YS, Bowling T, Nowack RL (2011) Probabilistic seismic-hazard assessment including site effects for Evansville, Indiana, and the surrounding region. *Bull Seismol Soc Am* 101:1039–1054
- ITACA Working Group (2016) Italian Accelerometric Archive, version 2.1. doi:[10.13127/ITACA/2.1](https://doi.org/10.13127/ITACA/2.1)
- Joyner WB, Boore DM (1981) Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. *Bull Seismol Soc Am* 71:2011–2038
- Kalkan E, Wills CJ, Branum DM (2010) Seismic hazard mapping of California considering site effects. *Earthq Spectra* 26:1039–1055
- Kramer SL (1996) *Geotechnical earthquake engineering*. Prentice-Hall, Englewood Cliffs
- Luzi L, Hailemikael S, Bindi D, Pacor F, Mele F, Sabetta F (2008) ITACA (ITalian ACcelerometric Archive): a web portal for the dissemination of Italian strong-motion data. *Seismol Res Lett* 79:716–722
- Luzi L, Bindi D, Puglia R, Pacor F, Oth A (2014) Single-station sigma for Italian strong-motion stations. *Bull Seismol Soc Am* 104:467–483
- MPS Working Group (2004) Redazione della mappa di pericolosità sismica prevista dall'Ordinanza PCM 3274 del 20 marzo 2003, Rapporto conclusivo per il dipartimento di Protezione Civile. INGV, Milano-Roma, Aprile 2004, 65pp. +5 appendici. http://zonesismiche.mi.ingv.it/documenti/rapporto_conclusivo.pdf
- Ministero delle Infrastrutture e dei Trasporti (2008) Norme tecniche per le costruzioni–NTC. D.M. 14 Gennaio 2008, Supplemento ordinario alla Gazzetta Ufficiale No. 29, 4 Febbraio 2008
- Musson RMW (2009) The use of Monte Carlo simulations for seismic hazard assessment in the U.K. *Ann Geophys* 43:1–9
- Pacor F, Paolucci R, Luzi L, Sabetta F, Spinelli A, Gorini A, Nicoletti M, Marcucci S, Filippi L, Dolce M (2011) Overview of the Italian strong motion database ITACA 1.0. *Bull Earthq Eng* 9:1723–1739
- Pelli F, Mangini M, Bazzurro P, Eva C, Spallarossa D, Barani S (2006) PSHA in northern Italy accounting for non-linear soil behaviour and epistemic uncertainty. In: *Proceedings of the 1st European conference on earthquake engineering and seismology—ECEES*, Geneva, Switzerland, Paper 1464
- Petersen MD, Bryant WA, Cramer CH, Reichle MS, Real CR (1997) Seismic ground-motion hazard mapping incorporating site effects for Los Angeles, Orange, and Ventura Counties, California: a geographical information system application. *Bull Seismol Soc Am* 87:249–255
- Rabinowitz N, Steinberg DM (1991) Seismic hazard sensitivity analysis: a multi-parameter approach. *Bull Seismol Soc Am* 81:796–817
- Rodriguez-Marek A, Montalva GA, Cotton F, Bonilla F (2011) Analysis of single-station standard deviation using the KiK-net data. *Bull Seismol Soc Am* 101:1242–1258
- Rodriguez-Marek A, Rathje EM, Bommer JJ, Scherbaum F, Stafford PJ (2014) Application of single-station sigma and site-response characterization in a probabilistic seismic-hazard analysis for a new nuclear site. *Bull Seismol Soc Am* 104:1601–1619
- Romeo R, Paciello A, Rinaldis D (2000) Seismic hazard maps of Italy including site effects. *Soil Dyn Earthq Eng* 20:85–92

- Strasser FO, Bommer JJ, Abrahamson NA (2009) Sigma: issues, insights, and challenges. *Seismol Res Lett* 80:40–56
- Task 4 Working Group (2013) D4.1—Site-specific hazard assessment in priority areas. Deliverable of the DPC-INGV S2 Project “Constraining observations into Seismic Hazard”. <https://sites.google.com/site/ingvdpc2012progettos2/deliverables/d4-1>. Accessed 14 April 2016