Site-specific probabilistic seismic hazard analysis for the Western area of Naples, Italy

Hossein Ebrahimian[1], Fatemeh Jalayer[2], Giovanni Forte[3], Vincenzo Convertito[4], Valeria Licata[5], Anna d'Onofrio[6], Antonio Santo[7], Francesco Silvestri[8], Gaetano Manfredi[9]

[1] Assistant Professor, Department of Structures for Engineering and Architecture (DIST), University of Naples Federico II, Naples, Italy, ebrahimian.hossein@unina.it .

[2] Associate Professor, Department of Structures for Engineering and Architecture (DIST), University of Naples Federico II, via Claudio 21, 80125 Naples, Italy, fatemeh.jalayer@unina.it, phone: (+39) 0817685959, (**Corresponding Author**).

[3] Assistant Professor, Department of Civil, Environmental and Architectural Engineering (DICEA), University of Naples Federico II, Naples, Italy, giovanni.forte@unina.it .

[4] Ph.D., Istituto Nazionale di Geofisica e Vulcanologia (INGV), Osservatorio Vesuviano, via Diocleziano 328, 80124, Naples, Italy, vincenzo.convertito@ingv.it .

[5] Ph.D., Anas SpA, Rome, Italy, v.licata@stradeanas.it .

[6] Associate Professor, Department of Civil, Environmental and Architectural Engineering (DICEA), University of Naples Federico II, Naples, Italy, anna.donofrio@unina.it .

[7] Associate Professor, Department of Civil, Environmental and Architectural Engineering (DICEA), University of Naples Federico II, Naples, Italy, antonio.santo@unina.it .

[8] Professor, Department of Civil, Environmental and Architectural Engineering (DICEA), University of Naples Federico II, Naples, Italy, francesco.silvestri@unina.it .

[9] Professor, Department of Structures for Engineering and Architecture (DIST), University of Naples Federico II, via Claudio 21, 80125 Naples, Italy, gamanfre@unina.it .

Abstract
Probabilistic Seismic Hazard Analysis (PSHA) encompasses quantitative estimation of seismic hazard at a site by considering all plausible earthquake scenarios. The outcome of a PSHA is often reported as the mean rate of exceeding a specific ground motion intensity measure at a given site. This study attempts to perform PSHA for the western area of the city Naples (southern Italy) by employing the most advanced methods and new databases; namely, DISS3.2 (Database of Individual Seismogenic Sources) and CPTI15 (Parametric Catalogue of Italian Earthquakes). Seismogenic models include individual seismogenic structures/faults liable to generating major earthquakes with magnitude greater than 5.5, and background areal and point source models to evaluate the effect of earthquakes with magnitude less than 5.5. The PSHA is built up based on the long-term earthquake recurrence on seismogenic tectonic faults and the spatial distribution of historical earthquakes. Site amplification is considered based on seismic microzonation maps derived for the western area of Naples. The microzonation maps delineate expected levels of ground motion amplification based on reliable geological and geotechnical subsoil models. Hazard maps are derived for a number of return periods for ground-shaking in terms of peak ground acceleration and 5%-damped pseudo-spectral acceleration at a range of periods that are representative of the existing construction within the area. Detailed comparisons of the PSHA results with Italian national hazard maps and the code-based design spectra emphasize the importance of performing site-specific PSHA with explicit consideration of site effects.
Keywords Site-specific PSHA, Seismogenic source model, Seismic microzonation, Ground-motion prediction equation, Hazard map, Uniform hazard spectrum

1 Introduction
The term *Seismic Hazard* has been used ambiguously for addressing both the potentially disastrous future earthquake events as well as the uncertainty in determining their magnitude, time and location (McGuire 2008). In this context, *Seismic Hazard Analysis* (SHA) encompasses the quantitative estimation of seismic hazard for a designated site. The *Probabilistic SHA* (PSHA) considers all plausible earthquake scenarios that can affect a given site combined with the conditional distribution of ground motion intensity for a prescribed earthquake scenario. This conditional distribution is usually represented through the *ground motion prediction equations* (GMPEs). The GMPEs predict the ground motion intensity as a function of variables such as the earthquake's magnitude, distance, fault mechanism, near-surface site conditions, the potential presence of directivity effects, etc.
Italy is one of the most seismically active regions in the Mediterranean. The first attempt to provide seismic hazard maps for Italian territory, following the standard PSHA method proposed by Cornell (1968), was carried out by Slejko et al. (1998). This effort led to 10% exceedance probability in 50 years maps for peak ground acceleration (*PGA*) and the macro-seismic intensity. The seismic hazard in this work was based on the seismic source zone model ZS4 (Meletti et al. 2000), the earthquake catalog NT4.1 (Camassi and Stucchi 1997), and the European GMPE (Ambraseys 1995). This work was further extended (Albarello et al. 2000) by using an updated seismic zonation in Italy (Gruppo di Lavoro 1999), the CPTI99 catalog (Gruppo di Lavoro CPTI 1999), and including the pseudo-spectral acceleration (S_a) as the intensity measure. Two equally-weighted GMPEs, namely, a European GMPE (Ambraseys et al. 1996) and an Italian GMPE (Sabetta and Pugliese 1996, SP96), were used. Meanwhile, Slejko et al. (1999) employed the same methodology and data provided previously in Slejko et al. (1998) using a more recent GMPE (Ambraseys et al. 1996) to perform the PSHA for Adriatic region within the Global Seismic Hazard Assessment Program (GSHAP, Giardini 1999). The next updating of seismic hazard maps for Italy (Romeo and Pugliese 2000, Romeo et al. 2000) was conducted by using the NT4.1 catalog and adding the instrumental earthquakes with magnitude threshold 4.6 in the time span 1981-1996. The SP96 attenuation relation was used to compute the maps with 10% exceedance probability in 50 years for *PGA*, peak ground velocity (*PGV*), and S_a. Using the same data for GSHAP, a unified seismic hazard modelling was also conducted separately for Mediterranean area (Jiménez et al. 2001) by employing a modified version of seismic zone model ZS4. In 2004, the reference seismic hazard map for Italy (MPS04, Gruppo di Lavoro 2004; see also Stucchi et al. 2011) was released by Istituto Nazionale di Geofisica e Vulcanologia (INGV) following the standard Cornell's PSHA approach. The input parameters adopted by INGV for computing MPS04 were the CPTI04 earthquake catalog (Gruppo di lavoro CPTI 2004), the seismogenic source model ZS9 (Meletti et al. 2008), and the GMPEs of Ambraseys et al. (1996), SP96, and a set of regional GMPEs (see Montaldo et al. 2005 for a detailed description of the ground-motion models). MPS04 provides the seismic hazard maps for *PGA* with a probability of exceedance of 10% in 50 years. The MPS04 was consequently improved by carrying out the national research Project S1 (2004-2006, see Meletti et al. 2007) funded by the Italian Department of Civil Protection (DCP). The project provided maps associated with different exceedance probabilities in 50 years for *PGA*, and S_a at the period range of 0.1-2.0 seconds (Meletti and Montaldo 2007; Montaldo and Meletti 2007).
The seismic hazard maps for Italy are developed based on seismotectonic zoning and historical catalogues without exploiting the geometries and seismicity rates derived directly from geological and paleoseismological observations (Pace et al. 2006). It is noteworthy that the average recurrence time for large earthquakes is sometimes significantly longer than the time window covered by the historical catalogue (Stucchi et al. 2011).

That is, important earthquakes might be overlooked in the PSHA calculations based on historical catalogues (Stucchi et al. 2011). The seismic sequences in central Italy starting from the 1997–1998 Umbria-Marche sequence raised the problem of seismogenic source characterization and seismicity rate estimation for faults that have been silent or unknown during historical times (Peruzza and Pace 2002). In this context, the Database of Individual Seismogenic Sources (DISS) has been first developed in 2000 and updated in the next years (Basili et al. 2008). Nevertheless, the PSHA calculations in Italy (at national level) have not utilized so far the DISS data (Basili et al. 2008). These calculations are usually based on vast areal seismic source zones (e.g., Gruppo di Lavoro 2004). An alternative has been to use kernel estimation method governed by concepts of fractal geometry and self-organized seismicity; concepts that do not require the definition of seismogenic zoning (Zuccolo et al. 2013). The direct use of seismogenic sources, provided by an early version of DISS (i.e., DISS 2.0, Valensise and Pantosti 2001), was not believed to be viable at that time because the mapping of the seismogenic sources was incomplete and affected by large uncertainties (Stucchi et al. 2011). Nevertheless, the individual and composite seismogenic areas, recently introduced in the latest version 3.0 of the DISS database (Basili et al. 2008) provide an interesting alternative to areal seismic zoning to be explored in future Italian national seismic hazard maps (Stucchi et al. 2011). A typical composite seismogenic source may span several individual faults (DISS Working Group, 2015). The importance of database such as DISS is further emphasized due to the fact that the national earthquake catalogues may not adequately represent the seismogenic sources of medium-to-strong events. In an alternative take to PSHA, individual sources are held responsible for major earthquakes, which can be supported by detailed geological evidence (Pace et al. 2006). By addressing the individual sources, the estimation of the seismicity rate and the time- and memory-dependent earthquake recurrence can be considerably enhanced (Pace et al. 2002; Faenza et al. 2003; Marzocchi et al. 2003; Boncio et al. 2004; Convertito et al. 2006; Pace et al. 2006; Akinci et al. 2009; Peruzza et al. 2010; Akinci et al. 2016; Zimmaro and Stewart 2017; Vanini et al. 2017, Faenza et al. 2017).

This works lays out the details of a novel site-specific PSHA for the western area of Naples in southern Italy. The aim has been to use all the available seismological, geological, geophysical, and geotechnical information for the designated area. The only seismic hazard maps available for Naples are those provided by INGV in terms of MPS04 national hazard maps (not based on seismogenic individual faults). The seismogenic source modelling herein is based on a bi-layer model (Pace et al. 2006). The individual structures, capable of producing major earthquakes (magnitude greater than 5.5), are referred to as *Seismogenic Boxes* (SBoxes) herein. The SBoxes are nothing but the surface projection of active faults. According to the segmentation model defined by Boncio et al. (2004), a SBox is a major structure that can be considered substantially continuous through its depth for several kilometres. The magnitude recurrence relation and seismicity rate associated with each SBox are estimated based on the DISS 3.2.0 (DISS Working Group 2015). For magnitudes less than 5.5, a *Background Source model* is characterized based on the latest INGV-released Parametric Catalogue of Italian Earthquakes (CPTI15; Rovida et al. 2016), which includes macroseismic and instrumental data and parameters for Italian earthquakes with magnitude greater than or equal to 4.0 in the time period 1000-2014. The western area of Naples has a high potential of being affected by site response amplification, since it rests on soft soil of alluvial and volcanic origin. In the present study, a detailed seismic microzonation has been conducted on the zone in order to properly evaluate the local soil-site effects. The PSHA calculations are performed by using SBoxes, point sources and background area, and the results are combined to provide seismic hazard maps for *PGA* and S_a at different hazard levels according to the National Technical Code for seismic design in Italy (NTC 2018). Three most recent Italian, European and global models are adopted as GMPE's herein; namely, ITA10 (Bindi et al. 2011), BND14 (Bindi et al. 2014a, b) and BSSA (Boore et al. 2014). Since ITA10 and BND14 use the geometric mean of the two horizontal components of ground motion, we have modified these two GMPEs in order to account for an arbitrary horizontal component of ground shaking. The seismic hazard results obtained in this study are compared (in terms of the uniform hazard spectrum and hazard curves) with the national hazard data provided by INGV and NTC 2018 for the designated site.

2 A general description of the zone of interest

The areal extent of the zone of interest is nearly 7.5 km^2 and is located along the Southern margin of the Phlaegrean fields, in the western area of Naples in Southern Italy (see Fig. 1; the region is highlighted in red in the upper-left corner of the figure). Naples is the third-largest municipality in Italy after Rome and Milan. The case-study area encompasses the zones of Bagnoli and Fuorigrotta (see the red-dashed rectangle in Fig. 1 and the area highlighted as yellow in Fig. 1's sub-figure). The area is an active volcanic field, whose eruptive history is characterized by the emplacement of several pyroclastic deposits and lava flows (Di Vito et al. 1999; Orsi et al. 2004). Powerful eruptions in the past has led to the creation of geological formations such as the Campanian Ignimbrite (CI; ~39 ky; De Vivo et al. 2001; Fedele et al. 2008; the black-colored border in the sub-figure of Fig. 1), and the Neapolitan Yellow Tuff (NYT; ~15 ky; Deino et al. 2004; yellow-colored border in the sub-figure of Fig. 1). The recent volcanic activity has taken place completely within the NYT caldera in three different phases. The first phase was formed by the NYT eruption with nearly 37 explosive eruptions (from 15 ky to 9.5 ky); the second one consisted in 6 events in the Northern part of the designated area (from 8.6 ky to 8.2 ky); finally, the third phase was made

of 20 explosive and 3 effusive eruptions (from 4.8 ky to 3.8 ky). The last stasis was interrupted historically by the Monte Nuovo eruption in 1538. The morphological setting of the area is shaped by the presence of several volcanic landforms and embedded craters which host a geothermal system. The changes in the pressure and temperature in this geothermal system have led to frequent episodes of ground uplift and subsidence, called *bradyseism*. The most important bradyseismic crises occurred between 1969–1972 and 1982–1985 (Chiodini et al. 2010), with a maximum uplift of 1.79m in Pozzuoli area in 1985 (Del Gaudio et al. 2010) (see its location in the sub-figure of Fig. 1).

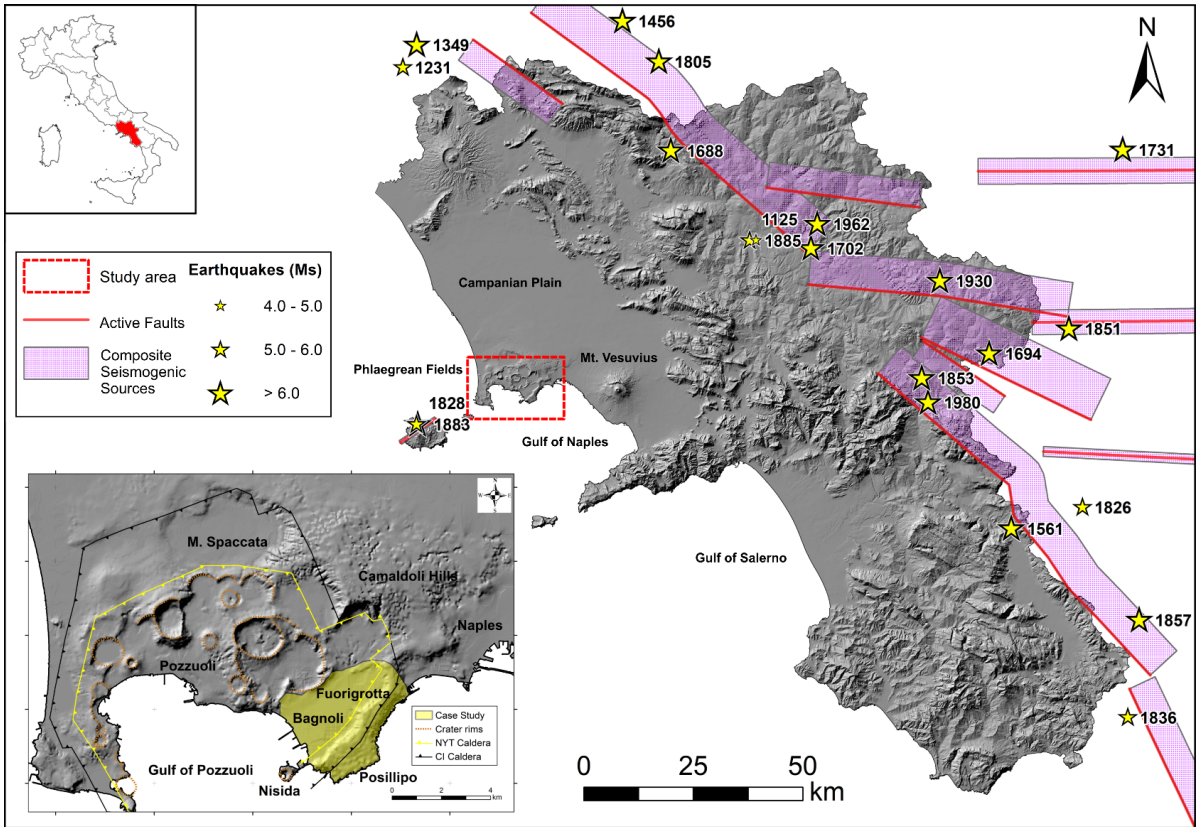


Figure 1: Tectonic and seismological setting of Campania Region; the location of the case-study area, main faults, and historical earthquakes.

This brief description draws attention to the fact that the area under study rests on soft soil of alluvial and volcanic origin that can strongly affect the local seismic response. The urban fabric is highly heterogeneous in this area. It consists of masonry and reinforced concrete constructions of different ages ranging from 1919 (and even older) up to 2001 –the majority of buildings are constructed in the period between within 1946-1971. An urban requalification masterplan, mainly aimed at restoring its original touristic attraction, involves the district of Bagnoli-Fuorigrotta. The plan envisions the rehabilitation of the area occupied by an abandoned steel factory and construction of new transportation infrastructure. Fig. 1 also shows the distribution of the composite seismic sources within and around the Campania region (Basili et al. 2008) together with few important historical events (highlighted with asterisks of different size according to their corresponding magnitudes in the figure caption). It should be noted that the seismic sources will be discussed in detail in Section 4 of this paper.

3 Geological Characteristics and Microzonation

The geological and geotechnical subsoil model adopted for the case-study area is described in detail in Licata et al. (2016) and Licata et al. (2019a, 2019b). It is based on nearly 330 stratigraphic logs (with depths ranging from 30-100m), coupled with the study of the geomorphological, structural and hydrogeological setting of this area. The analysis of these stratigraphic logs led to the recognition of volcanic formations and Holocene deposits of anthropic, aeolian, alluvial, transitional and marine origin. In Fig. 2a, the areas characterized by homogenous stratigraphic patterns are displayed as a synthetic geolithological map. The location of the stratigraphic logs are highlighted with white circles, and the Cone Penetartion Tests (CPTs) are highlighted with blue inverted triangles. Moreover, black contour lines represent the topographic elevations. As it can be seen in Fig. 2a, the area consists of 7 geolithological complexes made of pyroclastic deposits, which are briefly described as follows:

- The Neapolitan Yellow Tuff (NYT, geological formation 7 in Fig. 2a) represents the engineering seismic bedrock of the area. It should be noted that the NYT is not the real bedrock. Due to the complex geological formation of the area, the estimation of the real bedrock depth is not straightforward. The NYT reveals a complex geometry; that is, it crops out at the SE side on the Posillipo hill and it is lowered in the central area of the plain, down to more than 120 m below sea level. The buried morphology of the lithic portion of NYT has been estimated by interpolating the logs, whose NYT (engineering bedrock) depths are shown as red-dotted contours (negative values) in Fig. 2a.
- Along the NW side, the NYT deposits do not crop out and are buried in the plain as a pyroclastic soil whose depth increases toward the NW direction (see geological complexes 1-6 in Fig. 2a). In the center of the plain, tephra and tuff (geological formation 6 in Fig. 2a) remnants of the Santa Teresa Volcano (STT, 12.7 ky) crop out. The other complexes (geological formations 1-5 in Fig. 2a) are mainly constituted of a succession of pyroclastic soils coming from the most recent eruptions of Phlaegrean Fields in primary deposition. The distinction among these sectors is based on the depth of the tuff (i.e. bedrock depth); namely, (<20m) for geological formation 5, (> 20m) for geological formation 4, and undefined bedrock (tuff) depth for geological formation 2. Geological formations 1 and 3 reveal the presence of aeolian sands and clayey silts with peat layers.

The seismic microzonation was carried out based on 30 measurements of the shear wave velocity (V_S) obtained through several geophysical tests (21 down-hole, DH; 5 cross-hole, CH; 4 MASW tests, see Licata et al. 2019a, 2019b). The locations of V_S-measurement tests are highlighted with yellow circles on Fig. 2a. Accordingly, the average shear wave velocity of the upper 30m, denoted as V_{S30}, is calculated for each geolithological complex. The seismic soil classification map, shown in Fig. 2b, is based on the Eurocode 8 (EC8; CEN 2004) and the National Technical Code for seismic design (NTC 2008) site classes, using the estimated spatial average values for V_{S30} (see Forte et al. 2017). More specifically, the outcropping NYT along the SE sector is classified as class A (with V_{S30}>800m/s representing the engineering bedrock); whereas the main part of the plain (Fuorigrotta and Bagnoli zones) is classified as soil-site class C (V_{S30}=180-360m/s, with bedrock depth higher than 20m). The transition between these two zones (i.e., geological formation 5, see Fig. 2a), characterized by V_{S30} ranging between 180-360m/s and a relatively shallow (<20 m) NYT depth, is classified as type E (soil class E is defined as 5–20m of C- or D-type alluvium lying upon stiffer material with V_{S30}≥800m/s). The deposits below the dune sands (geological formation 1), with a mean V_{S30} between 180-360m/s but highly susceptible to liquefaction, are classified as soil type S_2. Finally, the pyroclastic soils below STT (geological formation 6) are left as unclassified. This is because the stratigraphic profile reveals significant inversions of the shear wave velocity with the depth; and hence, the soil classification cannot be applied. The results of microzonation of the studied area are summerized in Table 1 as follows:

Table 1: Seismic Microzonation

Geological Characterization*	Soil Classification (EC8)**	Suggested V_{S30} (m/s)	Description
(1) Pyroclastic soils below sand dune	C	302	Aeolian sands above pyroclastic deposits made of clayey silts with peat layers, highly susceptible to liquefaction, classified as ground type S_2 in Fig. 2b
(2) Pyroclastic soils on unlithified NYT	C	295	Pyroclastic deposits lying on the soil of the NYT eruption. The depth of the bedrock is not defined.
(3) Pyroclastic and marine soils with peats on NYT	C	295	Pyroclastic deposits made of clayey silts with peat layers; the depth of the bedrock is >20m.
(4) Pyroclastic soils on NYT	C	295	Pyroclastic deposits made of clayey silts resting on the bedrock (NYT). The depth of the layer is >20m.
(5) Pyroclastic soils on NYT	E	196	Pyroclastic deposits made of clayey silts resting on the bedrock (NYT). The depth of the layer is <20m..
(6) Pyroclastic soils below STT	C	295	Pyroclastic deposits below the tuff of STT, characterized by V_S inversions. The class cannot be defined (ND); it is assumed to be of class C herein.
(7) NYT outcrop	A	800	Represents the engineering seismic bedrock of the whole area

* See the geological characterization map in Fig. 2a., ** See the site classification map in Fig. 2b.

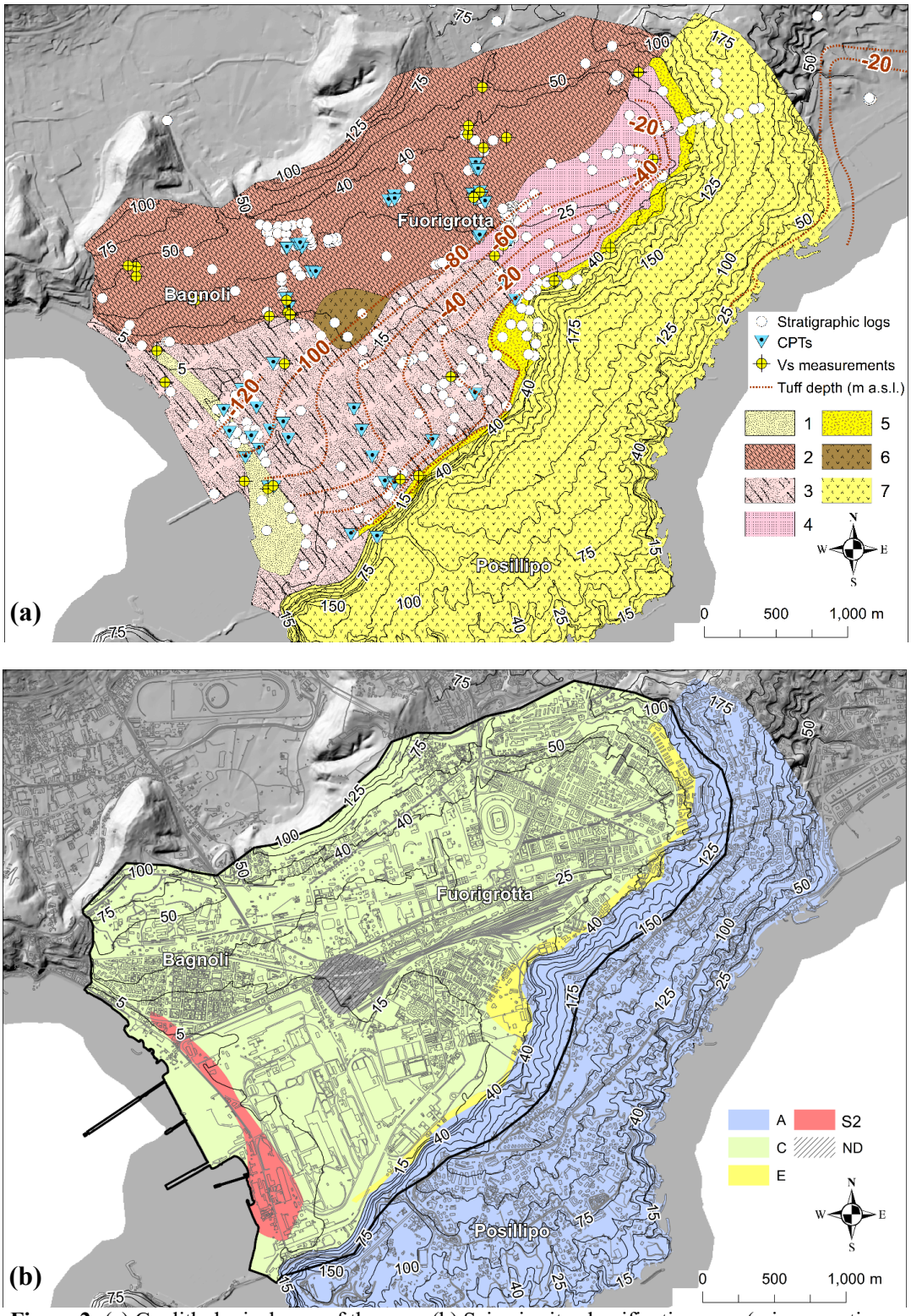


Figure 2: (a) Geolithological map of the area, (b) Seismic site classification map (microzonation map)

4 Seismogenic Source Characterization

The first step in a PSHA procedure is the characterization of the seismogenic source model(s). The model should incorporate the available seismological, geological, and geophysical information for the active area together with historical and instrumental earthquake catalogues. As mentioned before, the approach adopted herein attributes the major earthquakes to the individual fault structures and the low-to-medium earthquakes to the background seismicity. Section 4.1 describes the finite-fault source model, for which the list of key input parameters are: SB=Seismogenic Boxes or SBox; $\mathbf{M}_{\max}$=vector of alternative calculated/observed maximum moment magnitude estimates; M_{char}=the average value of the vector $\mathbf{M}_{\max}$; σ_m=standard deviation of M_{char}; M_l^{SB}=lower magnitude of

a SB; M_u^{SB}=upper magnitude of a SB; f_M^{SB}=truncated normal probability density function, PDF, of magnitudes assigned to a SBox; v^{SB}=annual seismicity rate assigned to a SB. Section 4.2 charachterizes the background source model with the following key input parameters: M_l^{BG}=lower magnitude of the background area; f_M^{BG}=truncated Exponential distribution of magnitudes for background source; M_c=the completeness magnitude; β=the slope of the Gutenberg-Richter (GR) earthquake rate model (β=bln10??); v^{BG}=annual seismicity rate within the background area.

4.1 Finite-fault Source Models

A segmented belt of large normal faults running along the crest of the Apennines are mostly responsible for strong earthquakes taken place in the southern Apennines (see e.g., Valensise and Pantosti 2001; DISS Working Group 2015; Basili et al. 2008). These large faults dip predominantly to the southwest in the central Apennines and to the northeast in the southern Apennines (DISS Working Group, 2015). The rectangular area located in southern Apennines, highlighted in Fig. 3a (and Fig. 3b), is adopted herein as the "background area". Fig. 3b shows also the Italian national seismogenic zones (ZS9, Gruppo di Lavoro 2004) surrounding the background zone. It can be observed that the background area overlaps significantly with seismic zones Z927 and Z928 (partial overlap with zones 924 and 925). The case-study site is located in the western part of the city of Naples and is shown with cyan-colored area in Fig. 3a. The background area is extended through the northeast around 130 km, extending to three neighboring provinces. It extends in the southeast direction around 160 km. The yellow-colored rectangles represent the Seismogenic Boxes (SBox) within the background area. As mentioned before, the SBoxes are the surface projection of individual seismogenic active faults (the geometry of a typical dipping fault is shown in Fig. 3c). The name assigned by DISS 3.2.0 (DISS Working Group 2015) to each SBox is shown right next to the corresponding SBox, together with an identification number in the parenthesis (referred to in Table 2 for identifying the SBox). The closest edge of each seismic box to the ground surface is highlighted with thick gray line.

The spatial distribution of historical earthquakes based on CPTI15 catalog in the time interval from 1000 to 2014 (Rovida et al. 2016) and with magnitude greater than 4.0 is also shown in Fig. 3a. The small gray-colored dots indicate historical events with moment magnitude (M_w) interval of [4.0, 5.50); the blue dots represent events with $M_w \in [5.50, 6.0)$; the magenta triangles represent events with $M_w \in [6.0,7.0)$; finally, the red stars show the very rare $M_w \geq 7.0$ historical events. The historical earthquakes shown in Fig. 3a took place within the background area or are located at a distance within 10 km outside of it. For each event with $M_w \geq 5.5$ in Fig. 3a, two lines of information are provided: the first line specifies M_w and the time of occurrence (year) of the event; the second line identifies the record's label in CPTI15 catalog and the seismogenic fault (i.e., SBox) to which it is attributed (if available). The 14 SBoxs falling within the background area are reported in Table 2 by summarizing their description in DISS 3.2.0 (name, fault type, time of the latest associated event, and the moment magnitude) and listing of the historical earthquake(s) assigned to each SBox (according to CPTI15 catalog). Finally, for each SBox, a brief discussion (with relative references) of each source according to DISS commentary is reported.

A quick survey of events in Fig. 3a reveals that over the past seven centuries, the background area has been struck by several M_w (+)6 earthquakes: September 1349 (moment magnitude M_w 6.8, multiple events), December 1456 (M_w 6.9-7.2, multiple events), June 1688 (M_w 7.1), September 1694 (M_w 6.7), March 1702 (M_w 6.6), November 1732 (M_w 6.7), July 1805 (M_w 6.7), August 1851 (M_w 6.5), July 1930 (M_w 6.7), August 1962 (M_w 6.1), and November 1980 (M_w 6.2-6.8, multiple events).

One of the main assumptions in source modelling herein is the attribution of the earthquakes of $M_w \geq 5.5$ events to the SBoxes (see Table 2). The threshold of M_w 5.5 is also recommended in the DISS database where it is highlighted that DISS is a compilation of potential sources for earthquakes larger than 5.5 in Italy and surrounding areas. According to DISS, this assignment is done based on field mapping, landscape evolution, geophysics, and paleoseismology data. The PSHA formulation for the calculation of hazard does not necessarily state that the sources should be mutually exclusive; rather, they should be independent. However, the fact that the background source and the SBoxes are overlapping (at the locations of the SBoxes) implies that the minimum magnitude for SBoxes and the maximum magnitude for the background source need to be equal. Otherwise, the PSHA calculation can double count the possible events (see also Akinci et al. 2016). On the other hand, in a layered modeling in which the SBoxes are modelled as characteristic, the background needs to overlap the SBoxes. Otherwise, the events with magnitude less than the minimum magnitude of the SBoxes (5.5 in this case) are not going to be assigned neither to the SBoxes and nor to the background. With reference to both Table 2 and Fig. 3a, out of 28 earthquakes with $M_w \geq 5.5$ highlighted in Fig. 3a, the following issues can be highlighted:

(1) The southeast of the background area is close to the debated epicentral locations of the 19 August 1561 (M_w 6.7) event and includes the 31 July 1561 (M_w 6.3) event. These two events are assigned to a neighboring seismogenic fault outside the zone of interest (see DISS Working Group 2015), and hence are not considered in PSHA calculations (see also the discussion in Section S.M.4 in the electronic supplementary material of this manuscript).

(2) The events 05 May 1990 (M_w 5.8) and 17 July 1361 (M_w 6.0) in the right side are also assigned in DISS to neighboring seismogenic faults outside the zone of study (see DISS Working Group 2015). Therefore, these events are not considered in the list of historical earthquakes affecting this site (see also the discussion in Section S.M.4 in the electronic supplementary material of this manuscript).
(3) Three events with $5.5<M_w<6.0$ (marked with blue dots) are not assigned to any known individual source including: the 05 December 1499 (M_w 5.6 – event 238 in CPTI15 database whose epicenter is located near the city of Nola), the 31 July 1561 (M_w 5.6 – event 348 in CPTI15 with epicenter located in the Sorrento peninsula), both in the middle of the background area, and the 27 December 1978 (M_w 5.9 – event 3208 in CPTI15 database whose epicenter is located in central Tyrrhenian sea) on the left-hand side. With reference to CPTI15 and also the International Seismological Centre (ISC) on-line bulletin (Di Giacomo et al. 2014), the hypocentral depth of the recent event 3208, taken place in 1978 (with M_w 5.9), is 391.8 km, indicating that it originated from a subduction zone. Therefore, it is not considered within the PSHA calculation herein as our hazard estimates do not take into account the deep subduction earthquakes. Regarding historical events 238 and 348 (with M_w 5.6) taken place in 1499 and 1561 respectively, and in lieu of any known individual or composite sources located nearby, we assume two individual point sources, called herein "*PS Nola*" and "*PS Sorrento peninsula*", as shown in Fig. 3a. This assumption is supported through the studies conducted by Camassi et al. (2011) on Italian historical events, as well as Castelli et al. (2008) and the fact that for both the far-field assumption is valid. According to Camassi et al. (2011), some earthquakes in Campania, located both in the Apennines and in the Vesuviana area have remained unknown until now because they occurred in periods in which compilation of seismic events are incomplete (e.g., 1499, Nola) or they have been somewhat overshadowed by more important events (e.g., 1561, Sorrento peninsula). Specifically, the source of M_w 5.6 Sorrento peninsula (1561) is also studied by Castelli et al. (2008) while revisiting the 1561 complex seismic sequence in southern Italy. They have concluded that the earthquake around Vietri sul Mare in 31 July 1561 (the approximate epicenter of Sorrento peninsula event) should be connected to a locally triggered source, and not to the Apennine faults (that were responsible for 31 July to 19 August 1561 seismic sequence in southern Italy), since no other localities between the coast and the Apennine chain were damaged.
(4) Out of a total of 28 events with $M_w \geq 5.5$, 7 events have already been discarded based on the considerations in (1) to (3). Among the remaining 21 events (see also Table 2), 13 events are not attributed to any known causative fault according to DISS 3.2.0 commentary (see the descriptions in the last column of Table 2). These events are highlighted in Table 2 in blue under the title "Assigned Event (CPTI15)". A tentative event-to-fault attribution is carried out for these 13 events based on criteria such as nearest source to the desired event (see the last column in Table 2). Clearly, the validity of such attribution depends on the quality of macroseismic data used to estimate the earthquake location in the CPTI15 database (see also Vannoli et al. 2015 and Akinci et al. 2016 for more details about the assignment of the causative fault).
(5) The 1456 destructive event consists of at least two large mainshocks (Fracassi and Valensise 2007), the first event occurred on 5 December with M_w 6.90 assigned to individual source Ariano Irpino (ITIS092, as denoted in Table 2, No. 5), and the second event occurred on 30 December with M_w 7.0 assigned to individual source Frosolone (ITIS095, see Table 2 No. 6). The location of the former event is confirmed in CPTI15 (presents an offset with respect to the causative fault, see Fig. 3); the location of the second event is not available.
(6) The CPTI15 catalog does not contain the second and third shock of the 1980 Irpinia earthquake with M_w 6.2 (took place within 20 seconds from each other, see the description in Table 2, No. 10-12). However, according to DISS database, these two events are assigned to SBoxes San Gregorio Magno (ITIS078), and Pescopagano (ITIS078).
(7) Among all the individual seismogenic sources defined in Table 2, the last row (No. 14) represents a composite source and is associated to two historical events with $M_w \geq 5.5$. The M_w 6.9 assigned by DISS to this source is based on the strongest earthquake occurred in the composite source. The events of 8 September 1694 (M_w 6.9 in DISS and M_w 6.73 in CPTI15, see Table 2) and the 7 June 1910 (Mw 5.80, see Table 2) are assigned according to the DISS commentary (DISS Working Group 2015) to this composite source. In the absence of any detected individual seismic source by DISS, we considered the whole composite source herein within our source characterization. The source has a complex geometry and it cannot be treated as a dipping plane. Therefore, the length of the expected rupture length is poorly defined or unknown.

In order to give a better perspective of the case-study zone, the SBoxes, the background source, the two point sources, the Italian seismic zonation around the background zone, and the earthquakes taken place within and surrounding the background area, two figures are provided in the electronic supplementary material of this manuscript (see Section S.M.4, Fig. SM5 and Fig. SM7).

Table 2: Individual seismogenic sources according to DISS 3.2.0 database (DISS Working Group, 2015) and their assigned historical events based on DISS database and CPTI15 catalog (Rovida et al. 2016)

Seismogenic Source (SBox) (DISS 3.2.0)				DISS 3.2.0		Assigned Event (CPTI15)			Notes and References (based on DISS 3.2.0)
No.	DISS ID	Name	Fault Type	Latest Event	M_w	ID	Time	M_w	
1	ITIS005	Tammaro Basin	Normal	Unknown	6.60 (ER)	524 58	05/06/1688 04/09/1293	7.06 5.80	Although no clear evidence exists about the actual location and geometry of the source of the M_w 7.06, 1688 earthquake, the assignment of the event to this source is based on Boschi et al. (1995), and Fracassi and Valensise (2007). In addition, the event M_w 5.8, 1293 is assigned to this source as the nearest source.
2	ITIS004	Boiano Basin	Normal	26/07/1805	6.60*	994	26/07/1805	6.68	Di Bucci et al. (2002)
3	ITIS089	Carpino-Le Piane	Normal	Unknown	5.80 (ER)	3395	07/05/1984	5.86	The area is of crucial importance in the seismotectonics of central and southern Italy especially for the M_w 5.86, 1984 earthquake (see D'Addezio et al. 2001, Fracassi & Valensise 2007). There is no sufficient evidence that the fault is responsible for the 1984 earthquake (which occurred very close to two nearby northwest sources); however, this event is assigned to this source herein.
4	ITIS138	Aquae Iuliae	Normal	09/09/1349	6.50	95 13	09/09/1349 25/03/1120	6.80 5.80	Galli and Naso (2009). The (multiple) 9 September 1349 earthquake stands out for its puzzling epicentral location(s). This seismic event includes two to three principal events occurred on September 9. In addition, the M_w 5.8, 1120 event is also assigned to this source as the nearest source.
5	ITIS092	Ariano Irpino	Normal	05/12/1456	6.90	169	05/12/1456	7.19	Fracassi and Valensise (2007). The 5 to 30 December 1456 earthquake is the largest destructive event occurred in the Italian peninsula and one of the strongest earthquakes of the entire seismic history of Italy. The area affected by the 1456 sequence extends from Central Abruzzo to Southeastern Italy) and from the Adriatic to the Tyrrhenian coasts. It had caused multiple destructive shocks across a wide area by the combination of faulting on three or more parallel E-W strands.
6	ITIS095	Frosolone	Normal	30/12/1456	7.00	NP			Fracassi & Valensise (2007). See also the notes for Source No. 5
7	ITIS006	Ufita Valley	Normal	29/11/1732	6.60*	692 590 2847 2848	29/11/1732 14/03/1702 21/08/1962 21/08/1962	6.75 6.56 5.68 6.15	According to the discussions in DISS 3.2.0 commentary, it has been decided to assign also the events of 1702 and 1962 to this source herein.
8	ITIS088	Bisaccia	Normal	23/07/1930	6.70	2422	23/07/1930	6.67	Pino et al. (2008)

9	ITIS081	Melfi	Strike-Slip	14/08/1851	6.30*	1217 1218 431	14/08/1851 14/08/1851 08/1625	6.52 5.48 5.80	Alessio et al. (1995) described the M_w 6.30, 1851 earthquake as consisting of two sub-events. We also assigned the M_w 5.8, 1625 event to this source.
10	ITIS077	Colliano	Normal	23/11/1980	6.80	3256 1227 546 185	23/11/1980 09/04/1853 04/03/1692 15/01/1466	6.81 5.60 5.88 5.98	First shock at 18:34:53 GMT. The 1980 Irpinia earthquake is made up by a series of three mainshocks occurred within 40 seconds. They are generally referred to as 0 sec, 20 sec and 40 sec shocks. For a complete description and the associated references, see DISS 3.2.0 commantary references (e.g., Westaway 1993). Moreover, with reference to Pantosti and Valensise (1988), the event M_w 5.6, 1853 is also assigned to this source. In addition, we assigned the events M_w 5.98, 1466, and M_w 5.88, 1692 to this source.
11	ITIS078	San Gregorio Magno	Normal	23/11/1980	6.20				Second shock at 18:35:13 GMT
12	ITIS079	Pescopagano	Normal	23/11/1980	6.20				Third shock at 18:35:33 GMT
13	ITIS068	Casamicciola Terme	Normal	28/07/1883	5.40 (ER)	1481	28/07/1883	4.26	
14	ITCS063	Andretta-Filano	Strike-Slip	Based on the strongest earthquake occurred in the region	6.90	566 2015	08/09/1694 07/06/1910	6.73 5.76	This is a composite source. With reference to DISS 3.2.0 commentary, the 1694 earthquake was not caused by the faults responsible for the 1980 Irpinia events. Hence, this composite source is assumed herein to be the causative complex system for the 1694 event. We also assigned the event M_w 5.76, 1910 to this source.

Note: ER = Empirical Relationship, NP = Location not defined,
* Value adopted from the older historical earthquake catalogue CPTI04 (Gruppo di lavoro CPTI 2004).

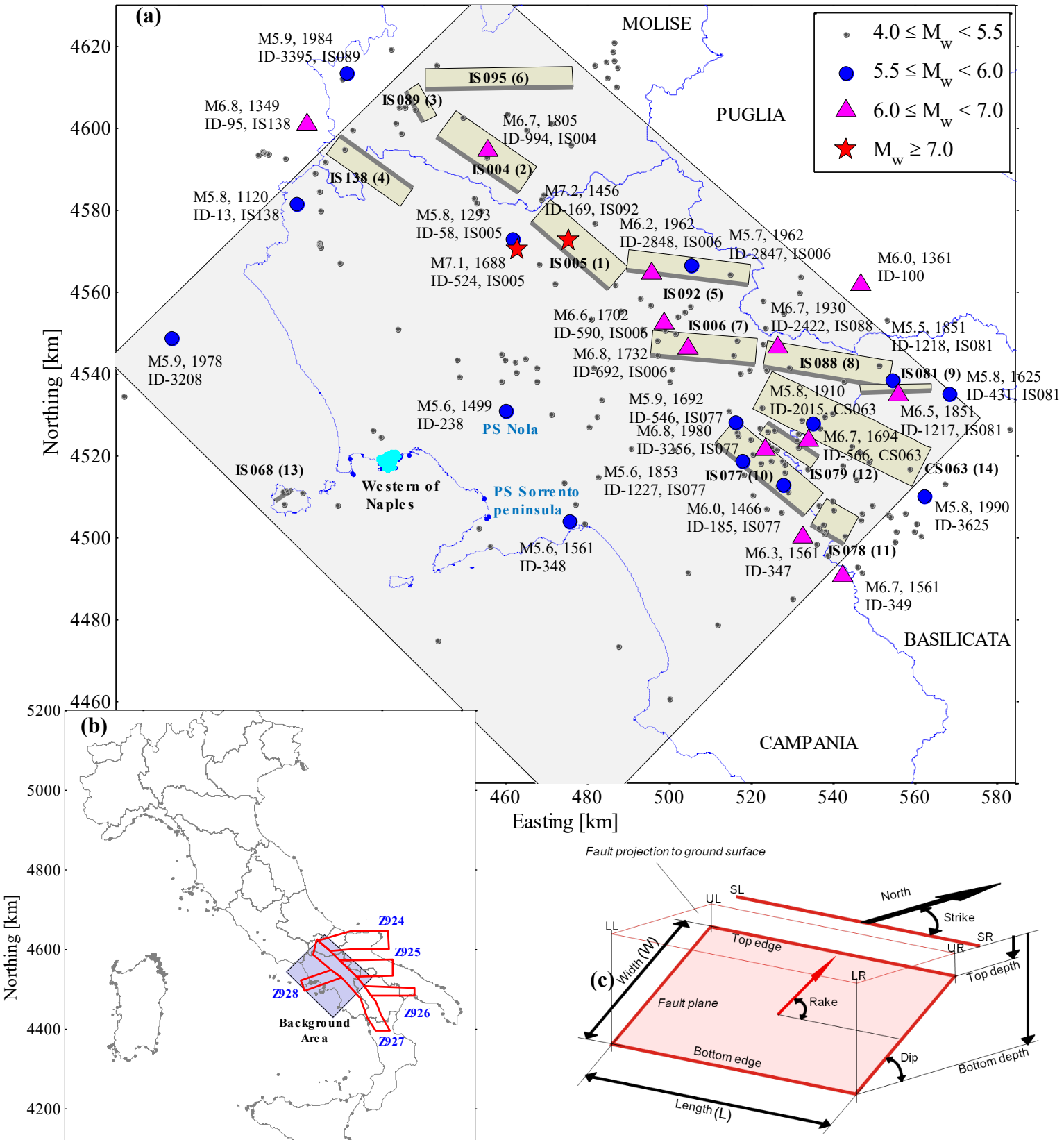


Figure 3: (a) The seismogenic sources: The zone of background seismicity is filled with light-blue colour; the SBoxs are shown with yellow-colour rectangles; the historical events according to CPTI15 catalogue with magnitude greater than or equal to 4.0 are marked with points, circles, triangles and stars (to differentiate their magnitude); (b) the background area overlaid with the Italian seismic zonation (ZS9); (c) the geometry of a typical seismogenic source http://diss.rm.ingv.it /diss /index.php/help/15-individual-seismogenic-sources).

4.1.1 Characterizing the distribution of maximum characteristic magnitude for seismic sources

The problem of the maximum magnitude estimation is widely debated in seismology (see e.g. Kagan and Jackson 2013; Zöller et al. 2013). To estimate the distribution of maximum characteristic magnitude on a fault segment, the data collected/assumed for the fault, such as size, rheological properties and empirical size-magnitude relationships, are employed. With reference to previous studies (Peruzza and Pace 2002, Pace et al. 2006, and Pace et al. 2016), the following sources for obtaining maximum magnitude (M_{max}) estimates (and the corresponding confidence band) are considered herein:

1. **Observed magnitude from DISS:** It corresponds to the maximum observed magnitude in the DISS database and is denoted herein as $M_{max,DISS}$ (see M_w reported in the 6th column of Table 2). $M_{max,DISS}$ is the magnitude of the strongest event observed in the region and attributed to the corresponding source. In case no such event has been registered, empirical relationships (denoted by ER in Table 2 and explained later on in points 3 to 5) are employed by DISS to get an estimate of the maximum magnitude. It is noteworthy that the type of employed empirical relation is not defined in the DISS commentary. In lieu of case-specific values for the standard deviation of $M_{max,DISS}$, a fixed value of 0.30 is set (reported right next to the expected value of $M_{max,DISS}$ in the 9th column of Table 3).
2. **Observed magnitude from CPTI:** The maximum observed event(s) based on CPTI15 catalog is denoted herein as the vector $\mathbf{M}_{max,CPTI}$ (see M_w's reported in the 9th column of Table 2). The standard deviations of the observed magnitudes in $\mathbf{M}_{max,CPTI}$ are taken from the CPTI15 catalog (reported right next to the expected value of $\mathbf{M}_{max,CPTI}$ in the 11th column of Table 3). In case the maximum magnitude values for the same event (distinguished by its origin time in the 8th column of Table 2) are reported both in CPTI and DISS catalogs, two different situations are encountered: (a) The maximum magnitudes are numerically close; (a) The maximum magnitudes are not numerically close. In case (a), the $M_{max,DISS}$ values, underlined with red color, is not taken into account. In case (b), both $M_{max,DISS}$ and $\mathbf{M}_{max,CPTI}$ are taken into account (as they represent the uncertainty in the estimation of maximum magnitude). The observed values for $\mathbf{M}_{max}$ (from DISS and CPTI) and the corresponding standard deviations are reported in Table 3 (columns 8 to 11). Note that no $M_{max,DISS}$ is assigned to SBoxes 2, 7, 8, 9, and 10 (the red-underlined values in column 6, Table 2). Moreover, $\mathbf{M}_{max,CPTI}$ does not exist for the following SBoxes: (a) ITIS095 (No. 6) as the location of the corresponding event is not defined in CPTI15 catalog (denoted as NP in Table 2). (b) ITIS078 and ITIS079 (No.'s 11 and 12) since the second and third shocks of the Irpinia earthquake are not listed in CPTI15 catalog. (c) ITIS068, the low-magnitude historical earthquake (M4.26, ID#1481) is not considered because it was significantly lower than empirically-estimated value from DISS (M5.4),
3. **Calculated as a function of seismic moment:** An expression based on scalar seismic moment (M_0, [N.m]) is used for calculating M_{max} which is denoted herein as (see Hanks and Kanamori 1979):

$$M_{\max,M_0} = \frac{2}{3}\left(\log_{10} M_0 - 9.1\right) \quad (1)$$

   The seismic moment M_0 is the most physically meaningful way to describe the strength of an earthquake caused by fault slip. It is expressed as:

$$M_0 = \mu LWD \quad (2)$$

   where μ is the rigidity or shear modulus (usually taken to be ~3×10^{10} Pa), L is along-strike rupture length, W is down-dip width, and D is the average slip on the fault (these parameters are extracted from the DISS database for individual SBoxes). The standard deviation for $M_{\max,M_0}$ is set to 0.30 (see Pace et al. 2016).
4. **Calculated as a function of seismic moment with modified rupture length:** The maximum magnitude estimate denoted herein as $M_{max,ASP}$ (where ASP stands for the aspect ratio) is calculated in the same manner as in part 3 using Eq. (1). The sole difference is that the along-strike rupture length denoted herein as L_{ASP} is calculated in terms of W as (Peruzza and Pace 2002):

   (3)

   By substituting L_{ASP} in Eq. (2) in order to obtain M_0, $M_{max,ASP}$ is estimated as a function of M_0 using Eq. (1). A value of 0.25 is set for the standard deviation of $M_{max,ASP}$ (Pace et al. 2016).
5. **Calculated based on the capacity of the fault:** The maximum magnitude of a fault can also be determined based on a number of empirical relationships (denoted as the vector $\mathbf{M}_{max,ER}$). These relationships are essentially regression models predicting the magnitude expected for a given fault as a function of physical parameters such as surface rupture length, subsurface rupture length, down-dip rupture width, rupture area, and average displacement per event. Two well-known empirical relationships for tectonic contexts are Wells and Coppersmith (1994, a.k.a. WC94), and Leonard (2010). Herein, two alternative WC94 models that predict the maximum moment magnitude as a function of logarithm of the down-dip rupture width (i.e., $\log_{10}W$) and the logarithm of maximum rupture area $RA=W\cdot L$ (i.e., $\log_{10}RA$), respectively, are employed. The linear regression coefficients are fault-style-dependent. The standard deviation reported for each

component of $\mathbf{M}_{\max,\mathrm{ER}}$ is the standard error of the corresponding predictive regression model (WC94 and reported in Wells and Coppersmith 1994).

The empirically-estimated $M_{\max,M_0}$, $M_{\max,\mathrm{ASP}}$ and $\mathbf{M}_{\max,\mathrm{ER}}$ and their corresponding standard deviations are listed in Table 3 (each magnitude estimate is followed by its standard deviation, since two WC94 models were used for estimating $\mathbf{M}_{\max,\mathrm{ER}}$ two rows are dedicated to the results). Note that for the composite source ITCS063, no empirical magnitudes $\mathbf{M}_{\max,\mathrm{ER}}$ are assigned since the required physical parameters of this source (i.e., down-dip rupture width *W* and maximum rupture area *RA)* were not defined in DISS database. Recalling that a composite source may span various individual sources, the maximum magnitude capacity cannot be calculated from empirical relationships that are expressed as a function of geometrical source parameters (i.e., points 3 to 5 above). Therefore, as shown in Table 3, only the observed magnitude values $M_{\max,\mathrm{DISS}}$ and $\mathbf{M}_{\max,\mathrm{CPTI}}$ are assigned to the composite sources.

The distribution of the maximum magnitude for a given SBox is characterized based on the maximum magnitude estimates and the corresponding dispersion values obtained based on empirical relationships and observations (methods in 1-5). In this context, M_{char} denotes the average value of the maximum magnitude and σ_m represents its standard deviation. Thus,

$$M_{\mathrm{char}} = \overline{\mathbf{M}_{\max}} \tag{4}$$

$$\sigma_m = \sqrt{s^2_{\mathbf{M}_{\max}} + \frac{1}{n^2_{\mathbf{M}_{\max}}} \sum_{i=1}^{n_{\mathbf{M}_{\max}}} \sigma^2_{\mathbf{M}_{\max}(i)}} \tag{5}$$

where $\overline{\mathbf{M}_{\max}}$ is the average of $\mathbf{M}_{\max} = \left[M_{\max,\mathrm{DISS}}, \mathbf{M}_{\max,\mathrm{CPTI}}, M_{\max,M_0}, M_{\max,\mathrm{ASP}}, \mathbf{M}_{\max,\mathrm{ER}}\right]$, the latter being the vector of alternative calculated/observed maximum moment magnitude estimates (itemized above in 1-5, bold indicates that the quantity is a vector itself; referring to cases in which more than one estimate is available; see the multiple rows reported for some SBoxs in Table 3), $n_{\mathbf{M}_{\max}}$ is the length of the vector $\mathbf{M}_{\max}$, $s_{\mathbf{M}_{\max}}$ is the sample standard deviation of the calculated/observed values in $\mathbf{M}_{\max}$, and finally $\sigma_{\mathbf{M}_{\max}(i)}$ is the standard deviation reported for each individual component in $\mathbf{M}_{\max}$ (as described in points 1-5 above). Note that the expression in Eq. 5 encompasses not only the uncertainty in the estimation of $\overline{\mathbf{M}_{\max}}$, but also the dispersion due to the calculated/observed values in $\mathbf{M}_{\max}$. Eq. 5 is derived assuming that the different components of $\mathbf{M}_{\max}$ are uncorrelated.

A truncated normal distribution (with M_{char} as mean and σ_m as the standard deviation) can be generated for each SBox, which leads to a magnitude domain consistent with the variability in the empirical and observed data on each individual seimogenic source. The normal distribution is truncated at a lower magnitude equal to M_l^{SB}=5.50 in one side and at an upper magnitude M_u^{SB} lying two standard deviation above the mean value, $M_{\mathrm{char}}+2\sigma_m$ (for the upper-bound truncation, see Abrahamson 2000). Given that the maximum moment magnitude assigned in ZS9 to areal zones in the Italian territory is 7.29, and to avoid unrealistic magnitude assignments to SBoxes, an upper-bound threshold of 7.50 is assigned to all seismic sources (i.e., $M_u^{\mathrm{SB}} \leq 7.50$). The last three columns in Table 3 list the values assigned to M_{char}, σ_m, and M_u^{SB} for all sources based on all the observed and empirical magnitude values. The truncated normal probability density function (PDF) of magnitudes assigned to a SBox, and denoted herein as f_M^{SB}, can then be expressed as (see Abrahamson 2000),

$$f_M^{\mathrm{SB}}(m) = \frac{\phi\left(\frac{m - M_{\mathrm{char}}}{\sigma_m}\right)}{\Phi\left(\frac{M_u^{\mathrm{SB}} - M_{\mathrm{char}}}{\sigma_m}\right) - \Phi\left(\frac{M_l^{\mathrm{SB}} - M_{\mathrm{char}}}{\sigma_m}\right)} \tag{6}$$

where $\phi(\cdot)$ is the standard Normal PDF, and $\Phi(\cdot)$ is the standard Normal cumulative density function (CDF).

Fig. 4 illustrates graphically the information provided in Table 3 for the seismogenic sources. The red stars reflect the components of vector $\mathbf{M}_{\max}$ (observed/calculated) reported in Table 3 and the corresponding characteristic distribution is shown with gray-solid line. The distribution is mainly truncated Normal (see Eq. 6) with the lower and upper magnitudes M_l^{SB}=5.50, and $M_u^{\mathrm{SB}} \leq 7.50$ (see Table 3 for M_u^{SB} values) indicated with black vertical dotted lines. Note that if the difference between the upper and lower magnitudes is less than unity (i.e., $M_u^{\mathrm{SB}} - M_l^{\mathrm{SB}} < 1.0$), a Uniform distribution is assigned to the corresponding source (see the case of SBox 3, SBox 12, and SBox 13 in Fig. 4). It is to note that the maximum magnitude characteristic model with truncated Normal/Uniform distribution as the magnitude recurrence relation allows for a narrow range of high magnitude and does not account for small-to-moderate size earthquake occurrences along the fault. Hence, it is assumed that all the seismic energy is released through characteristic earthquakes whose magnitude is described with the truncated distributions shown in Fig. 4 in gray-colored lines. On the other hand (as noted also previously), the events with magnitudes

smaller than the threshold M_l^{SB}=5.50 are assigned to the background source (as shown in the last sub-plot of Fig. 4 with cyan-colored circles; for more details see the next Section 4.2).

Characterizing the distribution of maximum characteristic Magnitude for the two point sources:
The two points sources PS, namely "PS Nola" and "PS Sorrento peninsula" (see Fig. 3a, and also the 3rd remark in Section 4.1) are assumed to have a Uniform distribution of magnitudes in the range of [M_l^{SB}=5.50, 5.56+0.46≅6.0]. The upper-bound magnitude of 6.0 is the sum of the maximum magnitude of the assigned historical event to each source (i.e., 05 December 1499 with M_w 5.56, and 31 July 1561 with M_w 5.56) plus the error associated to each event (= 0.46) based on CPTI15.

Discussion: It is worth mentioning that the source characterization carried out in this study is more sophisticated than the areal source characterization which is the basis of the INGV national hazard maps (see Section 1, Introduction). Nevertheless, this study can be improved as more data become available about the individual boxes. For instance, if it was possible to have complete catalogues for each SBox for smaller magnitudes, there would have been probably no need for using the layered approach to source characterization adopted in this study. In such case, we could have attributed a Youngs and Coppersmith recurrence model (Youngs and Coppersmith 1985, see also Convertito et al. 2006, and Gülerce and Ocak 2013) to each SBox and the background (excluding the SBoxes) could have been attributed to a larger maximum magnitude.

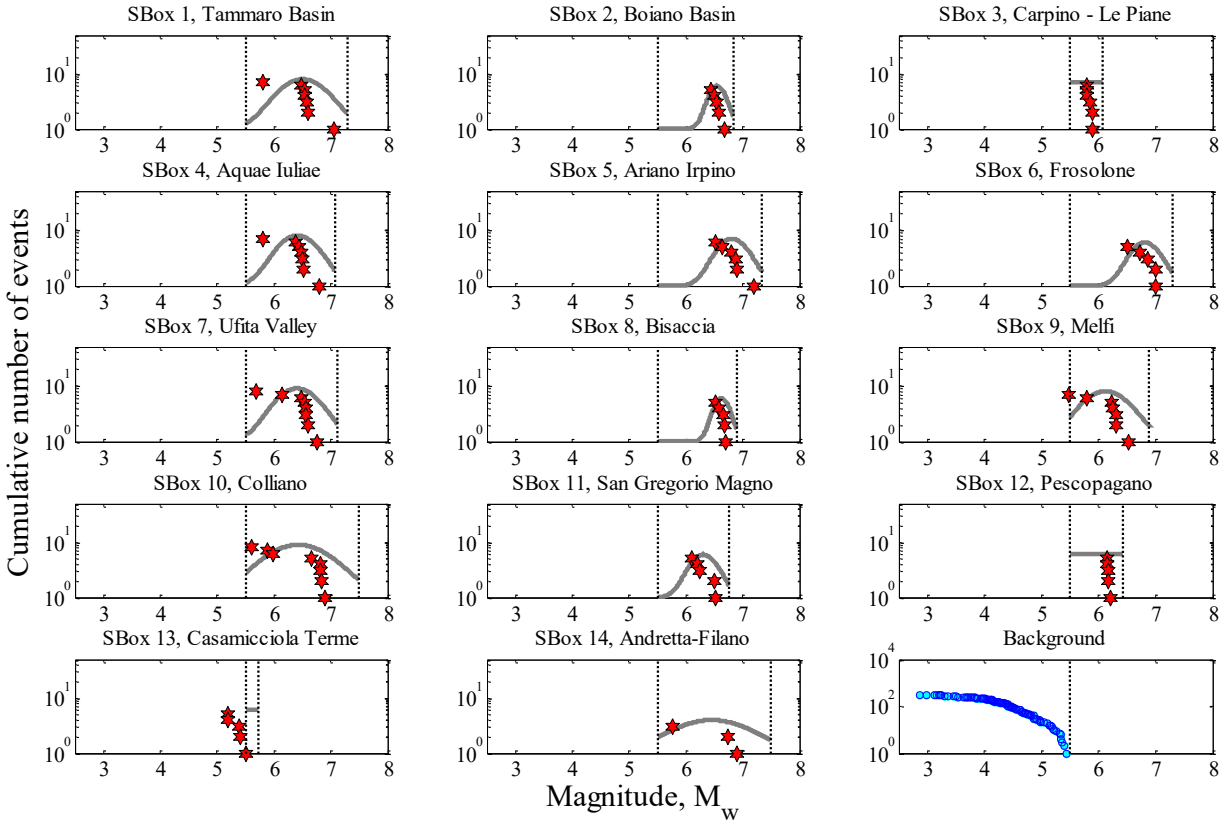


Figure 4: Distribution of maximum characteristic magnitude for each individual seismic source (gray-solid line), together with the observed/calculated maximum moment magnitudes (red stars); the last sub-plot shows all the events with M_w≤5.50 assigned to the background source. Note that the truncated normal distribution for each SBox is not necessarily the best fit to the observed / calculated maximum moment magnitudes (defined by red stars). This distribution is characterized through the detailed approach described through Section 4.1.1 and tabulated in Table 3.

Table 3: Distribution of maximum characteristic magnitude for different seismic sources

No.	DISS ID	Name	L (km)	L_{ASP} (km)	W (km)	D (m)	Observed $\mathbf{M}_{max}$ values and their standard deviations				Calculated $\mathbf{M}_{max}$ values and their standard deviations						M_{char}	σ_m	M_u^{SB}
							DISS		CPTI		M_0		ASP		ER				
1	ITIS005	Tammaro Basin	25	21.0	14.3	0.90	6.6	0.3	7.06 5.80	0.10 0.46	6.6	0.3	6.5	0.25	6.5 6.5	0.31 0.25	6.5	0.39	7.3
2	ITIS004	Boiano Basin	24	20.3	13.8	0.97			6.68	0.10	6.6	0.3	6.5	0.25	6.4 6.5	0.31 0.25	6.6	0.14	6.8
3	ITIS089	Carpino-Le Piane	8.9	12.5	7.5	0.30	5.8	0.3	5.86	0.10	5.8	0.3	5.9	0.25	5.9 5.8	0.31 0.25	5.8	0.12	6.1
4	ITIS138	Aquae Iuliae	22	19.3	13	0.90	6.5	0.3	6.80 5.80	0.17 0.46	6.5	0.3	6.5	0.25	6.4 6.4	0.31 0.25	6.4	0.32	7.1
5	ITIS092	Ariano Irpino	30	21.7	14.9	2.00	6.9	0.3	7.19	0.10	6.9	0.3	6.8	0.25	6.5 6.6	0.31 0.25	6.8	0.26	7.3
6	ITIS095	Frosolone	36	21.7	14.9	2.50	7.0	0.3			7.0	0.3	6.9	0.25	6.5 6.7	0.31 0.25	6.8	0.24	7.3
7	ITIS006	Ufita Valley	25.6	21.2	14.5	0.90			6.75 6.56 5.68 6.15	0.10 0.11 0.14 0.08	6.6	0.3	6.5	0.25	6.5 6.6	0.31 0.25	6.4	0.35	7.1
8	ITIS088	Bisaccia	31	21.8	15	0.95			6.70	0.08	6.7	0.3	6.6	0.25	6.5 6.7	0.31 0.25	6.6	0.14	6.9
9	ITIS081	Melfi	17.2	16.8	11	0.66			6.52 5.48 5.80	0.11 0.27 0.46	6.3	0.3	6.3	0.25	6.2 6.3	0.31 0.25	6.1	0.38	6.9
10	ITIS077	Colliano	28	37.6	15	1.65			6.81 5.60 5.88 5.98	0.10 0.16 0.36 0.17	6.8	0.3	6.9	0.25	6.8 6.7	0.45 0.23	6.4	0.53	7.5
11	ITIS078	San Gregorio Magno	9	21.8	15	0.70	6.2	0.3			6.2	0.3	6.5	0.25	6.5 6.1	0.31 0.25	6.3	0.23	6.8
12	ITIS079	Pescopagano	15	15.6	10	0.50	6.2	0.3			6.2	0.3	6.2	0.25	6.2 6.1	0.31 0.25	6.2	0.13	6.4
13	ITIS068	Casamicciola Terme	5	7.5	3.5	0.30	5.4	0.3			5.4	0.3	5.5	0.25	5.2 5.2	0.31 0.25	5.3	0.19	5.7
14	ITCS063	Andretta-Filano					6.9	0.3	6.73 5.76	0.10 0.08							6.5	0.62	7.5

4.1.2 Calculating the activity rate for different seismic sources

The activity rate can be expressed in terms of mean annual rate of events with magnitude greater than or equal to M_l^{SB}=5.50 for each individual seismic box. After quantifying the distribution of maximum characteristic magnitude of individual fault segments, i.e. f_M^{SB} from Eq. (6), the corresponding activity rate can be obtained using the geometric and kinematic parameters assigned to each SBox. To this end, one can estimate the mean occurrence time of the characteristic magnitude, T_{mean}, assigned to each SBox by means of the methodology presented by Peruzza et al. (2010) and implemented by Pace et al. (2016). The procedure uses the criterion of "segment seismic moment conservation" proposed by Field et al. (1999) by dividing the seismic moment that corresponds to $M_{\max}$ by the moment rate given a slip rate S as follows:

$$T_{\mathrm{mean}} = \frac{10^{1.5M_{\max}+9.1}}{\mu LWS} \tag{7}$$

where T_{mean} is known as the mean recurrence time in year. Therefore, $1/T_{\mathrm{mean}}$ can be interpreted as the mean annual rate of occurrence of the characteristic event. To estimate the uncertainty in T_{mean}, one needs to consider the uncertainties in $M_{\max}$ (truncated Normal distribution with mean M_{char} and standard deviation σ_m or Uniform distribution as shown in Fig. 4), the slip rate S (Uniform distribution as given in DISS database, see Table 4) and the uncertainties in μ, L, and W (if such uncertainties are taken into account, herein they are assumed to be deterministic quantities). While Pace et al. (2016) employs a mean-value first-order second-moment (MVFOSM) approach for estimating the uncertainty in T_{mean}, we use herein the Monte Carlo simulation method considering zero correlation between the parameters $M_{\max}$ and S. Fig. 5 shows the distribution of T_{mean} for SBox 1 up to SBox 13 based on Eq. (7). In addition, to have an estimate for the mean recurrence time for each seismic source from the distribution of T_{mean}, we have selected the 16$^{\mathrm{th}}$ percentile (marked with red star) as the best-estimate value. Note that the inverse of this value corresponds to the 84$^{\mathrm{th}}$ percentile of the mean annual rate of the characteristic events associated with each SBox denoted as $v_{M_{\max}}^{84\mathrm{th}}$. For all the individual seismic sources, the assigned $v_{Mmax}^{84\mathrm{th}}$ is reported in Table 4.

Discussion about the composite source ITCS063 (SBox 14): For the composite source ITCS063, estimating the seismicity rate with respect to Eq. (7) is not possible, as the geometric and kinematic parameters of the source are not identified in DISS. In order to obtain a rough estimate for the seismicity of this composite source in terms of the annual rate of events with $M \geq M_l^{\mathrm{SB}}$, one can refer to Table 2. This source is shown to be responsible for two events with $M > M_l^{\mathrm{SB}}$; i.e., M_w 6.73 in 1694 and M_w 5.76 in 1910. Based on the Italian seismic zonation, this composite source is mainly located within the seismic zone Z927 (see Fig. SM7 in Section S.M.4 of the electronic supplementary material of this manuscript). With reference to the completeness period for Z927 (Gruppo di Lavoro 2004), for 5.68≤M_w=5.76≤5.91, the completeness period of the zone starts from the year 1787. Moreover, for 6.60≤M_w=6.73≤6.83, the completeness period starts from 1530. As a result, the completeness period for events with $M > M_l^{\mathrm{SB}}$=5.50 can be assumed to start from 1787. Thus, considering that the time extent of CPTI15 catalog is up to 2014, the approximate annual seismicity rate of this composite source, denoted herein as $v_{app}(M_l^{\mathrm{SB}})$, can be calculated as the number of events with $M > M_l^{\mathrm{SB}}$ within the completeness period divided by the time span; i.e., $\frac{1}{2014-1787} = 0.0044$, as outlined in Table 4 for this composite source.

Finally, the assigned annual seismicity rate to each SBox, denoted as v^{SB}, is assumed herein to be $v_{Mmax}^{84\mathrm{th}}$ for individual SBoxes and $v_{app}(M_l^{\mathrm{SB}})$ for the composite source, as reported in Table 4. As a final note, it is also possible to estimate $v_{app}(M_l^{\mathrm{SB}})$ for individual sources; however, the number of events assigned to each SBox is too small for making meaningful verifications of the completeness time interval for the calculation of $v_{app}(M_l^{\mathrm{SB}})$. Hence, relying on the geometric and kinematic parameters of the individual SBoxes to estimate the seismicity rate (by means of Eq. 7) sounds more rational.

Calculating the seismicity rate for the two point sources:

To have an estimate for the seismicity rate of the two point sources, namely "PS Nola" and "PS Sorrento peninsula", it is to recall the assigned characteristic event of M_w 5.56 in 1499 and 1561 respectively (events 238 and 348 in CPTI15). The point source Nola is located within the Italian seismic zone Z928 and the source Sorrento peninsula is close to this seismic zone (see Fig. SM7 in Section S.M.4 of the electronic supplementary material of this manuscript). For 5.45≤M_w=5.56≤5.68, the completeness period of this zone starts from the year 1787 (Gruppo di Lavoro 2004), which does not include the two characteristic events. Alternatively, to have a rough estimate of seismicity rate, i.e. annual rate of events with $M \geq M_l^{\mathrm{SB}}$=5.50 denoted as $v\left(M \geq M_l^{\mathrm{SB}}\right)$, for these two point sources, we employ directly the filtered Poisson rate (based on the seismicity rate of zone Z928) as follows: $v\left(M \geq M_l^{\mathrm{SB}}\right) = v(M \geq M_l^{\mathrm{zone}}) \times \exp\left[-\beta\left(M_l^{\mathrm{SB}} - M_l^{\mathrm{zone}}\right)\right]$. Accordingly, M_l^{zone}=4.76 is the lower-bound magnitude, β=2.40, and the rate $v(M \geq M_l^{\mathrm{zone}})$=0.21 for zone Z928 (Gruppo di Lavoro 2004); thus, we have $v\left(M \geq M_l^{\mathrm{SB}}\right)$=0.0357.

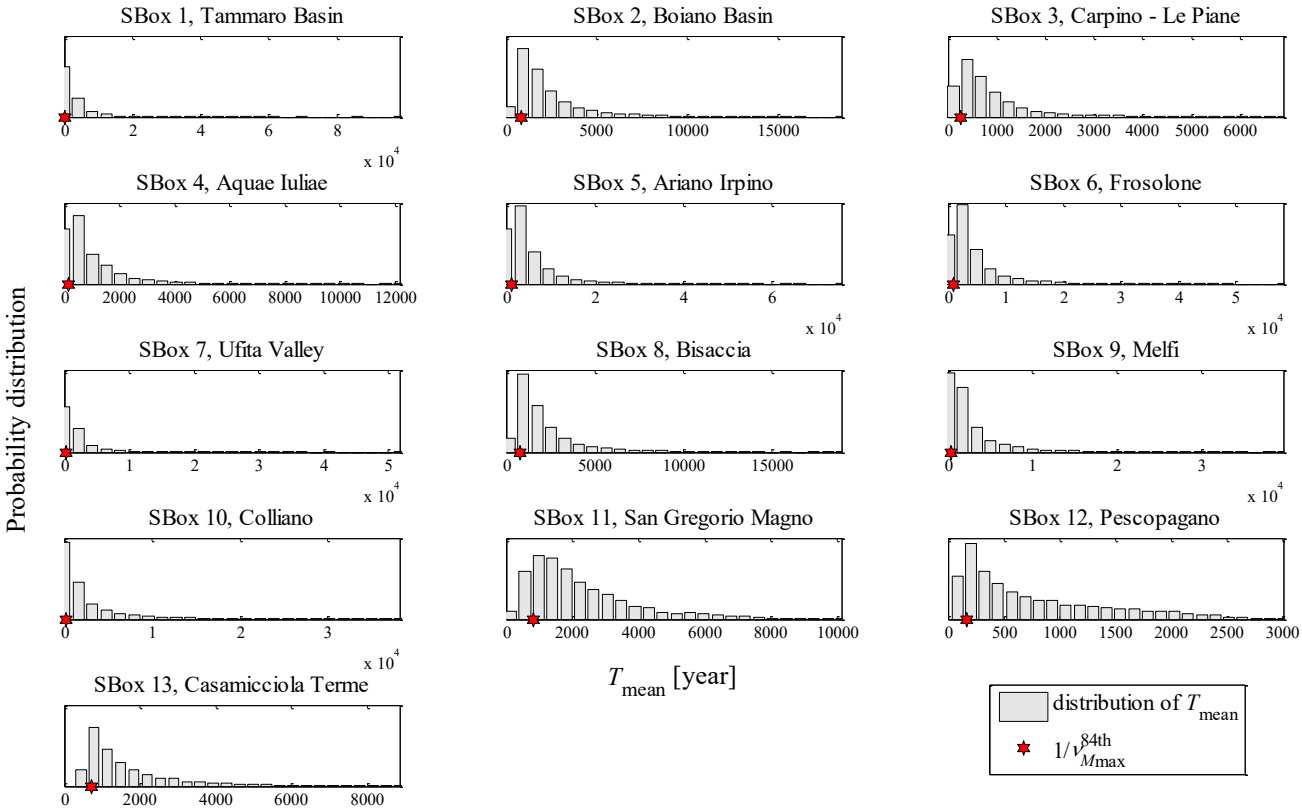


Figure 5: The distribution of mean recurrence time of the characteristic event, T_{mean}, for all seismic sources (see Eq. 7), and the 16th percentile of T_{mean} ($1/\nu_{M_{max}}^{84th}$) marked with red star

Table 4: The mean annual seismicity rate for different individual seismic sources (SBoxes)

No.	DISS ID	Name	Slip rate S (mm/ys)	$\nu_{M_{max}}^{84th}$s	$\nu_{app}(M_l^{SB})$	ν^{SB}
1	ITIS005	Tammaro Basin	[0.1, 1]	0.0029	–	0.0029
2	ITIS004	Boiano Basin	[0.1, 1]	0.0012	–	0.0012
3	ITIS089	Carpino-Le Piane	[0.1, 1]	0.0039	–	0.0039
4	ITIS138	Aquae Iuliae	[0.45, 1.90]	0.0059	–	0.0059
5	ITIS092	Ariano Irpino	[0.1, 1]	8.91e-4	–	0.0009
6	ITIS095	Frosolone	[0.1, 1]	0.0010	–	0.0010
7	ITIS006	Ufita Valley	[0.1, 1]	0.0038	–	0.0038
8	ITIS088	Bisaccia	[0.1, 1]	0.0012	–	0.0012
9	ITIS081	Melfi	[0.1, 0.5]	0.0026	–	0.0026
10	ITIS077	Colliano	[0.4, 0.6]	0.0056	–	0.0056
11	ITIS078	San Gregorio Magno	[0.4, 0.6]	0.0012	–	0.0012
12	ITIS079	Pescopagano	[0.4, 0.6]	0.0059	–	0.0059
13	ITIS068	Casamicciola Terme	[0.1, 1]	0.0014	–	0.0014
14	ITCS063	Andretta-Filano		NA	0.0044	0.0044

4.2 Background Source Model

The background source is shown in Fig. 3a with a light-blue rectangular area. The magnitude distribution for the background source, denoted as f_M^{BG} follows a truncated Exponential distribution as (see also Jalayer et al. 2011, Ebrahimian et al. 2014):

$$f_M^{BG}(m) = \frac{\beta e^{-\beta m}}{e^{-\beta M_l^{BG}} - e^{-\beta M_l^{SB}}} \tag{8}$$

where $M_l^{BG} \leq m \leq M_l^{SB}$=5.50, M_l^{BG} is the lower cut-off magnitude of the background area, β is the slope of the Gutenberg-Richter (GR) earthquake rate model (Gutenberg and Richter 1949, 1956). This frequency-magnitude

model has the log-linear relationship of the form $\ln N(M \geq m) = \alpha - \beta \cdot m$ where N is the number of events with magnitudes greater than or equal to the value m.

4.2.1 Estimating the lower cut-off magnitude of the background, M_l^{BG}

The first step toward quantifying f_M^{BG} (see Eq. 8) is to have an estimate of M_l^{BG}. As shown in Fig. 3b, the background area is located mostly within the seismic zones Z927 and Z928 according to Italian Seismogenetic Zonation (ZS9, Gruppo di Lavoro 2004). The minimum magnitude of both seismic zones according to ZS9 is 4.76. Nevertheless, with reference to the historical events within the background area, the issue of catalog incompleteness should be explicitly considered (Stucchi et al. 2004). Specifically, the designated lower cut-off magnitude of the background area M_l^{BG} should be greater than or equal to M_c, i.e., $M_l^{BG} \geq M_c$, where M_c denotes the completeness magnitude. There exist different approaches in the literature for estimating M_c (see e.g., Marzocchi et al. 2016). Herein, we have employed two different methods described in the following:

1. **Direct use of frequency-magnitude distribution plot:** with reference to the GR relationship, the frequency-magnitude curve should have an approximately exponential decrease as the magnitude increases (i.e., linearly decreasing in logarithmic scale). In case the data is incomplete, a flattening in a certain lower magnitude range (having higher frequencies) can be detected. In such cases, M_c can be visually marked as the point where the magnitude-frequency curve becomes approximately linear in the logarithmic y-scale (see Ebrahimian et al. 2014). Fig. 6a shows the frequency-magnitude scatter plot (the same previously shown in the last sub-plot of Fig. 4) consisting of data pairs $[m, \ln N(M \geq m)]$ shown with cyan-colored circles (according to CPTI15 catalog). The upper magnitude threshold M_l^{SB}=5.50 is marked with a black-dotted line. The scatter plot reaches saturation at a certain lower magnitude range; this is while it demonstrates approximately linear behaviour for higher magnitudes. M_c is marked at the point where the frequency-magnitude curve becomes (roughly) a line. As it can be observed in the figure, it is not straightforward to visually pin-point the M_c corner. More specifically, M_c can vary in the range of magnitudes between 4.0 and 4.8 (herein it is set conservatively to 4.8; compatible with ZS9 data). The next method can lead to quantified estimates of M_c.
2. **Bayesian updating approach for calculating the β-value versus various magnitude thresholds denotes as m_l:** the method detects the magnitude threshold where the maximum likelihood of the posterior probability distribution of β (mode of the distribution), denoted as $\hat{\beta}$ herein, becomes roughly invariant (i.e., the estimate rendered by the Bayesian updating for β parameter is quasi invariant with respect to the adopted m_l). This magnitude threshold can be interpreted as M_c (see Ebrahimian et al. 2014; Ebrahimian and Jalayer 2017). The posterior probability distribution for β given the data **D** and a magnitude threshold m_l is denoted herein as $p(\beta|\mathbf{D},m_l)$. The data **D** consist of i=1:N^{BG} historical earthquakes from the catalog within the background area with magnitudes $m_i \geq m_l$. The probability $p(\beta|\mathbf{D},m_l)$ can be determined according to the Bayes's theorem as (see also Ebrahimian et al. 2014, Ebrahimian and Jalayer 2017):

$$p\left(\beta|\mathbf{D},m_l\right) = c^{-1} p\left(\mathbf{D}|\beta,m_l\right) p\left(\beta|m_l\right) = c^{-1}\left(\prod_{i=1}^{N^{BG}} p\left(m_i|\beta,m_i \geq m_l\right)\right) p\left(\beta\right) = c^{-1}\left(\prod_{i=1}^{N^{BG}} \beta e^{-\beta(m_i-m_l)}\right) p\left(\beta\right) \quad (9)$$

where c^{-1} is the normalizing constant of the Bayes's expression; $p(\mathbf{D}|\beta,m_l)$ is the likelihood function for data **D** given β and m_l; $p(\beta|m_l) \approx p(\beta)$ is the prior probability distribution. The term $p(m_i|\beta, m_i \geq m_l)$ is the conditional probability of having an event with magnitude equal to m_i that is greater than or equal to m_l. Herein, we use a Lognormal probability distribution to define the prior $p(\beta)$ having median equal to ln10 and COV equal to 0.30. Fig. 6b illustrates $\hat{\beta}$-value calculated as the maximum likelihood (mode) of posterior probability distribution $p(\beta|\mathbf{D},m_l)$, with respect to various magnitude thresholds, m_l. It can be seen that $\hat{\beta}$ increases monotonically with respect to m_l up to a value around 4.40, after which it become roughly invariant; thus, $M_c \approx 4.40$. This value is highlighted with a red-dotted vertical line in both Fig. 6a and Fig. 6b. According to Fig. 6b, we set M_l^{BG}=4.60 (highlighted with cyan-dashed line) to bypass the initial fluctuation in $\hat{\beta}$-value above M_c. The GR regression line, starting from M_l^{BG}=4.60, is also plotted with gray (thick) solid line in Fig. 6a to highlight the linear trend in the frequency-magnitude scatter plot.

Once M_l^{BG} is estimated, the next (and last) parameter required for calculating the background magnitude distribution, f_M^{BG}, is β (i.e., the GR seismicity slope in the logarithmic scale).

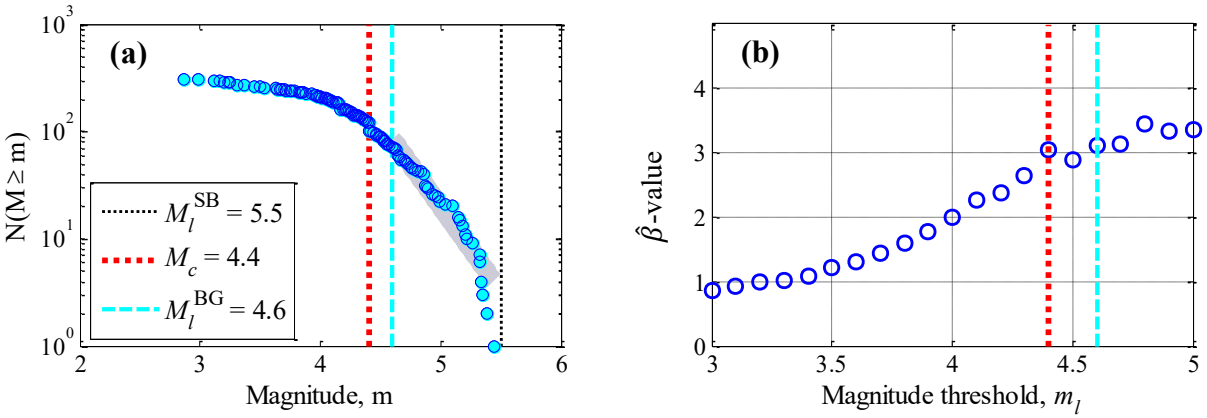


Figure 6: (a) Frequency–magnitude distribution of the background area based on CPTI15 catalog, (b) lower magnitude versus the estimated $\hat{\beta}$-value through Bayesian updating

4.2.2 Estimating the completeness interval of the background for $M_l^{BG} \leq m \leq M_l^{SB}$

To make a sound estimate of the seismicity rate, it is essential to identify a time interval in which the catalogue is complete. This section describes how such interval, known as the *completeness interval* is identified. More specifically, the completeness interval is the time interval in which the magnitude range assigned to the background, i.e., M_l^{BG}=4.60 ≤ m ≤ M_l^{SB}=5.50, is likely to be completely reported. In this study, we use two different methods to address this issue: (a) the *Visual Cumulative Method* proposed by and Mulargia et al. (1987); and (b) the statistical approach called herein as *Stepp Method* and proposed by Stepp (1972). These two methods have been widely applied in regional PSHA studies (e.g., Benito et al. 2010, Ornthammarath et al. 2011, El-Hussain et al. 2012, Ksentini and Romdhane 2014, Woessner et al. 2015, Şeşetyan et al. 2016, Mihaljević et al. 2016, Parra et al. 2016, Akinci et al. 2016, Kadirioğlu et al. 2016, Danciu et al. 2016, Demircioğlu 2017, Sokolov et al. 2017).

In the visual cumulative method, the cumulative number of events in the desired magnitude class is plotted as a function of time. The time interval with the largest apparent slope is visually identified and taken as completeness period. Fig. 7a illustrates the plot of the cumulative number of seismic events in the magnitude interval [M_l^{BG} =4.60, M_l^{SB} =5.50], located within the background region, versus time corresponding to the entire time interval of CPTI15 catalog which is [1000, 2014]. It is not straightforward to pick a point, as there are a few apparent increases in slope due to the change in seismicity rate. The first marked increase is around 1700 and the second one is around 1900. However, the sharp slope is more pronounced in the time interval 1900-2014.

The Stepp's method relies upon the statistical characteristics of a Poisson distribution for the occurrence of earthquakes in the complete time interval. For a given magnitude, starting backward in time from the end of catalog (herein, from the year 2014), assume that the number of events occurring in the unit time intervals (say, e.g., 1-year) are λ_i where i=1:T≤2014-1000=1014 (years); T is the length of the time interval. Let λ denote the average of the random variables λ_i in the time interval T:

$$\lambda = \frac{1}{T}\sum_{i=1}^{T}\lambda_i = \frac{N_{\text{obs},T}}{T} = \lambda_{\text{obs}} \tag{10}$$

where λ_i is the number of events in the unit time interval [i, i+1] and $N_{\text{obs},T}$ is the total number of events in the time interval T. The variance of λ denoted as $\sigma^2(\cdot)$, considering a Poissonian occurrence of events, can be estimated as:

$$\sigma^2(\lambda) = \frac{1}{T^2}\sigma^2\left(\sum_{i=1}^{T}\lambda_i\right) = \frac{1}{T^2}\left(\sum_{i=1}^{T}\sigma^2(\lambda_i)\right) = \frac{1}{T^2}\left(\sum_{i=1}^{T}\lambda_i\right) = \frac{1}{T^2}(\lambda\cdot T) = \frac{\lambda}{T} \tag{11}$$

Eq. (11) indicate that if the occurrence of seismic events in a given time interval follows a homogenous Poisson distribution with rate λ, the standard deviation σ varies proportional to $1/\sqrt{T}$. As a result, if the σ estimate does not show $1/\sqrt{T}$-proportional behavior, this can be used as a sign that the corresponding time interval is not properly modelled by a homogenous Poisson process with parameter λ. The time interval length T, in which no deviation from the curve $1/\sqrt{T}$ occurs, defines the completeness time interval for the corresponding magnitude range. Fig. 7b shows the curve $\sigma_\lambda = \sqrt{\lambda_{\text{obs}}}/\sqrt{T}$ with black hollow circles and the curve proportional to $1/\sqrt{T}$ with blue dashed line (in logarithmic scale). The first part of σ_λ illustrates a sharp jump exactly at T=34 years (i.e., corresponding to the interval [1980, 2014]). This is due to the sharp increase in the number of registered events in 1980 due to events related to the Irpinia Earthquake. Fig. 7a also registers an abrupt jump in seismicity at 1980. Careful visual detection of Fig. 7b reveals a deviation from $1/\sqrt{T}$-consistent behavior after an interval of

$T \approx 110$ years corresponding to the year 1904. This is in line with the results obtained from the visual cumulative method. Moreover, it agrees with the Italian seismic zonation (Gruppo di Lavoro 2004) where completeness intervals starting from 1895 and 1871 for seismic zones Z927 and Z928 for magnitudes greater than or equal to 4.76 (the minimum magnitude assigned to both seismic zones) are assigned, respectively. Thus, the time interval of 1900-2014 is considered as the completeness period of the catalog for the desired background area. This completeness interval (i.e., [1900, 2014]) is also shown with black dash-dotted line in Fig. 7.

It is noteworthy that σ_λ can also be estimated based on a Bayesian approach by calculating the posterior probability distribution of λ, denoted as $p(\lambda|data)$, given **data** consisting of $N_{\text{obs},T}$ observed events in time interval T. This posterior probability distribution can be estimated using Bayes's theorem as follows:

$$p\left(\lambda|\mathbf{data}\right) = p\left(\lambda|N_{\text{obs},T},T\right) = c^{-1}p\left(N_{\text{obs},T}|\lambda T\right)p\left(\lambda|T\right) = c^{-1}\left(\frac{\left(\lambda T\right)^{N_{\text{obs},T}}e^{-\lambda T}}{N_{\text{obs},T}!}\right)p\left(\lambda\right) \tag{12}$$

where c^{-1} is a normalizing constant; $p(N_{\text{obs},T}|\ \lambda T)$ is the likelihood of observing $N_{\text{obs},T}$ events during the time interval T with the rate λ (i.e., $p(\mathbf{data}|\lambda)$) which can be expressed by a Poisson distribution; the term $p(\lambda|T) \approx p(\lambda)$ is the prior probability distribution for rate λ. Herein, a Uniform prior is adopted for $p(\lambda)$. Fig. 7b illustrates the standard deviation of the posterior probability distribution $p(\lambda|\mathbf{data})$, denoted as $\sigma_{\lambda|data}$. The trend in the curve $\sigma_{\lambda|data}$, shown in Fig. 7b with a red line, reveals that $\sigma_{\lambda|data}$ and σ_λ are quite close; especially for $T \geq 34$ years. This results in the identification of the same completeness period by Bayesian inference.

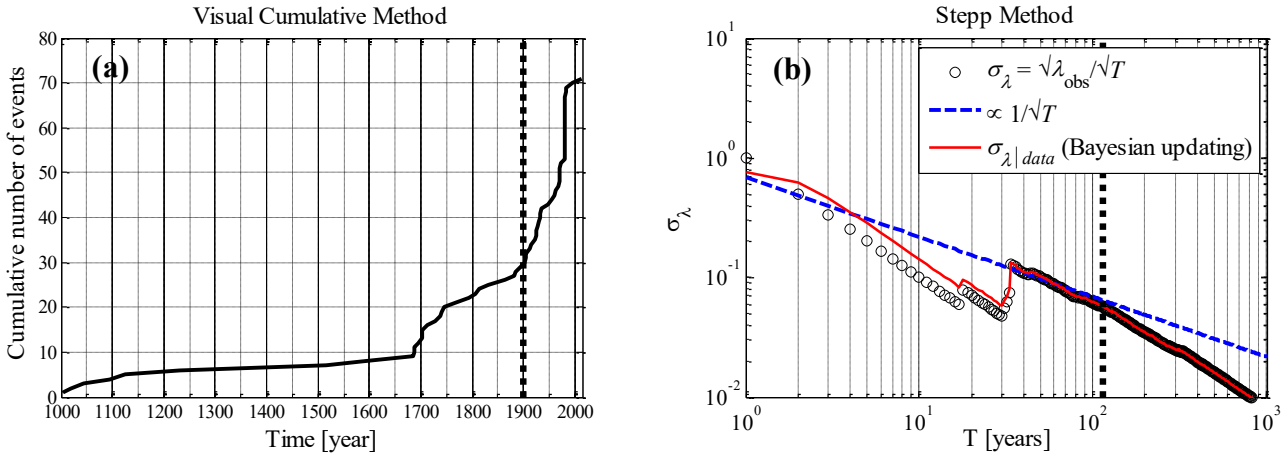


Figure 7: Analysis of completeness of the catalog for events in the background area within the magnitude range $[M_l^{\text{BG}}=4.60, M_l^{\text{SB}}=5.50]$ using (a) the visual cumulative method, (b) Stepp's method

4.2.3 Calculating the activity rate for the background source

After assigning [1900, 2014] as the completeness interval with $T=114$ (year), corresponding to the magnitude range [4.60, 5.50] for the background zone, one can estimate the annual rate of events within the background area as $v\left(M_l^{\text{BG}} \leq M \leq M_l^{\text{SB}}\right) \triangleq v^{\text{BG}} = 0.37$ based on Eq. (10). Note that the Italian seismic zonation (Gruppo di Lavoro 2004) assigns the values $v(4.76 \leq M \leq 7.06)=0.43$ to the seismic zone Z927 and $v(4.76 \leq M \leq 5.91)=0.21$ to Z928 (the minimum magnitude of 4.76 is assigned to both seismic zones while the upper magnitudes are different).

4.2.4 Calculating the GR slope β

Fig. 8a shows a frequency-magnitude scatter plot of events (cyan circles) occurring in the completeness period [1900, 2014] within the background area having $M \geq M_l^{\text{BG}}$. A GR model in semi-logarithmic scale is fitted to the frequency-magnitude scatter plot. The slope of the linear regression model is $\beta_{\text{reg}}=3.61$. Alternatively, the slope parameter can be estimated by employing the Bayesian inference (described in part 2 of Section 4.2.1, Eq. (9)). Fig. 8b demonstrates the prior probability distribution function (PDF) of β, denoted in Eq. (9) as $p(\beta)$, in gray dashed line. $p(\beta)$ is set as a Lognormal probability distribution with median equal to ln10 and COV equal to 0.30. The figure also shows the posterior PDF $p(\beta|\mathbf{D}, m_l=M_l^{\text{BG}})$ with black solid line. The maximum likelihood (mode) of the posterior distribution $p(\beta|\mathbf{D},m_l=M_l^{\text{BG}})$ is marked by a red asterisk on the posterior PDF at $\beta_{\text{ML}}=2.99$. In this study, we have used β_{ML} for estimating f_M^{BG} in Eq. (8). The Italian seismic zonation (Gruppo di Lavoro 2004) sets $\beta=1.70$ for seismic zone Z927 with $4.76 \leq M \leq 7.06$ and $\beta=2.40$ for Z928 with $4.76 \leq M \leq 5.91$. Table 5 shows the seismicity parameters corresponding to the background area (as delineated in this study in Fig. 3b) and those corresponding to Italian seismic zones Z927-Z928.

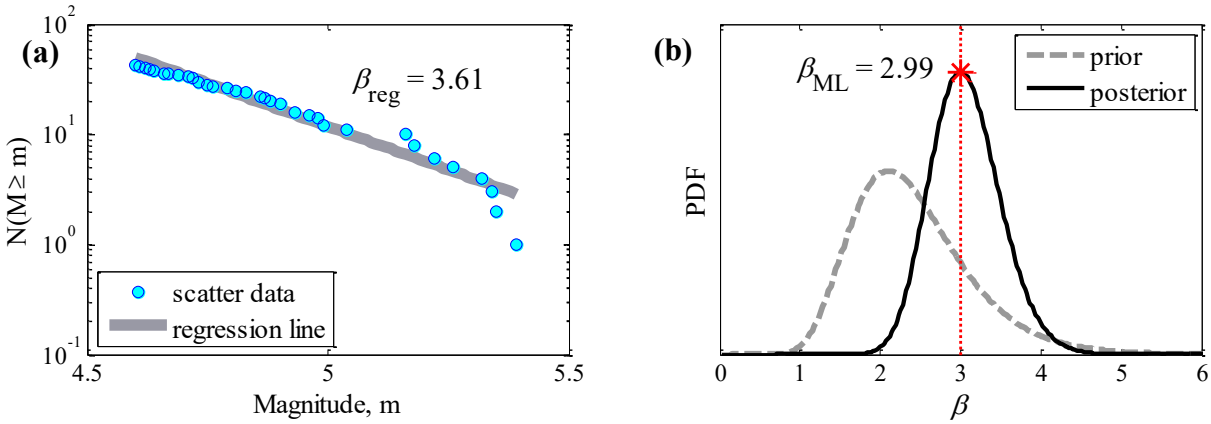


Figure 8: Calculation of GR-model slope β of the events and lying in the completeness period of [1900, 2014] within the background area with magnitudes greater than or equal to M_l^{BG}=4.60; (a) using frequency-magnitude scatter plot, (b) using the mode of the posterior distribution of β

Table 5: The key seismicity parameters corresponding to the background area in comparison with the Italian Seismic zonation

Zone	Lower Magnitude, M_l	Upper Magnitude, M_u	Gutenberg-Richter slope β	Annual rate of events with $M_l \leq M \leq M_u$
Background Area	4.60 (M_l^{BG})	5.50 (M_l^{SB})	2.99	0.37 (υ^{BG})
Z927*	4.76	7.06	1.70	0.43
Z928*	4.76	5.91	2.40	0.21

* Italian seismic zonation (Gruppo di Lavoro 2004)

5 The Ground Motion Prediction Models

In the context of PSHA, ground-motion prediction equations (GMPEs) play a key role. A GMPE represents the probability distribution for a given ground shaking parameter at a designated site as a function of source/site characteristics such as magnitude, source-to-site distance, style of faulting, and soil-site conditions. The increasing availability of strong-motion data has led to significant improvements in ground-motion prediction worldwide and especially in Italy. With reference to the detailed studies recently performed on scoring of GMPEs for probabilistic seismic hazard analysis in Italy (see Barani et al. 2017a, b), the overall performance of selected GMPEs in the evaluation of seismic hazard in Italy has been ranked. Accordingly, the following GMPEs are selected for this study (Table 6 summarizes basic features of these GMPEs):

(1) **ITA10** (Bindi et al. 2011): This GMPE is derived based on the improved ITalian ACcelerometric Archive (ITACA; Luzi et al. 2008, Pacor et al. 2011; http://itaca.mi.ingv.it) and is the updated version of the GMPE **ITA08** (Bindi et al. 2010). Although ITA08 is slightly more accurate at larger spectral periods, it is shown (Barani et al. 2017b) that the model based on ITA10 is more accurate for higher frequencies (i.e., PGA), possibly due to the anelastic attenuation term included in the distance function.

(2) **BND14** (Bindi et al. 2014a, b): This Pan-European GMPE is derived from the reference database for seismic ground-motion prediction in Europe (RESORCE, Akkar et al. 2013) consisted of the earthquakes occurred in Europe and Middle East from 1967 to 2011. BND14 has been found to be the most effective one among a pool of selected Italian, European and global GMPEs (see Barani et al. 2017a, b).

(3) **BSSA** (Boore et al. 2014): This Global GMPE is derived within the PEER Next Generation Attenuation-West project (NGA-West, Bozorgnia et al. 2014, see also Gregor et al. 2014) from the reference NGA-West2 database (Ancheta et al. 2014). This model, derived for finite fault sources, provides regional adjustments for Italy. BSSA did not perform as well as the other two models, i.e. ITA10 and BND14, in the ranking done by Barani et al. (2017a, b). Nevertheless, it benefits from a more sophisticated site amplification function compared to other two models (will be discussed subsequently).

Table 6: The ground-motion prediction equations (GMPEs) used in this study

GMPE	Number of Events	Number of Recordings	Magnitude range	Distance type and range (km)	Ground-motion parameters	Horizontal Component	Site Classification	Style of Faulting	Region
ITA10	99	769	$4.0 \leq M_w \leq 6.9$	R_{JB}, $R_{epi} \leq 200$ [4]	PGA (cm/s^2), PGV (cm/s), and 5%-damped $S_a(T)$ [cm/s^2] over 20 periods $T \in$ [0.04s, 2s] [6]	G.M.$_{xy}$ [7]	Based on EC8 [9]	normal (N), reverse (R), strike slip (SS), and unknown (U)	Italy
BND14	255 [1] 365 [2]	1224 2126	$4.0 \leq M_w \leq 7.6$	R_{JB}, $R_{epi} \leq 300$ [5]	PGA (cm/s^2), PGV (cm), and 5%-damped $PS_a(T)$ [cm/s^2] over 23 periods $T \in$ [0.02s, 3s]	G.M.$_{xy}$	Based on EC8[2, 9] and also V_{S30} [1]	normal (N), reverse (R), strike slip (S), and unknown (U, only for dataset *DS-EC8* [2])	Europe and the Middle East
BSSA	> 300	> 12000	$3.0 \leq M_w \leq 8.5$[3] $3.0 \leq M_w \leq 7.0$	$R_{JB} \leq 400$	PGA (g), PGV (cm/s), and $PS_a(T)$ [g] over 105 periods $T \in$ [0.01s, 10s]	RotD50 [8]	Based on V_{S30} [10]	normal-slip (NS), reverse-slip (RS), strike-slip (SS), and unspecified (U)	Global (with regional adjustment for Italy)

[1] BND14 employs two data sets (DS's), namely *DS-EC8* and *DS-VS30*. The number of events equal to 255 corresponds to the *DS-*V_{S30} data set that composed of waveforms recorded by stations with measured average shear-wave velocity of the uppermost 30 m, V_{S30}, mainly in the range 100–1000m/s and corresponding to events with known focal mechanism.

[2] This number of events corresponds to *DS-EC8* data set composed of waveforms recorded by stations characterized by EC8 site classes (CEN 2004). Waveforms from events with unspecified focal mechanism (U) are included herein. EC8 classes[10] A-D are accounted for in the BND14 model, as only few stations (less than 5) are classified as class E in the RESORCE database.

[3] The magnitude range for Strike-Slip and Reverse-Slip earthquakes is 3.0-8.5; it is 3.0-7.0 for Normal-Slip earthquakes.

[4] The site-to-source distance R for ITA10 is the Joyner-Boore distance (R_{JB}), or the epicentral distance (R_{epi}) when $M_w < 5.5$ (i.e., the fault geometry is unknown)

[5] The site-to-source distance R for BND14 is the Joyner-Boore distance (R_{JB}), or the epicentral distance (R_{epi}) when $M_w < 5.0$ and $R_{epi} \geq 10$ km

[6] *PGA*: peak ground acceleration, *PGV*: peak ground velocity, $S_a(T)$: acceleration response spectra at period T, $PS_a(T)$: pseudo-acceleration response spectra at period T (this is the spectral acceleration used in structural engineering applications); note that the differences between PS_a and S_a is small for short-period systems and is of some significance only for long-period systems with large values of damping. Thus, for a wide range of conditions the pseudo-acceleration may be treated as an approximation to the true acceleration (Chopra 2012).

[7] G.M.$_{xy}$: the geometrical mean of the two as-recorded horizontal components x and y

[8] RotD50 (Boore 2010): the median single-component horizontal ground motion across all azimuths

[9] Site class A: $V_{S30} > 800$ m/s; class B: $V_{S30} = 360-800$m/s; class C: $V_{S30} = 180-360$m/s; class D: $V_{S30} < 180$m/s; class E: 5-20m of C- or D-type alluvium underlain by stiffer material with $V_{S30} > 800$m/s. In the new release of the ITACA database (used in ITA10 model), site classes D and E are poorly sampled, while the other classes are well represented.

[10] The database contains V_{S30} in the range from 150-1500m/s.

The functional form of both ITA10 and BND14 are as follows:

$$\log_{10} IM = \underbrace{\left[e_1 + F_D\left(R,M\right) + F_M\left(M\right) + F_S + F_{sof} \right]}_{\mu_{\log_{10} IM}} + \varepsilon_n \sigma_{\log_{10} IM} \quad (13)$$

where *IM* denotes the ground-motion intensity measure of interest (see Table 6); e_1 is a constant term, $F_D(R, M)$ is the distance function; F_M is the magnitude scaling; F_S is the site amplification function; F_{sof} is the style-of-faulting correction; ε_n is the fractional number of logarithmic standard deviations of a single predicted value of $\log_{10} IM$ away from the mean value (indicated in the bracket in Eq. 13, and denoted as $\mu_{\log_{10} IM}$); finally, $\sigma_{\log_{10} IM}$ is the total logarithmic standard deviation of the model.

Both $F_D(R, M)$ and F_M have the same functional form in ITA10 and BND14 models; $F_D(R, M)$ features a term decreasing linearly with distance (geometric and anelastic attenuation), and $F_M(M)$ considers a fixed *hinge magnitude* M_h=6.75 (see Bindi et al. 2011 and Bindi et al. 2014a for the functional forms). The model derived for F_S in ITA10 and BND14 features a linear site amplification term although nonlinear site effects are expected to be important in soft soil sites near the seismic rupture area with low V_{S30} values (classes C and D of EC8). Regarding the linear site amplification term, the two GMPEs suggest the expressions reported in Table 7:

Table 7: Linear site amplification terms in ITA10 and BND14

GMPE	Expression	Site Classification	Definitions
ITA10	$F_s = s_j C_j$	Based on EC8	for j=1:5 defining classes A-E, s_j are regression coefficients, and C_j are dummy variables used to denote the different site classes
BND14	$F_s = s_j C_j$	Based on EC8	For j=1:4 defining classes A-D
BND14	$F_s = \gamma \log_{10}\left(V_{S30}/V_{ref}\right)$	Based on V_{S30}	V_{ref}=800m/s and γ is a regression coefficient

The functional form for F_{sof}, which represents the style of faulting correction in the GMPEs ITA10 and BND14, is given by $F_{sof} = f_j E_j$ for j=1:4 where f_j are coefficients and E_j are dummy variables used to denote the different style of faulting defined in Table 6. Note that for BND14, class U is considered only for *DS-EC8* (see (2) in Table 6).

The functional form of BSSA is as follows:

$$\ln IM = \underbrace{\left[F_E\left(M, mech\right) + F_P\left(R_{JB}, M, region\right) + F_S(V_{S30}, R_{JB}, M, region, z_1) \right]}_{\mu_{\ln IM}} + \varepsilon_n \sigma_{\ln IM}\left(M, R_{JB}, V_{S30}\right) \quad (14)$$

where F_E is the source effect function, F_P is the path effect function, and F_S is the site effect function (including linear and nonlinear site amplifications as well as basin depth effect), and their sum in the bracket represents the mean of ln*IM*, denoted as $\mu_{\ln IM}$. The term $\sigma_{\ln IM}$ is the total standard deviation of the model. The predictor variables are *M*, R_{JB} (in km), V_{S30} (in m/s), basin depth z_1 (depth from the ground surface to the 1.0 km/s shear-wave velocity, in km), *mech* = 0, 1, 2, and 3 for U, SS, NS, and RS style of faulting, respectively (see Table 6), and finally *region* is for regional correction (including China, Turkey, Italy and Japan). The term $F_E(M, mech)$ has functional form identical to BND14 and ITA10's $F_{sof}+F_M$; except that the hinge magnitude M_h is not constant. The expression for F_P (with regional adjustment) is quite similar to the functional form of F_D in BND14 and ITA10 models. The site amplification function F_S (based on the work by Seyhan and Stewart 2014) is much more complicated than that of the two other GMPEs. It composes of 3 expressions: a linear term (generally similar to BND14 site function based on V_{S30} as noted in Table 7, with V_{ref}=760m/s and period-dependent parameters), a nonlinear term, and an adjustment factor to account for basin depth effect; where the last two expressions are region-dependent. The three GMPEs, i.e. ITA10, BND14 (for both *DS-EC8* and *DS-*V_{S30}) and BSSA, are compared in terms of their predictions of pseudo-acceleration response spectrum $S_a(T)$ in the electronic supplementary material (Section S.M.1, and Fig. SM1) of this manuscript (note that for the rest of the manuscript, we interchangeably use S_a instead of PS_a for pseudo spectral acceleration, see also caption 6 of Table 6)

5.1 *IM* of arbitrary horizontal component

In the engineering application of seismic hazard, it is often of interest to use the GMPE for an arbitrary horizontal component with the desired intensity IM_{arb} instead of the geometric mean of the two horizontal components of ground motion, $IM_{g.m.}$. While ITA10 and BND14 use $IM_{g.m.}$, BSSA employs the single-component horizontal ground-motion. This inconsistency in definitions is typically not recognized when GMPEs are combined to provide the seismic hazard of the site of interest. For instance, integrating seismic hazard calculated based on $IM_{g.m.}$ GMPE and structural analysis based on IM_{arb} leads to inaccurate and unconservative estimates of the seismic risk (see Baker and Cornell 2006). One possible solution is to modify the GMPEs, which are based on $IM_{g.m.}$, in order to account for IM_{arb}. To this end, the marginal mean and variance of ln$IM_{g.m.}$, denoted as $\mu_{\ln IM_{g.m.}}$ and $\sigma^2_{\ln IM_{g.m.}}$, can be derived as (see also Baker and Cornell 2006):

$$\mu_{\ln IM_{g.m.}} = \mathrm{E}\left(\frac{\ln IM_x + \ln IM_y}{2}\right) = \frac{1}{2}[\mathrm{E}\left(\ln IM_x\right) + \mathrm{E}\left(\ln IM_y\right)] = \mu_{\ln IM_{arb}} \quad (15)$$

where $\mu_{\ln IM_{arb}}$ is the mean of lnIM_{arb}, and IM_x and IM_y are the intensity measures corresponding to the two horizontal components of the ground motion. By similar reasoning, we can obtain:

$$\sigma^2_{\ln IM_{g.m.}} = \mathrm{Var}\left(\frac{\ln IM_x + \ln IM_y}{2}\right) = \frac{1}{4}[\sigma^2_{\ln IM_x} + \sigma^2_{\ln IM_y} + 2\rho_{\ln IM_x, \ln IM_y}\sigma_{\ln IM_x}\sigma_{\ln IM_y}] = \sigma^2_{\ln IM_{arb}}\frac{1+\rho_{\ln IM_x, \ln IM_y}}{2} \quad (16)$$

where $\sigma^2_{\ln IM_{arb}}$ is the variance of lnIM_{arb}, and $\rho_{\ln IM_x, \ln IM_y}$ is the correlation coefficient between lnIM_x and lnIM_y. According to Eq. (15) and Eq. (16), the median of the desired GMPE for both $IM_{g.m.}$ and IM_{arb} is identical, while the dispersion of the GMPE based on IM_{arb} is larger than of $IM_{g.m.}$(the factor multiplied by $\sigma^2_{\ln IM_{arb}}$ is always less than or equal to unity). The correlation coefficient $\rho_{\ln IM_x, \ln IM_y}$ is dependent on the record set in use, but typically falls between 0.8 and 0.9, depending on the range of magnitudes and distances of the records (Baker and Cornell 2006). Alternatively, $\sigma_{\ln IM_{arb}}$ can be estimated directly (without the need to estimate the correlation term separately) as a function of $\sigma_{\ln IM_{g.m.}}$(see Appendix 1 for the derivation):

$$\sigma_{\ln IM_{arb}} = \sqrt{\sigma^2_{\ln IM_{g.m.}} + \sigma^2_C} \quad (17)$$

where σ_C is equal to (Boore 2005, see also Campbell and Bozorgnia 2008):

$$\sigma^2_C = \frac{1}{4N}\sum_{j=1}^{N}\left(\ln IM_{x,j} - \ln IM_{y,j}\right)^2 \quad (18)$$

where $IM_{x,j}$ and $IM_{y,j}$ are the ground motion intensities for two horizontal components of jth recording in the database and N is the total number of recordings. The logarithmic standard deviation for IM related to the arbitrary component is calculated for both ITA10 and BND14 in Appendix 2 (see the electronic supplementary material of this manuscript, Section S.M.3, Tables A1, A2 and A3). For these GMPEs, which provide the functional form in term of $\log_{10}IM$ (instead of lnIM), Eq. (15) to Eq. (18) should be modified by substituting ($\log_{10}$) instead of (ln).

6 General formulation of the Probabilistic Seismic Hazard Assessment

Probabilistic seismic hazard analysis is the most appropriate approach for considering various sources of uncertainty (e.g., in magnitude, distance, rate of recurrence of earthquakes and in the variation of ground motion intensity measure with earthquake magnitude and distance) to be explicitly considered for the evaluation of seismic hazard. By employing PSHA, instead of hunting for a worst-case ground motion intensity, one can consider all plausible earthquake scenarios along with their associated probabilities of occurrence, to find the rate of exceeding a specific ground-motion intensity measure, IM, in a given interval of time and at a given site. Generally, for estimating the hazard due to strong ground-motions, the "annual" rate of exceedance is estimated. The annual rate of exceeding a specified level of IM equal to x is denoted as $\lambda(IM > x)$. Assume that a Poisson distribution for the events with magnitude greater than a lower-bound threshold, $M > M_l$, has a rate $\lambda(M > M_l)$; the rate $\lambda(IM > x)$ can be interpreted as the rate of events with the probability of success $P(IM > x|\mathrm{EQ})$ (read as the probability of $IM > x$ given that an EQ of interest with $M > M_l$ takes place). Thus, $\lambda(IM > x)$ can directly be estimated by assuming a filtered Poisson process. The term $\lambda(IM > x)$ can be expanded as the sum of the exceedance rates of $IM > x$ for all seismic sources (assuming that earthquake occurrence for these alternative sources can be expressed as independent Poisson processes) including: (a) finite-fault sources (a.k.a. SBoxes; for brevity in the formulation, we show it as SB_i, where i=1,⋯,n_{SB}=14; see Section 4.1, Table 2) identified with rate $\lambda_{\mathrm{SB}_i}(IM > x)$; (b) areal source (i.e., background, BG, see Section 4.2) with rate $\lambda_{\mathrm{BG}}(IM > x)$. Thus, the rate $\lambda(IM > x)$ can be calculated as follows:

$$\begin{aligned}\lambda\left(IM > x\right) &= \sum_{i=1}^{n_{\mathrm{SB}}}\lambda_{\mathrm{SB}_i}\left(IM > x\right) + \lambda_{\mathrm{BG}}\left(IM > x\right) \\ &= \sum_{i=1}^{n_{\mathrm{SB}}}\underbrace{\lambda_{\mathrm{SB}_i}\left(M > M_l^{\mathrm{SB}}\right)}_{\nu^{\mathrm{SB}_i}\ \text{(Section 4.1.2, Table 4)}}\cdot P\left(IM > x\middle|\mathrm{EQ}_{\mathrm{SB}_i}\right) + \underbrace{\lambda_{\mathrm{BG}}\left(M_l^{\mathrm{SB}} \geq M \geq M_l^{\mathrm{BG}}\right)}_{\nu^{\mathrm{BG}}\ \text{(Section 4.2.3)}}\cdot P\left(IM > x\middle|\mathrm{EQ}_{\mathrm{BG}}\right)\end{aligned} \quad (19)$$

The term $\lambda_{\mathrm{SB}_i}(M > M_l^{\mathrm{SB}})$, i=1:n_{SB}, is the annual seismicity rate of each SBox, which was denoted as $\nu^{\mathrm{SB}i}$ in Table 4 (Section 4.1.2). Consequently, $\lambda_{\mathrm{BG}}\left(M_l^{\mathrm{SB}} \geq M \geq M_l^{\mathrm{BG}}\right)$ is annual rate of events within the background area (activity rate), and is defined in section 4.2.3 as ν^{BG}. Finally, $P\left(IM > x\middle|\mathrm{EQ}_{\mathrm{SB}_i}\right)$ and $P(IM > x|\mathrm{EQ}_{\mathrm{BG}})$ denote the probability of $IM > x$ given that an EQ of interest takes place in SB_i and BG, respectively. These exceedance probabilities are described by a lognormal distribution whose statistical parameters (logarithmic mean and standard deviation) are provided by an appropriate attenuation relation; i.e., a GMPE, depending on the type of seismic source. With reference to Section 5, the (logarithmic mean and standard deviation) are a function of

magnitude m, R_{JB} (or R_{epi}), soil-site condition, as well as the style-of-faulting. For brevity, we use the general function $\varphi_{s,mech}$ for expressing the two latter terms; where "s" stands for soil-site effect (functional form F_S for ITA10 and BND14 models; directly expressed as V_{S30} and z_1 in BSSA), and "*mech*" denotes style-of-faulting (functional form F_{sof} in ITA10 and BND14; *mech* in BSSA).

The exceedance probability $P(IM > x|\text{EQ})$ can be defined as the expected value for a prescribed GMPE taking into account the joint probability distribution for the model parameters $\boldsymbol{\theta}$, denoted as $f(\boldsymbol{\theta})$. By using total probability theorem, Eq. (19) can expanded as:

$$\lambda\left(IM > x\right) = \sum_{i=1}^{n_{\text{SB}}} \nu^{\text{SB}_i} \int_{\boldsymbol{\Omega}_{\boldsymbol{\theta}_{\text{SB}_i}}} P\left(IM > x \middle| \text{EQ}_{\text{SB}_i}, \boldsymbol{\theta}_{\text{SB}_i}\right) f\left(\boldsymbol{\theta}_{\text{SB}_i}\right) \text{d}\boldsymbol{\theta}_{\text{SB}_i} + \nu^{\text{BG}} \int_{\boldsymbol{\Omega}_{\boldsymbol{\theta}_{\text{BG}}}} P\left(IM > x \middle| \text{EQ}_{\text{BG}}, \boldsymbol{\theta}_{\text{BG}}\right) f\left(\boldsymbol{\theta}_{\text{BG}}\right) \text{d}\boldsymbol{\theta}_{\text{BG}} \tag{20}$$

where $\boldsymbol{\Omega}_{\boldsymbol{\theta}_{\text{SB}_i}}$ and $\boldsymbol{\Omega}_{\boldsymbol{\theta}_{\text{BG}}}$ are the domains of the model parameters $\boldsymbol{\theta}_{\text{SB}_i}$ and $\boldsymbol{\theta}_{\text{BG}}$, respectively.

6.1 Model Parameters $\boldsymbol{\theta}_{\text{SB}}$ for finite fault sources

The model parameters for fault sources can be expressed as $\boldsymbol{\theta}_{\text{SB}}$=[$m$, w, l, Np], where m is the magnitude; w is down-dip rupture width (see Right sub-figure in Fig. 1); l is the subsurface horizontal fault rupture length (a.k.a. fault rupture length, see Fig. 3c), and Np denotes the location of the nucleation point within the ruptured surface (fault plane). With reference to Eq. (20), the term $P\left(IM > x \middle| \text{EQ}_{\text{SB}_i}\right)$ in Eq. (19) can thus be expanded as:

$$P\left(IM > x \middle| \text{EQ}_{\text{SB}_i}\right) = \int_{\boldsymbol{\Omega}_{\boldsymbol{\theta}_{\text{SB}_i}}} P\left(IM > x \middle| \text{EQ}_{\text{SB}_i}, \boldsymbol{\theta}_{\text{SB}_i}\right) f\left(\boldsymbol{\theta}_{\text{SB}_i}\right) \text{d}\boldsymbol{\theta}_{\text{SB}_i} = \int_{\boldsymbol{\Omega}_{\boldsymbol{\theta}_{\text{SB}_i}}} P\left[IM > x \middle| m, R_{JB}\left(m,w,l,Np\right), \varphi_{s,mech_i}\right] \underbrace{f_M^{\text{SB}_i}\left(m\right) \underbrace{f\left(w|m\right) f\left(l|w,m\right)}_{f\left(w,l|m\right)} f\left(Np|w,l,m\right)}_{f\left(m,w,l,N_p\right)} \text{d}Np\, \text{d}l\, \text{d}w\, \text{d}m \tag{21}$$

where $f_M^{\text{SB}_i}$ is the truncated Normal/Uniform PDF for magnitude m calculated based on Eq. (6) as a function of the lower cut-off magnitude M_l^{SB}=5.50 for all fault sources, and the source-dependent upper magnitude $M_u^{\text{SB}_i}$ (see Section 4.1.1 and Table 3). The PDF $f(w|m)$ is the distribution of down-dip rupture width w given a value of magnitude m; we employ herein the well-known WC94 (Wells and Coppersmith 1994) empirically-estimated lognormal distribution with a (logarithmic) mean of the general form $\mu_{\log_{10}w|m} = a + bm$ and a (logarithmic) standard deviation, $\sigma_{\log_{10}w|m}$, derived from linear regression model of $\log_{10}w$ versus m for various styles of faulting. The PDF $f(l|w)$ is the distribution of subsurface rupture length l with respect to w. In order to estimate $f(l|w)$, one can use the linear relationship proposed by Peruzza and Pace (2002) in the form $l=a_{ASP}+b_{ASP}\cdot w$ for different style of faulting (see also Section 4.1.1, part 4, assuming that, conditioned on the width of faulting, the faulting length is independent of magnitude). Leonard (2010) proposes a w versus l relationship in the form $w = C_1 l^{\beta}$, where $\beta \approx \frac{2}{3}$ is a constant value, and C_1 is a variable depending on the style of faulting. Herein, we use the latter relationship by assuming that $f(l|w)$ is a lognormal distribution with the (logarithmic) mean $\mu_{\log_{10}l|w}$ and logarithmic standard deviation $\sigma_{\log_{10}l|w}$ equal to (taking the logarithm of both side and calculating l in terms of w):

$$\mu_{\log_{10}l|w} = \frac{1}{\beta}\left(\log_{10} w - \mu_{\log_{10}C_1}\right), \quad \sigma_{\log_{10}l|w} = \frac{1}{\beta}\sigma_{\log_{10}C_1} \tag{22}$$

where $\mu_{\log_{10}c_1}$ and $\sigma_{\log_{10}c_1}$ are the logarithmic mean and standard deviation of C_1 (Leonard, 2010). Finally, the PDF $f(Np|w,l,m)$ denotes the distribution of the nucleation point of an event with magnitude m and rupture area $w\times l$ within the fault plane (i.e., individual seismic source having total down-dip ruptured width w and ruptured length l, see the typical seismoigenic source in Fig. 3c). According to this sub-figure, Np is assumed herein to be located along the Bottom edge (along LL-UL projection on the Bottom depth) of the fault plane either at the ends (LL or UL projection) or in the middle (in this case, the rupture is assumed to symmetrically spread out). Thus, due to the three possible locations, the probability $f(Np|w,l)\text{d}Np \approx \frac{1}{3}$ (assuming that the location of the nucleation point is independent of the magnitude and rupture area dimensions). It is to mention that the Joyner-Boore distance, R_{JB}, which is a function of m, w, l and Np, is estimated as the nearest distance from the designated site to the projection to the ground surface of the ruptured plane $w\times l$.

With reference to Eq. (21), the exceedance probability $P\left[IM > x \middle| m, R_{JB}, \varphi_{s,mech_i}\right]$ is estimated for each SB$_i$ (with the style-of-faulting $mech_i$, see Table 2) based on the weighted average of the exceedance probability for the four considered GMPEs in Section 5 including ITA10, BND14 (both *EC8*- and V_{S30}-based models), and BSSA as shown in the following Fig. 9:

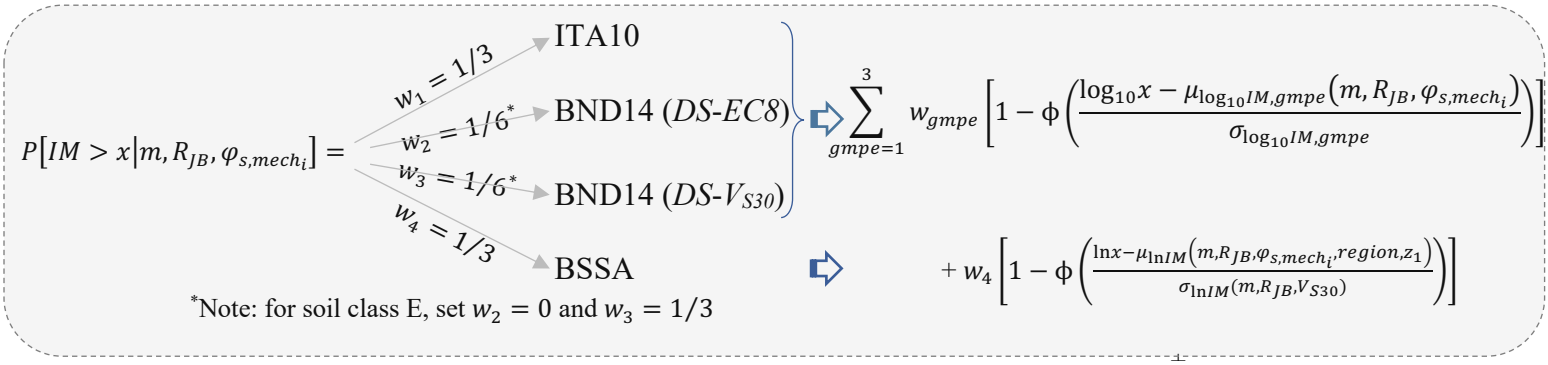


Figure 9: Implementation of the GMPEs for finite fault sources within the PSHA framework

According to Fig. 9, w_{gmpe} (where *gmpe*=1:4) denotes the weight associated with each GMPE. The probability $IM > x$ for each GMPE is calculated as a function of a Lognormal CDF (ϕ is the standard Normal CDF), whose logarithmic mean and standard deviation are estimated based on Eq. (13) for ITA10 and two BND14 models (defined with $\mu_{\log_{10}IM,gmpe}$ and $\sigma_{\log_{10}IM,gmpe}$ where *gmpe*=1:3) and Eq. (14) for BSSA (shown with $\mu_{\ln IM}$ and $\sigma_{\ln IM}$ where *gmpe*=4). Herein, the term *region* for the BSSA model is set for Italy to account for regional effect. Moreover, the basin depth z_1 for the BSSA model is treated as unknown for the case-study site due to the complicated geological formation of basin; thus, its effect on BSSA's site amplification function is ignored in this case-study. It is worth mentioning that for Soil Class E, the *EC8*-based BND14 model is not included (as noted in Fig. 9, the weight w_3 is modified accordingly; see also footnote 2 of Table 6).

As far as it regards the composite source (SB_{14}, see Fig. 3a and Table 2), it is noteworthy that the length of the expected earthquake ruptures is unknown while the source itself spans an unspecified number of Individual Sources. The seismic potential of this source is estimated by assigning a magnitide ditribution $f_M^{SB_{14}}$ (see Fig. 4 and Table 3), and activity rate $v^{SB_{14}}$ (see Fig. 5 and Table 4). Considering the fact that the surface projection of this complex source is determined (as shown in Fig. 3a), we assign a Uniform distribution to $f(w,l|m)$ in Eq. (21), i.e. we assume that all the possible (w,l) pairs within the domain $\boldsymbol{\Omega}_{\boldsymbol{\theta}_{SB_{14}}}$ are equally likely to occur. Other terms in Eq. (21) are estimated in the same way as already explained.

6.2 Model Parameters $\boldsymbol{\theta}_{BG}$ for the background area

The model parameters for the background area is composed of $\boldsymbol{\theta}_{BG}=[m, x_g, y_g]$, where (x_g, y_g) are the Cartesian coordinates of the epicenter of an event, occurred within the Cartesian Background plane, with respect to the location of the desired site. Thus, the probability term $P(IM > x|\mathrm{EQ}_{BG})$ in Eq. (19) can be expressed as:

$$\begin{aligned} P\left(IM > x\middle|\mathrm{EQ}_{BG}\right) &= \int_{\boldsymbol{\Omega}_{\boldsymbol{\theta}_{BG}}} P\left(IM > x\middle|\mathrm{EQ}_{BG},\boldsymbol{\theta}_{BG}\right) f\left(\boldsymbol{\theta}_{BG}\right) \mathrm{d}\boldsymbol{\theta}_{BG} \\ &= \int_{\boldsymbol{\Omega}_{\boldsymbol{\theta}_{BG}}} P\left[IM > x\middle|m, R_{epi}\left(x_g, y_g\right), \varphi_{s,mech}\right] f_M^{BG}\left(m\right) f\left(x_g, y_g\right) \mathrm{d}x_g\, \mathrm{d}y_g\, \mathrm{d}m \end{aligned} \tag{23}$$

where f_M^{BG} is the truncated Exponential PDF of magnitude m based on Eq. (8) for the magnitude range $M_l^{BG} = 4.60 \le m \le M_l^{SB} = 5.50$ (see Section 4.2, and Table 5). The joint PDF for the epicenter location of an event within the BG zone with the area A_{BG}, denoted as $f\left(x_g, y_g\right)$ in Eq. (23), is calculated assuming equiprobable occurrence of seismic events (and independence of magnitude); thus, $f\left(x_g, y_g\right)\mathrm{d}x_g\mathrm{d}y_g = f\left(x_g, y_g\right)\mathrm{d}A_g \approx \frac{1}{A_{BG}}\mathrm{d}A_g$. Alternatively, one can consider the distribution of past events within the BG zone. In such case the PDF $f\left(x_g, y_g\right)$ will no longer be a Uniform distribution. However, we have preferred the equiprobability assumption herein as it is coherent with the approach followed by INGV to develop the national seismic hazard maps (http://esse1-gis.mi.ingv.it/s1_en.php, last accessed 20/02/2018). The epicentral distance R_{epi} is calculated as $R_{epi} = \sqrt{{x_g}^2 + {y_g}^2}$.

With reference to Eq. (23), the probability $P\left[IM > x\middle|m, R_{epi}, \varphi_{s,mech}\right]$ is estimated based on the weighted average of the exceedance perobabilitues associated with the ITA10 and *EC8*-based BND14 models. It is worth mentioning that the focal mechanism of the equiprobable earthquakes within the BG zone is unknown (i.e., *mech* is "unspecified"); this is while the BND14 model based on *DS-*V_{S30} data set is developed for events with known focal mechanism (see footnote 1 of Table 6). Although both ITA10 and BND14 prediction models are based on the site-to-source distance measure R_{JB}, ITA10 employs the epicentral distance (Bindi et al. 2011, see also footnote 4 of Table 6) when the fault geometry is unknown (generally the case when $m < 5.50$). Similarly, BND14 allows

substituting R_{JB} with the epicentral distance R_{epi} when the latter is unspecified --only when $m \leq 5$ and $R_{epi} \geq 10$ km (Bindi et al. 2014, see also footnote 5 of Table 6). By relaxing the restriction on magnitude threshold (i.e., 5.0) to become 5.50, and considering that the number of points with $R_{epi} < 10$ km are small in $\boldsymbol{\Omega}_{\boldsymbol{\theta}_{\mathrm{BG}}}$, we have employed both ITA10 and *EC8*-based BND14 models for predicting the exceedance probability in Eq. (23). On the other hand, there are no specified and tectonically-developed fault sources in the vicinity (<10 km) of the case-study area (with the exception of volcano-tectonic faults of the Campi Flegrei which are not considered in this study). Last but not least, for Soil Class E, the *EC8*-based BND14 model is not included. The implementation of the GMPEs (*gmpe*=1:2) for the background area is illustrated in Fig. 10.

$$P\left[IM > x \middle| m, R_{epi}, \varphi_{s,mech}\right] = \begin{cases} w_1 = 1/2 \rightarrow \text{ITA10} \\ w_2 = 1/2^{*} \rightarrow \text{BND14 (}DS\text{-}EC8\text{)} \end{cases} \Rightarrow \sum_{gmpe=1}^{2} w_{gmpe}\left[1 - \phi\left(\frac{\log_{10}x - \mu_{\log_{10}IM,gmpe}\left(m, R_{epi}, \varphi_{s,mech}\right)}{\sigma_{\log_{10}IM,gmpe}}\right)\right]$$

*Note: for soil class E, set $w_2 = 0$ and $w_1 = 1.0$

Figure 10: Implementation of the GMPEs for areal source (background area) within the PSHA framework

6.3 Model Parameters for the point sources

To perform PSHA for each of the two point sources, namely "PS Nola" and "PS Sorrento peninsula", the expression is similar to Eq. (23) considering only one model parameter, m, (the epicenters are known). The PDF of magnitude m is a Uniform distribution, defined in Section 4.1.1, in the range of $[M_l^{\mathrm{SB}}$=5.50, 6.0] for both sources. The implementation of the GMPEs for the two point sources within the PSHA framework of Eq. (23) follows the instructions given previously in Section 6.2 (see also Fig. 10).

7 The seismic hazard analysis maps

The site-specific seismic hazard maps display the earthquake ground-shaking intensity for various probability levels across the case study located in the western Naples and can be directly applied in seismic risk assessment. The resulting maps are derived from seismic hazard curves calculated for a grid of sites covering the studied area. The PSHA curves are derived considering the spectral acceleration at the first-mode period, i.e. IM=$S_a(T)$, for a vector of periods T(sec) = {0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.25, 1.50} as the intensity measure (T=0.0 denotes IM=PGA). There is also the possibility to consider other IMs (see Ebrahimian et al. 2015) in case of their computability through a GMPE.

7.1 Distribution of Buildings' fundamental periods in the western Naples

As noted in Section 2, the case-study area hosts a wide variety of masonry and reinforced concrete (RC) buildings of different ages (the older buildings' construction dates back to the turn of twentieth century). The height distribution (in terms of the number of stories) for RC and Masonry buildings within the study area and the corresponding number of buildings are reported in Table 8 (Dr. Carlo Del Gaudio 2018, University of Naples Federico II, *Personal Communication*):

Table 8: Number of stories, the number of buildings and their associated fundamental periods for RC and Masonry buildings within the study area (Del Gaudio 2018)

Number of stories	Number of RC buildings	Number of Masonry buildings	Period range (sec) lower and upper limits			
			RC buildings		Masonry buildings	
1	39	27	0.11	0.24	0.13	0.16
2	14	65	0.22	0.41	0.22	0.27
3	28	122	0.33	0.59	0.29	0.37
4	28	127	0.43	0.78	0.36	0.46
5	77	53	0.54	0.98	0.43	0.54
6	213	14	0.65	1.17	0.49	0.62

The empirical formulations proposed by Ricci et al. (2011) and Eurocode 8 (CEN 2004) are employed herein for estimating the fundamental period of RC buildings (with infills) whereas the empirical relation proposed by Eurocode 8 (CEN 2004) is employed for estimating the fundamental period of masonry buildings. The story heights (needed as a parameter in the empirical relations) is assumed to vary in the range 3.5m to 4.76m, considering the age distribution of the buildings. The last two columns of Table 8 report the period range (the

lower and upper values) for each building category --according to the empirical relationships. The histogram of the periods associated with each building type is shown Fig. 11. With reference to the period distribution of the buildings in the western area of Naples, we have selected five target *IM*s including *PGA*, S_a(*T*=0.3sec), S_a(*T*=0.50sec), S_a(*T*=0.70sec), and S_a(*T*=1.0sec) for constructing the seismic hazard maps.

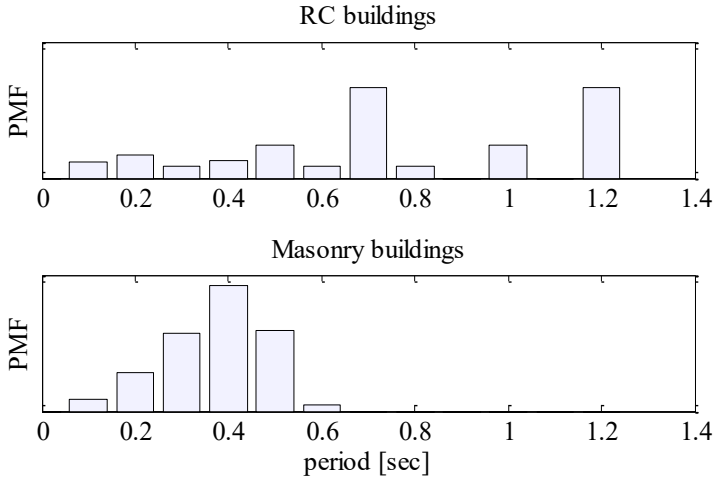


Figure 11: The distribution of the fundamental periods associated with the two predominant building types in the western Naples

7.2 Seismic hazard maps

With reference to the updated National Technical Code for seismic design in Italy (NTC 2018), the performance objectives are categorized in four tiers as outlined in Table 9:

Table 9: The NTC Performance objectives

Performance Level: limit states	Hazard Level P_{V_R}: Exceedance probability in the reference time period of V_R^* years	Return Period T_R (years): $T_R = \frac{-V_R}{\ln(1-P_{V_R})}$
Operational (SLO)	81%	30
Damage (SLD)	63%	50
Life Safety (SLV)	10%	475
Collapse (SLC)	5%	975

*Based on the building classification according to NTC (2018) in the designated area, V_R can be set to 50 years.

The seismic hazard maps calculated for five target *IM*s defined in Section 7.1 including *PGA*, S_a(0.3sec), S_a(0.50sec), S_a(0.70sec), and S_a(1.0sec) and four hazard levels with T_R (inverse of mean annual rate of exceedance) of [30, 50, 475, 975] years. Fig. 12 and Fig. 13 show the hazard maps for *PGA*, and S_a(0.5sec) associated with the four aforementioned hazard levels (please note that the hazard maps for other intensity measure including S_a(0.3sec), S_a(0.70sec), and S_a(1.0sec) are presented in the electronic supplementary material of this manuscript, Section S.M.2, Fig. SM2 to Fig. SM4).

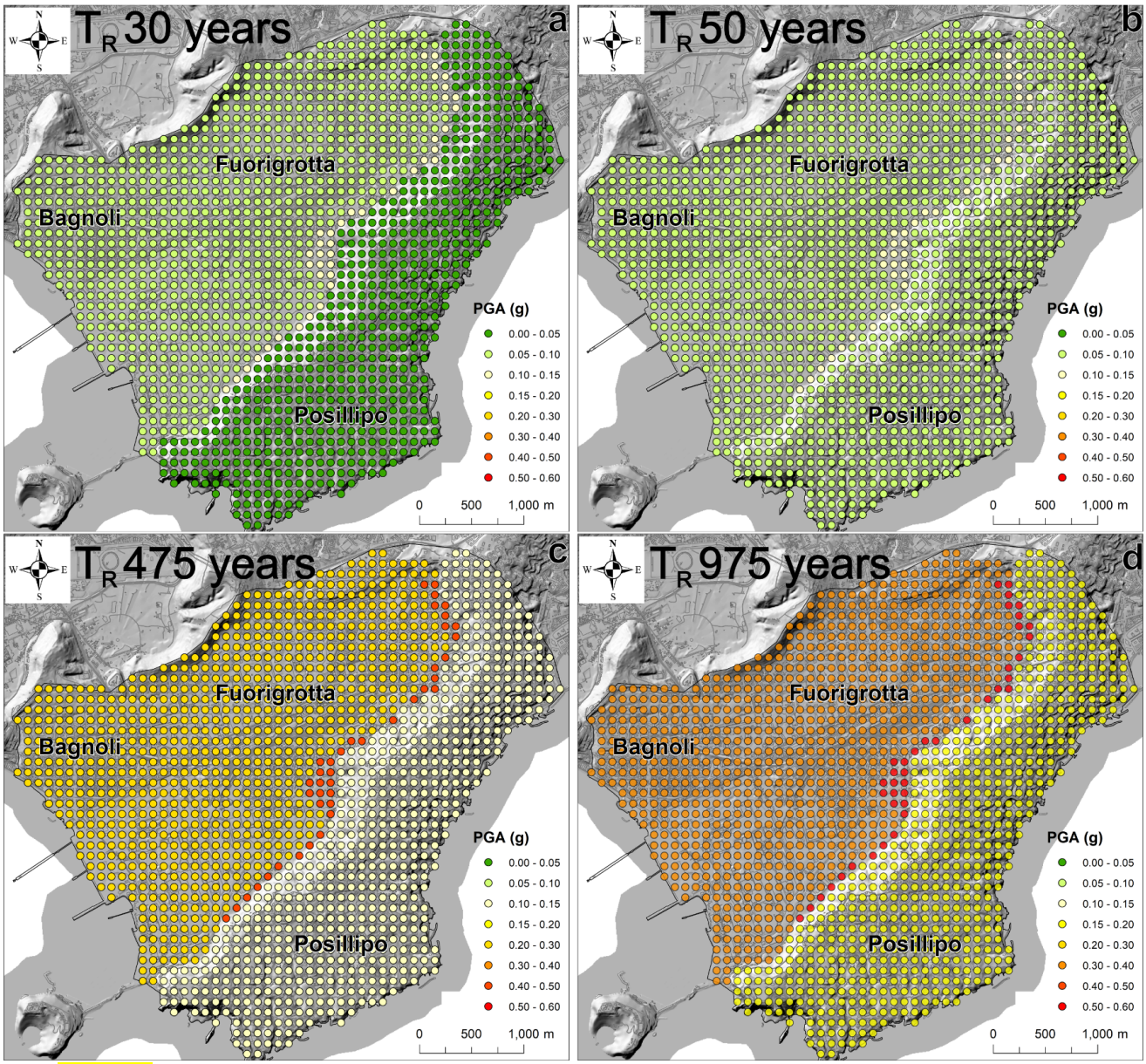


Figure 12: Seismic Hazard maps for *PGA* associated with 4 hazard levels with the return periods of [30, 50, 475, 975] years

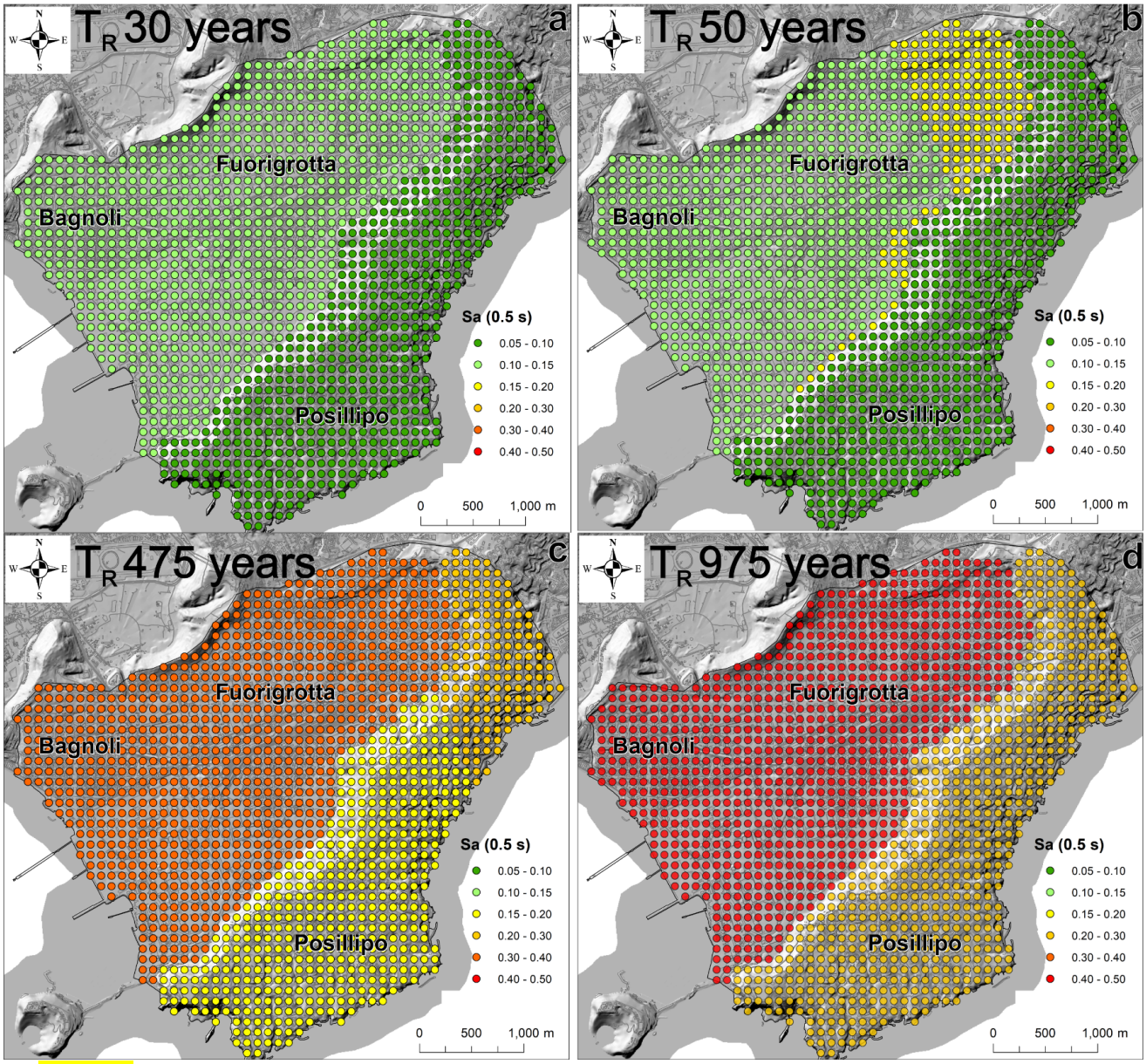


Figure 13: Seismic Hazard maps for $S_a(T=0.5\text{sec})$ associated with 4 hazard levels with the return periods of [30, 50, 475, 975] years

7.3 Site-specific Uniform Hazard Spectra compared with the INGV and code spectra

In this section, the PSHA estimates in terms of Uniform Hazard Spectrum (UHS, a by-product of site-specific PSHA that expresses pseudo spectral acceleration values for a range of periods given a uniform hazard level) are compared with the national hazard map data (Meletti et al. 2007, http://esse1.mi.ingv.it/) released by INGV as well as the code-based response spectra (NTC 2018). The UHS are based on the four hazard levels indicated in the previous section corresponding to the four exceedance probabilities in 50 years for $IM=S_a$ [g] at the period range of 0.0 sec to 1.50 sec (T=0.0 denotes *PGA*). Fig. 14 illustrates the case-study area and the four nearby grid points of INGV hazard maps surrounding the desired area (green circles; note that the INGV maps are also provided for a finer mesh grid only for *PGA* at 475 years return period). We select three grid points (belonging to the gridded fishnet map) located on three soil types A, C, and E, with the highest hazard value (calculated through the site-specific PSHA herein), as highlighted in Fig. 14 with small colored circles. The rationale behind this choice was to highlight the intra-grid variations in hazard that are not going to be detected by INGV for each soil type. The closest INGV grid point to these reference points has latitude and longitude of [40.83, 14.22] (INGV grid number 33200, see Fig. 14). The UHS estimated by this study are shown for the four designated hazard levels (see Table 9) in Fig. 15 with thick gray line. Accordingly, the INGV-based spectra (Meletti et al. 2007, http://esse1.mi.ingv.it/) for the nearest grid point 33200 are illustrated with dashed blue line (for the four defined hazard levels). The code-based spectra (NTC 2018) are also drawn with black solid line. It is to note that the INGV spectra are derived for stiff-soil condition (i.e., soil type A) and they should be later modified for site and topographical amplification effects by multiplying the spectrum ordinates by the coefficient $S = S_S \cdot S_T$, where S_S denotes the site amplification and S_T is the topographical amplification (NTC 2018). For the sake of comparison, we set the topographical effects S_T=1.0 (for INGV and code spectra), as our PSHA calculations did not take the topographical effects into account. The coefficient S_S, for different soil types, is estimated based on the code provision (NTC 2018, Table 3.2.IV) as a function of the a_g (i.e., *PGA*) that can be extracted form INGV data for grid point 33200 for each hazard level, and the coefficient F_o (that quantifies the maximum spectral amplification on a horizontal rigid reference site). The latter coefficient is extracted from the commentary of the National Italian code (NTC 2018) for the four considered hazard levels for the grid point 33200. The code-based pseudo acceleration response spectrum is derived based on the following expression (see NTC 2018):

$$S_a(T) = \begin{cases} a_g \cdot S \cdot \eta \cdot \mathrm{F_o}\left(\dfrac{T}{T_B} + \dfrac{1}{\eta \cdot \mathrm{F_o}}\left(1 - \dfrac{T}{T_B}\right)\right) & 0 \le T < T_B \\ a_g \cdot S \cdot \eta \cdot \mathrm{F_o} & T_B \le T < T_C \\ a_g \cdot S \cdot \eta \cdot \mathrm{F_o} \cdot \dfrac{T_C}{T} & T_C \le T < T_D \\ a_g \cdot S \cdot \eta \cdot \mathrm{F_o} \cdot \dfrac{T_C \cdot T_D}{T^2} & T_D \le T \end{cases} \quad (24)$$

where η stands for the modification of damping coefficient if it is different from 5% (η = 1.0 herein); $T_C = C_C \cdot T_C^*$ where the coefficient C_C is calculated for different soil types a as a function of T_C^*, and the period T_C^* is extracted from the commentary of NTC 2018 (depending on the hazard level); $T_B = T_C/3$ and finally $T_D = 4\,a_g/g + 1.6$ (g is the gravitational acceleration).

With reference to Fig. 15, the following remarks can be addressed:

- For periods greater than 1.0 sec, regardless of the soil type and hazard level, the derived UHS in this study are higher than those estimated by INGV, and lower than the code-based spectrum. Hence, the high-rise RC buildings in the area (5 to 6 story) as well as long-period infrastructures (bridges) may be subjected to seismic actions that are higher than those estimated by INGV (although the code-based spectrum performs on the conservative side).
- For stiff-soil sites (i.e., soil type A), the spectral ordinates estimated by INGV and NTC code are quite similar, and are fairly larger than that of the UHS derived in this study for periods smaller than 0.5 sec. It is worth noting that the trend of spectral ordinates on soil type A between the INGV and the predictions of this study somehow reflects the objective differences between the source characterization and the adopted GMPE's.
- For sites with soil type C (see Table 1), the UHS ordinates proposed by INGV, especially for periods greater than 0.5 sec, are lower than those calculated by this study. However, NTC spectrum provides conservative spectral values (especially at smaller hazard levels) with respect to our study. It is to note that the trend of the code spectrum is more similar to the UHS of this study especially for T>0.50 sec. Considering the fact that most of the built environment in the case-study site are constructed on soil type C, following the NTC 2018 spectrum leads to reasonable intensity estimates for the designated hazard levels.

- For sites with soil type E, the UHS is considerably underestimated both by INGV (modified to account for soil type E) and national code when T<0.30 sec. The code spectrum performs conservatively for T>0.30 sec while the INGV-based UHS falls below the UHS in this study as well as that of the code. Since most of the buildings in the case-study area have fundamental period less than 0.5 sec (see Fig. 11), there is a need for a more detailed study on amplification factors related to soil type E in the national Italian code .
- Considering the *PGA* as the intensity measure instead of S_a, the values proposed in this study are very similar to the INGV-based ones for soil type A and C. However, for soil type E, the values calculated in this study are higher than those estimated by INGV.
- Careful revisiting should be done by national Italian code for site amplification on soft soil (e.g., soil type E). Moreover, the use of more advanced GMPEs with improved site amplification terms are highly recommended. It is to note that the GMPEs used in INGV-based approach (Meletti et al. 2007, http://esse1.mi.ingv.it/) are different from the ones used herein.

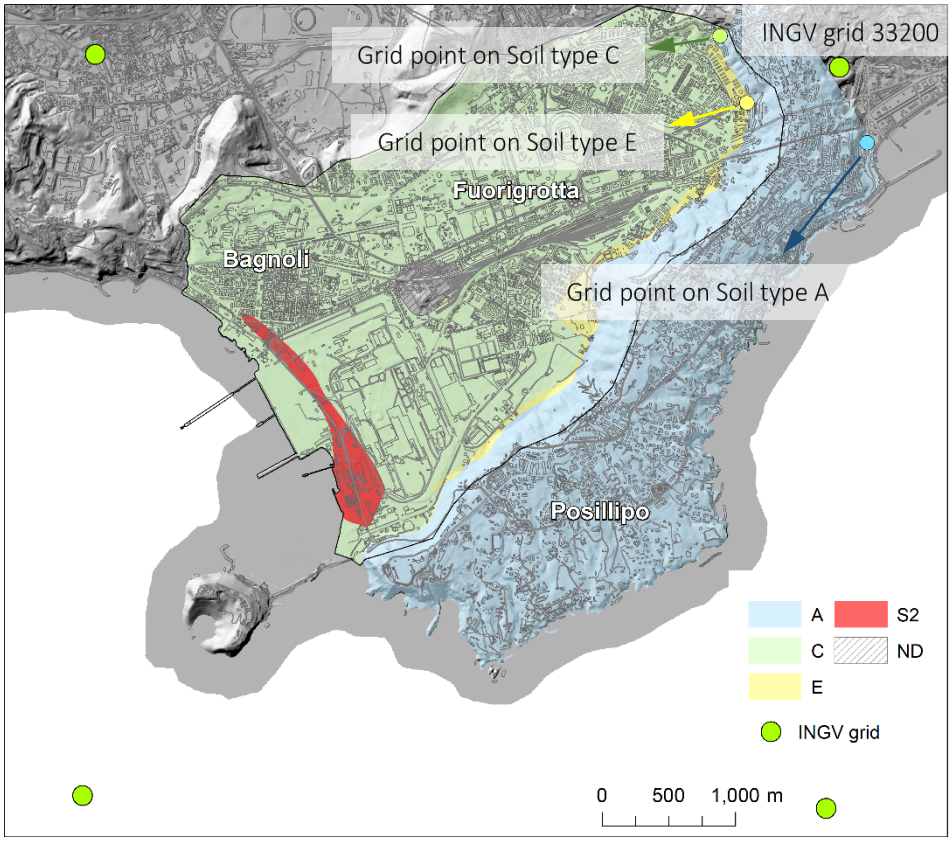
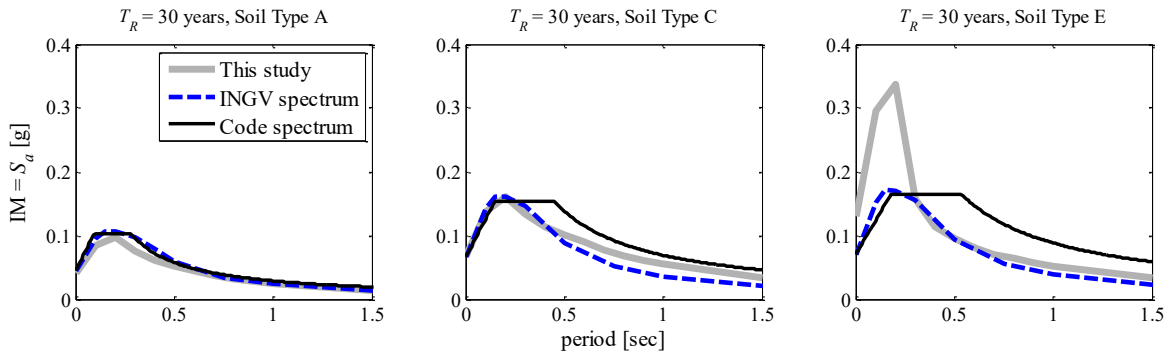


Figure 14: The location of the INGV grid points surrounding the case-study area; the representative grid points within the case-study area on soil types A, C and E; the INGV grid point 33200 as the benchmark for comparing the code-based and INGV-based UHS with the UHS calculated in this study



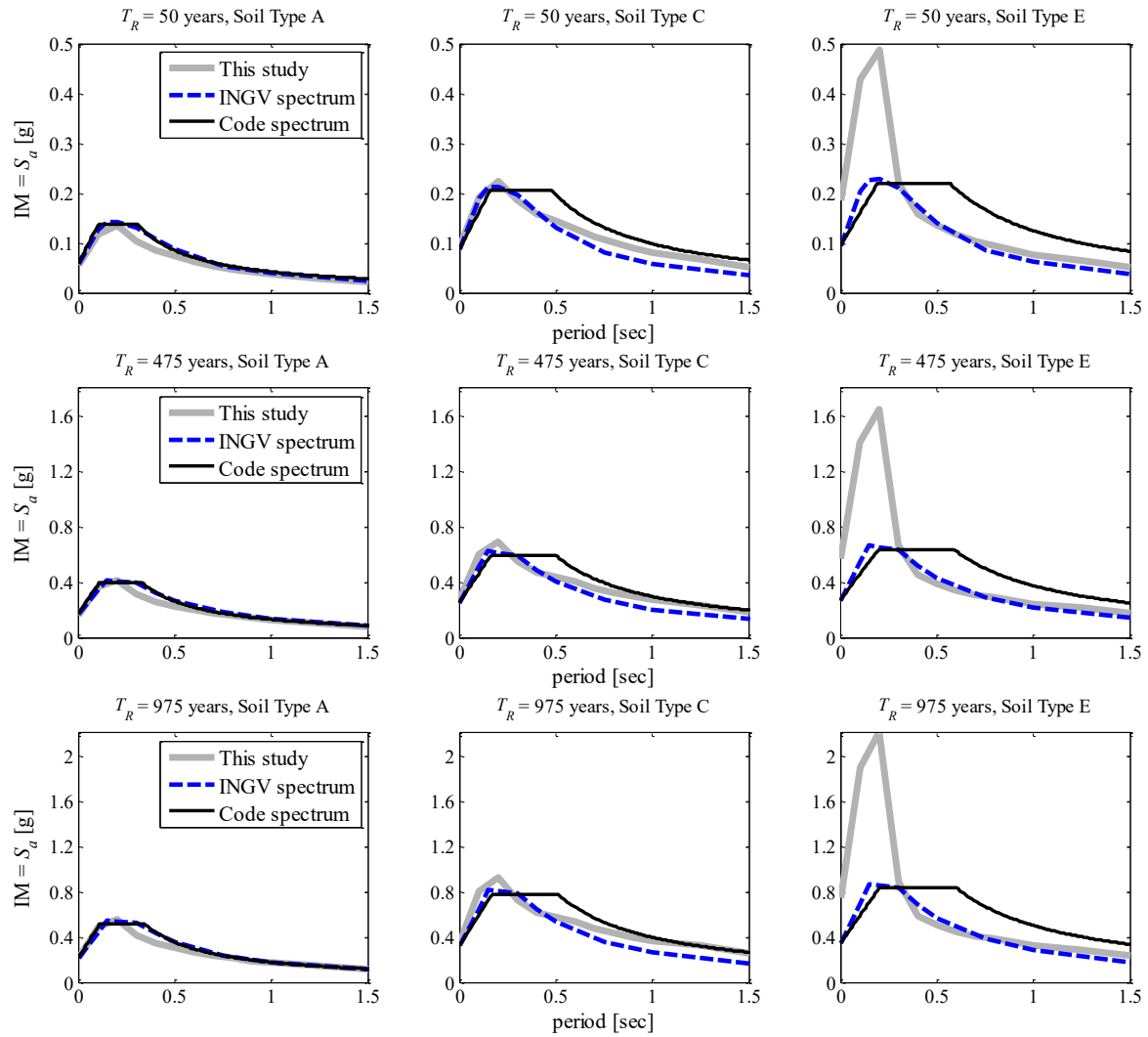


Figure 15: Comparison of the UHS calculated in this study for the four designated performance objectives (see Table 9) with the INGV-based and code-based (NTC 2018) spectra considering different soil types of the case-study area

7.4 Site-specific hazard curves compared with the INGV hazard data

The UHS is a by-product of the PSHA that is useful not only in design but also in assessment of building structures. However, there is wide-spread use of PSHA curves (in terms of intensity measure versus the annual rate of exceedance) in probabilistic performance-based earthquake engineering (Jalayer and Cornell 2009, see also Jalayer and Ebrahimian 2017, and Jalayer et al. 2017 for a comprehensive literature review) in quantifying the safety margin for newly designed buildings as well as existing ones located in seismic areas. To this end, the PSHA curves constructed by this study for reference points (shown in Fig. 14) are compared with the INGV hazard curves (Meletti et al. 2007, http://esse1.mi.ingv.it/). As illustrated in Fig. 16, the comparison is done for five target *IM*s; namely, *PGA*, $S_a(T$=0.3sec), $S_a(T$=0.50sec), $S_a(T$=0.70sec), and $S_a(T$=1.0sec) and for different soil types of the designated area. The following observations can be made with reference to Fig. 16:

- For sites with soil type A, the INGV hazard curves lead to a fine agreement with the results of this study, especially at higher periods. Some deviation (with INGV on the conservative side) can be observed at the small/intermediate period range.
- For sites with soil type C, INGV-based hazard curves, to various degrees, underestimate the exceedance frequencies compared to our study for T>0.5sec (of interest to mid- to high-rise buildings and bridge infrastructures).
- For sites with soil type E, INGV-based hazard curves underestimate the exceedance frequencies considerably for PGA and T<0.3sec; this can lead to an underestimation of seismic risk.
- It is to note that the INGV hazard curves (Meletti et al. 2007, http://esse1.mi.ingv.it/) are defined on stiff-soil sites (V_{S30}>800m/sec). Hence, in order to account for soft soil conditions, the same approach proposed

in Section 7.3 should be employed by multiplying the *IM*s by the site and topographical amplification factor *S* (NTC 2018).

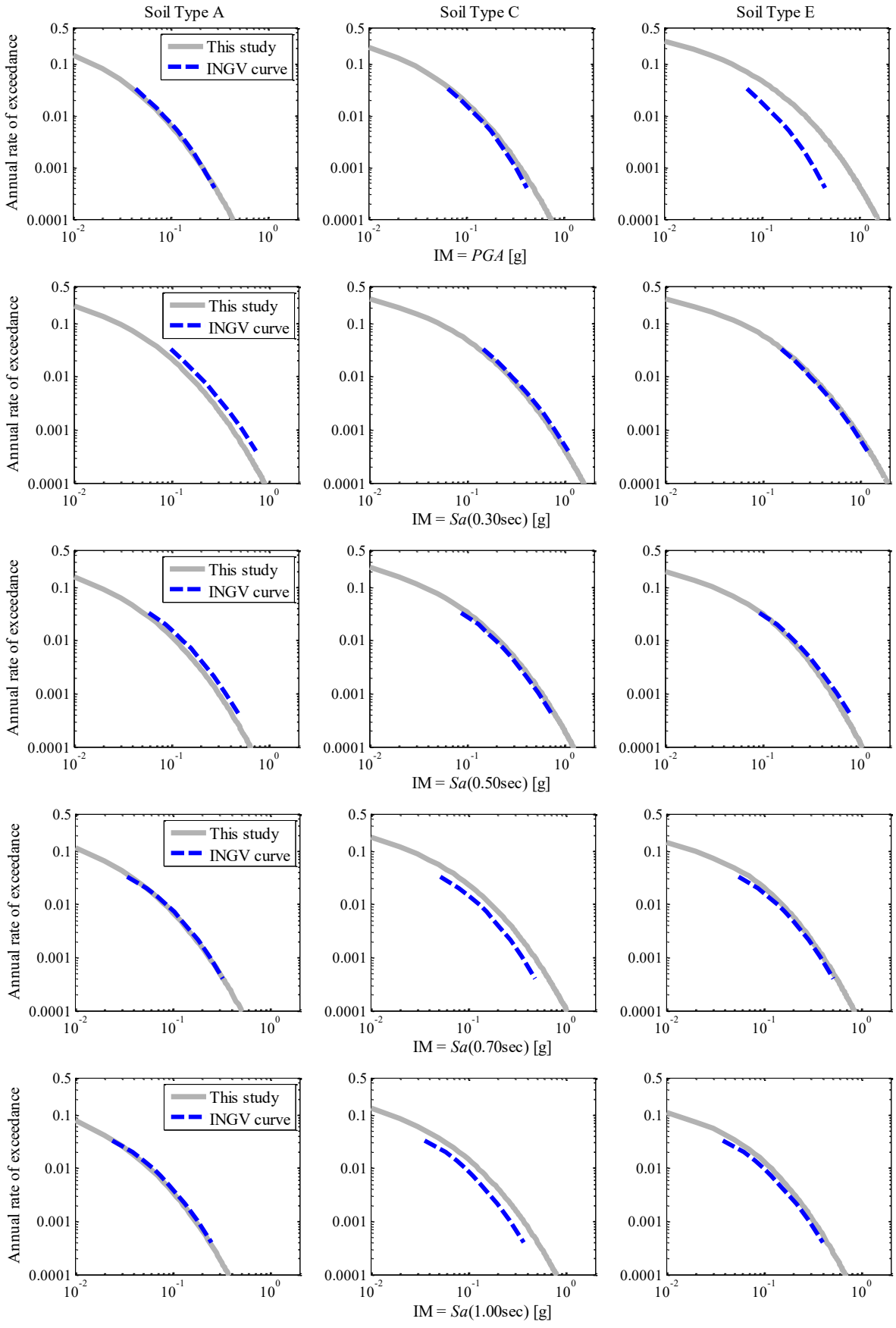


Figure 16: Comparison of the PSHA curves estimated by this study with the INGV derived hazard curves for five target intensities and considering different soil types of the case-study area

8 SUMMARY AND CONCLUSIONS

A Probabilistic seismic hazard analysis (PSHA) is performed for the western area of the city Naples (southern Italy) using the most recent parametric catalog of Italian earthquakes (CPTI15) and the enhanced database of Individual Seismogenic Sources (DISS 3.2). The seismogenic models used herein comprise of (1) individual seismogenic structures/faults (a.k.a. seismogenic boxes) liable to generate major earthquakes with magnitudes greater than 5.5; (2) background areal source model which allows to evaluate earthquakes with magnitude less than 5.5 and two point sources. Three most recent Italian, European and global ground motion prediction equations (GMPE's) are adopted herein; namely, ITA10 (Bindi et al. 2011), BND14 (Bindi et al. 2014a, b) and BSSA (Boore et al. 2014). Since both ITA10 and BND14 use the geometric mean of the two horizontal components of ground motion, these two GMPEs are modified in order to account for an arbitrary horizontal component of ground-shaking.

As a result, the PSHA is performed based on a bi-layer model of seismogenic tectonic faults, together with the background spatial model. The site amplification is considered through a detailed seismic microzonation study derived for the western area of Naples based on reliable geological and geotechnical subsoil models. It is characterized by different expected amplification levels of ground motion and instability due to liquefaction. PSHA results obtained from fault and background source models are presented for a number of return periods (corresponding to prescribed probability of exceedance in 50 years) for peak ground acceleration (*PGA*) and 5% damped pseudo spectral accelerations (S_a) at periods representative of the existing masonry and reinforced concrete (RC) structures of the case-study area. The seismic hazard results obtained in this study are compared (in terms of the uniform hazard spectra, UHS, and the hazard curves) with the national hazard maps provided by INGV and the national Italian code (NTC 2018) for the designated site. The following major observations are made:

- The seismic intensities suggested by INGV for structures with period greater than 0.5 sec are underestimated especially for medium soil sites (NTC soil types C) compared to the study performed herein. The code-based spectrum overestimates the seismic intensities for periods greater than 0.5 sec for all soil types in the area.
- The UHS is considerably underestimated both by INGV and national code for periods smaller than 0.30 sec for soft soil type E. Moreover, the *PGA* estimates based on INGV on soft soil type E are underestimated.
- For medium soil sites (soil types C), INGV-based hazard curves underestimate the annual exceedance frequencies compared to our study, especially for S_a at periods greater than 0.5sec. The level of underestimation is more accentuated for soil type E; the INGV-based *PGA* hazard curve and that for S_a at periods smaller than 0.3 sec are considerably underestimated.
- It is to note that the INGV hazard data (Meletti et al. 2007, http://esse1.mi.ingv.it/) are defined on stiff-soil sites (V_{S30}>800m/sec). Hence, in order to account for soft soil conditions, the intensities should be multiplied by the code-based site amplification factor. This leads to an underestimation of ground shaking for soft soil type E at small periods. Therefore, it is more suitable that the nonlinear site response be considered properly within the PSHA by incorporating it in the hazard calculation at the GMPE level.
- More attention needs to be focused on the consideration of local amplification on soft soil (especially soil type E).

The following limitations should be kept in mind:

- The topographical effect is not included within the site amplification factor. This is due to the fact that the current GMPEs used in this study do not include this effect.
- The effect of volcano-tectonic earthquakes (better to be called earthquake swarms) are not included herein (Convertito and Zollo, 2011). These type of low-magnitude earthquakes can affect the low-intensity region of the seismic hazard curve significantly, and hence, affect the seismic risk estimates.
- With reference to seismic microzonation map Fig. 1(b), a part of the case-study area is susceptible to liquefaction (denoted as S2 in the corresponding figure and outlined in Table 1). The liquefaction triggering and its consequences are not evaluated herein.
- Further studies should be carried out for characterizing the zone with unknown soil type defined as ND in Fig. 2b. As outlined in Table 1, we have assigned type C to the subsoil for this zone within the study area.
- The uncertainty in the soil amplification model parameters and its effect on the PSHA results should be explored and verified. In addition, employing a more sophisticated amplification model may address the site effects more properly.
- It should be underlined that the current study does not provide seismic hazard deaggregation results which are useful for construction of conditional spectra (see Ebrahimian et al. 2012).
- The layered source model adopted herein cannot consider events larger than 5.5 occurring in the background zone. However, the catalog of historical earthquakes seems to support the possibility of having events with

magnitude larger than 5.5 in the background area. These events, which are not attributed by DISS data base (Basili et al. 2008) to any known causative fault structure, are treated within the PSHA framework herein as point sources (with rough assumptions of their characteristic magnitudes and seismicity rates). These sources need further investigation.

AKNOWLEDGEMENT

This work was supported in part by Project METROPOLIS (**Me**todologie e **T**ecnologie Integ**r**ate e S**o**stenibili **P**er L'adattament**o** e **L**a S**i**curezza di **S**istemi Urbani). This support is gratefully acknowledged. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the sponsor. The authors would also like to gratefully acknowledge Dr. Dino Bindi (Helmholtz Centre Potsdam) for the fruitful discussions on the database of ground motions used for ITA10 and BND14 ground-motion prediction models. In addition, the authors would like to gratefully acknowledge Dr. Carlo Del Gaudio (University of Naples Federico II) for providing information and statistics on the buildings in the Western Naples including the zones of Bagnoli and Fuorigrotta. Last but not least, the authors would like to acknowledge the scientific coordinator Prof. Gerardo M. Verderame (University of Naples Federico II) and STRESS Scarl staff for their invaluable support and help through the METROPOLIS Project. The authors would also like to acknowledge the anonymous reviewers who have contributed significantly to improving and enriching the paper.

REFERENCES

Abrahamson NA (2000). State of the practice of seismic hazard evaluation. *International Society for Rock Mechanics (ISRM), International Symposium*, 19-24 November, Melbourne, Australia, ISRM-IS-2000-014.

Akkar S, Sandıkkaya MA, ¸Senyurt M et al (2013) Reference database for seismic ground-motion in Europe (RESORCE). *Bull Earthq Eng*, 12(1): 311-339.

Akinci A, Galadini F, Pantosti D, Petersen M, Malagnini L, Perkins D (2009) Effect of time dependence on probabilistic seismic-hazard maps and deaggregation for the central Apennines, Italy. *Bull Seismol Soc Am* 99(2A): 585-610.

Akinci A, Vannoli P, Falcone G et al (2016) When time and faults matter: towards a time-dependent probabilistic SHA in Calabria, Italy. *Bull Earthq Eng*, 15(6): 2497-2524.

Albarello D, Bosi V, Bramerini F, Lucantoni A, Naso G, Peruzza L, Rebez A, Sabetta F, Slejko D (2000) Carte di pericolosità sismica del territorio nazionale, *Quaderni di Geofisica* 12, Roma, 7 pp.

Alessio G, Esposito E, Gorini A, Porfido S. (1995) Detailed study of the Potentino seismic zone in the Southern Apennines, Italy. *Tectonophysics*, 250(1): 113-134.

Ambraseys NN (1995) The prediction of earthquake peak ground acceleration in Europe, *Earthq Eng Struct Dyn* 24(4): 467-490.

Ambraseys NN, Simpson KA, Bommer JJ (1996). Prediction of horizontal response spectra in Europe. *Earthq Eng Struct Dyn* 25(4): 371-400.

Ancheta TD, Darragh RB, Stewart JP et al (2014) NGA-West2 Database. *Earthquake Spectra*: August 2014, Vol. 30, No. 3, pp. 989-1005.

Baker JW, Cornell CA (2006) Which spectral acceleration are you using?, *Earthquake Spectra* 22(2): 293-312.

Barani, S., Albarello, D., Spallarossa, D., & Massa, M. (2017a). Empirical scoring of ground motion prediction equations for probabilistic seismic hazard analysis in Italy including site effects. *Bull Earthq Eng*, 15(6): 2547-2570.

Barani S, Albarello D, Massa M, Spallarossa D (2017b) Influence of twenty years of research on ground motion prediction equations on probabilistic seismic hazard in Italy. *Bull Seismol Soc Am*, 107(1): 240-255.

Basili R, Valensise G, Vannoli P, Burrato P, Fracassi U, Mariano S, Tiberti MM, Boschi E (2008) The Database of Individual Seismogenic Sources (DISS), version 3: summarizing 20 years of research on Italy's earthquake geology. *Tectonophysics* 453:20-43. doi:10.1016/j.tecto.2007.04.014.

Benito MB, Navarro M, Vidal F et al (2010) A new seismic hazard assessment in the region of Andalusia (Southern Spain). *Bull Earthq Eng*, 8 (4): 739-766.

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(5): 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(6):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(1): 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(1): 431-448.

Boncio P, Lavecchia G, Pace B (2004) Defining a model of 3D seismogenic sources for seismic hazard assessment applications: The case of central Apennines (Italy). *J Seismol* 8(3): 407–425.

Boore DM (2005) Erratum: equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: a summary of recent work, *Seismol Res Lett* 76(3): 368–369.

Boore DM (2010) Orientation-independent, non geometric-mean measures of seismic intensity from two horizontal components of motion. *Bull Seismol Soc Am*, 100(4): 1830-1835.

Boore DM, Stewart JP, Seyhan E, Atkinson GM (2014) NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. *Earthq Spectra*, 30(3): 1057-1085.

Boschi E, Pantosti D, Valensise G (1995) La valutazione del potenziale sismogenetico in Italia: progressi metodologici e conoscitivi. *Proc. Meeting Terremoti in Italia: previsione e prevenzione dei danni*, Accademia Nazionale dei Lincei, Rome, 1-2 December 1994: 133-138.

Bozorgnia Y, Abrahamson NA, Al Atik L et al. (2014) NGA-West2 Research Project. *Earthq Spectra*, 30(3): 973-987.

Camassi R, Stucchi M (1997). NT 4.1.1, un catalogo parametrico di terremoti di area italiana al di sopra della soglia di danno, GNDT Technical Report, Milano, available at http://emidius.mi.ingv.it/NT/home.html (last accessed March 2018), 66 pp.

Camassi R, Castelli V, Molin D, Bernardini F, Caracciolo C H, Ercolani E, Postpischl L (2011) *Materiali per un catalogo dei terremoti italiani: eventi sconosciuti, rivalutati o riscoperti*. Quaderni di Geofisica 96 INGV, Roma, 53pp (in Italian).

Campbell KW, Bozorgnia Y (2008) NGA Ground Motion Model for the Geometric Mean Horizontal Component of PGA, PGV, PGD and 5% Damped Linear Elastic Response Spectra for Periods Ranging from 0.01 to 10 s. *Earthq Spectra* 24(1): 139-171.

Castelli V, Galli P, Camassi R, Caracciolo CH (2008) The 1561 earthquake(s) in Southern Italy: New Insights into a Complex Seismic Sequence. *J Earthq Eng* 12(7): 1054-1077.

CEN - European standard EN1998-1 (2004) *Eurocode 8: design of structures for earthquake resistance - Part 1: general rules, seismic actions and rules for buildings*. Comité Européen de Normalisation, Brussels.

Chiodini G, Caliro S, Cardellini C, Granieri D, Avino R, Baldini A, Donnini C, Minopoli C (2010). Long-term variations of the Campi Flegrei, Italy, volcanic system as revealed by the monitoring of hydrothermal activity. *J Geophysical Res Solid Earth* 115(B3): 1-17.

Chopra AK (2012) *Dynamics of structures. Theory and application to earthquake engineering*. 4th Edition, Prentice Hall, Upper Saddle River.

Convertito V, Emolo A, Zollo A (2006) Seismic-hazard assessment for a characteristic earthquake scenario: an integrated probabilistic–deterministic method. *Bull Seismol Soc Am*, 96 (2): 377-391.

Convertito V, Zollo A (2011) Assessment of pre-crisis and syn-crisis seismic hazard at Campi Flegrei and Mt. Vesuvius volcanoes, Campania, southern Italy. *Bull of Volcanology*, 73(6): 767-783.

Cornell CA (1968) Engineering seismic risk analysis. *Bull Seismol Soc Am*, 58 (5): 1583-1606.

D'addezio G, Masana E, Pantosti D (2001) The holocene paleoseismicity of the Aremogna-Cinque Miglia fault (central Italy). *Journal of Seismology* 5(2): 181-205.

Danciu L, Şeşetyan K, Demircioğlu M et al (2018) The 2014 earthquake model of the Middle East: seismogenic sources. *Bull Earthq Eng*, 16(8): 3465-3496.

Di Giacomo D, Storchak DA, Safronova N, Ozgo P, Harris J, Verney R, Bondár I (2014) A New ISC Service: The Bibliography of Seismic Events. *Seismol Res Lett* 85 (2): 354-360.

Deino AL, Orsi G, de Vita S, Piochi M (2004) The age of the Neapolitan Yellow Tuff caldera-forming eruption (Campi Flegrei caldera - Italy) assessed by 40Ar/39Ar dating method. *J Volcanol Geotherm Res* 133 (1-4): 157-170.

Del Gaudio C, Aquino I, Ricciardi GP, Ricco C, Scandone R (2010) Unrest episodes at Campi Flegrei: A reconstruction of vertical ground movements during 1905–2009. *J Volcanol Geotherm Res* 195(1):48-56.

Demircioğlu MB, Şeşetyan K, Duman TY et al (2018). A probabilistic seismic hazard assessment for the Turkish territory: part II—fault source and background seismicity model. *Bull Earthq Eng* 16(8): 3399-3438.

De Vivo B, Rolandi G, Gans PB, Calvert A, Bohrson WA, Spera FJ, Belkin HE (2001) New constraints on the pyroclastic eruptive history of the Campanian Volcanic Plain (Italy). *Mineral Petrol* 73 (1-3): 47-65.

Di Bucci D, Corrado S, Naso G (2002) Active faults at the boundary between Central and Southern Apennines (Isernia, Italy). *Tectonophysics*, 359(1): 47-63.

Di Vito MA, Isaia R, Orsi G, Southon J, De Vita S, D'Antonio M, Pappalardo L, Piochi M (1999) Volcanism and deformation since 12,000 years at the Campi Flegrei caldera (Italy). *J Volcanol Geotherm Res* 91 (2-4): 221-246.

DISS Working Group (2015). Database of Individual Seismogenic Sources (DISS), Version 3.2.0: A compilation of potential sources for earthquakes larger than M 5.5 in Italy and surrounding areas. *Istituto Nazionale di Geofisica e Vulcanologia*; doi:10.6092/INGV.IT-DISS3.2.0, http://diss.rm.ingv.it/diss/.

Ebrahimian H, Azarbakht AR, Tabandeh A, Golafshani AA (2012) The exact and approximate conditional spectra in the multi-seismic-sources regions. *Soil Dyn Earthq Eng* 39 (1): 61-77.

Ebrahimian H, Jalayer F, Asprone D, Lombardi AM, Marzocchi W, Prota A, Manfredi G. (2014) Adaptive daily forecasting of seismic aftershock hazard. *Bull. Seism. Soc. Am.*, 104 (1): 145-161.

Ebrahimian H, Jalayer F, Lucchini A, Mollaioli F, Manfredi G (2015) Preliminary ranking of alternative scalar and vector intensity measures of ground shaking. *Bull Earthquake Eng* 13(10): 2805-2840.

Ebrahimian H, Jalayer F (2017) Robust seismicity forecasting based on Bayesian parameter estimation for epidemiological spatio-temporal aftershock clustering models. *Scientific Reports*, Nature, 7(9803): 1-15. doi: 10.1038/s41598-017-09962-z.

El-Hussain I, Deif A, Al-Jabri K et al (2012) Probabilistic seismic hazard maps for the sultanate of Oman. *Natural hazards* 64 (1): 173-210.

Faenza L, Marzocchi W, Boschi E (2003) A non-parametric hazard model to characterize the spatio-temporal occurrence of large earthquakes; an application to the Italian catalogue. *Geophys J Int* 155(2): 521-531.

Faenza L, Pierdominici S, Hainzl S, Cinti FR, Sandri L, Selva J, Tonini R, Perfetti P (2017) A Bayesian seismic hazard analysis for the city of Naples. *J. Geophys. Res. Solid Earth*, 122 (3), 1990-2012.

Fedele L, Scarpati C, Lanphere M, Melluso L, Morra V, Perrotta A, Ricci G (2008) The Breccia Museo formation, Campi Flegrei, southern Italy: geochronology, chemostratigraphy and relationship with the Campanian Ignimbrite eruption. *Bull Volcanol* 70 (10): 1189-1219.

Field EH, Johnson DD, Dolan JF (1999) A mutually consistent seismic-hazard source model for southern California, *Bull Seismol Soc Am* 89(3): 559-578.

Forte G, Fabbrocino S, Fabbrocino G, Lanzano G, Santucci de Magistris F, Silvestri F (2017). A geolithological approach to seismic site classification: an application to the Molise Region (Italy). *Bull Earthquake Eng*, 15 (1): 175–198.

Fracassi U, Valensise G (2007) Unveiling the sources of the catastrophic 1456 multiple earthquake: hints to an unexplored tectonic mechanism in southern Italy. *Bull Seismol Soc Am*, 97(3): 725-748.

Galli PAC, Naso JA (2009) Unmasking the 1349 earthquake source (southern Italy): paleoseismological and archaeoseimological indications from the Aquae Iuliae fault. *J. Struct. Geol.*, 31(2): 128-149.

Giardini D (1999). The Global Seismic Hazard Assessment Program (GSHAP) - 1992/1999. *Ann. Geophys* 42(6): 957-974.

Gregor N, Abrahamson NA, Atkinson GM et al (2014) Comparison of NGA-West2 GMPEs. *Earthq Spectra*, 30(3): 1179-1197.

Gruppo di Lavoro (1999). Proposta di riclassificazione sismica del territorio nazionale, *Ingegneria Sismica* 16(1): 5-14.

Gruppo di Lavoro CPTI (1999). *Catalogo Parametrico dei Terremoti Italiani*, ING, GNDT, SGA, SSN, Bologna. doi: 10.6092/INGV.IT-CPTI99.

Gruppo di lavoro CPTI (2004). *Catalogo Parametrico dei Terremoti Italiani, versione 2004 (CPTI04)*, INGV, Bologna. doi: 10.6092/INGV.IT-CPTI04.

Gruppo di Lavoro (2004). *Redazione della mappa di pericolosità sismica prevista dall'Ordinanza PCM 3274 del 20 marzo 2003*. Rapporto Conclusivo per il Dipartimento della Protezione Civile, INGV, Milano-Roma, April 2004: 65 pp. + 5 appendixes (in Italian).

Gutenberg B, Richter CF (1949) Seismicity of the earth and associated phenomena. Princeton UniversityPress, Princeton.

Gutenberg B, Richter CF (1956) Magnitude and energy of earthquakes. *Ann Geophyis*, 9: 1-15.

Gülerce Z, Ocak S (2013). Probabilistic seismic hazard assessment of Eastern Marmara Region. Bulletin of Earthquake Engineering, 11(5): 1259-1277.

Hanks TC, Kanamori H (1979) A moment magnitude scale. *J Geophys Res* 84(B5), 2348-2350.

International Seismological Centre, On-line Event Bibliography, http://www.isc.ac.uk/event_bibliography, Internatl. Seis. Cent., Thatcham, United Kingdom, 20yy.

Jalayer F, Cornell CA (2009) Alternative non-linear demand estimation methods for probability-based seismic assessments. *Earthq Eng Struct Dyn*, 38 (8): 951-972.

Jalayer F, Asprone D, Prota A, Manfredi G (2011) A decision support system for post-earthquake reliability assessment of structures subjected to aftershocks: an application to L'Aquila earthquake, 2009. *Bull Earthq Eng*, **9** (4): 997-1014.

Jalayer F, Ebrahimian H (2017) Seismic risk assessment considering cumulative damage due to aftershocks. *Earthq Eng Struct Dyn* 46(3): 369-389.

Jalayer F, Ebrahimian H, Miano A, Manfredi G, Sezen H (2017) Analytical fragility assessment using un-scaled ground motion records. *Earthq Eng Struct Dyn*, 46(15): 2639-2663.

Jiménez MJ, Giardini G, Grünthal G, et al. (2001) Unified seismic hazard modelling throughout the Mediterranean region. *Boll Geof Teor Appl* 42(1-2): 3-18.
Kadirioğlu FT, Kartal RF, Kılıç T et al (2018) An improved earthquake catalogue (M≥ 4.0) for Turkey and its near vicinity (1900–2012). *Bull Earthq Eng* 16(8): 3317-3338.
Kagan YY, Jackson DD (2013) Tohoku Earthquake: A Surprise?. *Bull Seismol Soc Am* 103 (2B): 1181-1194.
Ksentini A, Romdhane NB (2014) Updated seismic hazard assessment of Tunisia. *Bull Earthq Eng* 12 (2): 647-670.
Leonard M (2010). Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release. *Bull Seismol Soc Am* 100 (5A): 1971-1988.
Licata V, Forte G, d'Onofrio A, Evangelista L, Jalayer F, Santo A, Silvestri F. (2016). Microzonation Study on the Western area of Napoli. *Procedia Engineering*, 158: 511-516.
Licata V, Forte G, d'Onofrio A, Santo A, Silvestri F (2019a). A multi-level study for the seismic microzonation of the Western area of Napoli (Italy). *Bull Earthq Eng* (under review).
Licata V, Forte G, Ebrahimian H, d'Onofrio A, Jalayer F, Santo A, Silvestri F (2019b). Evaluation of the seismic ground amplification considering the variability of the depth of the bedrock and random shear wave velocity profiles. *7th International Conference on Earthquake Geotechnical Engineering (VII ICEGE)*, Rome, Italy, 17-20 June 2019.
Luzi L, Sabetta F, Hailemikael S, Bindi D, Pacor F,Mele F (2008) ITACA (ITalian ACcelerometric Archive): a web portal for the dissemination of Italian strong motion data. *Seismol Res Lett*, 79(5):716-722. doi: 10.1785/gssrl.79.5.716.
Marzocchi W, Sandri L, Boschi E (2003). On a validation of earthquake-forecasting models: the case of pattern recognition algorithms. *Bull Seism Soc Am* 93(5): 1994-2004.
Marzocchi W, Sandri L, Heuret A, Funiciello F (2016) Where giant earthquakes may come? *J. Geophys. Res. Solid Earth*, 121(10): 7322–7336.
McGuire RK (2008) Probabilistic seismic hazard analysis: Early history. *Earthq Eng Struct Dyn*, 37(3):329-338.
Meletti C, Patacca E, Scandone P (2000). Construction of a seismotectonic model: The case of Italy. *Pure Appl Geophys* 157: 11-35.
Meletti C, Calvi GM, Stucchi M (2007) *Project S1 - continuation of assistance to DPC for improving and using the seismic hazard map compiled according to the Prime Minister "Ordinanza" 3274/2003 and planning future initiatives - final report*, INGV (in Italian). Interactive maps of seismic hazard (WebGis). Available at http://esse1.mi.ingv.it/.
Meletti C, Montaldo V (2007) *Stime di pericolosità sismica per diverse probabilità di superamento in 50 anni: valori di ag*. Project DPC-INGV S1, Deliverable D2 (in Italian). Istituto Nazionale di Geofisica e Vulcanologia – Sezione di Milano-Pavia. Available at http://esse1.mi.ingv.it/d2.html
Meletti C, Galadini F, Valensise G, Stucchi M, Basili R, Barba G, Vannucci G, Boschi E (2008) A seismic source model for the seismic hazard assessment of the Italian territory. *Tectonophysics* 450(1): 85-108.
Mihaljević J, Zupančič P, Kuka N et al (2017) BSHAP seismic source characterization models for the Western Balkan region. *Bull Earthq Eng* 15(10): 3963-3985.
Montaldo, V, Faccioli E, Zonno G, Akinci A, Malagnini L (2005) Treatment of ground-motion predictive relationships for the reference seismic hazard map of Italy. *J Seismol* 9(3): 295-316.
Montaldo V, Meletti C (2007) *Valutazione del valore della ordinata spettrale a 1 sec e ad altri periodi di interesse ingegneristico*. Project DPC-INGV S1, Deliverable D3 (in Italian). Istituto Nazionale di Geofisica e Vulcanologia – Sezione di Milano-Pavia. Available at http://esse1.mi.ingv.it/d3.html
Mulargia F, Gasperini P, Tinti S (1987) Contour mapping of Italian seismicity. *Tectonophysics* 142: 203–216.
NTC (2008). *Norme Tecniche per le Costruzioni*. Gazzetta Ufficiale 29, 4 February 2008.
NTC (2018). *Aggiornamento delle «Norme tecniche per le costruzioni»*. Gazzetta Ufficiale 42, 20 February 2018.
Orsi G, Di Vito MA, Isaia R (2004) Volcanic hazard assessment at the restless Campi Flegrei caldera. *Bull. Volcanol.* 66 (6): 514-530.
Ornthammarath T, Warnitchai P, Worakanchana K et al (2011) Probabilistic seismic hazard assessment for Thailand. *Bull Earthq Eng* 9 (2): 367-394.
Pace B, Peruzza L, Lavecchia G, Boncio P (2002). Seismogenic sources in Central Italy: from causes to effects. *Mem Soc Geol Ital* 57: 419-429.
Pace B, Peruzza L, Lavecchia G, Boncio P (2006). Layered seismogenic source model and probabilistic seismic-hazard analyses in central Italy. *Bull. Seismol. Soc. Am.* 96 (1): 107-132.
Pace B, Visini F, Peruzza L (2016). FiSH: MATLAB Tools to Turn Fault Data into Seismic-Hazard Models. *Seismol Res Lett*, 87(2A): 374-386. doi: 10.1785/0220150189.
Pacor F, Paolucci R, Ameri G, Massa M, Puglia R (2011) Italian strong motion records in ITACA: overview and record processing, *Bull Earth Eng*, 9(6): 1741-1759.
Pantosti D, Valensise G (1988) La faglia sud-appenninica: identificazione oggettiva di un lineamento sismogenetico nell'Appennino meridionale. *Proceedings 7° Meeting G.N.G.T.S.*, Rome 1988: 205-220.

Parra H, Benito MB, Gaspar-Escribano JM (2016) Seismic hazard assessment in continental Ecuador. *Bull Earthq Eng* 14 (8): 2129-2159.
Peruzza L, Pace B (2002). Sensitivity analysis for seismic source characteristics to probabilistic seismic hazard assessment in central Apennines (Abruzzo area). *Boll. Geofis. Teor. Appl.* 43: 79-100.
Peruzza L, Pace B, Cavallini F (2010) Error propagation in time-dependent probability of occurrence for characteristic earthquakes in Italy. *J Seismol* 14 (1): 119-141.
Pino NA, Palombo B, Ventura G, Perniola B, Ferrari G (2008) Waveform modeling of historical seismograms of the 1930 Irpinia earthquake provides insight on 'blind' faulting in Southern Apennines (Italy). *J. Geophys. Res.*, 113(B05303). doi: 10.1029/2007JB005211.
Ricci P, Verderame GM, Manfredi G (2011) Analytical investigation of elastic period of infilled RC MRF buildings. *Eng Struct* 33(2): 308-319.
Romeo, R., and A. Pugliese (2000). Seismicity, seismotectonics and seismic hazard of Italy. *Eng Geol* 55(4): 241–266.
Romeo R., Paciello A, Rinaldis D (2000). Seismic hazard maps of Italy including site effects, *Soil Dyn. Earthq Eng* 20(1-4): 85–92.
Rovida A, Locati M, Camassi R, Lolli B, Gasperini P (eds) (2016). *CPTI15, the 2015 version of the Parametric Catalogue of Italian Earthquakes*. Istituto Nazionale di Geofisica e Vulcanologia http://doi.org/10.6092/INGV.IT-CPTI15 (accessed on 24 March 2017)
Sabetta F, Pugliese A (1996) Estimation of response spectra and simulation of nonstationary earthquake ground motions. *Bull Seismol Soc Am* 86(2): 337-352.
Sesetyan K, Demircioglu MB, Duman TY et al (2016) A probabilistic seismic hazard assessment for the Turkish territory—part I: the area source model. *Bull Earthq Eng* 16(8): 3367-3397.
Seyhan E, Stewart JP (2014) Semi-empirical nonlinear site amplification from NGAWest2 data and simulations. *Earthq Spectra* 30(3): 1241–1256.
Seyhan E, Stewart JP, Ancheta TD, Darragh RB, Graves RW (2014) NGA-West2 site database. *Earthq Spectra* 30(3): 1007–1024.
Slejko D, Peruzza L, Rebez A (1998). The seismic hazard maps of Italy. *Ann Geophys* 41(2): 183-214.
Slejko D, Camassi R, Cecic I, et al. (1999) Seismic hazard assessment of Adria, *Ann Geophys* 42(6): 1085-1107.
Sokolov V, Zahran HM, Youssef SEH et al (2017). Probabilistic seismic hazard assessment for Saudi Arabia using spatially smoothed seismicity and analysis of hazard uncertainty. *Bull Earthq Eng* 15(7): 2695-2735.
Stepp JC (1972) Analysis of completeness of the earthquake sample in the Puget Sound area and its effect on statistical estimates of earthquake hazard. In: Proceedings of first microzonation conference, Seattle, USA, pp 897–909.
Stucchi M, Albini P, Mirto C, Rebez A (2004) Assessing the completeness of Italian historical earthquake data. *Ann Geophys* 47(2/3): 659-673.
Stucchi M, Meletti C, Montaldo V, Crowley H, Calvi GM, Boschi E (2011) Seismic hazard assessment (2003–2009) for the Italian building code. *Bull Seismol Soc Am* 101(4): 1885-1911.
Valensise G, Pantosti D (Editors) (2001) Database of potential sources for earthquakes larger than M 5.5 in Italy. *Ann. Geophys.*, 44(4): 797-964, with CD-Rom.
Vanini M, Corigliano M, Faccioli E, Figini R, Luzi L, Pacor F, Paolucci R (2018) Improving seismic hazard approaches for critical infrastructures: a pilot study in the Po Plain. *Bull Earthq Eng* 16(6): 2529-2564.
Vannoli P, Burrato P, Valensise G (2015) The seismotectonic of the Po Plain (northern Italy): tectonic diversity in a blind faulting domain. *Pure Appl Geophys* 172(5):1105–1142.
Wells DL, Coppersmith KJ (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bull Seismol Soc Am* 84 (4): 974-1002.
Westaway R (1993) Fault rupture geometry for the 1980 Irpinia earthquake: a working hypothesis. *Ann. Geophys.*, 36(1): 51-69.
Woessner J, Laurentiu D, Giardini D et al (2015). The 2013 European seismic hazard model: key components and results. *Bull Earthq Eng* 13 (12): 3553-3596.
Youngs RR, Coppersmith KJ (1985) Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates. *Bull Seismol Soc Am*, 75 (4): 939-964.
Zimmaro P, Stewart JP (2017) Site-specific seismic hazard analysis for Calabrian dam site using regionally customized seismic source and ground motion models. *Soil Dyn Earthq Eng* 94: 179-192.
Zöller G, Holschneider M, Hainzl S (2013) The Maximum Earthquake Magnitude in a Time Horizon: Theory and Case Studies. *Bull Seismol Soc Am* 103 (2A): 860-875.
Zuccolo E, Corigliano M, Lai CG (2013) Probabilistic seismic hazard assessment of Italy using kernel estimation methods. *J Seismol* 17(3): 1001-1020.

Appendix 1

Consider the standard deviation of the residuals of a GMPE that provides the geometric mean associated with the two horizontal components of ground motion, be denoted as $\sigma_{\ln IM_{g.m.}}$. To obtain the standard deviation for the arbitrary horizontal ground motion, $\sigma_{\ln IM_{arb}}$, we calculate the variance of the finite record numbers N in the database as follows:

$$\begin{aligned}\sigma^2_{\ln IM_{g.m.}} &= \frac{1}{N}\sum_{j=1}^{N}\left(\ln IM_{g.m.,j} - \mu_{\ln IM_{g.m.}}\right)^2 = \frac{1}{N}\sum_{j=1}^{N}\ln IM^2_{g.m.,j} - \mu^2_{\ln IM_{g.m.}} \\ &= \frac{1}{4}\left[\frac{1}{N}\sum_{j=1}^{N}\ln IM^2_{x,j} + \frac{1}{N}\sum_{j=1}^{N}\ln IM^2_{y,j} + \frac{1}{N}\sum_{j=1}^{N}2\ln IM_{x,j}\ln IM_{y,j}\right] - \mu^2_{\ln IM_{g.m.}}\end{aligned} \tag{A1}$$

where $IM_{g.m.j}$ is the ground motion intensity for the geometric mean of the two horizontal components of jth recording, $j\epsilon[1,\ldots,N]$, $\mu_{\ln IM_{g.m.}}$ is the logarithmic mean of the GMPE (see Eq. 14). Knowing that

$$\sigma^2_{\ln IM_{arb}} = \frac{1}{N}\sum_{j=1}^{N}\left(\ln IM_{arb,j} - \mu_{\ln IM_{arb}}\right)^2 = \frac{1}{N}\sum_{j=1}^{N}\ln IM^2_{arb,j} - \mu^2_{\ln IM_{arb}} \tag{A2}$$

Eq. (A1) can be written as,

$$\begin{aligned}\sigma^2_{\ln IM_{g.m.}} &= \frac{1}{4}\left[2\sigma^2_{\ln IM_{arb}} + 2\mu^2_{\ln IM_{arb}} + \frac{1}{N}\sum_{j=1}^{N}2\ln IM_{x,j}\ln IM_{y,j}\right] - \mu^2_{\ln IM_{g.m.}} \\ &= \frac{1}{4}\left[2\sigma^2_{\ln IM_{arb}} + 2\mu^2_{\ln IM_{arb}} + \frac{1}{2N}\left(\sum_{j=1}^{N}\left(\ln IM_{x,j} + \ln IM_{y,j}\right)^2 - \sum_{j=1}^{N}\left(\ln IM_{x,j} - \ln IM_{y,j}\right)^2\right)\right] - \mu^2_{\ln IM_{g.m.}} \\ &= \frac{1}{4}\left[2\sigma^2_{\ln IM_{arb}} + 2\mu^2_{\ln IM_{arb}} + 2\sigma^2_{\ln IM_{g.m.}} + 2\mu^2_{\ln IM_{g.m.}} - \frac{1}{2N}\sum_{j=1}^{N}\left(\ln IM_{x,j} - \ln IM_{y,j}\right)^2\right] - \mu^2_{\ln IM_{g.m.}} \\ &= \frac{1}{2}\sigma^2_{\ln IM_{arb}} + \frac{1}{2}\sigma^2_{\ln IM_{g.m.}} - \frac{1}{8N}\sum_{j=1}^{N}\left(\ln IM_{x,j} - \ln IM_{y,j}\right)^2\end{aligned} \tag{A3}$$

Hence, we have,

$$\sigma^2_{\ln IM_{arb}} = \sigma^2_{\ln IM_{g.m.}} + \underbrace{\frac{1}{4N}\sum_{j=1}^{N}\left(\ln IM_{x} - \ln IM_{y}\right)^2}_{\sigma_C^2} \tag{A4}$$