



Università degli Studi di Camerino

School of Advanced Studies

Doctoral course in
Physics, Earth and Materials Sciences

Curriculum
Physical and Chemical Processes in Earth Systems

Cycle XXXVII

Calibration of correction factors for source, attenuation and site effects to predict ground shaking across Italy

PhD Student
Giulio Brunelli

Supervisor
Prof. Emanuele Tondi

Co-supervisors
Dr. Lucia Luzi
Dr. Sara Sgobba

Coordinator of the PhD Programme

Prof. Roberto Gunnella

Date of award __/__/2025

Acknowledgments

Abstract

This doctoral research focuses on improving the spatial prediction of ground shaking across Italy by developing empirical correction factors for source and attenuation effects and creating a new national-scale Vs30 map to account for site conditions. These tools are designed to support and enhance the generation of empirical ground-shaking scenarios at both regional and local scales.

The work is framed within the fields of applied seismology and seismic hazard assessment. Its goal is to improve the accuracy of ground motion predictions and reduce the uncertainty associated with the classical ergodic assumption still widely used in empirical Ground Motion Models (GMMs).

Ground motion is the result of the combination of three fundamental components: the earthquake source, the wave propagation path, and local geological conditions (site effects). While traditional empirical GMMs allow for fast and reliable ground motion estimates, they rely on the ergodic assumption, which implies that variability observed across multiple sites is representative of the variability expected at a single site over time. However, practically, this assumption tends to overestimate uncertainty and fails to capture regional geological and tectonic peculiarities.

In this context, the regionalization of source and attenuation effects is based on the analysis of residuals, defined as the logarithmic differences between ground motion observations and model predictions, and their decomposition into repeatable components following the methodology of Al Atik et al. (2010): between-event residuals (δBe), representing source effects not captured by the GMMs; site-to-site residuals ($\delta S2S$), associated with site effects not explained by the reference model; and within-event residuals (δWes) corrected for source and site, which may include additional effects not accounted for by the model, such as regional propagation phenomena or source directivity.

In this thesis, the δBe and δWes components are spatially modeled to identify homogeneous zones and derive regional correction factors for source and attenuation.

The entire process is supported by a robust dataset, in particular the ITACAext database (Brunelli et al., 2022), which includes more than 37,000 accelerometric and velocimetric recordings from over 1,800 earthquakes across Italy. As a predictive model, the ITA18 GMM (Lanzano et al., 2019), developed for shallow crustal earthquakes in Italy, is selected as the national reference.

A large part of the work of this thesis has been spent on the evaluation of site effects, through a new nationwide map of the shear wave velocity in the uppermost 30 m (V_{s30}), constructed through parametric linear regressions and geostatistical analysis. To this end, approximately 15,000 shear wave velocity (V_s) profiles were collected and organized into a georeferenced archive, derived from regional microzonation studies and surveys conducted by freelancers. Furthermore, several site proxies, both raster and vector-based, were compiled and analyzed for their correlation with V_{s30} at multiple spatial resolutions. A new lithological classification was developed and used, together with slope, as predictors in the final V_{s30} model.

The result is a comprehensive framework for generating regionalized empirical ground motion scenarios applicable to various purposes such as the generation of ground shaking maps, the reconstruction of historical events without recordings, and the production of scenarios for possible future events. The methodology represents a robust and efficient alternative to physics-based simulations, providing faster computation and improved consistency with observations.

In conclusion, this thesis contributes significantly to the evolution of ground motion prediction models, promoting a data-driven, non-ergodic, and scientifically sound approach capable of leveraging Italy's rich geological and seismological data to meet the challenges of more accurate and realistic seismic modeling.

Keywords (EN)

Seismic source; attenuation of seismic waves; site effects; ground motion model; shaking scenario.

Scientific field of the dissertation (SSD):

GEO/10

School the Ph.D. Student belongs to at UNICAM:

School of Advanced Studies

Other (e.g. Co-funding Programs or Institutions):

Istituto Nazionale di Geofisica e Vulcanologia (INGV)

Table of Contents

Acknowledgments	ii
Abstract	iii
1. Introduction and State of the Art	1
1.1. Introduction, Objectives of the Thesis	1
1.2. State of the art on seismic scenarios	5
2. Data Collection	9
2.1. Recording Dataset	9
2.2. Dataset of shear wave velocity profiles	13
2.3. Lithological Maps	16
2.4. Digital Elevation Models	19
2.5. Map of soil and sedimentary deposit thickness proposed by Pellettier et al. (2016) ..	20
3. Methodology for regionalization of the ground motion model	21
3.1. Introduction to Empirical Ground Motion Models	23
3.2. Reference ground motion model for Italy	24
3.3. Ground Motion Variability and Residual Decomposition	27
3.4. Application of Residual Analysis in Italy	29
4. Data-driven zonations for modelling the regional source and propagation effects into Ground Motion Models in Italy	31
4.1. Tectonic setting of Italy	31
4.2. Spatial variability of residuals related to source and path effects	34
4.3. Regionalization of source effects	34
4.4. Zonation for regionalization of attenuation effects into GMMs	48
4.5. Conclusions	56
5. Evaluation of Site Effects for Shaking Scenarios	60
5.1. Quality control, extension of profiles that do not reach 30m and development of the parametric table	62
5.2. Spatial data processing (DEM analysis)	70
5.3. Correlation between Vs30 and site proxies	80
5.4. Geological reclassification	90
5.5. Parametric Linear Regression	93
5.6. Geostatistical Analysis	99
5.7. Model Construction	106
5.9. Comparison with Vs30 Maps from literature	109
6. Shaking scenarios	114

6.1 Simulation of the Mw 5.7 Colfiorito 1997 earthquake scenario.....	115
6.2 Simulation of the Mw 6.6 Norcia 2016 earthquake scenario	121
7. Discussion and General Conclusions	131
References	137
Data and Resources.....	156
APPENDICES.....	157
Appendix A.....	157
Appendix B.....	164
Appendix C.....	167
Appendix D.....	169
Appendix E.....	184

1. Introduction and State of the Art

1.1. Introduction, Objectives of the Thesis

This doctoral research aims to develop empirical correction factors for source and attenuation effects, as well as a new national-scale Vs30 map for site effects, to improve the spatial prediction of ground shaking across the Italian territory

Earthquake scenarios play a crucial role in seismic hazard assessment and in predicting the effects of earthquakes on infrastructure and populations, therefore the accurate estimation of ground motion is fundamental for engineering design and effective seismic risk mitigation.

The core of a seismic scenario is the evaluation of reliable ground motion, which can be broken down into three main components (Figure 1):

1. Source effects – how the earthquake rupture generates seismic waves.
2. Attenuation effects – how seismic waves lose energy as they travel through the Earth.
3. Site effects – how local geological conditions modify the arriving waves.

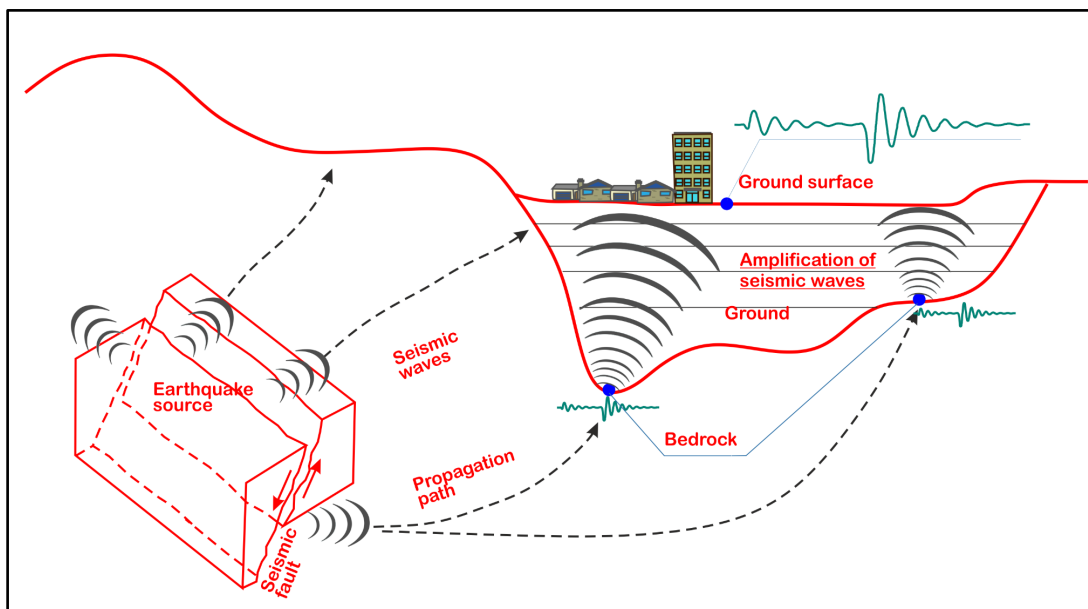


Figure 1. Schematic representation of the seismic source, propagation path, and stratigraphic amplification of seismic waves

The source of an earthquake refers to the physical process of fault rupture and the generation of seismic waves. It describes how stress is released along the fault plane, how the rupture

propagates, and how seismic energy is radiated outward (Aki & Richards, 2002). The characteristics of this process influence the frequency content and amplitude of ground motion, which are essential for seismic hazard assessment (Stein & Wyss, 2003).

One of the most widely used models for representing the spectral characteristics of an earthquake source is the ω^2 (omega-square) model. The ω^2 model has its origins in the work of Keiiti Aki (1967), who was one of the first seismologists to analyze the spectral characteristics of seismic sources.

However, the modern ω^2 model was formally developed and popularized by Brune (1970, 1971), who provided the widely used spectral formulation. Brune's model introduced a clear physical interpretation of the earthquake source spectrum, showing that the seismic spectrum transitions from a low-frequency growth to a high-frequency plateau in acceleration, which is directly related to the rupture dynamics and stress drop of the earthquake (Brune, 1970).

The ω^2 model describes the amplitude spectrum of seismic waves in terms of acceleration as:

$$A(f) \propto M_0 \cdot \frac{(2\pi f)^2}{1 + (f/f_c)^2} \quad [1]$$

where:

- $A(f)$ is the amplitude spectrum of the seismic source in acceleration,
- M_0 is the seismic moment ($N \cdot m$ or $\text{dyn} \cdot \text{cm}$), which represents the total energy released by the earthquake,
- f is the frequency (Hz),
- f_c is the corner frequency (Hz), which separates the low-frequency growth from the high-frequency plateau.

At low frequencies ($f < f_c$), the acceleration spectrum increases with frequency, while at higher frequencies ($f > f_c$), it remains approximately flat, as described by the ω^2 source model.

Several other spectral models have been developed to refine or modify the ω^2 model, addressing its limitations in representing finite-fault effects, rupture dynamics, and spectral scaling variations across different earthquake magnitudes. One of the earliest alternatives is Haskell's Finite Fault Model (Haskell, 1964), which accounts for the finite dimensions of a fault and the propagation of rupture along the fault plane, providing a more realistic representation of how seismic energy is radiated. Later, Boatwright's Modified Ω^2 Model (Boatwright, 1980)

was introduced to improve the characterization of high-frequency decay, offering a more accurate spectral representation of earthquake sources at short wavelengths. Another important contribution came from Kanamori and Anderson (1975), who proposed a spectral model that explicitly incorporates energy balance considerations, particularly for large earthquakes, helping to better describe the scaling of seismic events across different magnitudes (Kanamori & Anderson, 1975). These models represent key advancements in earthquake source modeling and continue to be widely used in seismic hazard analysis and ground motion prediction (Shearer, 2019).

As seismic waves propagate away from the earthquake source, they undergo attenuation, a process that leads to a gradual loss of energy. Attenuation occurs due to three main mechanisms:

- Geometrical Spreading, where energy is distributed over a larger area as waves travel outward (Aki & Richards, 2002).
- Anelastic attenuation, where part of the wave energy is converted into heat due to internal friction within the Earth's materials, is often described by the quality factor Q (Knopoff, 1964).
- Scattering, where waves interact with small-scale heterogeneities in the Earth's crust, redistributing energy in different directions (Sato et al., 2012).

The combined effect of these mechanisms is commonly described by a decay function, incorporating both geometric spreading and anelastic attenuation (Boore, 2003):

$$A(R, f) \propto \frac{1}{R^n} e^{-\frac{\pi f R}{QV}} \quad [2]$$

where:

- $A(R, f)$ is the amplitude of the seismic waves at distance R ,
- $1/R^n$ represents geometrical spreading, describing how wave energy is distributed over an expanding volume. The exponent n generally takes a value of 1 but may vary depending on wave type, propagation medium, and region-specific characteristics.
- $e^{-(\pi f R / QV)}$ accounts for anelastic attenuation, where seismic energy is lost due to internal friction and converted into heat. The quality factor Q controls the efficiency of this process:
 - Higher $Q \rightarrow$ Less energy loss (seismic waves travel farther with minimal attenuation).

- Lower $Q \rightarrow$ Greater energy loss (waves are strongly attenuated with distance).
- f is the wave frequency (Hz),
- V is the seismic wave velocity (m/s).

Attenuation plays a fundamental role in seismic hazard assessment, as it directly influences the amplitude and frequency content of ground motion with distance.

The site effect refers to how local geological and geotechnical conditions modify the amplitude, frequency content, and duration of ground motion. While seismic waves originate at the source and attenuate as they travel through the Earth's crust, their final characteristics at the surface are strongly influenced by near-surface materials (Bard, 1999). The response of a site depends on several factors, including the stiffness of the subsurface layers, the presence of sediment-filled basins, and nonlinear soil behavior (Stewart et al., 2008).

Several key mechanisms influence site response:

- Soil Amplification – Soft sediments tend to amplify ground motion compared to hard rock sites due to their lower shear-wave velocity (V_s). The contrast in impedance between soft soils and underlying bedrock can cause wave trapping, leading to stronger shaking at the surface (Borcherdt & Glassmoyer, 1994). The degree of amplification depends on the velocity contrast and thickness of the sediment layers (Dobry et al., 2000).
- Resonance Effects – In many cases, sites with a well-defined sediment-bedrock interface exhibit a fundamental frequency (f_0), which can amplify seismic waves if the incoming wave frequencies match this natural frequency (Ishihara, 1977). However, in complex geological settings, multiple resonance frequencies may exist, or the resonance effect may be poorly defined due to irregular layering and lateral variations in sediment properties (Bard & Bouchon, 1985). In rock sites, resonance effects are generally negligible (Stewart et al., 2008).
- Nonlinear Behavior – Under strong ground shaking ($PGA > 0.2-0.3g$), soils may exhibit nonlinear response, meaning that their shear stiffness decreases and damping increases as strain levels rise (Vucetic & Dobry, 1991).
- Permanent displacements, such as surface faulting, landslides, and liquefaction, are not addressed in this work.

Assessing site effects is critical in seismic hazard analysis, as local soil conditions can significantly alter expected ground motions and influence building response and infrastructure resistance (Stewart et al., 2008).

1.2. State of the art on seismic scenarios

There is a variety of approaches to calibrate a seismic scenario accounting for source attenuation and site effects. Currently, substantial efforts and resources are dedicated to physics-based ground motion simulations, since they aim to provide a physically accurate representation of finite fault rupture processes, the propagation and attenuation of seismic waves, and site effects. These simulations are commonly performed using advanced numerical codes such as EQSIM (McCallen et al. 2021a; McCallen et al. 2021b), SPECfEM (Komatitsch & Tromp 2002a; Komatitsch & Tromp 2002b), New Zealand CyberShake (Graves et al. 2010; Bradley et al. 2020), and SPEED (Mazzieri et al. 2013; Paolucci et al. 2018). Nevertheless, these methods require complex and precise modeling of both fault ruptures and the crustal medium, particularly when extending the frequency range of the simulations (Sgobba & Pacor, 2023). However, despite their accuracy, physics-based simulations present several limitations. One of the main drawbacks is their high computational cost, as they require significant processing power and extensive simulation times. Running high-resolution 3D simulations of earthquake rupture and wave propagation demands access to high-performance computing infrastructure, which is not always readily available. Moreover, these methods cannot be used to compute seismic scenarios in real-time, making them impractical for rapid response applications. Another critical issue is the difficulty in obtaining reliable input parameters that play a crucial role in determining the accuracy of the results, but they are often poorly constrained or unavailable for many regions.

An alternative to physics-based scenarios is represented by empirical shaking scenarios, which rely on empirical Ground Motion Models (GMMs). These models consist of simple equations that estimate the probability distribution of ground motion intensity measures based on fundamental explanatory variables. Among these, the event magnitude is used to characterize the seismic source (e.g. eq.1) where M_w is proportional to M_0 , the source-to-site distance accounts for wave attenuation effects (e.g. eq. 2, which includes the distance term), and local soil conditions describe the near-surface amplification or deamplification of seismic waves. A general formulation of GMMs can be expressed as:

$$\log Y = a + FM(M_w, \text{source parameters}) + FD(M_w, \text{distance}) + FS(\text{site parameters}) + \epsilon \quad [3]$$

where $\log Y$ represents the logarithm base 10 of a ground motion intensity measure, such as response spectral ordinates or Fourier amplitude spectra. The term $FM(M_w, \text{source parameters})$ accounts for the seismic source effects, $FD(M_w, \text{distance})$ describes attenuation effects due to wave propagation, $FS(\text{site parameters})$ describes site-specific modifications and epsilon represents the aleatory variability.

A comprehensive compilation of GMMs has been provided by Douglas (2021). In Italy, the model developed by Sabetta and Pugliese (1996) was adopted in the MPS04 (Stucchi et al., 2011) national seismic hazard assessment. This model was later updated and refined as more strong-motion data became available, leading to the development of newer models such as Bindi et al. (2011) and Lanzano et al. (2019), which reflect the continuous evolution of ground motion prediction equations based on expanded observational datasets.

A key advantage of GMMs is their computational efficiency, as they allow for rapid seismic scenario calculations, making them essential for real-time applications and probabilistic seismic hazard assessments (Douglas & Edwards, 2016). GMMs provide consistent predictions within the range of magnitude and distance they were developed although they can often be extrapolated with reasonable confidence beyond those ranges. Additionally, modern GMMs benefit from an increasing number of predictive variables, enabling better characterization of source, path, and site effects (Douglas & Edwards, 2016).

However, the main limitation of GMMs, especially the ones calibrated with worldwide datasets, lies in the ergodic assumption, which presumes that the variability at a single site from a specific source is assumed identical to that derived from multiple sites in different regions (Anderson & Brune, 1999). This assumption can lead to biased ground-shaking estimates and increased uncertainty. To address these issues, the research community has increasingly focused on developing regionalized GMMs, shifting part of the aleatory variability into epistemic uncertainty (Al-Atik et al. 2010; Anderson & Uchiyama, 2011). This is achieved by relaxing the ergodic assumption, a task made feasible by the growing availability of seismic records from increasingly dense global seismic networks. A fully non-ergodic approach requires numerous records at individual stations from various earthquakes, as well as multiple stations recording each event. This allows for a systematic decomposition of variability into event-specific, source-specific, site-specific, and path-specific effects using statistical methods (Al-Atik et al. 2010). Examples of fully non-ergodic GMMs can be found in studies such as (Landwehr et al. 2016; Parker et al. 2022).

In recent years, advancements in artificial intelligence and computing techniques have led to the development of numerous GMMs based on machine learning algorithms. As an alternative to traditional regression-based models, nonparametric machine learning (ML) approaches are emerging as a powerful tool to capture the complex nonlinear behavior of earthquake ground motions (Mohammadi et al., 2023). Unlike parametric empirical models, which rely on predefined functional forms, nonparametric ML models offer greater flexibility in handling the inherent complexity and high nonlinearity of seismic motions (Mohammadi et al., 2023). These models have been proposed to address the challenges associated with conventional

approaches, particularly in capturing intricate dependencies and improving predictive accuracy (Kubo et al., 2020; Lekshmy & Raghukanth, 2021; Li et al., 2023). Nevertheless, GMMs obtained with ML techniques are still not employed for seismological applications such as seismic scenarios or seismic hazard.

In this study the seismic scenarios are calibrated with GMMs, that are adjusted locally to take into account the region-specific source, attenuation and site effects. This goal is achieved by selecting a reference GMM for the Italian territory by Lanzano et al (2019), hereinafter termed as ITA18, the latest model developed for active shallow crustal events in Italy. It provides the equations and the coefficients for estimating peak ground acceleration and velocity (PGA, PGV) and acceleration response spectra ordinates (5% damping) for the horizontal components, over a period range from 0.01 to 10 seconds. Figure 2 shows a flowchart summarizing the steps of the thesis work.

Adjustment factors are calculated to take into account the regional variability of source and attenuation through the analysis of residuals, defined as the logarithmic differences between observed intensity measures and GMM predictions, and their decomposition into components related to source, path, and site (Chapter 3). Their spatial distribution is then examined in Chapter 4 to identify zones with homogeneous residuals associated with source and attenuation. The regional site effects are taken into account by the site term of the GMM in Chapter 5. For this purpose, a new map of Vs30 was produced at the national scale. The core focus of this doctoral research is the improvement of site effect representation through the development of a new Vs30 map, refined using statistical and geospatial modeling techniques. The correction factors enhance the model accuracy for applications such as Probabilistic Seismic Hazard Assessment (PSHA), site-specific hazard evaluations, empirical earthquake scenarios, and ShakeMap generation (Worden et al. 2018; Michelini et al. 2020).

This work requires extensive data collection, including:

- A dataset of ground motion recordings (ITACAext flat-file V.3.2 - Brunelli et al. 2022), described in Chapter 2;
- A dataset of Vs profiles at the Italian scale (Chapter 2) and a collection of proxies necessary for modeling site effects, including topographic and geological parameters (Chapter 5);
- A set of maps of the spatial distribution of physical parameters used to validate the source and attenuation zonations (e.g., heat flow maps for attenuation effects), described in Chapter 4.

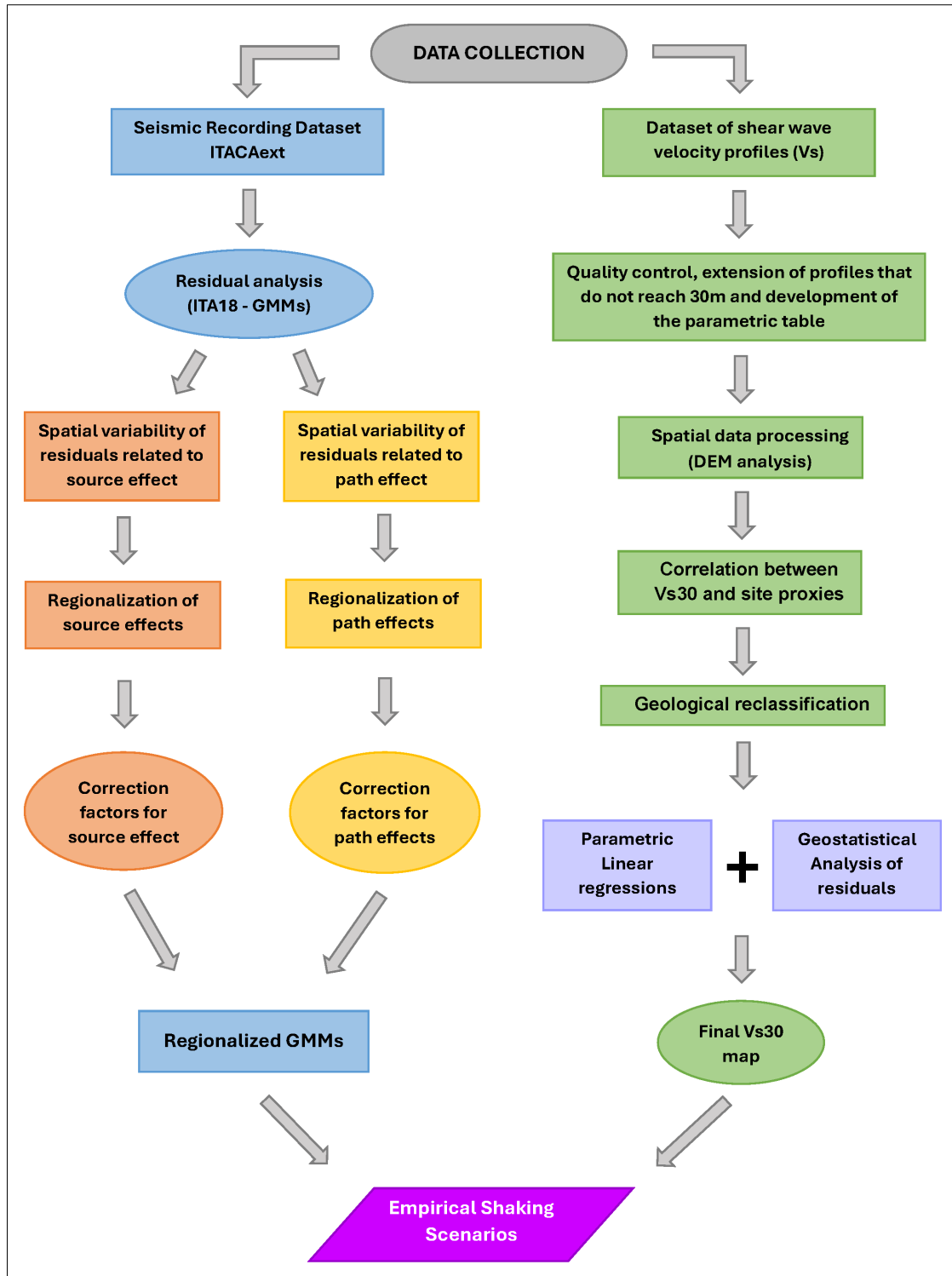


Figure 2. Flowchart summarizing the methodological workflow of the thesis. The left branch describes the analysis and regionalization of residuals computed from a reference ground motion model (GMM) to derive correction factors for source and path effects. The right branch illustrates the processing of Vs profiles, the correlation with site proxies, and the derivation of the final Vs30 map through both parametric regressions and geostatistical analysis. The outputs from both branches are used to produce the final empirical shaking scenarios

2. Data Collection

2.1. Recording Dataset

A dataset was compiled, including earthquakes with a magnitude equal to or greater than 3.0 that occurred between 1972 and 2020. The records were selected from the ITACA accelerometric database (Italian ACcelerometric Archive, http://itaca.mi.ingv.it/ItacaNet_32/#/home [Russo et al., 2022]).

In regions with limited seismic activity, such as Liguria, Piedmont, and Western Sicily, and for events with magnitudes below 4.0, accelerometric data were integrated with velocimetric recordings, excluding those affected by instrumental saturation. All waveforms in this dataset, referred to as ITACAext (Brunelli et al., 2022), were manually processed according to the standard ITACA methodology, as described in (Paolucci et al., 2011). The ITACAext flat-file, formatted consistently with the Engineering Strong-Motion (ESM) flat-file (Lanzano et al., 2019), is available on the ITACA v.3.2 website (http://itaca.mi.ingv.it/ItacaNet_32/#/products/itacaext_flatfile).

A further refinement was applied to select events from active crustal tectonic regimes (focal depth lower than 35 km), recorded by sensors installed at ground level and in free-field conditions. The maximum epicentral distance included in the analysis was limited to 220 km, which exceeds the 200 km threshold established by ITA18, the Ground Motion Model (GMM) used for subsequent residual analysis. In cases where accelerometers and velocimeters were co-located, the velocimetric records (if unsaturated) were prioritized due to their better resolution for weak motion.

The final dataset consists of 37,098 records from 1,863 earthquakes, recorded at 1,922 stations. Figure 3 includes maps that illustrate the spatial distribution of seismic events and recording stations. The station network covers almost the entire Italian territory (excluding Sardinia), with a higher density in areas affected by recent seismic sequences. Most seismic events occurred in the Northern and Central Apennines, while areas such as the northern Lombardy-Piedmont border, Salento, and Western Sicily are less represented. To support this work, we provide a map of the source-to-site paths (Figure 4), which highlights the dense and homogeneous spatial coverage across nearly the entire Italian territory. This study aims to analyze the relationship between the residuals of a reference Italian GMM and the physical parameters of seismic sources and wave propagation, and therefore this comprehensive sampling of source-to-site paths was a key consideration in selecting the dataset.

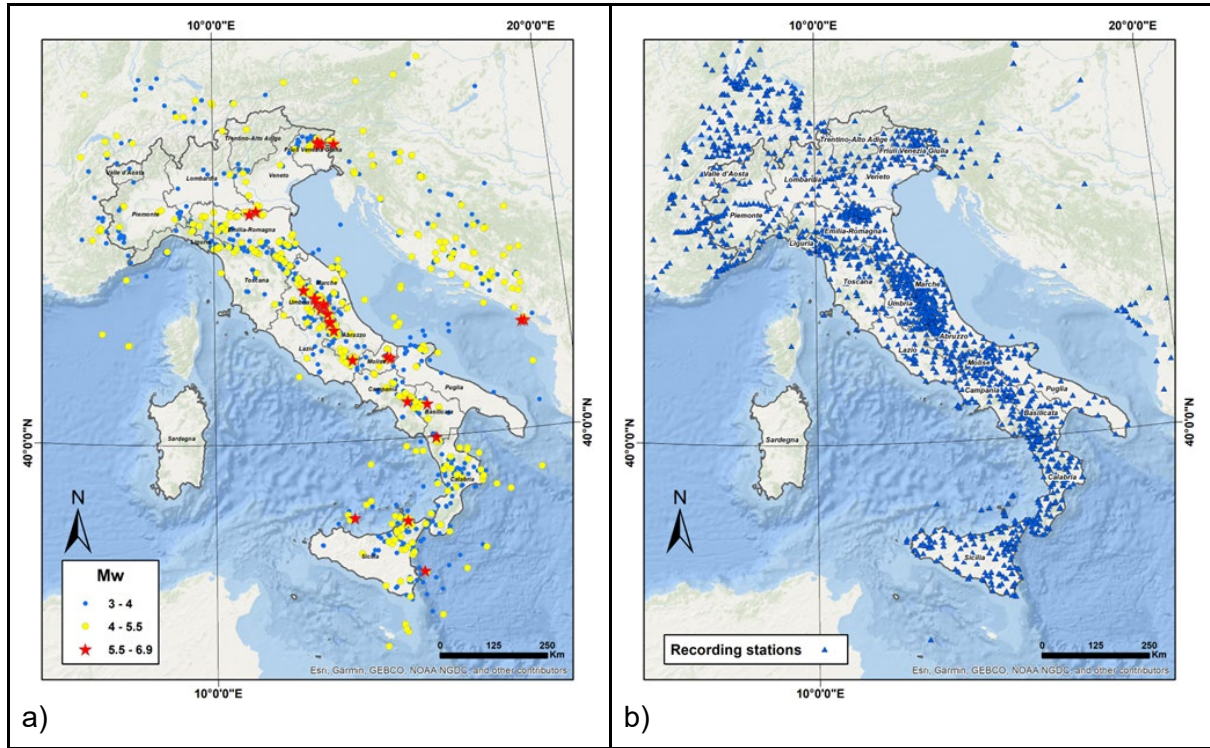


Figure 3. Subset used in this study: a) map of events; b) map of stations. Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses

To accurately interpret the residual analysis results, the moment magnitude (M_w) estimates for the events were standardized. Specifically, M_w values derived using the TDMT method (Time Domain Moment Tensor from Dreger et al., 2003) were aligned to those derived by the RCMT method (Regional Centroid Moment Tensor from Pondrelli et al., 2002). This conversion was necessary because ITA18 uses moment magnitudes from RCMT as the explanatory variable for magnitude scaling.

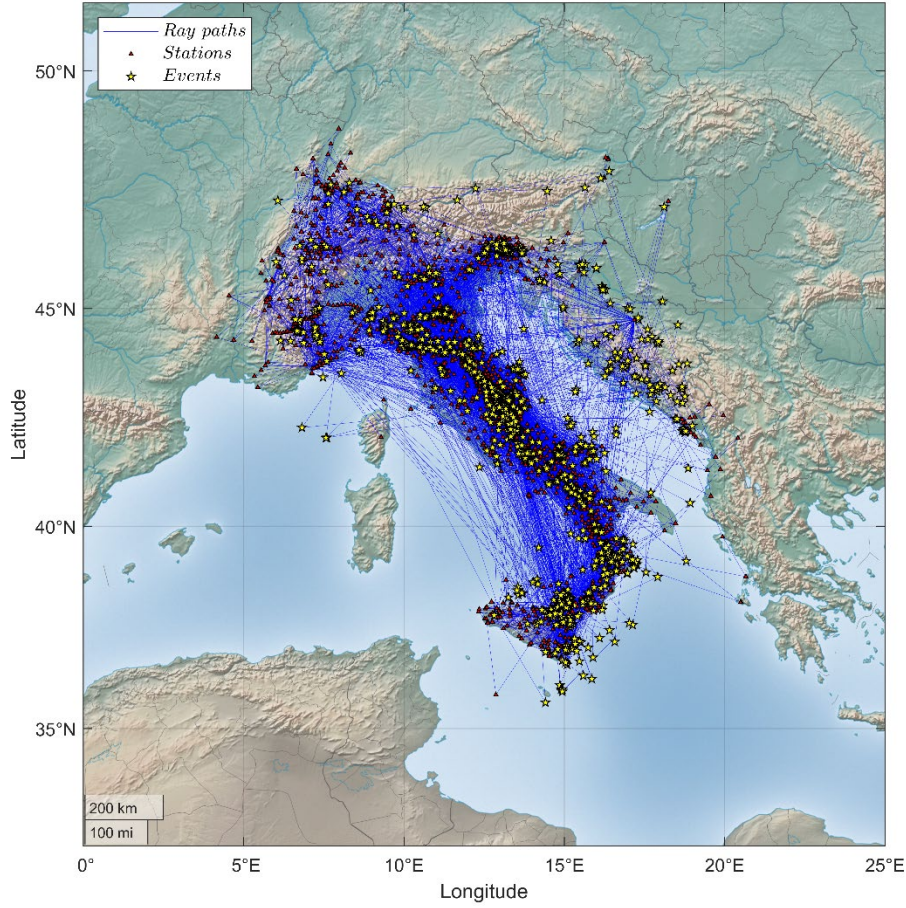


Figure 4. Source-to-site paths for the subset of ITACAext, used in this analysis. Reproduced from Brunelli et al. (2023), published in Soil Dynamics and Earthquake Engineering, Elsevier. Reused in accordance with Elsevier's policy for inclusion in doctoral theses

An empirical conversion equation was calibrated using 141 events for which magnitudes from both TDMT available from the INGV web-service (<http://webservices.ingv.it/>) and RCMT catalog (<https://rcmt2.bo.ingv.it/>) where available.

The derived equation is:

$$M_{WRCMT} = 0.6 + 0.917 M_{WTDMT} \pm 0.0909 \quad [4]$$

The coefficients in Eq. [4] are very similar to those suggested by Gasperini et al., (2012), although this study incorporates a significantly larger number of events. The local magnitude ML of smaller events (<http://terremoti.ingv.it/>), is converted to moment magnitude RCMT using the conversion formula provided by Gasperini et al., (2013). Figure 5, presents the regression line derived from Eq. [4] alongside the magnitude values. Figure 6, illustrates the relationship between RCMT-Mw (hereinafter Mw or moment magnitude) and the distance, defined as the

shortest distance to the fault plane (rupture distance, R_{rup}) for events with $M_w > 5.5$, while for smaller events, it corresponds to the hypocentral distance. Although the dataset lacks recordings near the source for very large events, the number of observations with rupture distance $R < 10$ km and moment magnitude $M \geq 6$ is still considerable, amounting to approximately 1000 records (3%).

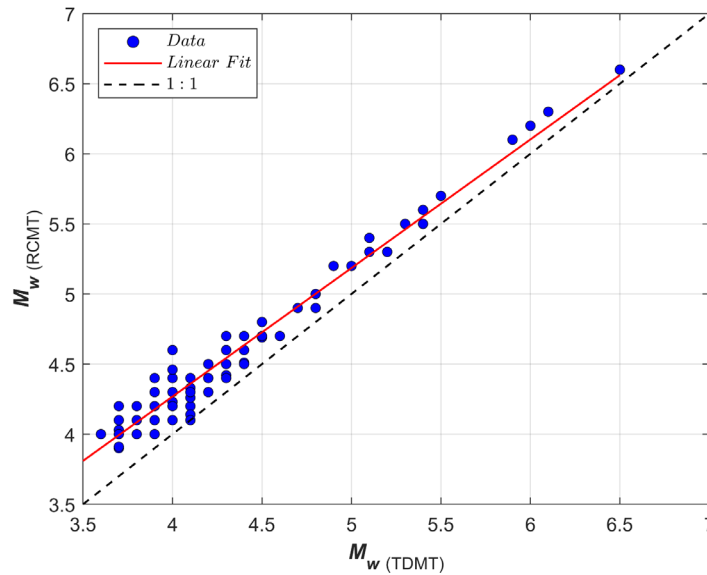


Figure 5. Relationship between $M_w(TDMT)$ and $M_w(RCMT)$. Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses

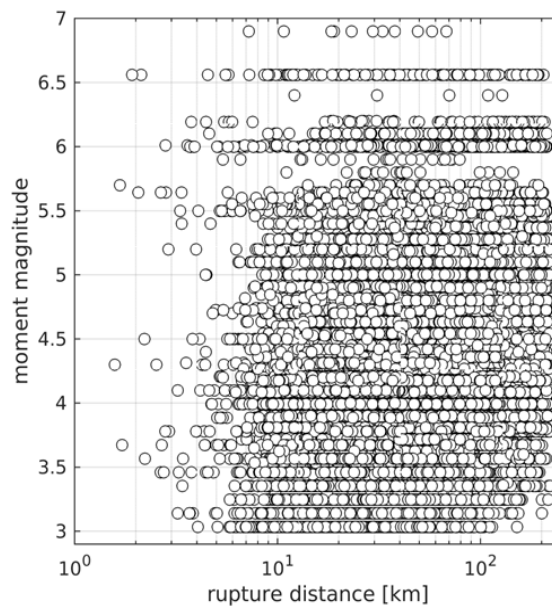


Figure 6. Magnitude-distance distribution. Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses.

2.2. Dataset of shear wave velocity profiles

As part of this doctoral thesis, approximately 15,000 shear wave velocity (V_s) profiles were collected throughout Italy. These profiles were obtained from geophysical measurements conducted using methods such as DOWN-HOLE, CROSS-HOLE, MASW, ESAC-SPAC, REMI, among others, through both active and passive surveys. The data were structured within an Access database organized into three relational tables (Site, Survey, and Parameters) following the framework used in the seismic microzonation (Commissione tecnica per la Microzonazione sismica, 2016), ensuring efficient and integrated data management.

The Access database, specifically, is composed of three tables (Site, Survey, and Parameter) for point-based investigations (e.g., Down-Hole, Cross-Hole, etc.) and three corresponding tables (Site, Survey, and Parameter) for line-based investigations (e.g., MASW, ESAC, Re.Mi., etc.).

In greater detail:

- The Site table contains information related to the identification of the investigation (e.g., address, ISTAT code for province and municipality, etc.); the location of the investigation (e.g., longitude and latitude coordinates expressed in WGS84 UTM 33N); the elevation above mean sea level (m) of the site; and additional notes such as the author of the investigation.
- The Survey table provides details primarily about the type of investigation (e.g., MASW, Re.Mi., or others).
- The Parameter table records the shear wave velocity (V_s) values for each layer of the profile.

These three tables are interconnected through primary keys (see Figure 7) that were modified from their original structure, as used in seismic microzonation database, to make the tables easier to integrate with new investigations. Specifically:

- In the Site table, the primary key was assigned to ID_SLN/ID_SPU (where SLN and SPU means linear or point investigation).
- In the Survey table, the primary key was assigned to ID_INDLN/ID_INDPU (depending on whether it refers to a linear or point-based site).
- In the Parameter table, the field ID_INDLN/ID_INDPU (depending on whether it refers to a linear or point-based site) was added to relate the Survey and Parameter tables.

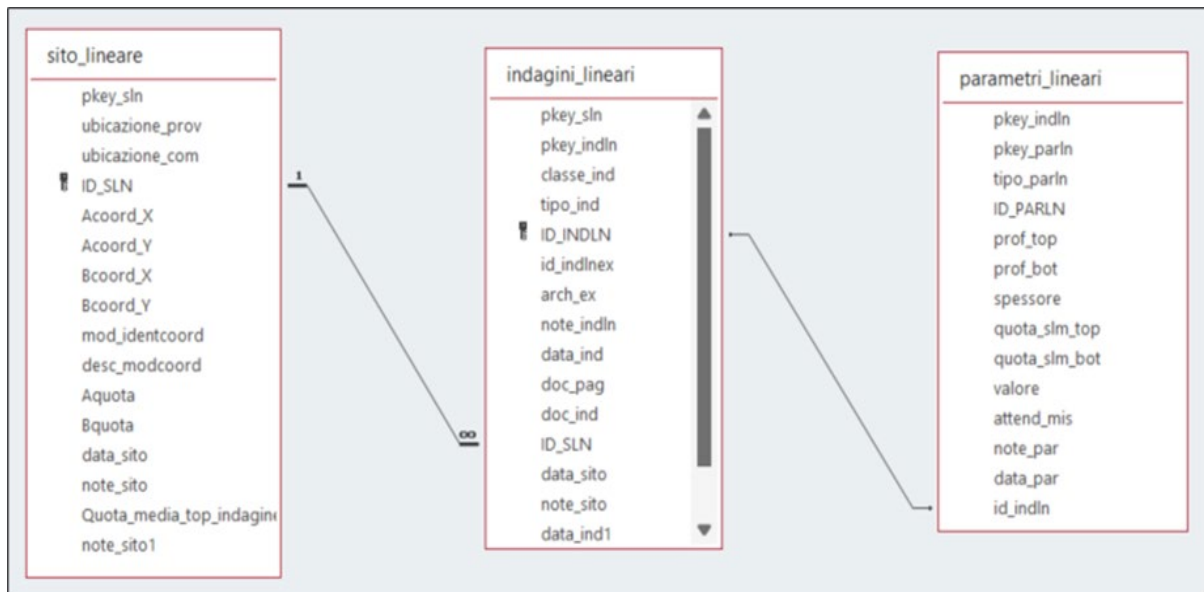


Figure 7. The figure illustrates the relationships between the tables in the Access database. In this case, the three tables shown (*sito_lineare* = Site, *indagini_lineari* = Survey, and *parametri_lineari* = Parameter) refer specifically to linear investigations

The spatial distribution of the dataset is illustrated in Figure 8.

The velocity profiles are derived from the following sources:

a) Seismic Microzonation studies:

- Central Italy following the 2016 seismic events (<https://sisma2016data.it/microzonazione/>);
- VEL Program of the Tuscany Region (Local Effects Evaluation, <https://www.regione.toscana.it/-/banca-dati-vel>);
- Basilicata Region (<https://www.webms.it/servizi/viewer.php>);
- Calabria Region (<https://www.webms.it/servizi/viewer.php>);
- Emilia-Romagna Region (Martelli, personal communication);
- Friuli Venezia Giulia Region (<https://www.webms.it/servizi/viewer.php>);
- Liguria Region (<http://srvcarto.regione.liguria.it/geoviewer2/pages/apps/geoportale-tecnico>);
- Molise Region (<https://www.webms.it/servizi/viewer.php>);
- Sicily Region (<https://www.webms.it/servizi/viewer.php>).

b) Territorial Information System of the Marche Region (<https://qmap-protciv.regione.marche.it/cs/>);

- c) Geological and Seismic Service of the Umbria Region (personal communication by the service representative);
- d) Professional studies available online;
- e) ITACA Database (data acquired within Annex A and B2 of the agreement between the Italian Department of Civil Protection and INGV for the characterization of accelerometric stations, <http://itaca.mi.ingv.it>);
- f) ARPA Piemonte Geophysical Database (<https://www.arpa.piemonte.it/temi/geologia-dissesto/geologia-regionale-geotecnica?pid=18>);
- g) Geophysical Database of the Umbria Region;
- h) Geoportal of the Aosta Valley Region (<https://geoportale.regione.vda.it/>).

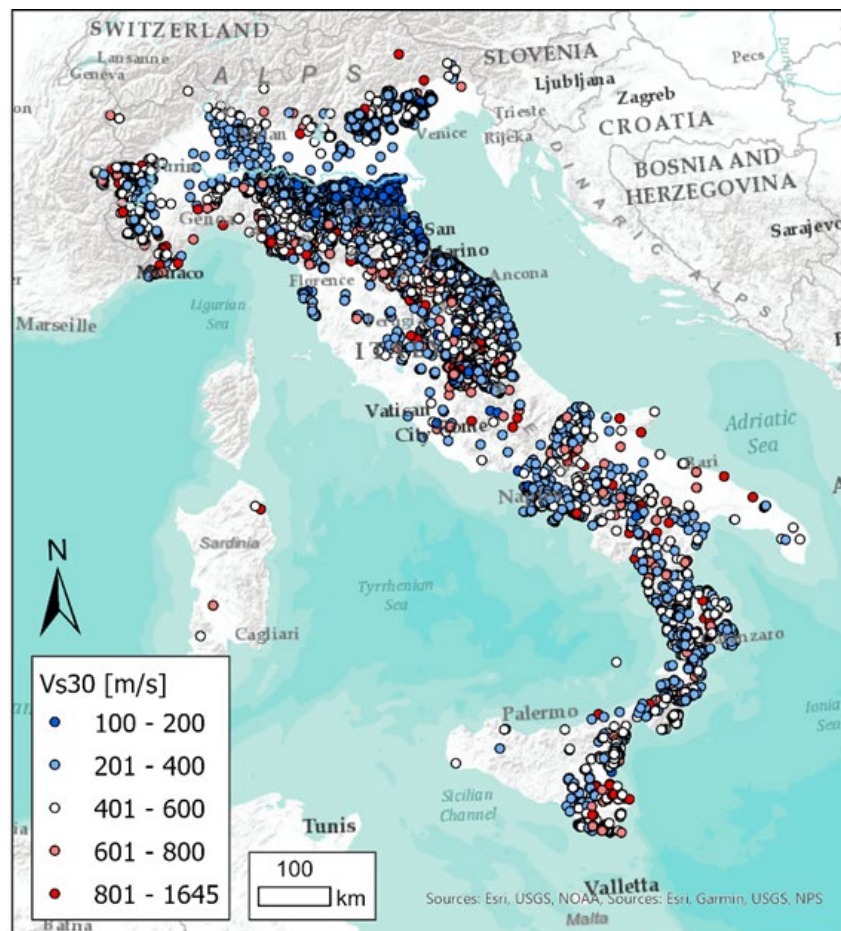


Figure 8. Spatial distribution of the Vs30 dataset

Figure 8 illustrates the spatial distribution of the dataset as a function of Vs30 values, where low values are represented in blue and high values in red. The final dataset contains few Vs30

profiles with high values (rock sites), which are primarily located along the Apennine arc, in the south-eastern Sicily, in the Prealps, and on the Puglia region. The lowest Vs30 values are predominantly found in the Po Plain and coastal areas, as expected.

Certain regions, such as Marche (in central Italy) and Emilia-Romagna, are well sampled, whereas others are underrepresented and the dataset exhibits a scarcity of profiles in mountainous areas.

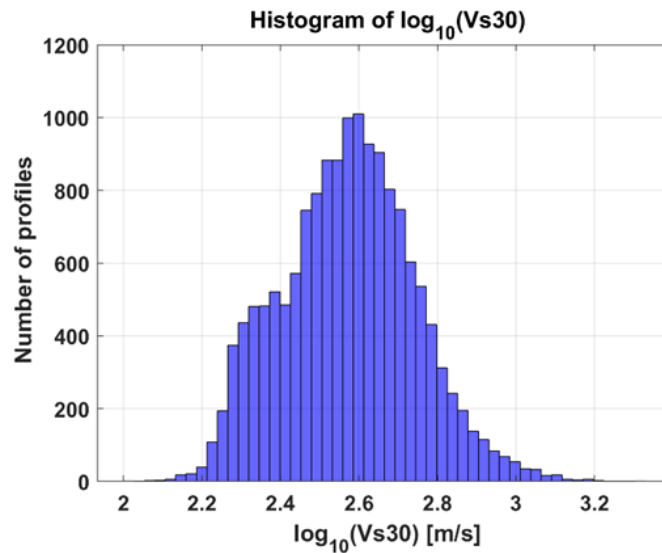


Figure 9. Histogram of $\log_{10}(Vs30)$ from the dataset

The histogram of $\log_{10}(Vs30)$, shown in Figure 9, reveals a bimodal distribution, with a first peak around 2.4 ($Vs30 \approx 250$ m/s) and a second peak near 2.6 ($Vs30 \approx 400$ m/s). This pattern reflects the presence of profiles acquired on fine-grained deposits and medium to coarse-grained deposits, including fractured rock. The distinction between these two geological settings underscores the importance of accounting for lithological differences when analyzing Vs30.

2.3. Lithological Maps

Lithology captures essential information about the composition and physical characteristics of rocks and soils, making it a crucial factor in understanding the surface and subsurface processes of the Earth. Its analysis plays a significant role in various fields, such as seismic amplification (Mori et al., 2020; Forte et al., 2019) and contributes to a deeper understanding of geological evolution.

Lithological data is typically derived from geological maps, and substantial efforts have been made in recent years to improve global accessibility to high-quality geological datasets.

In Italy, geological maps with nationwide coverage are available at scales of 1:1,250,000 (Bonomo et al., 2005), 1:1,000,000 (Pantaloni, 2011; Cipolloni et al., 2009; Compagnoni, 2004), 1:500,000 (Compagnoni et al., 1976–1983), and 1:100,000 (Servizio geologico d'Italia, 2004), all distributed by ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale–Delogu et al., 2012). Additionally, ISPRA coordinates and publishes the 1:50,000 geological map of Italy, which currently offers only partial coverage of the national territory.

Lithological Map of Italy at a 1:100,000 Scale Produced by ISPRA

Amanti et al. (2007, 2008) described ISPRA's initial effort to create a 1:100,000-scale lithological map of Italy, which covered about 65% of the country, excluding Sardinia, Sicily, and some map sheets. This preliminary version was not released in raster or vector format. The work was completed in 2018, resulting in a full national map at the same scale. The updated map, created by compiling data from the 277 sheets of the *Carta Geologica d'Italia* and standardizing the legend, provides a consistent lithological representation of the entire country. This combined effort resulted in a single shapefile, which is accessible for visualization via the geoportal (<https://sgi2.isprambiente.it/viewersgi2/>).

This map provides information on the lithologies of the Italian territory, as proposed by ISPRA, and is described through the following fields (Table 1):

Table 1. Attributes of the lithological map of Italy (by ISPRA) shapefile

Field	Description
Name	Description of the lithological unit
Nome_ulf	Name of the lithological unit
Eta_geol	Geological age of the lithological unit
Litologia	Lithology
Classi_lit	Lithology class
Cod_lito	ISPRA lithology code
Shape_Area	Polygon Area
Shape_Length	Polygon Perimeter

Refer to Appendix A for the detailed descriptions of all fields included in the dataset.

Lithological Map of Italy at a 1:100,000 Scale Proposed by Bucci et al. (2021)

The Lithological Map of Italy at a 1:100,000 scale, proposed by Bucci et al. (2021) shown in Figure 10, was developed using the 277 sheets of the Geological Map of Italy at the same scale, provided by ISPRA (ISPRA – Italian Geological Survey, 2004). The dataset contains approximately 300,000 georeferenced polygons, which were classified by expert evaluation into 5,456 unique descriptions, following compositional and geomechanical criteria.

The final map has 19 distinct classes, linked to a database that facilitates the interpretation of geomechanical properties and allows for future enhancements or refinements (Bucci et al., 2022).

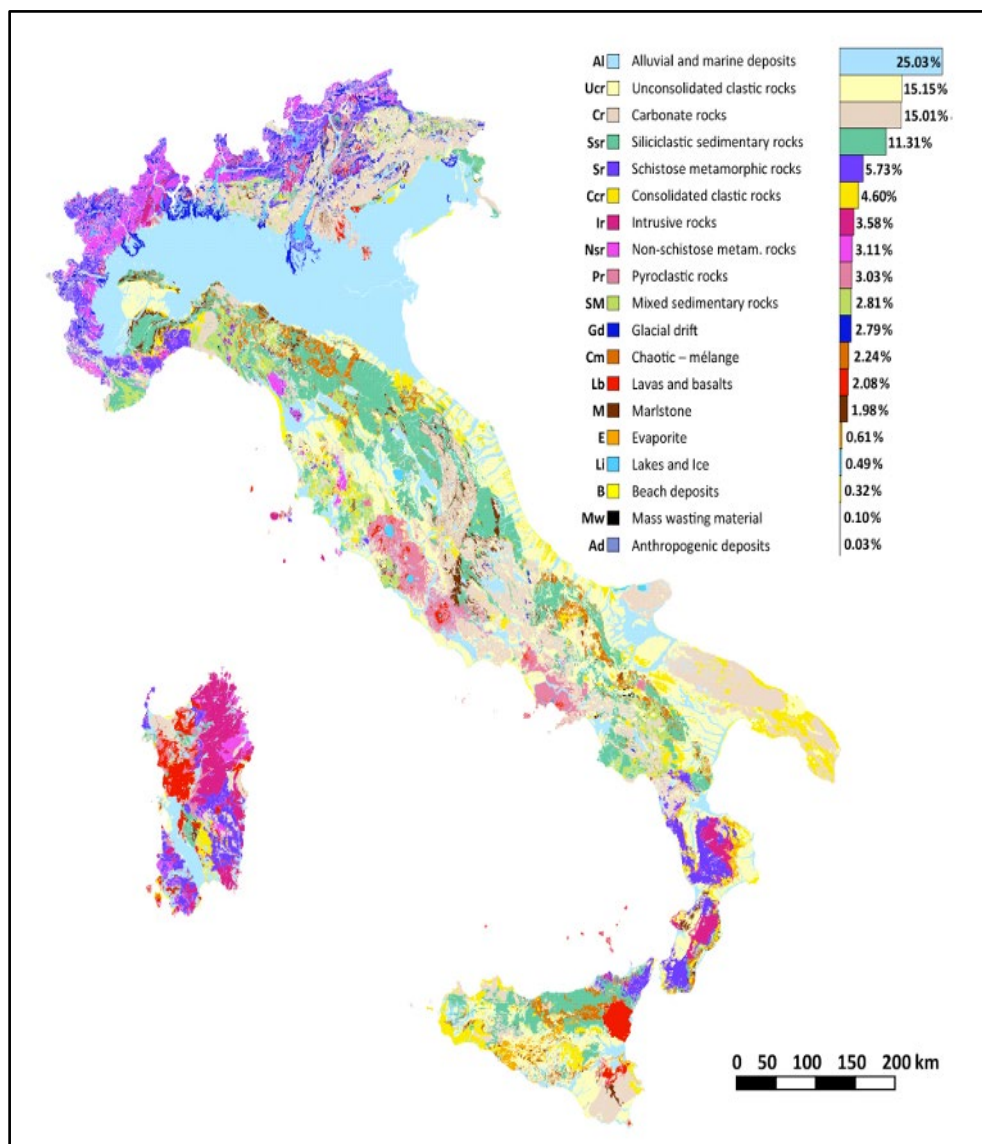


Figure 10. Lithological Map proposed by Bucci et al 2021. Figure from Bucci et al., 2022. Licensed under CC BY 4.0.

2.4. Digital Elevation Models

A Digital Elevation Model (DEM) is a representation of the elevation distribution over a specific surface. It can be understood as a statistical surface where elevation values (Z) are assigned to a set of defined X , and Y coordinates, reflecting a phenomenon that varies continuously across space.

The term DEM includes both DSMs and DTMs:

- DSM (Digital Surface Model): Represents the Earth's surface along with all the objects on it, such as buildings, tree canopies, and other structures.
- DTM (Digital Terrain Model): Represents the bare ground surface, filtered to exclude any above-ground features like vegetation or man-made structures.

In this thesis, the TINITALY/1.1 Digital Terrain Model (DTM) shown in Figure 11, developed by Tarquini et al. (2023) as an update to the previous version by Tarquini et al. (2007), was utilized. It features a spatial resolution of 10m x 10m and is available in GeoTIFF format, georeferenced using the UTM WGS 84 Zone 32N coordinate system (<https://tinality.pi.ingv.it/>).

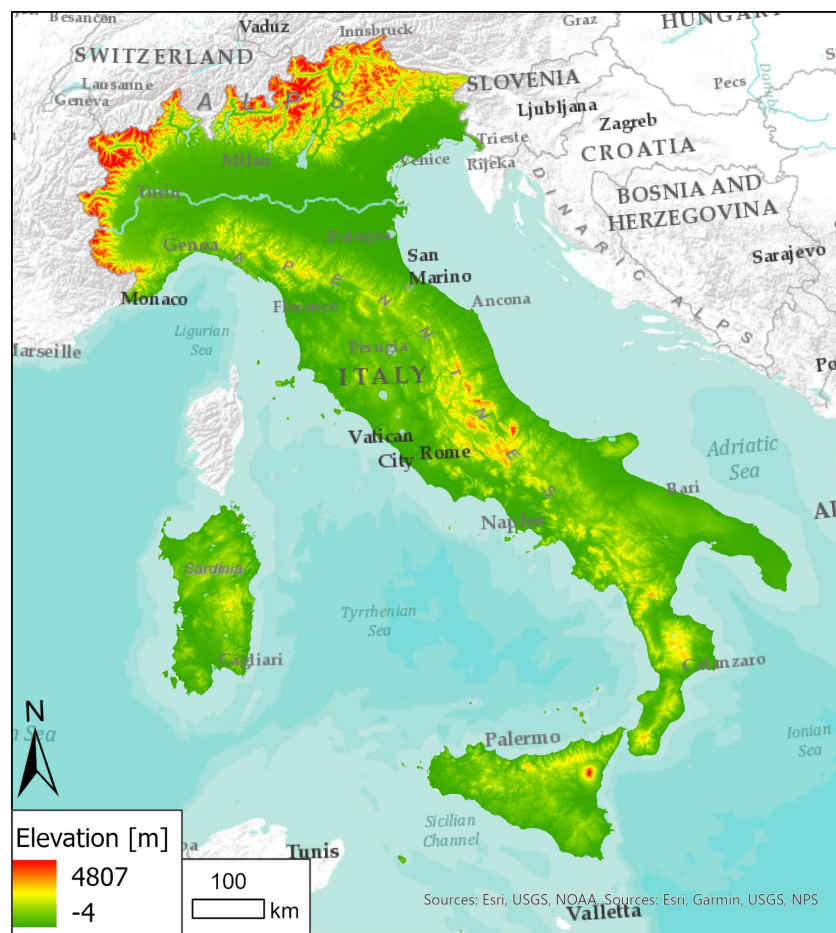


Figure 11. DEM TINITALY (Version 1.1) from Tarquini et al., 2023. Licensed under CC BY 4.0.

2.5. Map of soil and sedimentary deposit thickness proposed by Pelletier et al. (2016)

The study of Pelletier et al. (2016) presents a high-resolution global dataset (30 arcsec, ~1 km) of soil, intact regolith (heterogeneous material covering solid bedrock), and sedimentary deposit thicknesses above unweathered bedrock.

The dataset distinguishes between hillslopes, valley bottoms, and lowlands, using optimized models for each landscape type. Thickness estimates are calibrated and validated with independent data: soil thickness measurements from the U.S. and Europe for hillslopes, and depth-to-bedrock observations from groundwater wells in the U.S. for lowlands.

The main products of the work by Pelletier et al. (2016) include grids representing the average soil and regolith thickness on hillslopes, a grid depicting the average alluvial thickness in upland valley bottoms and lowlands, and a single integrated product that shows the average soil/alluvial thickness across both hillslopes and valley bottoms. Figure 12 presents this last product, specifically for the Italian territory, which has been used in the present research.



Figure 12. Average soil and sedimentary deposit thickness [m] as proposed by Pelletier et al. (2016)

3. Methodology for regionalization of the ground motion model

Over the past two decades, numerous studies have demonstrated that ground motion in Italy exhibits significant regional variability, largely influenced by tectonic and geological factors (Malagnini et al., 2002; Malagnini et al., 2008; Akinci et al., 2022).

The attenuation of ground motion is mainly driven by the distance from the site to the seismic source (decrease of wave amplitudes due to geometric attenuation) and by the composition, temperature, and pressure of the rocks through which the waves travel (decrease of wave amplitude due to energy loss or anelastic attenuation). Regional variations of the anelastic attenuation have been related to the anisotropic properties of the crust, such as lateral variations in S- and P-wave velocity in the lithosphere (Molinari et al. 2015; Di Stefano et al. 2020; Manu-Marfo et al. 2019) or to geothermal properties, including heat flow anomalies (Della Vedova et al., 1991). Crust anisotropy has been studied by Ameri et al. (2009) who analyzed the ground motion records from the 2009 L'Aquila sequence, that revealed a clear dependence on station backazimuth. Critical reflections at the Moho interface (SmS phases) can contribute additional energy to ground motion at distances between approximately 90 and 150 km from the source. Bragato et al. (2011) interpreted this effect as a constructive interference between Moho-reflected S waves and direct S waves, based on a dataset of peak ground acceleration recordings from weak to moderate earthquakes ($ML \leq 5.2$) in northern Italy, particularly within the Po Plain. Their analysis revealed a systematic increase in PGA, with amplification factors reaching 2.4–3.6 compared to standard ground motion prediction equations.

Although the effect is observed on PGA, the associated energy is mainly concentrated at frequencies below 5 Hz. This interpretation is further supported by Luzi et al. (2012), who observed consistently positive PGA residuals at distances of around 120 km during the 2012 Emilia earthquake sequence. Recent studies on the 2016–2017 Central Italy seismic sequence have identified a systematic difference in seismic wave attenuation depending on whether the waves propagate toward the Tyrrhenian Sea or the Adriatic Sea (Luzi et al., 2017; Gabrielli et al., 2022).

Another key factor that may influence the regional variability of ground motion is the earthquake stress drop which is the difference between the stress acting on a fault before and after an earthquake rupture. This parameter might vary due to fault strength, fault topography, the presence of fluids or rupture size, slip, and velocity which can have regional variability. In

particular, the ground motion of earthquakes with large stress drops is characterized by large high-frequency content.

In the early 2000s, Malagnini and co-authors conducted a systematic study on the regionalization of source properties such as stress drop ($\Delta\sigma$) and attenuation factors such as the quality factor (Q) in Italy applying the generalized inversion technique to small-magnitude earthquake recordings from past seismic sequences (Friuli 1976–1977, Irpinia 1980–1981, Umbria-Marche 1996–1997). They subdivided the territory into distinct structural domains: (i) the Apennines (Malagnini et al., 2000), (ii) the Eastern Alps (Malagnini et al., 2002), (iii) the Western Alps (Morasca et al., 2006), and (iv) Eastern Sicily (Scognamiglio et al., 2005). The findings of these studies were integrated into the Italian seismic hazard map MPS04 (Stucchi et al., 2011). Kotha et al. (2020) calibrated the empirical ground motion models for the Seismic Hazard map of Europe (Danciu et al., 2024) introducing regional correction factors with a statistical approach applied to accelerometric recordings. Finally, several authors (Akinci et al., 2004; Lombardi et al., 2005; Malagnini et al., 2011, Malagnini et al., 2008) have further investigated region-specific attenuation characteristics and source-related stress drop parameters, refining and confirming the findings of the early 2000s. In the following, a regionalization approach of the source and attenuation is performed based on the analysis of the residuals between the observed ground motion and the predictions from a GMM selected from literature.

The methodological approach consists of the following steps:

1. Selection of a representative GMM for Italy, that is identifying a suitable GMM from the literature that best represents seismic features in Italy
2. Residual analysis: Evaluating the differences between observed ground motion parameters and the predictions from the selected GMM
3. Modeling the spatial distribution of the residuals: Mapping and analyzing the spatial variability of residuals to identify regional patterns in source and attenuation characteristics

Steps 1 and 2 are described in this chapter, while the spatial distribution of the residuals (Step 3) is analyzed in Chapter 4.

3.1. Introduction to Empirical Ground Motion Models

Empirical GMMs estimate the median level of ground shaking and its associated uncertainty based on specific parameters, such as earthquake magnitude, source-to-site distance, style of faulting, and site effects.

GMMs are represented by functional forms that describe the median model as a function of the set of ground motion predictors, along with the associated standard deviation:

$$Y = \mu(X, \beta) + \epsilon \quad [5]$$

where Y represents a logarithmic of the ground motion intensity measure, μ is the median term modeled through a proper functional form, and σ is the standard deviation.

In the estimation of μ , X denotes the set of parameters describing the earthquake and site conditions, while β is the vector of coefficients that define the functional form.

It is also noted that in the equation above, ϵ represents the normally distributed residuals with zero mean and standard deviation σ .

Numerous studies in the literature highlight that the standard deviation can significantly affect the results of probabilistic seismic hazard analysis (Al-Atik et al., 2010). Consequently, considerable effort has been devoted in the last decades to developing GMM formulations that minimize standard deviation.

In ergodic ground motion models (Anderson and Brune, 1999), the assumption that the ground shaking is homogeneous and independent of the spatial location of the site significantly simplifies the calculation. The ergodic assumption is based on the principle that the variability of ground motion observed over a large geographical area is representative of the variability that might be observed at a single site over time. This definition implies that global ground motion models can be applied across a wide range of sites and geological conditions. However, this simplification introduces a critical limitation: it disregards the specific characteristics of individual sites, such as local soil response linked to regional geology, the specific features of the seismic source, and the effects related to wave propagation paths, which are highly dependent on the geometry and properties of the medium traversed. The simplified assumption of ergodicity may not adequately represent regional characteristics of ground shaking.

In recent years, the expansion of seismic datasets has enabled a relaxing of this assumption by decomposing model residuals into systematic contributions related to the earthquake source, propagation path, and local site effects. The availability of multiple recordings at a single station from different earthquakes across various paths has made it possible to quantify these effects and integrate them as systematic corrections to the median prediction in GMMs.

Within this revised framework, several studies (Lin et al., 2011; Anderson and Uchiyama, 2011; Stafford, 2014) have demonstrated that abandoning the ergodic assumption in GMMs formulations leads to a significant reduction in the standard deviation of GMMs.

To support the regionalization of GMMs at the Italian scale, this doctoral research aims to capture regional variations in source and attenuation effects, providing correction factors that adjust the median prediction of ground motion. These corrections are calibrated based on the decomposition of residuals, following the approach of Al-Atik et al. (2010), who distinguish repeatable contributions associated with the source, site, and path.

These correction factors along with the proposed Vs30 map, used to regionalize the site effects, are applied in conjunction with the national-scale GMM proposed by Lanzano et al. (2019) to generate ground shaking scenarios.

Before entering into the details of residual decomposition, the outline of the dependent and independent parameters included in the ITA18 model is provided with their definitions and geological significance.

3.2. Reference ground motion model for Italy

A scalar ground motion model for shallow crustal earthquakes in Italy developed by Lanzano et al. (2019), hereinafter termed ITA18, is the reference attenuation model for subsequent analyses in this thesis. This model represents an advancement of the model by Bindi et al. (2011) named ITA10, addressing and resolving some limitations identified in the formulation. Notably, the seismic sequences that occurred in Italy in 2012 and 2016-2017 allowed for an expansion of the dataset, covering a broader magnitude range and extending vibration periods up to 10 seconds. Additionally, in ITA18 the Vs30 has been introduced as a site proxy to characterize site conditions in place of the EC8 site categories included in ITA10. With the availability of a more comprehensive dataset, ITA18 is calibrated to predict peak motions (PGA and PGV) and spectral acceleration for a set of 35 discrete periods from 0.01 to 10 s. It employs a linear ordinary least-squares regression to separately fit N models corresponding

to different ground motion intensity measures. The median prediction can be presented as a functional form:

$$F_M(\text{SoF}) = f_j \text{SoF}_j, \quad [6]$$

where “a” is a constant, ϵ is the error associated with the median prediction, and the terms $FM(M_w, \text{SoF})$, $FD(M_w, R)$, and $FS(VS30)$ correspond to the *source*, *path*, and *site* terms, respectively. These terms are defined as follows:

$$F_M(M_w) = \begin{cases} b_1(M_w - M_h) & M_w \leq M_h \\ b_2(M_w - M_h) & M_w > M_h \end{cases}, \quad [7]$$

$$\log_{10} Y_j = a + FM(M_w, \text{SoF}) + FD(M_w, R) + FS(VS30) + \epsilon \quad [8]$$

$$F_D(M_w, R) = [c_1(M_w - M_{\text{ref}}) + c_2] \log_{10}(R) + c_3 R, \quad [9]$$

$$F_S(V_{S30}) = k \log_{10} \left(\frac{V_0}{800} \right) \quad [10]$$

The earthquake *source* is mainly defined by two parameters: magnitude (equation [7]) and faulting mechanism (equation [8]).

In equation [7], M_h represents the hinge magnitude that varies with period, while coefficients b_1 and b_2 control the magnitude scaling. Regarding magnitude, most studies, including Lanzano et al. (2019), use moment magnitude M_w as the preferred measure. This parameter quantifies earthquake size based on the total energy released. It is widely considered the most reliable criterion for ranking earthquakes by their strength. The advantage of using moment magnitude is twofold: first, it has a clear physical and seismological interpretation, and second, it remains effective for large earthquakes without reaching saturation, ensuring a consistent representation of energy release across all magnitude levels (Douglas, 2003).

In equation [8], the SoFjs are dummy variables to represent the fault mechanism, which describes how slip occurs along the fault plane during an earthquake: strike-slip ($j = 1$), reverse thrust ($j = 2$), and normal ($j = 3$). The corresponding coefficients, f_j , represent the SoF values for each fault type, with f_1 for strike-slip, f_2 for reverse thrust, and f_3 for normal faulting. While faulting style has been found to have a limited effect on the standard deviation of ground motion models, it is still considered a valuable parameter (Bommer, Douglas, and Strasser, 2003).

The *path* function, equation [9], is expressed with two main components: the first represents the magnitude-dependent geometrical spreading, while the second accounts for anelastic attenuation. Here, M_{ref} serves as the reference magnitude and c_1 – c_3 are the coefficients related to the path. The variable R is defined as:

$$R = \sqrt{R_i^2 + h^2}, \quad [11]$$

where R_i can refer to either the rupture distance (R_{rup}), which is the closest distance to the rupture plane, or the Joyner-Boore distance (R_{JB}), defined as the shortest horizontal distance from a recording station to the surface projection of the rupture plane. The parameter h represents the pseudo-depth (in km), which is determined through the regression analysis.

In the ITA18 model, the choice between using R_{rup} or R_{JB} depends on the level of detail available regarding the fault geometry. Specifically, R_{rup} is used when 3D information on the fault plane (including depth and dip angle) is available, while R_{JB} is adopted when only 2D surface projections of the fault trace are known.

The *site effect* is described in equation [10] where $V_0 = V_{s30}$ when $V_{s30} \leq 1500$ and $V_0 = 1500$ m/s otherwise. Local site conditions significantly influence ground motion recordings, making it essential to account for these effects. Two main approaches exist in the literature to integrate site effects into GMMs.

One approach is to use soil category classifications, such as the Eurocode 8 (CEN, 2004), which assigns soil categories based on predefined criteria. Each category is associated with a correction factor that enters the GMM as a categorical variable (Eurocode 8, 2004).

The other approach characterizes site conditions using a numerical parameter, in particular the shear-wave velocity (V_s). This parameter is advantageous because it has a clear physical meaning and does not rely on soil classification schemes. Specifically, ITA18 uses V_{s30} , which represents the average shear-wave velocity measured in the uppermost of 30 meters.

High Vs30 values are typically found in rock or stiff soil sites, where lower spectral acceleration amplitudes are expected. Conversely, lower Vs30 values indicate soft soils, which generally result in higher spectral acceleration amplitudes (Douglas, 2003).

To further clarify the model, some additional considerations regarding the response variable are necessary. Most accelerometric recordings capture three orthogonal components of ground motion intensity—two horizontal and one vertical. However, many ground motion models predict the geometric mean of the horizontal components that is independent of the sensor's orientation at the recording site. Among the various orientation-independent intensity measures available in the literature, ITA18 adopts the RotD50 metric, introduced by Boore (2010). This measure is derived by computing the median of the intensity measure distribution obtained from the combination of the two horizontal components across all unique azimuths (Lanzano et al., 2019).

3.3. Ground Motion Variability and Residual Decomposition

Reducing aleatory variability (σ , in equation [5]) requires moving away from the ergodic assumption and identifying components of ground-motion variability, through statistical decomposition of the total residuals, that are systematic and repeatable, rather than purely random (Al-Atik et al. 2010; Anderson & Uchiyama, 2011; Rodriguez-Marek et al., 2013; Villani & Abrahamson, 2015; Kotha et al., 2016; Baltay et al., 2017; Lanzano et al., 2017; Lanzano et al., 2018; Sgobba et al., 2021).

To better understand how these systematic components are identified, it is essential to first define the residuals (Res), which quantify the deviations between observed and predicted ground motions.

In particular:

$$R_{es} = \log(Y)_{obs} - \log(Y)_{pred} \quad [12]$$

Where $\log(Y)_{obs}$ represent the observed ground motion intensity measure, expressed in either log10 or natural logarithm (ln) depending on the chosen model, while $\log(Y)_{pred}$ is the corresponding predicted ground motion intensity measure.

In ergodic GMMs, residuals can be decomposed into two statistically independent components (Rodriguez-Marek et al., 2013):

$$Res = \delta Be + \delta Wes \quad [13]$$

where:

I. δBe (between-event residual) represents the systematic deviation observed between prediction and observation for a specific event, i.e., the median of the residuals for single events. Several authors (Bindi et al., 2018; Bindi & Kotha, 2020) show that δBe is correlated with source characteristics not captured by the standard magnitude scaling in GMMs, such as stress drop.

II. The within-event residual (δWes) quantifies site-specific deviations from the event-specific median prediction ($\mu + \delta Be$), influenced by wave propagation effects and local site conditions. This component represents record-to-record variability within a single earthquake (Al-Atik et al. 2010).

Since δBe and δWes are uncorrelated, the total standard deviation (σ) of a GMM can be expressed as:

$$\sigma = \sqrt{\tau^2 + \phi^2} \quad [14]$$

where τ represents the standard deviation of the between-events residual and ϕ represents the standard deviation of the within-event residual.

The within-event residuals can, in turn, be separated (Rodriguez-Marek et al., 2013):

$$\delta Wes = \delta S2S_s + \delta WS_{es} \quad [15]$$

where:

I. $\delta S2S_s$ (site-to-site residual) represents the systematic deviation of the observed ground motion at site s (i.e., the site term) from the median event-corrected ground motion predicted by the GMM (Rodriguez-Marek et al., 2013). The $\delta S2S$ can be related to the systematic effect of the site response not captured by the model.

II. δW_{Ses} (event- and site-corrected residual) is the residual corrected for both source and site effects, where any other effects not captured by the model may converge, such as directivity effects (Colavitti et al., 2022), regional propagation effects, aleatory variability.

The standard deviations of these terms are ϕ_{S2S} and ϕ_{ss} , known as the event-corrected single-station standard deviation, or simply single-station ϕ .

Although these components contribute to variability, they also include systematic effects related to the earthquake source, propagation path, and site conditions. Unlike ergodic models, which treat this variability as random, non-ergodic models incorporate these repeatable effects, reducing uncertainty and improving ground motion predictions.

3.4. Application of Residual Analysis in Italy

A common observation in the literature is that the residual terms of a GMM, particularly the within-event component, exhibit significant spatial correlation (Jayaram & Baker, 2010; Lin et al., 2011; Walling 2009). This correlation can be modeled using geospatial analysis and leveraged to estimate shaking intensities (Sgobba et al., 2021). Several studies have shown that spatial modeling is essential for generating correlated random fields of ground motion, as neglecting it can lead to substantial biases in loss estimation and seismic risk assessments (Park et al., 2007; Crowley et al., 2008; Schiappapietra et al., 2020).

However, only a limited number of studies in the literature address fully non-ergodic models and incorporate spatial correlation of the repeatable terms to generate ground motion fields. Most of these studies focus on well-instrumented regions, such as California (Landwehr et al., 2016; Sahakian et al., 2019; Kuehn & Abrahamson, 2019; Parker et al., 2020) and Europe (Lanzano et al., 2018; Sgobba et al., 2019; Menafoglio et al., 2020; Kotha et al., 2020).

In the Italian context, a fully non-ergodic methodology for shaking scenarios (referred to as NESK) has been introduced by Sgobba et al. (2019) for Northern Italy and Sgobba et al. (2021) for Central Italy. This approach corrects median ground motion estimates at the reference rock level (excluding site amplification) for between-event biases and integrates these with empirical maps of regional path- and site-response corrections. By doing so, it captures the characteristics of the source area and anisotropic propagation effects, which were previously associated only with methods based on numerical simulations (Sgobba & Pacor, 2023). As a result, this methodology enhances the accuracy of ground motion predictions within empirical shaking scenarios.

In this work, to capture both the regional source effects and regional attenuation effects not explained by the ergodic ITA18 the main step involved is the analysis of residuals.

The residuals Res (equation [12]) are calculated for the ordinates of the acceleration response spectra (5% damping) within the vibration period range of 0.01–10 seconds. They are decomposed into repeatable terms as in Al Atik et al. (2010):

$$R_{es} = \delta B_e |_{event} + \delta S_2 S_s |_{site} + \delta W_{es} \quad [16]$$

Refer to Chapter 3.3 for the description of repeatable terms. A random-effects regression model is applied to the residuals (Stafford, 2014; Bates et al., 2015). Given that the residuals are derived from a dataset that exceeds the validity limits of the ITA18 calibration, the studies by (Lanzano et al., 2022; Mascandola et al., 2022) ensure that no significant biases exist in the residual components relative to the model's explanatory variables. To establish a zonation for the source effect not captured by the GMM, the spatial distribution of δB_e was analyzed.

The zonation of ground motion attenuation depends on the distance term defined in the ITA18 model (Lanzano et al., 2019), which is expressed in Equation [6].

The coefficients c_1 and c_2 characterize the geometrical spreading, incorporating a magnitude-dependent saturation term, while c_3 accounts for the effects of anelastic attenuation. Assuming that the event- and site-corrected residuals are entirely attributable to regional path effects not captured by the model, the following relationship can be established:

$$\delta W_{es} = \delta c_3 \sqrt{R_{rup}^2 + h^2} \quad [17]$$

where δc_3 becomes:

$$\delta c_3 = \frac{\delta W_{es}}{\sqrt{R_{rup}^2 + h^2}} \quad [18]$$

where δc_3 represents a correction factor added or subtracted to the anelastic attenuation coefficient c_3 that compensates the δW_{es} residuals record-by-record. This approach is similar to the method used by Kotha et al. (2020) to refine the regionalized European Ground Motion Model (GMM), where variations in the propagation medium are attributed to changes in the anelastic attenuation coefficient.

4. Data-driven zonations for modelling the regional source and propagation effects into Ground Motion Models in Italy

4.1. Tectonic setting of Italy

A brief tectonic setting of Italy follows.

The north-south convergence of Eurasian and African plates, which started during the Cretaceous period, caused the subduction of the Tethys oceanic crust. Around 65 million years ago, this process culminated in the collision between the Adria and the Eurasian plate (Schmid et al, 1996), leading to the formation of the Alps. The Alpine chain stretches approximately 1000 km, forming a distinctive arc that encircles northern Italy.

The Alps are a thrust-and-fold belt composed of large sequences of cover nappes and basement rocks, which were transported toward the European foreland (Meletti et al, 2000). The range exhibits a double-verging growth structure. The Insubric Line (IL in Figure 13), a dextral transpressive fault zone (Rosenberg, 2004), marks the division between the Southern Alps, which predominantly verge to the south-southeast, and the other Alpine units that verge westward and northwestward toward the Eurasian plate.

The Apennines form an arc that stretches from northwestern Italy across the Italian peninsula, extending southwest toward Sicily and connecting with the Maghrebides in northwestern Africa. This fold-and-thrust belt, referred to in the literature as the Apennine-Maghrebian orogen (Malinverno et al 1986; Faccenna et al, 1997), began to develop during the Miocene as a result of the directed subduction of the Adriatic-Ionian lithospheric slab (Malinverno et al., 1986; Patacca et al., 1990).

The belt migrated southeast into Calabria and Sicily, moved eastward through the central and southern Apennines, and extended northeast into the Northern Apennines (Figure 13). Starting from the Tortonian stage (Late Miocene), extensional tectonics driven by the rollback of the Adriatic-Ionian segment of the subducting lithosphere led to ocean-floor spreading in the southern Tyrrhenian Sea (Sartori, 1991; Bassi et al., 1997). This rollback also triggered southwest-northeast extension within the inner Apennines, which are now marked by active NE- and SW-dipping normal and oblique-normal faults, often associated with continental basins (Patacca & Scandone, 1989).

The Apennine chain is divided into two arcs: the first, convex toward the Padanian-Adriatic-Ionian foreland, connects to the second arc via a significant transverse discontinuity known as the Ortona-Roccamonfina Line (ORL in Figure 13 (Locarni, 1982), which is interpreted as the surface expression of a deep lithospheric tear (Patacca & Scandone, 1989).

In the Calabrian-Peloritanian Arc (Figure 13), intermediate and deep earthquakes outline a Wadati-Benioff zone that descends beneath the Tyrrhenian Sea to depths reaching approximately 600 km (Anderson & Jackson, 1987, Neri et al., 2021). The Calabrian Arc (CA in Figure 13) subduction system is composed of two distinct lobes, representing separate domains (Polonia et al., 2011): the western lobe (WL) corresponds to areas where the lithospheric plate has already fragmented, while the eastern lobe (EL) features an unbroken plate that extends continuously into the mantle, as suggested by tomographic evidence (Polonia et al., 2011).

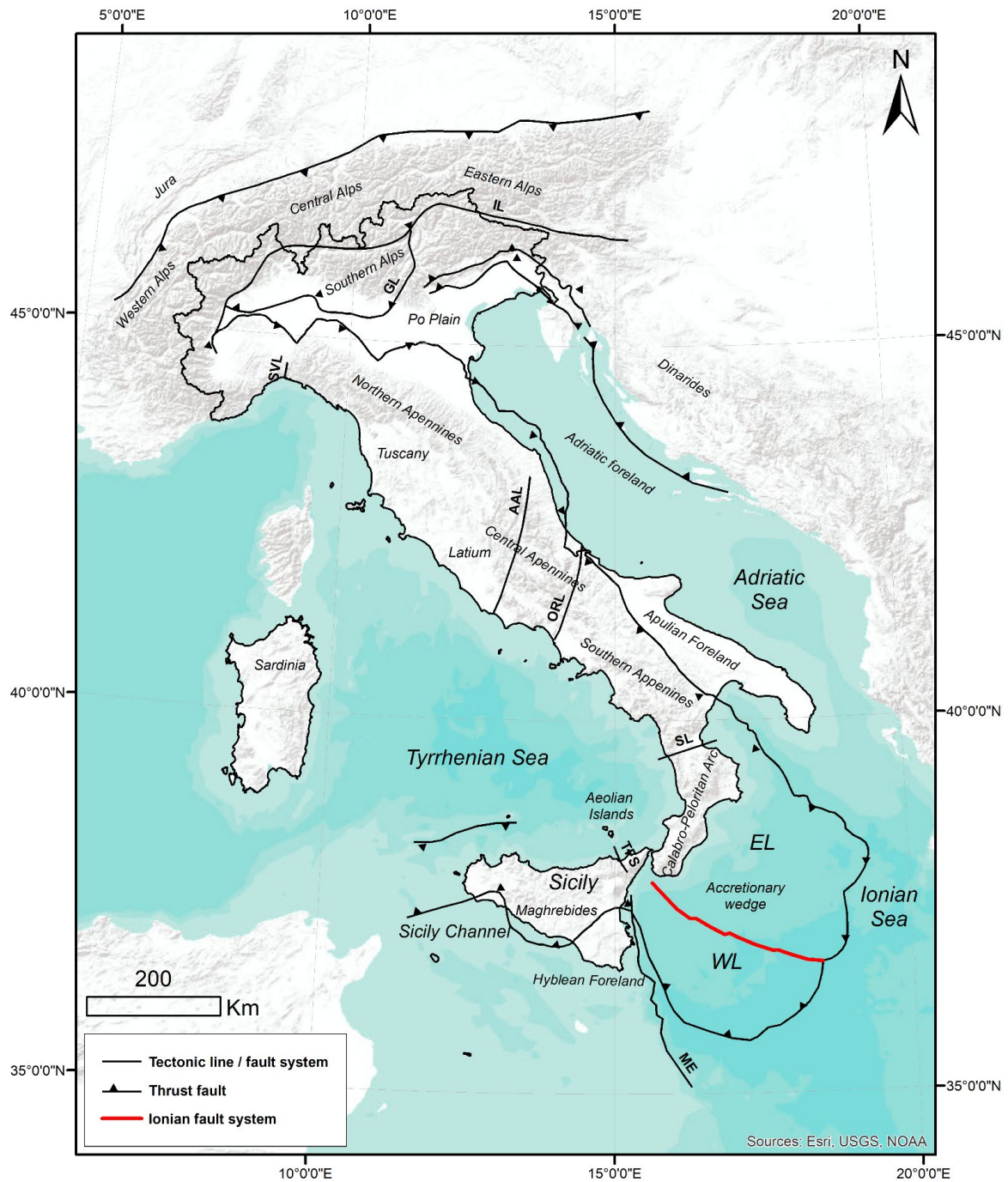


Figure 13. Simplified tectonic map of Italy (modified after Bigi et al.,1991). IL: Insubric Line; GL: Giudicarie Line; SVL: Sestri-Voltaggio Line; AAL: Ancona-Anzio Line; ORL: Ortona-Roccamonfina Line; SL: Sanginetto Line; EL: Eastern Lobe of the Ionian fault; WL: Western Lobe of the Ionian fault; ME: Maltese Escarpment; TFS: Tindari Fault System. Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses.

4.2. Spatial variability of residuals related to source and path effects

The regional variability of residuals is analyzed through a three-step process:

Step 1: Mapping the spatial distribution of the residual terms δB_e and δc_3 (Eq. 18) to identify spatial patterns and their variability across different regions. The residuals are averaged over two-period ranges: $T=0.07\text{--}0.4\text{ s}$ and $T=0.7\text{--}4\text{ s}$, to analyze spatial trends at short and long periods, respectively. These two ranges were selected due to their engineering significance, as they correspond to the resonance frequencies of various structural elements: high frequencies (14–2.5 Hz) and low frequencies (1.4–0.25 Hz), which are also relevant for seismic design according to the latest European standards (Paolucci et al., 2021). The intermediate period range ($T = 0.4\text{--}0.7\text{ s}$) was not examined in this study, as a preliminary analysis of the spatial distribution of the residual terms δB_e and δc_3 did not reveal any significant patterns in that range, similarly to what was observed for long periods. Nonetheless, the long-period range ($T = 0.7\text{--}4\text{ s}$) was retained in the analysis due to its engineering relevance for the dynamic response of tall structures, as reflected in seismic design standards.

Step 2: Validating the trends observed in Step 1 using independent physical parameters, providing a physical interpretation for the detected spatial variability.

Step 3: Propose a revised zonation based on the residual trends, integrating the existing seismic zonation, main structural features (Figure 13), and instrumental seismicity recorded since 1985.

4.3. Regionalization of source effects

In the *First step*, the between-event residuals, δB_e , are interpolated using a univariate Ordinary Kriging algorithm (Matheron G., 1965). This approach assumes the commonly accepted hypotheses of isotropy and stationarity, employing an exponential variogram model (Cressie, 1985) to fit the residuals. Figure 14 illustrates the interpolated δB_e values for two-period intervals, $T = 0.07\text{--}0.4\text{ s}$ and $T = 0.7\text{--}4\text{ s}$. In these maps, red regions indicate positive residuals, where observed values exceed predictions, while blue regions represent areas where the opposite occurs.

The spatial distribution for short periods, shown in Figure 14(a) suggests a correlation with the primary structural domains (Figure 13). The highest between-event residuals are observed in

regions such as the Alpine arc, particularly the Western Alps and the external Dinarides, as well as the Apulian platform, the Southern Ionian Sea, and the Hyblean plateau. Conversely, the lowest values are found in Northern Sicily and the Northern Ionian Sea. Along the Apennines, the residuals are near zero or slightly negative. For longer periods, these trends become less evident, and deviations from the reference model are significantly reduced. This indicates that the Ground Motion Model (GMM) at long periods does not show spatial variations.

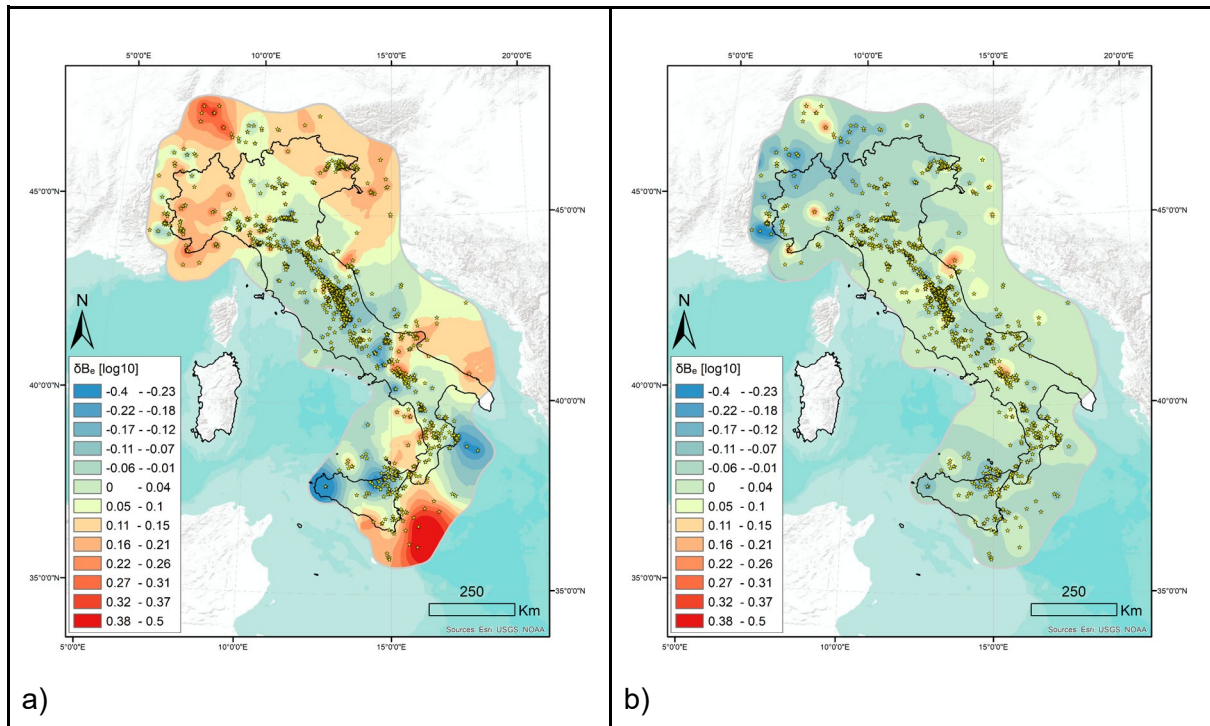


Figure 14. Interpolated maps of δB_e (in decimal log units) at a) short periods ($T = 0.07-0.4s$) and b) long periods ($T = 0.7-4s$). Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses

The deviations from the median trend predicted by the GMM at high frequencies (i.e., δB_e) can likely be attributed to the stress drop, a parameter that contributes to increasing ground motion values at high frequencies.

To further investigate this hypothesis, the *second step* involves comparing the δB_e estimates with Arias stress drop ($\Delta\sigma_{AI}$) values, as derived in (Bindi & Kotha, 2020) using the method outlined in (Baltay et al., 2019) that calculates stress drop from the Arias intensity. $\Delta\sigma_{AI}$ was determined for the events in the ESM flat-file (Lanzano et al., 2019), which includes records of pan-European earthquakes with magnitudes greater than 4.0. Figure 15 displays the stress

drop estimates for the 228 events common to the ITACAext dataset (out of 1863 total events), plotted against the δB_e estimates for short and long periods.

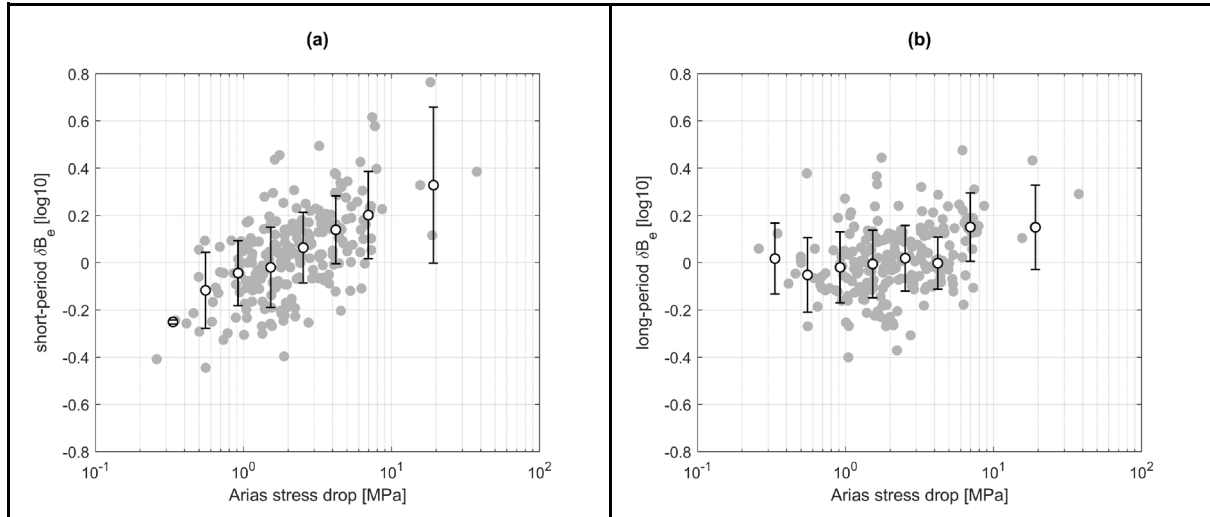


Figure 15. Arias stress drop values (Bindi & Kotha, 2020) vs. between-event residuals at (a) short and (b) long periods. Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses

The parameter $\Delta\sigma_{AI}$ demonstrates a strong correlation with the event residuals at short periods, as reported by several authors (Baltay et al., 2017; Bindi et al., 2017, Morasca et al., 2023). According to Brune's source model, stress drop governs the amount of energy radiated at high frequencies in the Fourier domain for a given seismic moment (Baltay et al., 2017, Anderson, 1997). Conversely, at long periods, the correlation weakens, which is consistent with prior observations (Morasca et al., 2023).

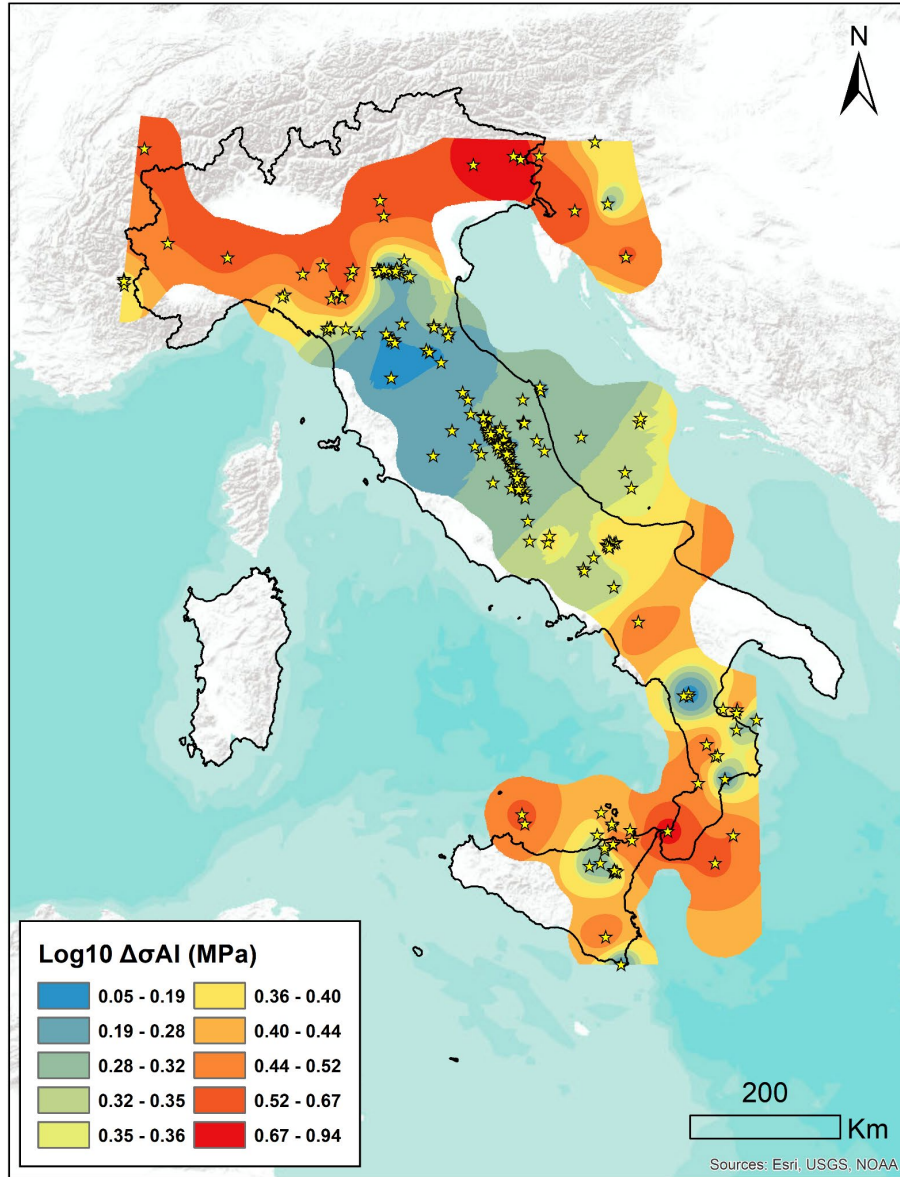


Figure 16. Interpolated map of the Arias stress drop estimates provided by (Bindi & Kotha, 2020). Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses

Figure 16 presents a map of the stress parameter obtained through spatial interpolation of $\Delta\sigma_{AI}$ estimates, allowing for a direct comparison with the map in Figure 14. As anticipated, the spatial distribution of $\Delta\sigma_{AI}$ aligns closely with the patterns observed for δB_e in the period range $T = 0.07\text{--}0.4$ s (Figure 14a). Specifically, lower $\Delta\sigma_{AI}$ values are observed from the Northern Apennines to Molise, corresponding to regions where the between-event residuals are negative. Conversely, higher $\Delta\sigma_{AI}$ values are found in Northern Italy and the southern Ionian Sea, matching areas where δB_e is positive. In Southern Italy, $\Delta\sigma_{AI}$ values are lower in central Calabria and northeastern Sicily, which also aligns with the trends observed in Figure 14a.

To regionalize the source contribution for ground motion modeling (*Third Step*), I use the seismogenic zonation ZS16 (Visini et al.; 2022) as a starting point. This zonation was specifically developed for the latest probabilistic seismic hazard model for Italy, MPS19 (Meletti et al.; 2021). ZS16 represents an updated version of the ZS9 model (Meletti et al., 2008), which served as the reference seismogenic framework for the MPS04 (Stucchi et al., 2011). The MPS04 model, in turn, was the basis for defining the design spectra in the current Italian building code (NTC18, NTC-Italian Building Code, 2018). Covering the entire national territory, ZS16 identifies 50 zones based on multiple criteria, including:

- i) updated epicentral locations provided by the Parametric Catalogue of Italian Earthquakes (CPTI15; Rovida et al., 2016);
- ii) the spatial distribution of seismogenic sources, as detailed in the most recent version of the Database of Individual Seismogenic Sources (DISS 3.2.1, (DISS Working Group, 2018); and
- iii) alignment with the geodynamic model proposed by (Meletti et al., 2000) for one of the earliest seismogenic zonations, known as ZS4.

The final definition of source zones is mainly driven by the spatial distribution of short-period δB_e , as illustrated in Figure 14a. Additionally, the spatial distribution and focal mechanisms of shallow crustal earthquakes since 1985, as well as the seismogenic faults cataloged in DISS, were considered. The proposed zonation for source correction is shown in Figure 17, with the criteria used for zone definition outlined in Table 2.

The Italian territory is divided into 21 zones; however, Table 2 provides estimates of the median event bias δB_e and its aleatory variability (the standard deviation of δW_{es}) only for 17 of these zones. Statistical values for five zones were deemed unreliable due to specific reasons:

- i) Background areas with sparse seismicity or unclear tectonic associations (zones #5, #9, and #18) are shown in Figure 17;
- ii) regions dominated by seismic activity of volcanic origin (#10 and #100) are shown in Figure 17.

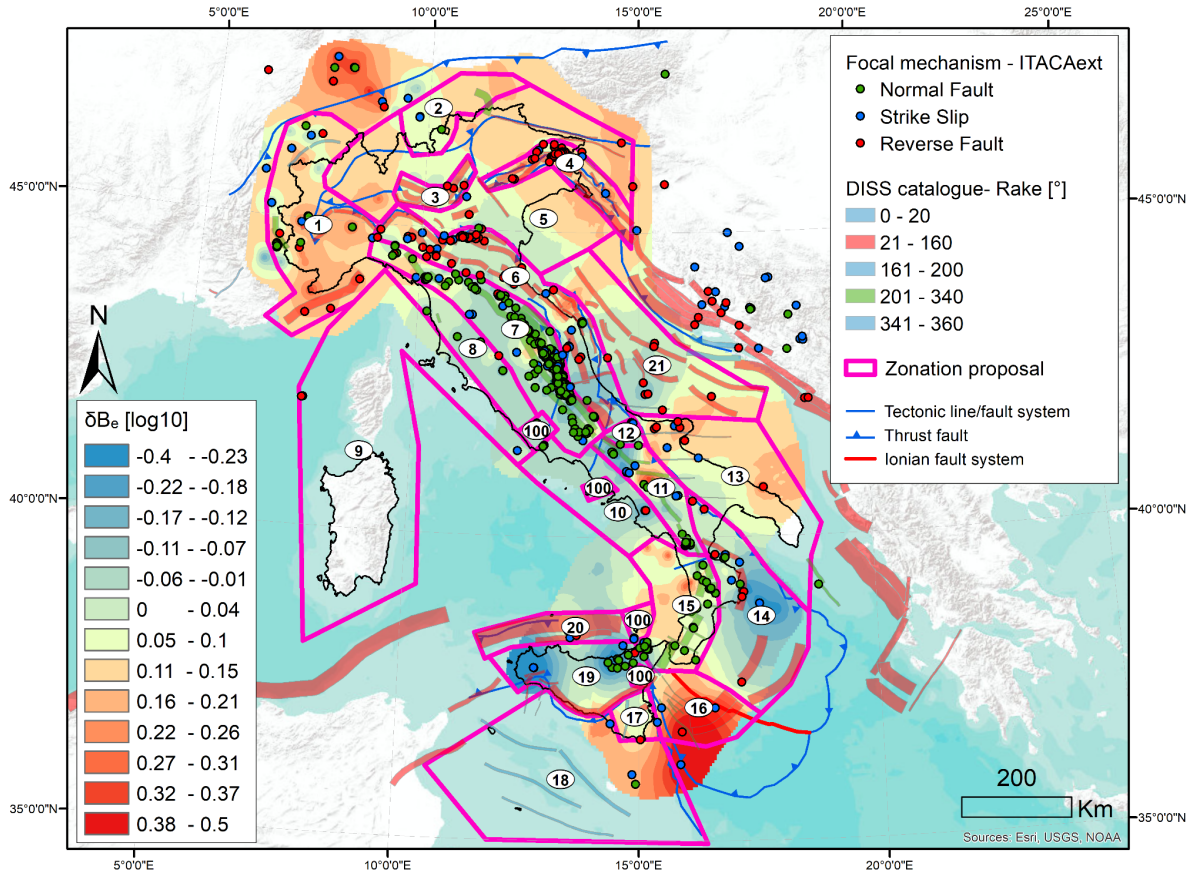


Figure 17. Zonation for regionalization of source effects into GMMs. The numbering of the zones is consistent with Table 2. Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses

Table 3 includes information for each zone: the number of events, the number of recordings, the number of events per km², the number of recordings per km², and the percentage variation with respect to the ITA18 prediction [%]. These details are provided to help assess the confidence level of the δB_e statistics.

For example, zones contain less than 10 events (zones #2 and #8) or have a spatial event density lower than 5×10^{-4} events/km² (zones #13 and #21). Consequently, it is recommended to exclude these low-sampled zones from calibrating a new regionalized model until more data becomes available to better constrain ground motion in these areas.

Regions where the between-event exceeds 0.15 log10 units are primarily found in Northern Italy, particularly along the Alpine chain in the Western Alps (zone 1). Similar biases are also observed in Southern Italy, especially in the Apulian Foreland (zone 13) and the Western Lobe of the Ionian Fault (zone 16). These areas are characterized by seismic motion that is

significantly higher than the predictions made by the GMM. Specifically, for zones 13 and 16, the between-event is 0.388 and 0.190. In contrast, regions where predicted ground motion overestimates observations are located in Southern Italy, such as Molise (zone 12), Northern Sicily (zone 19), and the Eastern Lobe of the Ionian Fault (zone 14). In zone 19, the between event is -0.166. In other cases, deviations from the reference model remain within $\pm 20\%$.

Zones with the highest variability of the between-event are Irpinia and the Southern Apennines (zone 11) and the Western Lobe of the Ionian Fault (zone 16). In the Irpinia region, this high variability is likely associated with the 1980 seismic sequence (Mw 6.9). While the mainshock was recorded by a broader network, many of the aftershocks were captured by a limited number of stations that had been deployed specifically to investigate site effects. As a result, the recordings for some of these aftershocks are dominated by local site amplification, introducing a potential trade-off between source and site terms. In such cases, when only a few stations affected by strong site effects are available, the between-event term may inadvertently reflect site-related contributions rather than true source characteristics. In zone 16, the variability may be influenced by a trade-off between source and path effects, because of the azimuthal gap of the recordings. This effect is particularly significant on the Hyblean plateau, where most recording stations are located.

I supposed that the variability of between-event term is primarily associated with small-magnitude events, which dominate the dataset. Considering the regional scaling laws between magnitude and stress drop, an important question arises: do deviations from the median for small-magnitude events indicate similar regional trends for large-magnitude events? To address this, I evaluate the representativeness of short-period δBe across a magnitude range that includes both small and large earthquakes. Four zones were selected, where significant earthquake sequences occurred, ensuring that the observations cover a sufficiently broad magnitude range.

The selected zones (Figure 17) include: zone #4, which encompasses the Friuli sequence (1976-1977, MW 6.4, reverse fault mechanism); zone #6, which covers the Po Valley sequence (2012, MW 6.0 and 6.1, reverse fault mechanism); zone #7, corresponding to the central Apennines, where at least three major sequences occurred (1997-1998 Umbria-Marche, 2009 L'Aquila, and 2016-2017 Amatrice-Norcia, all characterized by normal faulting); and zone #11, which includes the Irpinia sequence (1980, MW 6.9, normal fault mechanism).

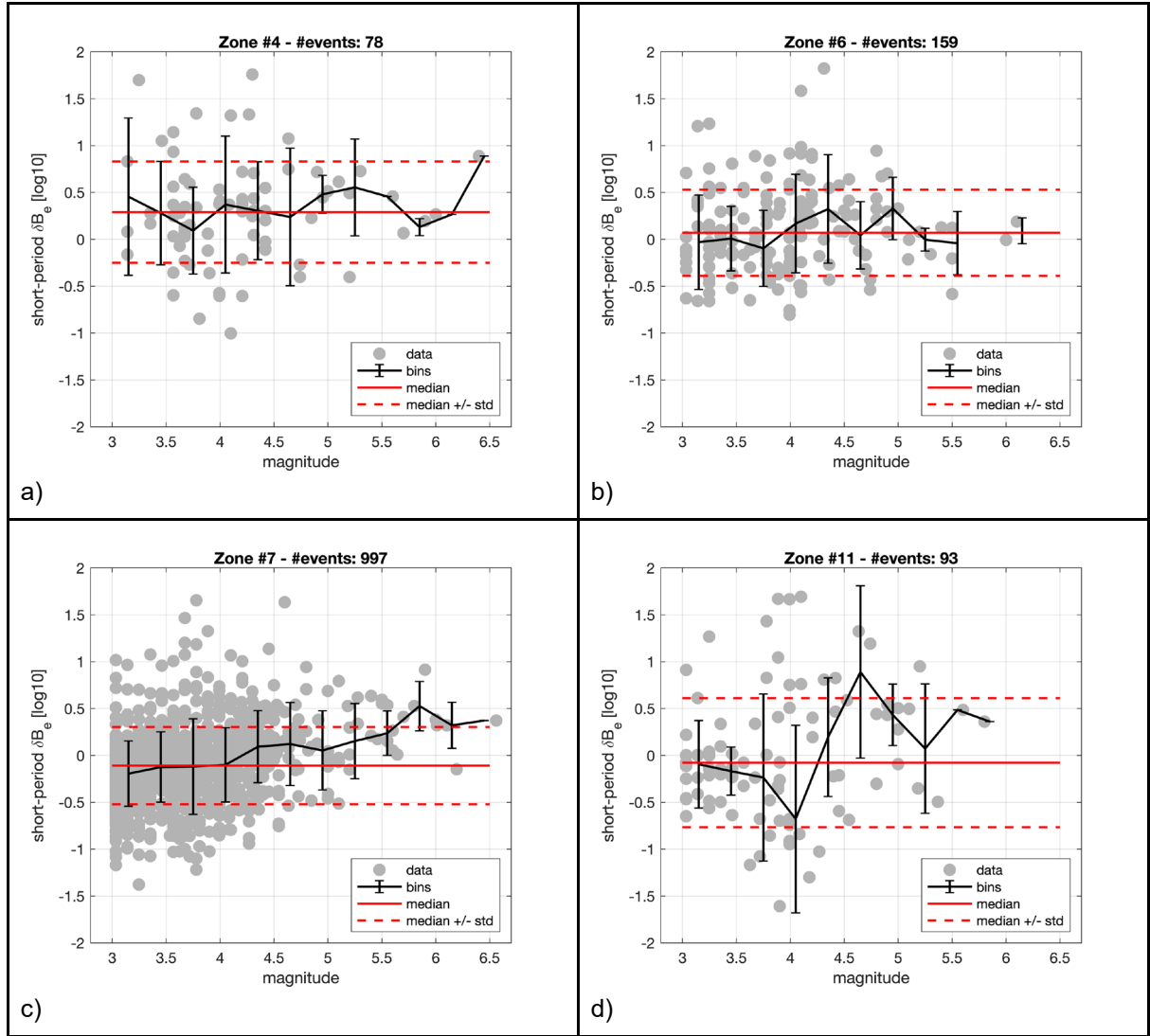


Figure 18. δB_e values for short periods ($T = 0.07\text{--}0.4$ s) as a function of moment magnitude, shown for different source zones: a) zone #4, b) zone #5, c) zone #7, d) zone #11. Reproduced from Brunelli et al. (2023), published in *Bulletin of Geophysics and Oceanography (BGO)*, under CC BY 4.0.

Figure 18 presents the δB_e values for each event within the selected zones (represented by grey circles), the median values calculated for magnitude bins along with their associated uncertainty (black line), and the overall median calculated across all data (red line).

For zones #4 and #6, a weak positive correlation with magnitude is observed. However, the median δB_e for short periods, computed over all data, effectively represents the entire data population, with fluctuations in the binned values remaining within the standard deviation range.

In zone #7, a clear dependence on magnitude and a positive correlation with δB_e are evident. This trend, supported by the availability of a large dataset, has been previously documented in the literature, including works by Bindi et al. (2018) and Morasca et al. (2023). In this case,

the median δB_e calculated across all data primarily reflects small-magnitude events, typically ranging between 3.5 and 4.

For zone #11, the data exhibit significant scatter due to the poor quality of the limited recordings available for the Irpinia sequence. Nevertheless, the observed trend is somewhat similar to that of zone #7.

Table 2. List and synthetic description of the zones for the regionalization of source effects. Reproduced from Brunelli et al. (2023), published in Soil Dynamics and Earthquake Engineering, Elsevier. Reused in accordance with Elsevier's policy for inclusion in doctoral theses.

Zone	Name	Description	Short-period median δB_e [log10]	Standard deviation δB_e [log10]
1	<i>Western Alps and the coast of Liguria region</i>	This zone is characterised by positive δB_e . It includes seismogenic structures with different focal mechanisms. The boundary line with zone 7 is represented by the Sestri-Voltaggio tectonic line.	0.200	0.236
2	<i>Northern Alps</i>	Merges two ZS16 zones with poor event sampling. Includes minor structures north of the Insubric line with predominantly normal mechanisms.	0.101	0.116
3	<i>Garda Lake</i>	It is one zone of ZS16, which includes structures located south of Garda Lake with predominantly reverse/thrust mechanisms. The southern margin corresponds to a tectonic alignment. Weakly positive δB_e .	0.031	0.199
4	<i>North-Eastern Alps</i>	This area includes the events of the 1976-77 Friuli sequence (mainshock Mw 6.4) and merges two zones of ZS16. The events have predominantly reverse mechanisms and, to a lesser extent, strike-slip. The tectonic alignments define the geometry of the area and the δB_e are positive.	0.126	0.235

5	<i>Northern Italy</i>	This zone merges areas where seismicity is sparse and with low magnitudes. Regardless of the structures, it is assumed that a robust estimate of δB_e is not possible for the area.	-	-
6	<i>Compressive Northern-Central Apennines</i>	The zone includes the eastern edge of the Apennine thrust belt from the northern Apennine to Molise region. It is mainly characterised by compressive tectonics and weakly positive values of δB_e . The observed focal mechanisms are predominantly reverse.	0.030	0.199
7	<i>Distensive Northern-Central Apennines</i>	The zones encloses the area where most of the seismicity of the Apennines is concentrated, including some of the most relevant sequences like the 1997-98 Mw 6.0 Umbria-Marche, 2009 Mw 6.1 L'Aquila and 2016-2017 Mw 6.5 Amatrice-Visso-Norcia. The prevalent style of faulting is normal. It is characterised by δB_e close to zero, mostly because the majority of the data used to calibrate the ITA18 GMM comes from this area. A tectonic line (named Sestri-Voltaggio) marks the northern boundary with zone 1.	-0.048	0.179
8	<i>Volcanic districts in Tuscany and Latium</i>	This area is characterised by extinct volcanism and infrequent seismicity. The event residuals are weakly negative.	-0.225	0.250
9	<i>Sardinia and Northern Tyrrhenian Sea</i>	It includes a very large area with very rare seismicity. A source correction is also not estimated for this area.	-	-

10	<i>Coast of Campania region</i>	It encloses some active volcanic districts. Active crustal seismicity is rare. The paucity of data does not allow estimation of δB_e .	-	-
11	<i>Irpina and Southern Apennines</i>	This area contains the seismicity of the southern Apennines (up to the Sangineto line), including the events of the 1980 Mw 6.9 Irpinia sequence. The Bradanic trough partially marks the north-eastern margin and the prevailing focal mechanism is normal.	-0.033	0.299
12	<i>Molise region</i>	The events in this region, including those in the 2002 Mw 5.7 sequence, have negative δB_e . It is a small zone characterized by predominant strike-slip mechanism.	-0.079	0.155
13	<i>Apulian Foreland</i>	The events that occurred in the Gargano and Salento peninsulas show remarkably positive values of δB_e , leading to the definition of a source zone corresponding to the Apulian Foreland, bounded to the south-west by the Bradanic trough. The focal mechanisms observed are reverse and strike-slip.	0.190	0.200
14	<i>Ionian fault Eastern Lobe</i>	This zone is characterized by negative residuals and predominantly reverse mechanisms. To the north the area is bounded by the Bradanian trench, to the south-west by the Ionian fault, and to the south-east by the eastern margin of the accretionary wedge. In particular, the Ionian fault marks a very clear transition between negative residuals in zone 14 and positive residuals in zone 16.	-0.088	0.229

15	<i>Calabrian Apennine and Southern Tyrrhenian Sea</i>	This zone covers the area where the Nubia plate, in subduction under the southern margin of the Eurasian plate, flexes and plunges within the asthenospheric mantle (Orecchio et al., 2021). Positive δB_e residuals are observed. The zone is located between the Sangineto line to the north and the Tindari fault system to the south-west. The focal mechanisms observed are predominantly normal.	0.086	0.181
16	<i>Ionian fault Western Lobe</i>	The area is bounded by the Eastern Lobe of the Ionian fault (north-east), the Malta escarpment (west) and the southern margin of the accretionary wedge (south-east). The residuals are remarkably positive.	0.388	0.261
17	<i>Hyblean Plateau</i>	The geometry of the zone is very similar to that proposed in ZS16 for the Iblean carbonate foreland. It is delimited to east by the Malta escarpment and to north by the Maghrebides thrust. The residuals are slightly positive. The focal mechanisms observed are mostly reverse and strike-slip.	0.010	0.152
18	<i>Sicily channel</i>	This large area includes few observations suggesting that a source correction factors should not be introduced.	-	-
19	<i>Northern Sicily</i>	It includes part of Sicily located north of the Maghrebid thrust and south of the Ustica fault. Residuals are negative. The focal mechanisms observed are predominantly normal in northeastern Sicily while reverse and strike-slip in the southern part of the zone.	-0.166	0.233

20	<i>Southern Tyrrhenian</i>	This zone delimits the tectonic alignment of the Ustica fault characterized by reverse focal mechanisms. Weakly positive δB_e .	0.007	0.197
21	<i>Central Adriatic Sea</i>	It includes the reverse faults system in the central Adriatic, with weakly positive or zero event residuals.	-0.043	0.177
100	<i>Volcanic areas</i>	It includes multiple areas in Italy (Mount Etna, Aeolian Islands, Ischia Island, Phlegrean Fields and Mount Vesuvius) where the seismicity is generated by the volcanic activity and where the ITA18 model is not applicable.	-	-

Table 3. Additional information on the zones is listed in Table 1. Reproduced from Brunelli et al. (2023), published in Soil Dynamics and Earthquake Engineering, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses

Zone	Name	Number of events	Number of records	Number of events/ Area [Km2]	Numbers of records/Area [Km2]	Percentage variation w.r.t. ITA18 prediction [%]
1	Western Alps and the coast of Liguria region	55	1458	1.00E-03	2.66E-02	58.66
2	Northern Alps	8	176	4.83E-04	1.06E-02	26.24
3	Garda Lake	12	253	1.72E-03	3.63E-02	7.45
4	Northeastern Alps	78	812	5.08E-03	5.29E-02	33.65
5	Northern Italy	16	370	1.78E-04	4.11E-03	69.97
6	Compressive Northern-	159	4384	3.79E-03	1.05E-01	7.16

	Central Apennines					
7	Distensive Northern-Central Apennines	997	22438	2.49E-02	5.60E-01	-10.38
8	Volcanic districts in Tuscany and Latium	7	112	3.71E-04	5.93E-03	-40.46
9	Sardinia and Northern Tyrrhenian Sea	3	18	2.27E-05	1.36E-04	37.87
10	Coast of Campania region	9	69	3.48E-04	2.67E-03	-22.91
11	Irpina and Southern Apennines	93	1217	6.43E-03	8.42E-02	-7.41
12	Molise region	43	810	1.69E-02	3.19E-01	-16.6
13	Apulian Foreland	22	463	4.09E-04	8.61E-03	54.75
14	Ionian fault Eastern Lobe	33	311	7.21E-04	6.79E-03	-18.39
15	Calabrian Apennine and Southern Tyrrhenian Sea	87	1383	3.16E-03	5.02E-02	22.02
16	Ionian fault Western Lobe	13	131	6.60E-04	6.65E-03	144.37
17	Hyblean Plateau	11	102	1.80E-03	1.67E-02	2.38
18	Sicily channel	9	156	9.75E-05	1.69E-03	46.69

19	Northern Sicily	34	737	1.33E-03	2.88E-02	-31.83
20	Southern Tyrrhenian	10	140	7.08E-04	9.91E-03	1.66
21	Central Adriatic Sea	14	235	2.99E-04	5.01E-03	-9.47
100	Volcanic areas	24	366	3.00E-03	4.57E-02	-20.05

4.4. Zonation for regionalization of attenuation effects into GMMs

To analyze the spatial distribution of δc_3 , it is important to consider that the characteristics of seismic wave propagation depend on the relative positions of the event and the station. This is quite challenging in ground motion modelling and was tackled in the literature in different ways:

- i) Ground motion models with spatially variable coefficients, as proposed by (Landwehr et al., 2016, Lanzano et al., 2021), consider a dependence on the event coordinates.
- ii) The grid-based method introduced by (Dawood & Rodriguez-Marek, 2013; Sgobba et al., 2021) maps each source-to-site path in the calibration dataset and incorporates it into the model as a random effect.
- iii) Kotha et al., 2020 propose a zonation approach based on station coordinates.

In this study, a different method is adopted: the parameter δc_3 is assigned to the midpoint of the event-to-station path, which is considered representative of the area influenced by wave propagation (see Figure 19). Specifically, the event origin corresponds to the nucleation point when the source geometry is known, or as the hypocenter coordinates when the source geometry is not available.

To investigate the regional variability of δc_3 , we focus on recordings where the source-to-site distance exceeds 80 km. This threshold is chosen because the effects of anelastic attenuation become significant at larger event-to-station distances (Sedaghati & Pezeshk, 2017). The dataset provides good spatial coverage across most of the national territory, with some exceptions, including the coastal areas of Tuscany and Latium, the Salento peninsula, western Sicily, and the Alpine regions of Piedmont and Trentino.

The use of a simplified 2D midpoint-based mapping approach, while practical, implies certain limitations when dealing with the inherently 3D nature of seismic wave propagation. Seismic rays associated with long source-to-site distances may travel through the crust at variable depths, depending on factors such as take-off angle and path geometry. As a result, the attenuation estimated for a given midpoint may reflect contributions from different crustal volumes, possibly with distinct physical properties.

In this context, δc_3 is interpreted as a corrective term for anelastic attenuation that also accounts for structural complexities along the wave path. Rather than representing pure anelastic attenuation, it reflects the combined influence of lateral and vertical heterogeneities in the subsurface. This interpretation allows for a consistent spatial comparison across the territory, acknowledging the integrated nature of seismic wave propagation effects.

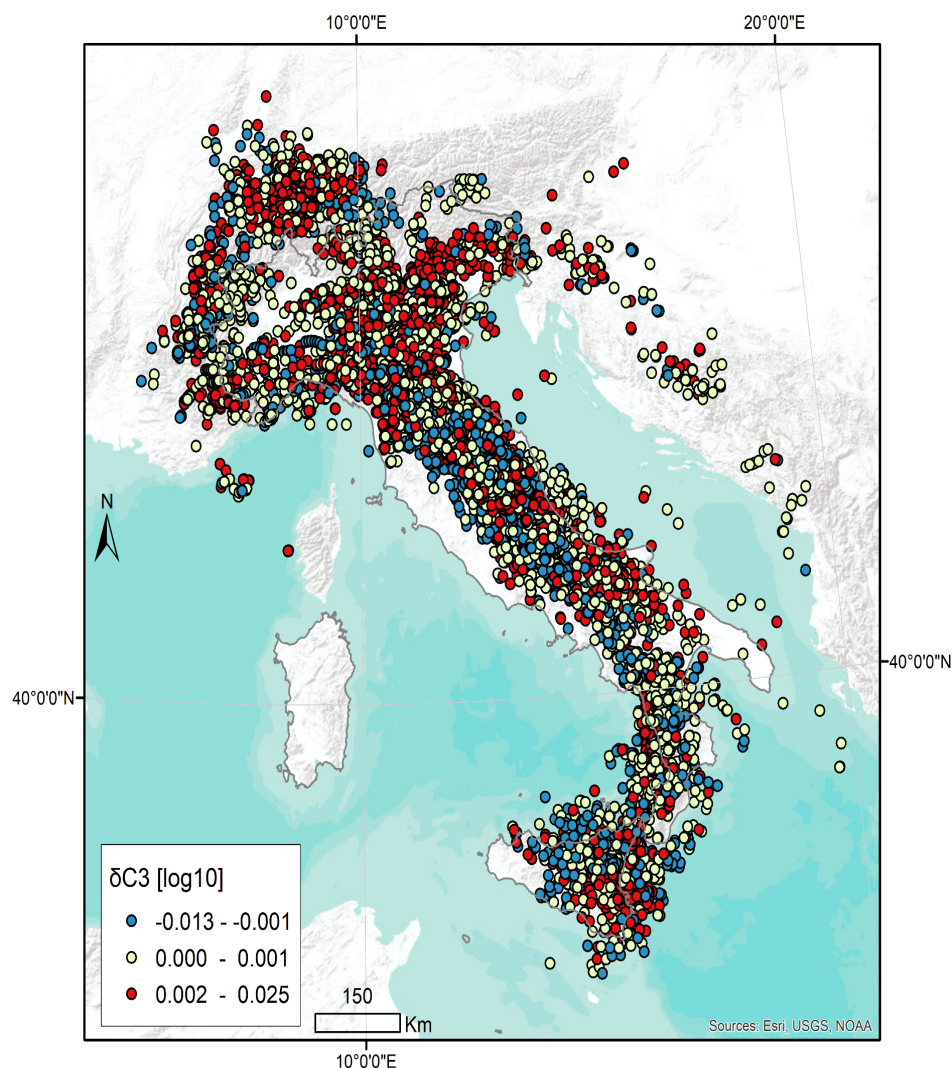


Figure 19. Spatial distribution of short period δc_3 computed from eq. [18] for distances $R_{rup} > 80$ km. Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses.

The δc_3 values are then spatially interpolated (*First Step*) using an ordinary univariate Kriging method, similar to the approach previously applied for δBe . The results of this interpolation are shown in Figure 20, representing short- and long-period ranges separately.

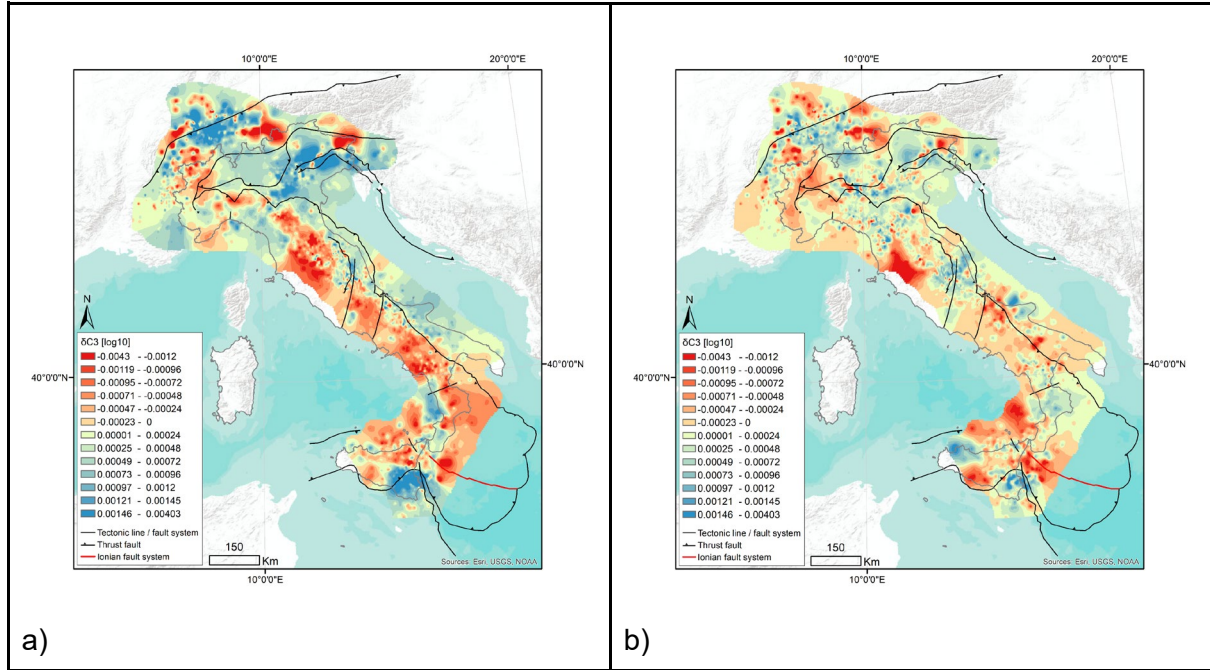


Figure 20. δc_3 of eq. [18] with $R > 80$ km: a) short and b) long periods intervals. Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses

Positive values, represented by blue areas, indicate slower attenuation compared to the reference model, while red areas signify faster attenuation. At short periods, a prominent feature is the stark contrast in attenuation between the Adriatic side (slower) and the Tyrrhenian side (faster) along the Apennine arc. In the Alps, attenuation varies between the eastern (slower) and western (faster) sides of the arc. In Sicily, the Hyblean foreland (south-eastern Sicily) distinctly shows slower attenuation compared to the rest of the island. Positive δc_3 values are also evident in the Po Valley, potentially influenced by SmS phases, that is, S waves reflected at the Moho interface. Although these phases are typically associated with energy below 5 Hz, previous studies (e.g., Bragato et al., 2011; Luzi et al., 2012) have shown that their constructive interference with direct S waves can lead to localized amplifications of short-period ground motion, including PGA. This effect may thus be detectable within the analyzed period range (0.07–0.4 s), particularly at distances exceeding 80 km where these reflections are expected to arrive.

On a national scale, at longer periods, the contribution of anelastic attenuation to distance scaling becomes almost negligible. This is partly because, in the reference Ground Motion Model (GMM), the c_3 parameter is set to zero for periods longer than 0.7 seconds.

To support our findings (*Second Step*), Figure 21 presents a national-scale map of geothermal properties, including the heat flow map of Italy from (Della Vedova et al., 1991). These parameters have also been considered on a wider European scale (Weatherill & Cotton, 2020) to classify stable cratonic regions. The observed variation in heat flow is consistent with the spatial patterns in Figure 20a) for the Apennine arc. Specifically, regions with negative δc_3 values (indicating faster attenuation) correspond to areas of higher heat flow, such as volcanic districts in Tuscany and Latium. Conversely, regions with positive δc_3 values (indicating slower attenuation) are located in the Po Plain and along the Adriatic coast, where heat flow is lower.

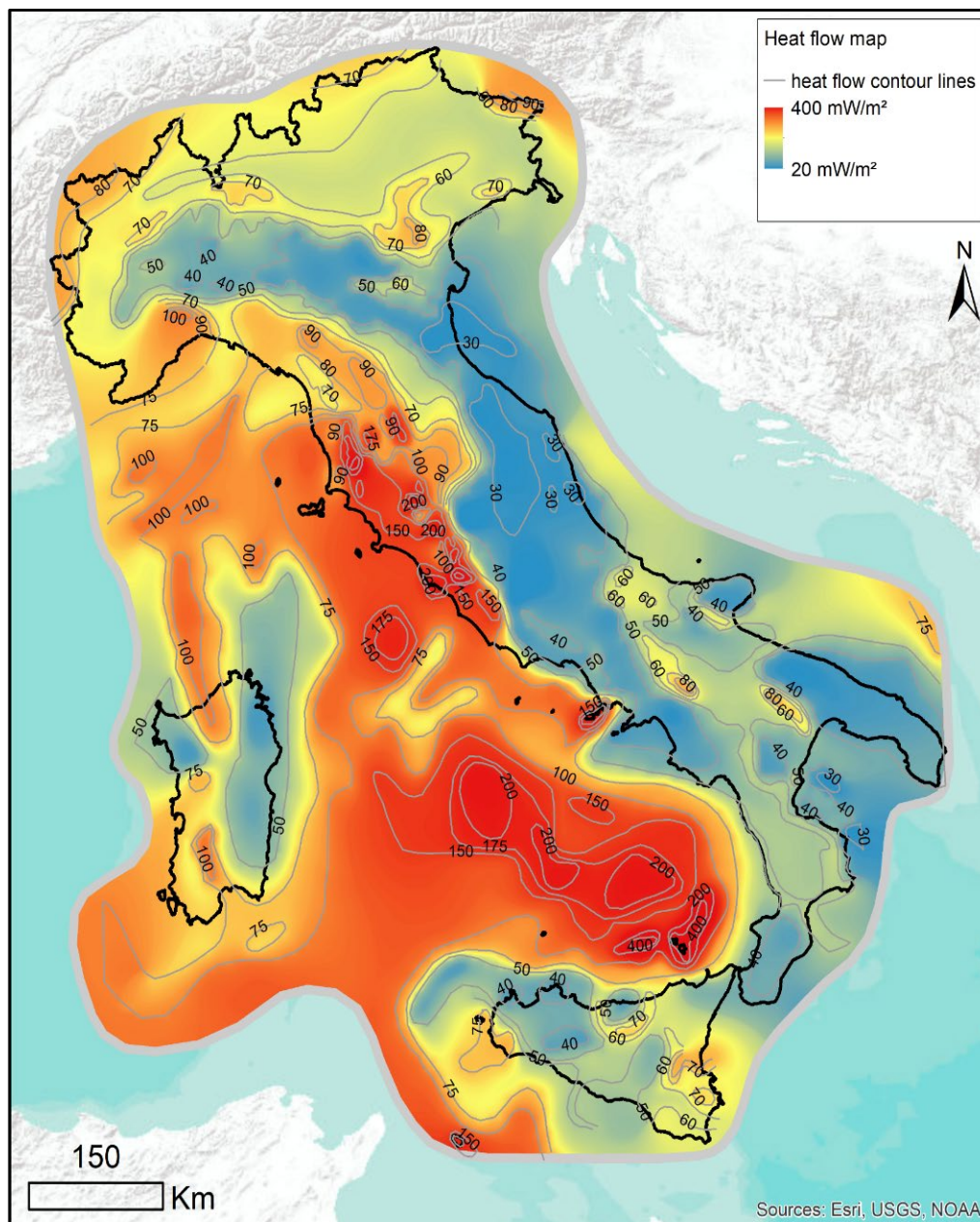


Figure 21. Heat flow map (modified from Della Vedova et al., 1991). Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses.

Further consistencies are observed with the attenuation parameter Q (quality factor) estimated by (Carletti & Gasperini, 2003), who analyzed non-parametric attenuation functions across various regions in Italy. They identified higher-than-average attenuation in the Central Apennines and Piedmont regions, while lower attenuation was observed in Eastern Sicily and Lombardy. More recently, Nardoni et al. (Nardoni et al., 2021) produced a map of coda Q_c , covering a significant portion of Italy. Their findings align with our observations, as they reported that “[...] found the main negative Q_c anomaly located in the region characterized by the shallowest Moho mapped by Manu-Marfo et al., 2019, in the Southern Tyrrhenian basin.” Moreover, “coda attenuation anomalies mapped using later onsets apparently resolve bodies associated with high absorption along the Italian peninsula, especially where geodynamic modelling infers the inflow of asthenospheric materials in the lowermost part of the crust, feeding magmatic systems. [...] The other high-attenuation anomalies across the Po Plain and the Adriatic coast are in agreement with the presence of thick sediment covers.”. Finally, the empirical patterns shown in Figure 20a align with the findings of (Carletti & Gasperini, 2003), who analyzed lateral variations in seismic attenuation across Italy using macroseismic data. They also compared their results with the heat flow map in Figure 21 and identified significant similarities.

In the *Third Step*, we use the zonation developed by (Basili et al., 2018) within the framework of the TSUMAPS-NEAM project as a reference for attenuation. This zonation was designed to support the first probabilistic tsunami hazard assessment induced by earthquakes (Seismic PTHA, S-PTHA) for the coasts of the northeast Atlantic Ocean, the Mediterranean Sea, and adjacent seas (NEAM). The model, referred to as TSUMAPS, defines a consistent zonation at the Euro-Mediterranean scale based on predominant tectonic processes. While originally developed to distinguish areas with different seismic sources, TSUMAPS was utilized by (Kotha et al., 2020) as the reference framework for adjusting the anelastic attenuation coefficient in the regionalized European GMM.

Figure 22 illustrates the subset of TSUMAPS polygons covering approximately 25 zones that include portions of the Italian territory, overlaid on the interpolated δc_3 map from Figure 20a. TSUMAPS proves to be effective in identifying areas with differing seismic wave propagation characteristics. For instance, it separates the Apennine chain from areas associated with active or inactive volcanism (zone 10). Additionally, it differentiates the Apulian and Hyblean forelands from the chain areas. Finally, it highlights the Po Plain, characterized by unique propagation features due to the significant sediment thickness and the effects of Moho reflections (Lanzano et al., 2016).

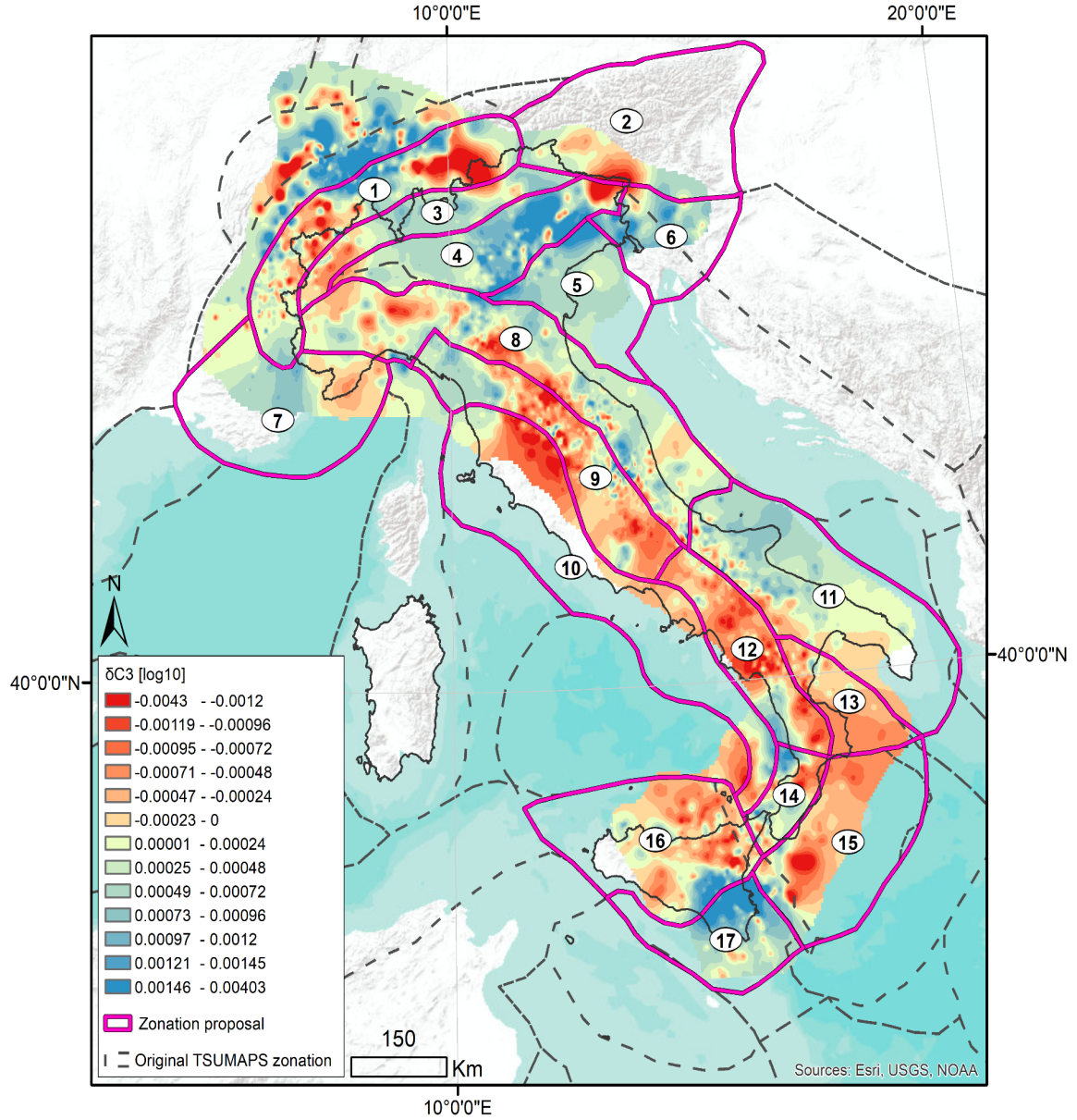


Figure 22. Zonation for regionalization of attenuation effects into GMMs. Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses

The final approach to incorporating local propagation effects into GMMs, as illustrated in Figure 22, involves a slight modification of the TSUMAPS zonation to account for the spatial trends in short-period δC_3 . The updated zones have been renumbered from 1 to 16, with only minor changes compared to TSUMAPS. These modifications include:

- The integration of the small TSUMAPS zone #98 into zone #19 due to their similar short-period δC_3 values, resulted in a new zone #4 (Figure 22).

- A redefinition of TSUMAPS zones #21 and #92 into a single zone that now includes only the northern portions of the original zones (new zone #6) (Figure 22).
- TSUMAPS zone #9, which outlines the Tindari fault, has been divided into two parts: the northern section is merged with zone #16, while the southern section forms part of zone #15 (Figure 22).
- TSUMAPS zone #13, which had very limited data, has been excluded.

Table 4 provides the final list of zones, including the median δc_3 values for the short-period interval within each zone, along with the associated standard error (SE) to quantify the epistemic uncertainty of δc_3 . Additional details, such as the number and density of records (#recs/area) for each zone, are available in Table 5.

Table 4. List of the zones for the regionalization of propagation effects. Reported data refer to the average values in the short-period range. Reproduced from Brunelli et al. (2023), published in Soil Dynamics and Earthquake Engineering, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses.

Zone	Name	Short period median δc_3 [$\cdot 10^{-4}$] [$\log_{10}$]	Standard error δc_3 [$\cdot 10^{-4}$] [$\log_{10}$]
1	North-western Alps	1.11	1.44
2	Eastern Alps	-0.18	2.99
3	Central Alps	3.24	1.81
4	Southern Alps	8.35	1.40
5	Adriatic North	8.55	1.60
6	Dinarides West	8.16	1.98
7	Ligurian-Provençal margin	3.76	1.61
8	Northern and central Apennines E	1.43	1.30
9	Northern and central Apennines W	-1.79	1.29
10	Tyrrhenian margin	-4.35	1.82
11	Adriatic South	4.97	1.48
12	Southern Apennines	-3.09	1.44
13	Calabrian Arc North	-1.92	2.06
14	Calabria	2.46	1.63
15	Calabrian inner wedge	-2.50	2.14
16	Sicily	0.21	1.48
17	Hyblean Plateau	15.03	1.99

Some zones exhibit notably positive δc_3 values, such as the Northwestern Alps (zone #1) and the Hyblean Plateau (zone #17). A gradual decrease in attenuation is observed moving from the Adriatic coast (zones #8 and #11, characterized by positive δc_3 values) towards the

Apennines (zones #9 and #12, with δc_3 values close to zero) and further to the Tyrrhenian coast (zone #10, with negative δc_3 values), consistent with the observations of (Kotha et al., 2020). The average standard error is approximately 4×10^{-5} , except the Eastern Alps (zone #2), where it is significantly higher, likely due to the limited dataset available for estimating the median δc_3 (Figure 22).

Table 5. Additional information on the zones listed in Table 4. Reproduced from Brunelli et al. (2023), published in Soil Dynamics and Earthquake Engineering, Elsevier. Reused in accordance with Elsevier's policy for inclusion in doctoral theses

Zone	Name	Numbers of records/Area [Km2]	Numbers of records
1	North-western Alps	1.44E-02	601
2	Eastern Alps	4.84E-04	29
3	Central Alps	7.89E-03	168
4	Southern Alps	2.52E-02	818
5	Adriatic North	1.08E-02	299
6	Dinarides West	4.96E-03	116
7	Ligurian-Provençal margin	5.82E-03	287
8	Northern and central Apennines E	4.02E-02	2797
9	Northern and central Apennines W	9.75E-02	3760
10	Tyrrhenian margin	2.19E-03	162
11	Adriatic South	7.22E-03	490
12	Southern Apennines	2.16E-02	592
13	Calabrian Arc North	6.26E-03	103
14	Calabria	2.52E-02	273

15	Calabrian inner wedge	1.89E-03	89
16	Sicily	9.46E-03	482
17	Hyblean Plateau	4.57E-03	115

4.5. Conclusions

The growing availability of high-quality accelerometric data has shown that ground motion models calibrated on national or global scales may fail to accurately describe small-scale seismic motion characteristics. As a result, there is a need to transition towards regionalized ground motion models, defined based on structural geology and the properties of the propagating medium (e.g., rock composition or thermal gradient), rather than purely geographical boundaries.

This approach was first introduced by (Kotha et al., 2020) to develop a European model, proposing zonations to account for differences in seismic sources and attenuation. My goal is to lay the foundation for applying a similar methodology in Italy by refining existing seismic zonations used in the Italian seismic hazard model (MPS19 (Meletti et al., 2021) and the European tsunami hazard model S-PTHA (Basili et al. 2018), to better capture regional variations in seismic source and attenuation characteristics.

To achieve this, we used the dataset of Italian seismic recordings ITACAext (Brunelli et al., 2022), which offers more comprehensive regional data compared to the ESM flatfile (Lanzano et al. 2019) utilized by (Kotha et al., 2020) for the regionalization of European ground motion models. A residual analysis was conducted to investigate the spatial variability of between-event residuals and event- and site-corrected residuals, calculated against a reference ground motion model for Italy (ITA18; Lanzano et al., 2019). Specifically, between-event residuals were mapped to assess the variability in source characteristics. Their spatial distribution was validated using independent physical parameters (e.g., stress drop), geological and structural data, leading to a new zonation for source effects through modifications to the existing seismic zonation ZS16 (Visini et al., 2022).

The same methodology was applied to zonation for ground motion attenuation, using event- and site-corrected residuals to explain variations in anelastic attenuation across the country. The findings were compared with the heat flow map and the final zonation was created by adjusting the TSUMAPS zonation. To ensure that proposed zonations are suitable for future use in ground motion models, I calculated the median values and variability of the event terms

(δB_e) and the corrections to the anelastic attenuation coefficient (δc_3) for each zone. The results align with trends observed from the spatial interpolation of residual terms. However, some zones still suffer from unreliable statistics due to limited data availability. To further validate these results, I compared the residuals with the regionalization proposed by Montaldo et al., (2005) for the MPS04 seismic hazard map (Stucchi et al., 2011). In the logic tree of MPS04, regional models collectively have a weight of 1/3, distributed across two regionalization schemes (Model A and Model B in Figure 23, each with a weight of 0.17). Each regionalization scheme involves assigning one of the available regional models for the Western Alps (REG1; Morasca et al., 2006), Eastern Alps (REG2; Malagnini et al., 2002), and Apennines (REG3; Malagnini et al., 2000) to the seismogenic zones defined by the ZS9 zonation (Meletti et al., 2008). Additionally, a specific regional model (REG4; De Natale et al., 1988; Patanè et al., 1997) is applied to volcanic zones. In areas lacking specific studies, regional models were attributed to other zones based on geological similarities and rupture mechanisms (Meletti et al., 2000). Importantly, this regionalization approach was applied to account for both source and propagation effects.

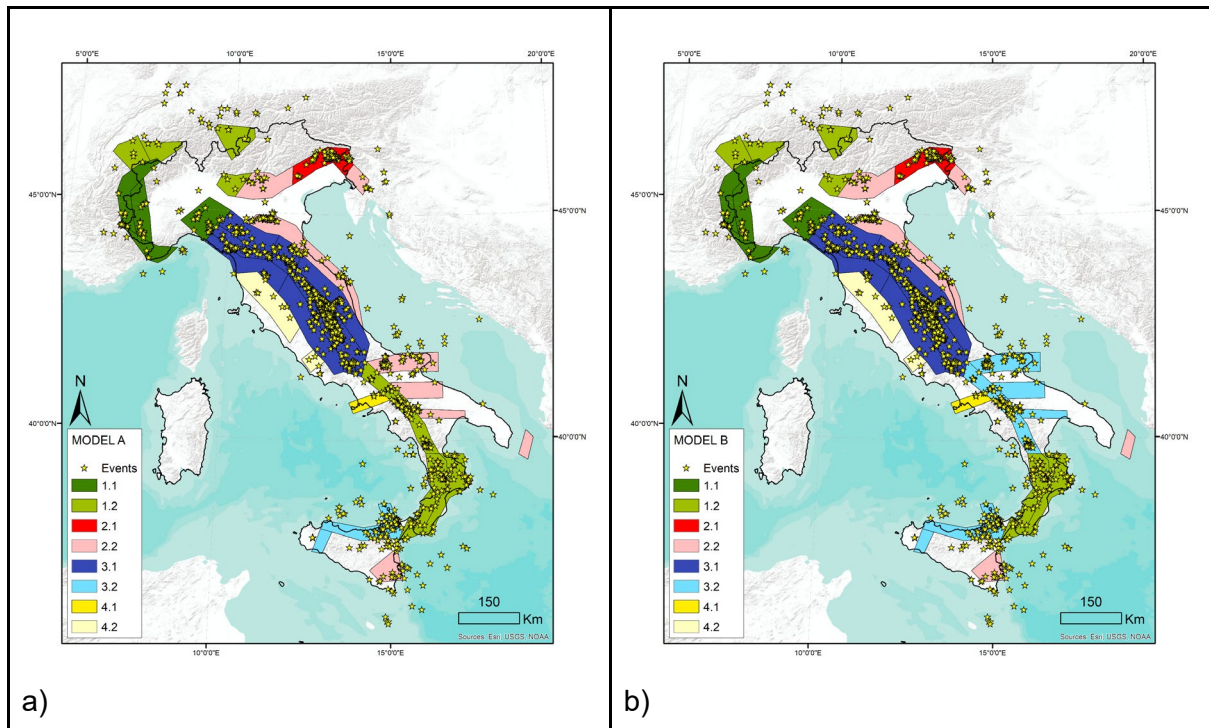


Figure 23. Model A (a) and B (b) by (Montaldo et al., 2005) used for the regionalization of GMMs in the MPS04: zones 1.1 are relative to REG1 (Morasca et al., 2006); zones 1.2 were attributed to REG1 due to geologic similarities; zone 2.1 is relative to REG2 (Malagnini et al., 2002); zones 2.2 were attributed to REG2 due to geologic similarities; zones 3.1 are relative to REG3 (Pondrelli et al., 2002); zones 3.2 were attributed to REG3 due to geologic similarities; zone 4.1 is relative to REG4 (De Natale et al., 1988); zones 4.2 were attributed to REG4 due to geologic similarities. Reproduced from Brunelli et al. (2023), published in *Soil Dynamics and Earthquake Engineering*, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses.

Using this framework, I calculated the log10 median values of short-period δB_e and δc_3 for each zone defined by the MPS04 regionalization (Table 6).

Table 6. Median short period δB_e and δc_3 ($\pm$ standard error) in log10 units for the zones considered in the regionalization of ground motion models proposed by (Montaldo et al., 2005) for the PSHA in Italy, MPS04 (Stucchi et al., 2011). Reproduced from Brunelli et al. (2023), published in Soil Dynamics and Earthquake Engineering, Elsevier. Reused by Elsevier's policy for inclusion in doctoral theses.

Model	Area	Specific ZS9 zones		REG. A		REG. B	
		δB_e [$\times 10^{-1}$]	δc_3 [$\times 10^{-4}$]	δB_e [$\times 10^{-1}$]	δc_3 [$\times 10^{-4}$]	δB_e [$\times 10^{-1}$]	δc_3 [$\times 10^{-4}$]
REG1	Western Alps	1.1 ± 0.3	-1.1 ± 0.8	0.6 ± 0.2	-1.6 ± 0.5	0.9 ± 0.2	-0.7 ± 0.6
REG2	Eastern Alps	1.2 ± 0.3	6.3 ± 2.8	0.4 ± 0.1	6.4 ± 0.8	0.6 ± 0.2	7.0 ± 1.1
REG3	Northern and Central Apennines	-0.5 ± 0.1	-0.9 ± 0.2	-0.5 ± 0.06	-1.0 ± 0.2	-0.5 ± 0.06	-0.9 ± 0.2
REG4	Volcanic (Phlegrean Fields)	–	-3.2 ± 15	–	-8.4 ± 1.8	–	-8.4 ± 1.8

These values were compared with seismological parameters such as stress drop and attenuation functions for each area. I report the residual statistics both for the zones where the REG1, REG2, and REG3 models were originally developed and for the extended zones aggregated by geological similarity (Figure 23). No δB_e estimates were made for volcanic areas, as the ITA18 model does not apply to volcanic earthquakes.

The higher δB_e values observed in the Western and Eastern Alps align with the higher average stress drop values estimated for REG1 and REG2, which are approximately 50–60 MPa. Conversely, for the Apennines (REG3), a lower stress drop of 20 MPa was assumed, consistent with the lower δB_e estimates. The differences between median event term estimates are smaller in the aggregated models, suggesting that a detailed source zonation, such as the one proposed in this study, is more effective in capturing local variations.

The median δc_3 values for the REG2 zones are notably positive, indicating slower attenuation compared to the REG1 and REG3 zones. This is consistent with the non-parametric attenuation functions derived by various studies for the Friuli region and the REG2 aggregated zones, including the Po Valley, the coastal area of Marche, and Southeastern Sicily (Akinci et al., 2022; Castro et al., 1999). In contrast, the δc_3 estimates for REG4 are negative, reflecting the fastest attenuation, which aligns with the propagation characteristics of volcanic regions where high frequencies diminish rapidly with distance (Lanzano & Luzi, 2020).

The proposed zonations are intended to develop an empirical regionalized model for Italy, which can be applied in various contexts, such as improving the Italian Shakemap (Michellini et al., 2020) or creating fully non-ergodic probabilistic seismic hazard maps (Abrahamson et al., 2019).

However, some additional considerations will need to be addressed for future applications:

- When using the zonation for modelling source effects, it is essential to consider fault scaling laws and the variability observed between small and large earthquakes. For instance, a magnitude-dependent correction and/or intra-zone heteroscedastic variability could be incorporated. In cases with insufficient data for strong earthquakes, it should be noted that these adjustments (and associated variability) may only apply to smaller events.
- Although the zonation for attenuation effectively captures the spatial patterns of δc_3 , some inconsistencies remain. These issues stem from the fact that the original TSUMAPS zonation was designed for source effects rather than propagation effects. While the effectiveness of this approach may be due to a trade-off between source and propagation effects at large scales, it would be more appropriate in the future to adopt a strategy better aligned with the parameters being modeled. An alternative approach could involve a zoneless correction, leveraging the dense spatial sampling of the δc_3 term, as shown in Figure 20.
- The incorporation of focal depth as an explicit explanatory variable, through the modulation of the pseudo-depth parameter, as suggested by Kotha et al. (2020), offers a means to adjust the extent of the area impacted by the strongest shaking around the epicenter based on the focal depth. However, implementing this parameter requires extreme caution, as it depends on the availability of accurate estimates of hypocentral depth (e.g., Chiaraluce et al., 2022).

The regionalization of site effects is the subject of Chapter 5.

5. Evaluation of Site Effects for Shaking Scenarios

Site effects play a crucial role in seismic hazard assessment, as local soil conditions can significantly amplify or attenuate ground motion during an earthquake (Kramer & Stewart, 2024). The mechanical properties of the subsurface influence wave propagation, leading to variations in seismic response across different locations.

The shear wave velocity in the top 30 meters of the subsurface (V_{s30}) is a widely used parameter in seismic regulations (NTC, 2008; CEN, 2004) to characterize seismic soil amplification. V_{s30} is calculated as a weighted average of soil layers using the formula:

$$V_{s30} = \frac{30}{\sum \frac{h_i}{V_{s_i}}} \quad [19]$$

where h_i represents the thickness of layer i , and V_{s_i} is the shear wave velocity in that layer.

The Italian (NTC, 2008) and European (CEN, 2004) standards define soil classes based on V_{s30} thresholds (categories from A to E). The new Italian standards (NTC, 2018) introduced an alternative classification system based on V_{seq} , a parameter that considers shear wave velocities down to either 30 meters or the engineering bedrock depth (H800), whichever is shallower:

- If the engineering bedrock (defined as $V_s = 800$ m/s) lies deeper than 30 meters, $V_{seq} = V_{s30}$.
- If the seismic bedrock is shallower than 30 meters, then V_{seq} is computed as the average shear wave velocity down to that depth.

These classifications are used to establish site-dependent design spectra, ensuring a standardized approach in seismic risk assessment (FEMA 273, 1997; CEN, 2004).

Beyond its role in seismic codes, V_{s30} plays a fundamental role in Ground Motion Models (GMMs) and is incorporated in three main ways:

1. Directly within the functional form of the GMM (e.g. Lanzano et al., 2019).
2. By classifying soil based on V_{s30} ranges (e.g. Akkar & Bommer, 2010).

Obtaining a large number of measurements that are continuously and evenly distributed across a territory is challenging; therefore, proxies used for modeling and spatially distributing Vs30 are widely adopted.

The first global Vs30 maps using topographic slope as a proxy were introduced by Wald and Allen (2007) and Allen and Wald (2009). Their studies established different predictive coefficients for tectonically active regions, which exhibit significant topographic variation, and for stable continental areas, where terrain changes are more subdued. These coefficients were later compared with site condition maps derived from geological data and direct Vs30 measurements.

In Italy, Michelini et al. (2008) developed a national-scale Vs30 map using the 1:100,000 geological map as a reference. Their approach classified the soil into five categories (A to E) based on Vs30 reference values established by the EC8 code (CEN, 2004), ranging from 1000 m/s for class A to 150 m/s for class D. Later, Di Capua et al. (2016) proposed an updated Vs30 map for Italy, relying on a revised 1:100,000 lithological map. This study aimed to group geological formations into litho-seismic classes based on their lithological characteristics to identify areas with homogeneous seismic responses.

Forte et al. (2019) integrated large-scale lithological maps with site-specific measurements to enhance Vs30 predictions. A further advancement was made by Mori et al. (2020), who developed a large-scale Vs30 map for Italy. Based on the geomorphological classes proposed by Iwahashi et al. (2018), they identified a lognormal linear regression between Vs30, slope, and elevation for each class, using a stepwise method. This approach incorporated a vast dataset consisting of approximately 35,000 borehole logs and 11,300 Vs profiles from the national seismic microzonation database (DPC, 2019). A further advancement was made by Mori et al. (2020), who produced a large-scale Vs30 map for Italy using the global morphological classes defined by Iwahashi et al. (2018) and incorporating an extensive dataset from the national seismic microzonation database (DPC, 2019). This approach proved to be more effective than previous models due to its stronger correlation between Vs30 values, morphological classes, and subsurface lithologies. For each morphological class, Mori et al. (2020) identified specific lognormal linear relationships between Vs30 and topographic proxies. Their method was comparable to those used in Japan (Matsuoka & Wakamatsu, 2006), Greece (Stewart et al., 2014), Central and Eastern North America (Parker et al., 2017), and the Pacific Northwest (Ahdi et al., 2017). Any inconsistencies with the Iwahashi et al. (2018) classification were resolved by incorporating lithological data from Amanti et al. (2008). Finally, statistical tests were conducted to validate the linear regression models, leading to the development of the final hybrid Vs30 map.

More recently, Gironelli et al. (2024) developed a Vs30 map for the Marche region (Central Italy) by applying a dual-proxy approach based on lithological classification and topographic slope. The model was trained using a subset of the national seismic microzonation database, including the same Vs profiles used in this thesis, and represents a first attempt to characterize Vs30 spatial variability at regional scale through regression analysis with random effects.

In this chapter, the work on the regionalization of site effects that led to the development of a new Vs30 map will be described.

5.1. Quality control, extension of profiles that do not reach 30m and development of the parametric table

This section describes the preparation of the Vs30 dataset, based on shear-wave velocity profiles collected from scientific studies, research projects, and contributions by freelancers. Many Vs profiles presented issues as they were manually entered by freelancers involved in seismic microzonation efforts. These issues were spotted using an automated procedure and subsequently corrected in a semi-automatic manner. Below is a summary of the quality control steps performed, which led to the exclusion or correction of profiles:

Excluded profiles

- Profiles lacking geographic coordinates.
- Profiles with depths <10 m that do not reach a Vs of 800 m/s.

Profiles Corrected (Semi-Automatic Procedure)

- Assigned the velocity of the second layer when the first layer's velocity was missing (NaN);
- Checked for layers with zero thickness, i.e., where the top and bottom depths were identical.
- Identified and corrected Vs profiles containing duplicate layers.
- Checked for duplicate Vs profiles in the initial databases.
- Reviewed profiles with anomalous Vs values, such as layers with Vs < 75 m/s and depth > 2 m, or Vs < 50 m/s within the first two meters.

At the end of this step, the dataset consists of approximately 15,000 profiles, of which around 3,000 do not reach a depth of 30 meters.

To recover these profiles, Vs30 was predicted from Vs10, Vs15, Vs20, and Vs25, which represent the average shear-wave velocity down to depths of 10, 15, 20, and 25 meters, respectively, similarly to what was done by Boore et al. 2011.

Two distinct approaches were employed: the first considers Vs30 as depending on Vsn, whereas the second incorporated a random effect based on the lithological classification proposed by Bucci et al. 2022. This allowed for robust and comparable Vs30 predictions and allows one to predict Vs30 when the lithology is unknown.

Profiles where Vs10 exceeds Vs30, suggesting a velocity inversion, were excluded. After this data-cleaning phase, the number of usable profiles was reduced from approximately 12,000 to 11,124, providing a robust dataset for performing regressions. Then, two functional forms were adopted to estimate Vs30 from Vs10, Vs15, Vs20, Vs25.

In particular:

1. I implemented a second-order polynomial regression calibrated for profiles lacking lithological information (see Table 9 for the regression coefficients).
2. An alternative version is calibrated with the random effect on the lithological classes proposed by Bucci et al. 2022 (see Table 7 for the regression coefficients).

The equation [20] describes the quadratic regression with random effects on the lithological classes proposed by Bucci et al. (2022), relating Vs30 to Vs10, 15, 20, and 25. If the lithological information is missing, the random effect term is set to zero ($u_{GEO} = 0$).

$$\log_{10}(\mathbf{VS30}) = \beta_0 + \beta_1 \cdot \log_{10}(\mathbf{VS}_{\text{depth}}) + \beta_2 \cdot \log_{10}(\mathbf{VS}_{\text{depth}})^2 + u_{GEO} + \epsilon \quad [20]$$

Where:

- β_0 , β_1 and β_2 are the regression coefficients that capture the relationship between $\log_{10}(\mathbf{VS30})$ and $\log_{10}(\mathbf{VS}_{\text{depth}})$. Vsdepth represents the Vs averaged up to depths of 10, 15, 20, and 25 meters.
- u_{GEO} represents the random effect associated with the lithological classes proposed by Bucci et al. (2022).
- ϵ is the residual error term of the model.

Table 7. Three coefficients of the quadratic regression β_0 , β_1 , β_2 for Vs10, Vs15, Vs20, and Vs25, respectively, for the regression model that includes the random effect on the lithological class proposed by Bucci et al. 2022. The standard deviation of residuals (σ_{RES}) is also shown for each model.

Vsx	β_0	β_1	β_2	Standard deviation of residuals σ_{RES}
Vs10	-0.6581	1.7455	-0.1686	0.0559
Vs15	-0.4084	1.4476	-0.0976	0.0381
Vs20	-0.265	1.2721	-0.057	0.0236
Vs25	-0.1293	1.1259	-0.0254	0.0113

Table 8. Random effect coefficients for each lithological class proposed by Bucci et al. 2022 for Vs10, Vs15, Vs20, and Vs25

Lithology	Fullname	VS10	VS15	VS20	VS25
B	Beache deposits	-0.0329	-0.0123	-0.0053	-0.0016
Ad	Anthropogenic deposits	-0.0285	-0.0161	-0.0092	-0.0041
Al	Alluvial and marine deposits	-0.0262	-0.0136	-0.0061	-0.002
Ucr	Unconsolidated clastic rocks	-0.0067	-0.0028	-0.0009	-0.0003
Lb	Lavas and basalts	-0.0065	-0.0025	-0.0007	-0.0002
Ccr	Consolidated clastic rocks	-0.006	-0.0037	-0.0024	-0.0012
E	Evaporite	-0.0015	-0.0009	-0.0003	-0.0001
Gd	Glacial drift	-0.0008	-0.0004	-0.0001	-0.0001
Li	Lakes and ice	-0.0002	-0.0002	-0.0002	-0.0004
Pr	Pyroclastic rocks	0.0011	0.001	0.0006	0.0003
Nsr	Non – schistose metamorphic rocks	0.0046	0.0012	0.0007	-0.0005
Ir	Intrusive rocks	0.004	0.0024	0.002	0.0009
CM	Chaotic–mélange	0.0095	0.006	0.0032	0.0012
Cr	Carbonate rocks	0.0113	0.0052	0.0031	0.0013
Sr	Schistose metamorphi crocks	0.0117	0.0063	0.0033	0.0017
M	Marls	0.0188	0.011	0.0056	0.0021
Ssr	Siliciclastic sedimentary rocks	0.0191	0.0117	0.0064	0.0028
SM	Mixed sedimentary rocks	0.029	0.0183	0.0117	0.0048

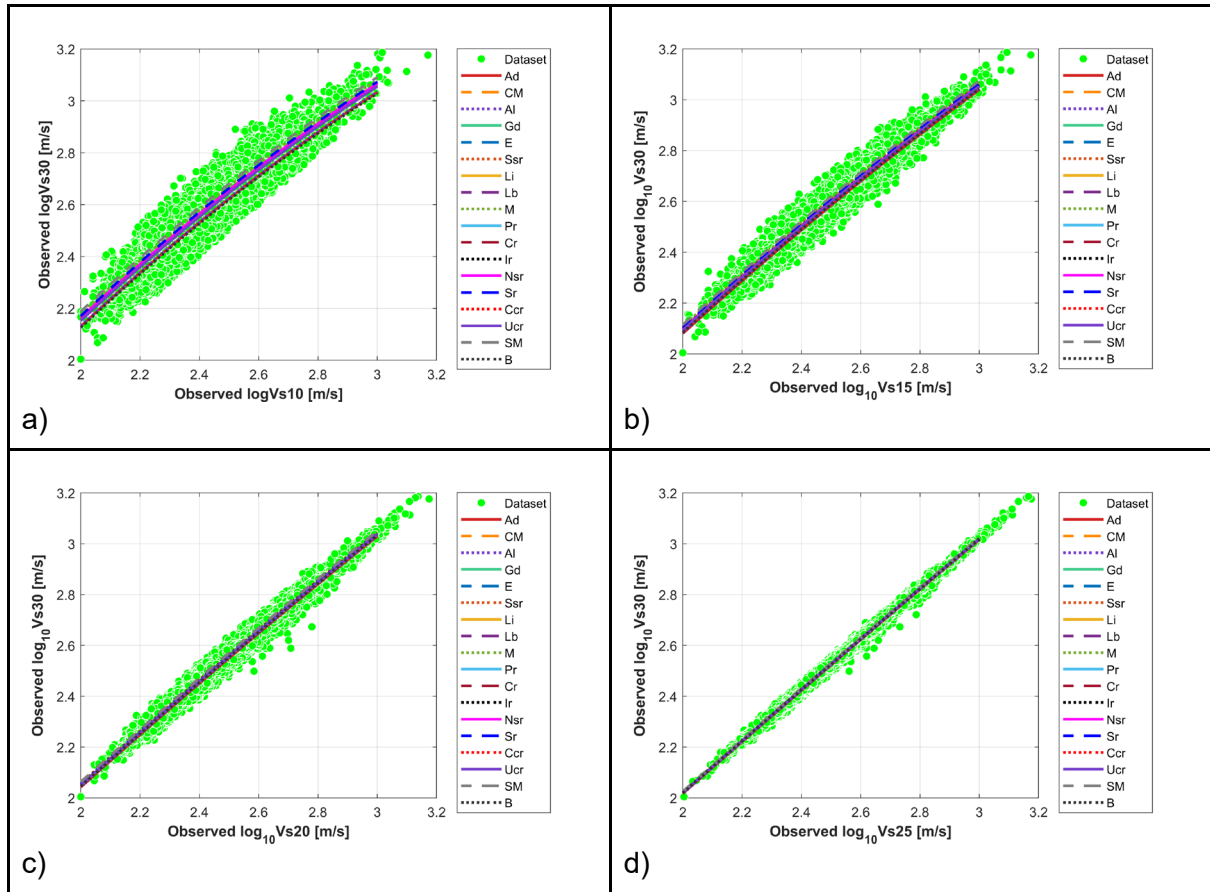


Figure 24. Relationship between observed and predicted $\log_{10}(V_s30)$ values using a quadratic regression model with random effects based on Bucci's lithological classes for different depths ($\log V_s10$ (a), $\log V_s15$ (b), $\log V_s20$ (c), and $\log V_s25$ (d), respectively)

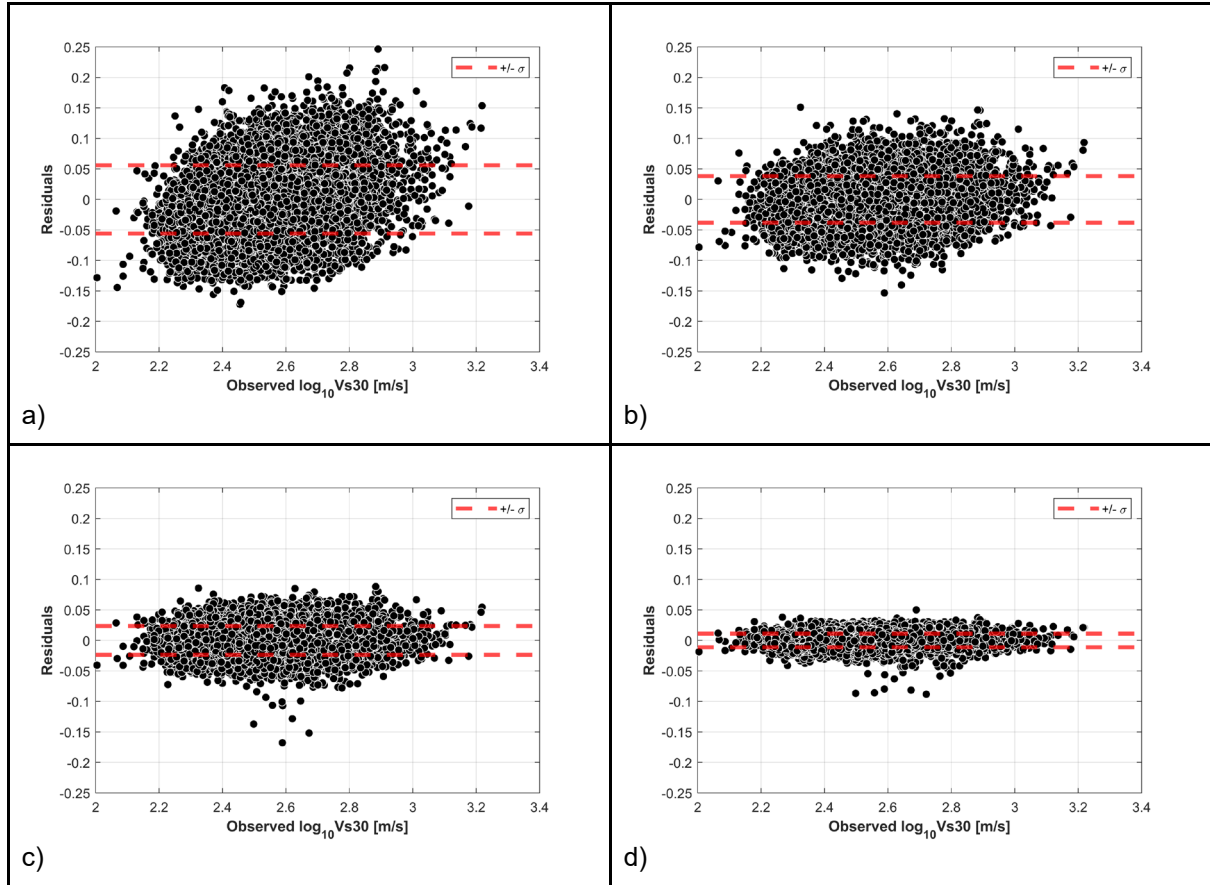


Figure 25. Residuals (observed – predicted $\log_{10}Vs_{30}$), obtained from quadratic regression models with random effects based on Bucci's lithological classes, plotted against observed $\log_{10}Vs_{30}$. The models are based on: (a) $\log Vs_{10}$, (b) $\log Vs_{15}$, (c) $\log Vs_{20}$, and (d) $\log Vs_{25}$

In Figure 24 the green points, representing the dataset, follow a clear quadratic trend, confirming the model's ability to capture the relationship between $\log_{10}(Vs_{30})$ and the predictor variables at different depths. The colored lines represent the various lithological classes, and they closely follow the trend of the observed data. Although these lines may appear visually similar due to overlapping trends, their vertical shifts reflect lithology-specific random effects estimated by the model. For example, predicting Vs_{30} from the same Vs_{10} value but including the random effect for either beach deposits (B) or mixed sedimentary rocks (SM) results in approximately a 15% difference in Vs_{30} . This percentage difference remains constant in logarithmic scale, meaning that the relative variation is preserved regardless of the specific Vs_{10} value used. Such a difference is not negligible and highlights the importance of including random effects associated with lithological classes, as they allow the model to adjust to the specific characteristics of each class, improving overall performance. However, the influence of the random effects associated with the lithological classes diminishes when predicting Vs_{30} using Vs_{20} and Vs_{25} , as the contribution of lithological variability becomes less relevant at greater depths (see Figure 24).

Table 8 highlights a clear distinction between deposits and rocks in terms of random effect: deposits generally show negative values, likely reflecting the incoherent or poorly consolidated nature of these materials, which tend to have lower Vs30 values compared to the global model average. In contrast, rocks tend to exhibit positive random effect values, with the largest values observed for consolidated units such as Marlstone, Siliciclastic sedimentary rocks, and Mixed sedimentary rocks.

Figure 25 shows the residual plots (observed – predicted $\log_{10}Vs_{30}$) as a function of the observed $\log_{10}Vs_{30}$ for four regression models, each using the average shear-wave velocity over increasing depths (Vs10, Vs15, Vs20, and Vs25) as a predictor. In panel (a), a clear trend is visible: the model tends to underestimate Vs30 at higher values of observed Vs30 and overestimate it at lower values. This behavior indicates that Vs10 is not an optimal predictor of Vs30, as also reflected by the larger spread of the residuals. In panel (b), the trend is less evident, but the dispersion remains relatively high. In contrast, panels (c) and (d) show residuals more symmetrically distributed around zero, with no significant systematic patterns. The spread of the residuals decreases slightly as depth increases, suggesting that the model's accuracy improves progressively from Vs10 to Vs25, as expected (see σ_{RES} in Table 9). For Vs25 the residuals are tightly clustered around zero, demonstrating that the model explains the variability of the observed data more effectively at this depth.

Overall, the inclusion of random effects for lithological classes plays a key role in refining the model's performance, particularly for lower mean values of Vs, such as Vs10 and Vs15. This analysis confirms the robustness and reliability of the model in predicting Vs30 values across different lithologies and mean Vs values at various depths.

Table 9. Shows the three coefficients of the quadratic regression β_0 , β_1 , β_2 for Vs10, Vs15, Vs20, and Vs25, respectively. The standard deviation of residuals (σ_{RES}) is also shown for each model

Vsx	β_0	β_1	β_2	Standard deviation of residuals σ_{RES}
Vs10	-0.916	1.9104	-0.1945	0.0583
Vs15	-0.5456	1.5322	-0.1103	0.0392
Vs20	-0.3294	1.3104	-0.0624	0.0241
Vs25	-0.1516	1.1383	-0.027	0.0115

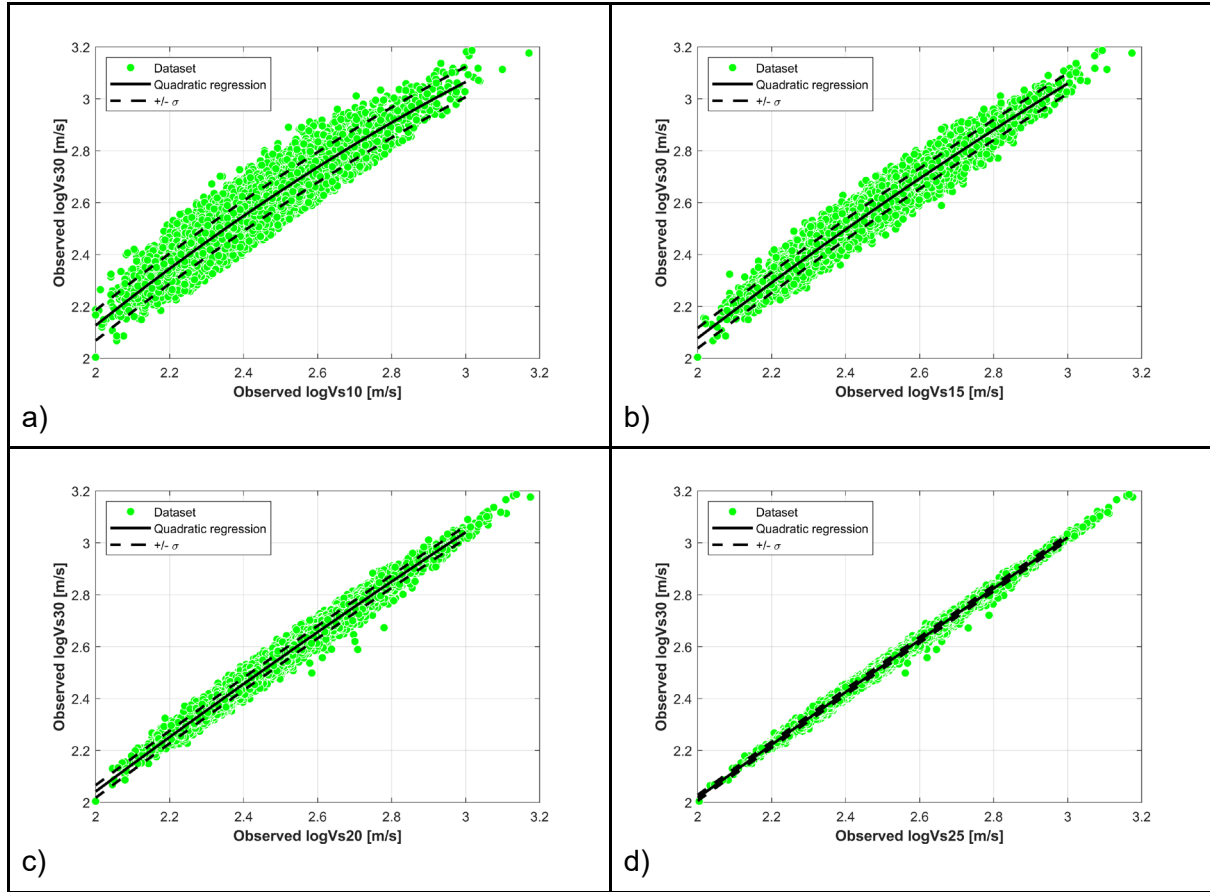


Figure 26. Relationship between observed and predicted $\log_{10}(V_{s30})$ values using a quadratic regression model for different depths ($\log V_{s10}$ (a), $\log V_{s15}$ (b), $\log V_{s20}$ (c), and $\log V_{s25}$ (d), respectively). Without including the random effect for lithological classes

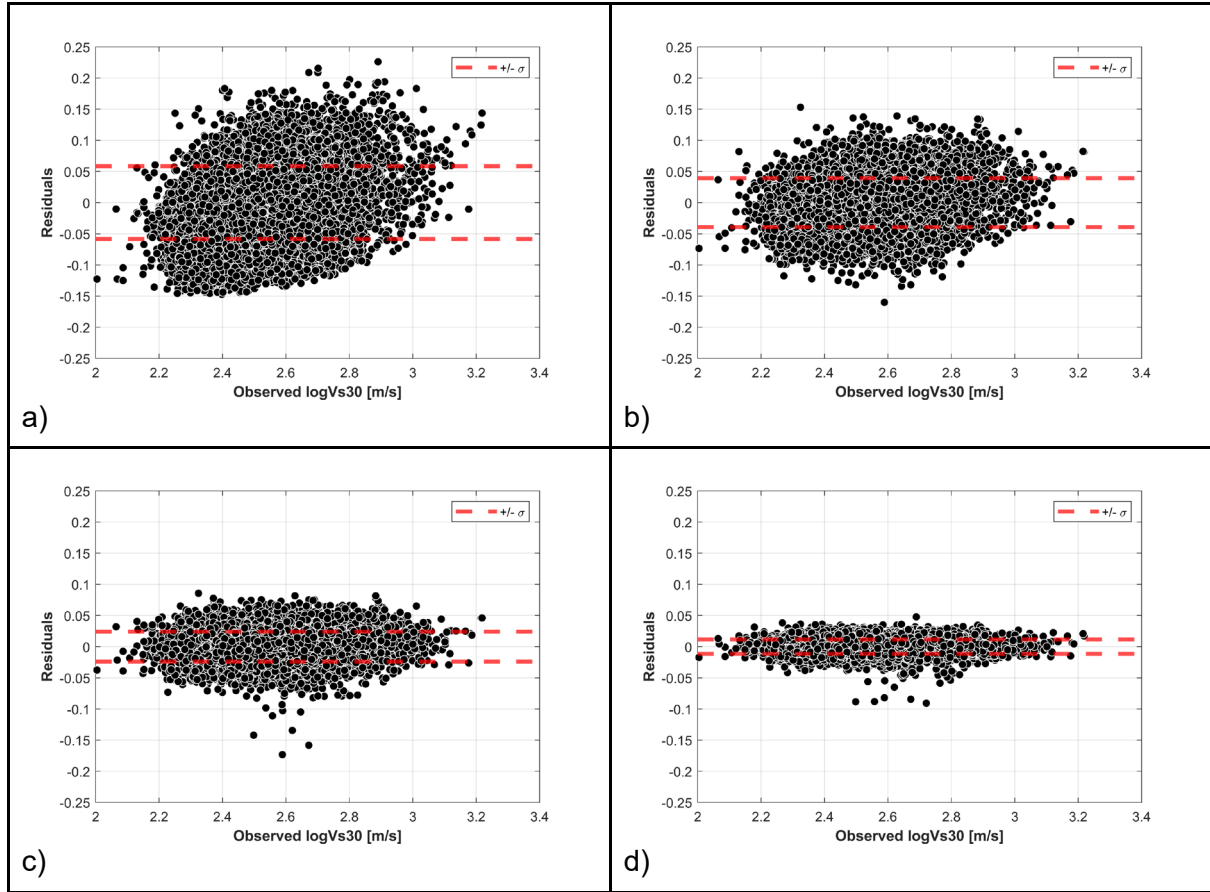


Figure 27. Residuals (observed – predicted $\log_{10}Vs_{30}$), obtained from quadratic regression models plotted against observed $\log_{10}Vs_{30}$. The models are based on: (a) $\log Vs_{10}$, (b) $\log Vs_{15}$, (c) $\log Vs_{20}$, and (d) $\log Vs_{25}$. Without including the random effect for lithological classes

In Figure 26, the solid black lines represent the fitted quadratic regression models, while the dashed lines indicate the ± 1 standard deviation (σ) confidence intervals (see Table 9). The regression curves closely follow the data distribution, effectively capturing the underlying quadratic trend.

Figure 27 shows residual plots, illustrating the differences between observed and predicted $\log_{10}(Vs_{30})$ values plotted against the observed $\log_{10}(Vs_{30})$. The red dashed lines indicate $\pm\sigma$, as reported in Table 9.

The residuals for the model using $\log Vs_{10}$ and $\log Vs_{15}$ show a broader spread, indicating a weaker fit. By contrast, the residuals for the models using $\log Vs_{20}$ and $\log Vs_{25}$ are more tightly concentrated around zero, suggesting an improved fit at these values, as expected.

Subsequently, for about 3,000 profiles that did not reach a depth of 30 meters, it was possible to calculate Vs_{30} using the equation [20] and the coefficient of Tables 7,8 and 9.

A parametric table was then compiled for all 15,333 profiles, consisting of the following fields:

- ID_SURV: Unique identifier for each point in the database.
- X: Longitude in UTM 33N – WGS84 coordinates.
- Y: Latitude in UTM 33N – WGS84 coordinates.
- Survey_type: Type of geophysical survey.
- sta_code: Station code for profiles extracted from the ITACA database.
- Vs10: Average shear wave velocity over the first 10 m.
- Vs15: Average shear wave velocity over the first 15 m.
- Vs20: Average shear wave velocity over the first 20 m.
- Vs30: Average shear wave velocity over the first 30 m.
- Vs_eq: Average shear wave velocity from the surface to the engineering bedrock depth (if the engineering bedrock is at a depth ≥ 30 m, Vs_eq equals Vs30).
- H800: Depth to engineering bedrock, if present.
- NTC2018cat: Subsoil category according to the 2018 Italian seismic code (NTC, 2018).
- EC8cat: Subsoil category according to Eurocode 8 (CEN, 2004).
- H800inv: A value of "1" indicates a significant Vs inversion in the profile. This occurs when the nth layer exceeds 800 m/s and has a thickness of at least 3 m, but the n+1th layer has a Vs < 800 m/s. In such cases, the top of the nth layer was not considered a reliable indicator of seismic bedrock depth.
- Pred= if the measurement of Vs30 will be predicted value =1 otherwise = 0.

5.2. Spatial data processing (DEM analysis)

Subsequently, to investigate the correlation between Vs30 and various site proxies, both raster and vector datasets were produced and compiled.

Regarding the generated raster datasets, the TINITALY Digital Elevation Model (DEM) with a spatial resolution of 10×10 m (Tarquini et al., 2023) was resampled to four different spatial resolutions (150 m, 250 m, 500 m, and 1000 m).

After resampling, the raster edges were extended in ArcGIS Pro to ensure complete coverage of the Italian borders. This was achieved by assigning each NoData cell at the edges the minimum value of its neighboring cells using a 3×3 moving window, through the Focal Statistics tool with the 'Minimum' option and iterative processing in the Raster Calculator.

For each resolution, the following topographic indices were computed using ArcGIS Pro 3.4.2 (developed by ESRI): Slope, Vector Ruggedness Measure (VRM), Terrain Ruggedness Index (TRI), Topographic Position Index (TPI), Profile Curvature, Plan Curvature, Standard Curvature, and Geomorphological Classes. Meanwhile, the Roughness index was calculated using the GDAL tool available in QGIS.

In addition, several raster and vector datasets were collected from the literature, including:

- The ISPRA lithological map at a 1:100,000 scale (Amanti et al. 2008), see Chapter 2;
- The lithological map produced by Bucci et al. (2022) at a 1:100,000 scale, see Chapter 2;
- The global geological bedrock depth map proposed by Pelletier et al. (2016), with a spatial resolution of 1000 × 1000 m, see Chapter 2.

The following sections provide a detailed description of the topographic proxies computed using ArcGIS Pro, a desktop GIS software developed by Esri.

The **Slope** parameter quantifies the rate of elevation change for each cell within a digital elevation model. As the first derivative of the DEM, it describes the steepness of the terrain at any given point (Figure 28).



Figure 28. Map of slopes expressed in degrees and calculated with a spatial resolution of 150m. The minimum value is 0°, the maximum value is 86°, and the average value is 9.4°

Roughness (Figure 29), as defined by Hobson, R.D. (1972), is calculated as the maximum elevation difference between a focal cell and its surrounding eight neighboring cells. High roughness values indicate a highly irregular or rugged terrain with significant elevation differences, such as mountainous or rocky areas. Low roughness values represent smoother, more homogeneous surfaces, such as plains or gently rolling landscapes. Roughness is an essential topographic metric used in geomorphology, hydrology, and ecological studies. This metric provides a simple yet effective way to quantify surface variability and is widely used in GIS-based terrain analysis.



Figure 29. Roughness with a spatial resolution of 150m. The raster has a minimum value of 0, a maximum value of 3480, and an average value of 74

Vector Ruggedness Measure (VRM) (Figure 30), introduced by Sappington et al. (2007), quantifies terrain ruggedness by analyzing the three-dimensional variability in the orientation of surface normals, which are vectors perpendicular to the terrain. For each grid cell in a digital elevation model (DEM), the slope and aspect are used to decompose the surface normal into its x, y, and z components. Within a defined neighborhood (e.g., a 3×3 moving window), these

components are summed to calculate a resultant vector, which represents the average direction of the normals in that area. The magnitude of this resultant vector, standardized by dividing it by the number of cells in the neighborhood, indicates the degree of alignment of the normals. Subtracting this standardized magnitude from 1 produces a dimensionless ruggedness index that ranges from 0 (flat surface) to 1 (highly rugged surface).

Larger neighborhoods capture broader-scale terrain features, while smaller ones focus on fine-scale details. This flexibility makes VRM a useful tool in many fields, including ecology (to identify habitats influenced by topographic complexity), hydrology (to study erosion and water flow patterns), and geomorphology (to map landforms).



Figure 30. VRM map calculated with a spatial resolution of 150m. The raster has a minimum value of 0, a maximum value of 0.4, and an average value of 0.006

The **Terrain Ruggedness Index (TRI)** (Figure 31), introduced by Riley et al. (1999), is a measure of terrain roughness that quantifies topographic heterogeneity in a DEM. TRI is widely used in ecological modeling, hydrology, and geomorphology to assess terrain complexity and habitat suitability. Higher TRI values indicate rugged or highly variable terrain, while lower values represent flat or smooth surfaces.

Mathematically, TRI is defined by the equation:

$$TRI = \sqrt{\sum (x_i - x_0)^2} \quad [21]$$

where:

- x_0 is the elevation of the central cell,
- x_i represents the elevation of each of the eight surrounding cells,

The TRI value for the central cell is obtained by taking the square root of the sum of the squared differences between the elevation of the central cell and its neighboring cells. This calculation is repeated for each cell in the DEM, generating a roughness map that reflects local topographic variations. TRI provides a quantitative indication of elevation changes within a defined area, making it a valuable index for many geospatial applications.



Figure 31. TRI map calculated with a spatial resolution of 150m. The raster has a minimum value of 0, a maximum value of 6768m, and an average value of 78m

The **Topographic Position Index (TPI)** (Figure 32), proposed by Weiss (2001), quantifies the relative elevation of a focal cell compared to the average elevation of its surrounding neighborhood. The TPI value can be positive, negative, or close to zero, reflecting different topographic features. A positive TPI indicates that the focal cell is higher than its surroundings, characterizing ridges, peaks, or elevated landforms. Conversely, a negative TPI suggests that the cell is lower than its surroundings, identifying depressions, valleys, or concave features. Values near zero imply that the cell has a similar elevation to the average of its neighborhood, corresponding to relatively flat or uniform slopes. Since TPI is based on elevation differences, it is particularly sensitive to the choice of neighborhood size and resolution, which influence whether small-scale or broad landforms are captured in the analysis.



Figure 32. TPI map calculated with a spatial resolution of 150m. The raster has a minimum value of -1691m, a maximum value of 1265m, and an average value of 2.52m

Different areas of a surface may exhibit concave or convex shapes, which can be determined by analyzing the curvature values. Profile curvature impacts the speed variations of flow, playing a key role in erosion and deposition processes. Meanwhile, planform curvature affects the way flow converges or diverges across the terrain. The Curvature tool in ArcGIS Pro fits a fourth-degree polynomial to the input surface for each cell using a 3×3 window. Once the polynomial is established, the second derivatives are computed at the central cell, taking into account the elevation values of the surrounding cells (Moore et al., 1991). This means that the curvature value at each cell is determined by how the elevation changes relative to its neighboring cells.

Three curvature outputs are available:

- **profile curvature**, which follows the direction of the steepest slope and affects flow acceleration and deceleration;
- **plan curvature**, which is calculated perpendicular to the direction of the maximum slope and affects the convergence and divergence of surface flow.
- **standard curvature** that combines both the profile and planform curvatures.

The values in the curvature raster are measured in units equivalent to one hundredth (1/100) of the elevation unit, which in the case of an input raster with elevation in meters corresponds to 0.01 meters⁻¹. In areas with moderate elevation changes, such as rolling hills, typical curvature values range between -0.5 and 0.5, while in rugged, mountainous terrain, they can vary from -4 to 4. However, in certain cases, these values may extend beyond these ranges depending on the characteristics of the raster surface.

Profile curvature (Figure 33), represents the curvature along the direction of the maximum slope in a terrain model. This measure describes how the shape of the terrain affects the acceleration or deceleration of surface water flow. A positive value indicates that the surface is concave upward in the slope direction, which accelerates flow and enhances erosion. Conversely, a negative value signifies a convex upward surface, causing flow to slow down and promoting deposition. A curvature value of zero means the surface is neither concave nor convex but remains flat along the slope direction.



Figure 33. Profile curvature was calculated with a spatial resolution of 150m. The raster has a minimum value of $-15.8 \times 10^{-2} \text{ m}^{-1}$, a maximum value of $15.2 \times 10^{-2} \text{ m}^{-1}$, an average value of $0.005 \times 10^{-2} \text{ m}^{-1}$ and a standard deviation of $0.13 \times 10^{-2} \text{ m}^{-1}$

Planform curvature (commonly known as **plan curvature**) (Figure 34) represents the curvature measured perpendicular to the direction of the maximum slope. Positive values indicate areas where surface flow diverges, while negative values correspond to zones of converging flow. When plan curvature is positive, the surface is convex in the direction perpendicular to the slope, whereas a negative value signifies concavity in the same orientation. A curvature value of zero indicates a flat surface in the analyzed direction.



Figure 34. Plan curvature was calculated with a spatial resolution of 150m. The raster has a minimum value of $-14.8 \times 10^{-2} m^{-1}$, a maximum value of $5.3 \times 10^{-2} m^{-1}$, an average value of $0.005 \times 10^{-2} m^{-1}$ and a standard deviation of $0.15 \times 10^{-2} m^{-1}$

Standard curvature (Figure 35) combines both profile and planform curvatures, offering a more detailed understanding of how flow interacts with the surface. Analyzing these two components together allows for a better interpretation of water movement across the terrain and, consequently, the dynamics of soil deposition.



Figure 35. Standard curvature was calculated with a spatial resolution of 150m. The raster has a minimum value of $-30 \times 10^{-2} \text{ m}^{-1}$, a maximum value of $15.3 \times 10^{-2} \text{ m}^{-1}$, an average value of $-8.4 \times 10^{-16} \text{ m}^{-1}$ and a standard deviation of $0.25 \times 10^{-2} \text{ m}^{-1}$

The Italian territory has been categorized into **geomorphological classes** following the approach proposed by Jasiewicz and Stepinski (2013) (Figure 36). These geomorphological classes provide a representation of the landscape based on variations in elevation within the local neighborhood of a target cell. In total, 498 distinct geomorphon patterns can be grouped into ten primary landform types: flat, peak, ridge, shoulder, spur, slope, hollow, footslope, valley, and pit (Jasiewicz & Stepinski, 2013). A threshold of 1° has been set, below which a cell is classified as flat. Additionally, a 1500 meter radius has been defined around the target cell to identify the geomorphon pattern in the terrain analysis.

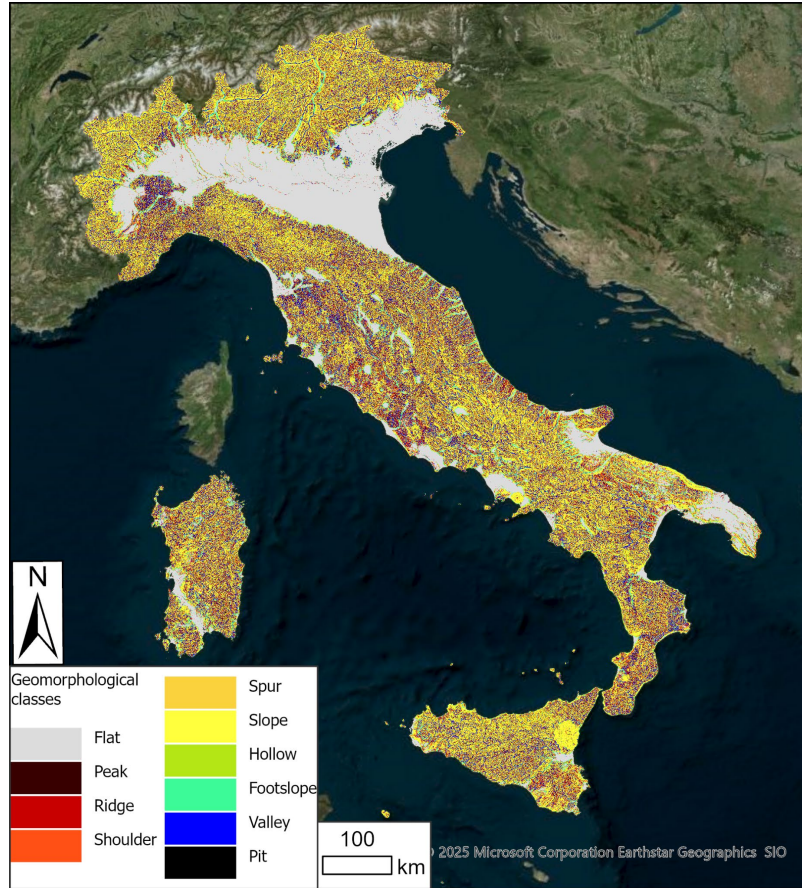


Figure 36. Geomorphological classes are calculated with a spatial resolution of 150m

5.3. Correlation between Vs30 and site proxies

The next step was the intersection of the point file containing Vs30 values with the shapefiles of lithological classifications at the Italian scale, described in the previous chapter 2, and with the raster datasets of topographic indices described in the previous chapter and in chapter 2.

Next, we explored the correlation between Vs30 and the various numerical and categorical proxies collected and processed using a cross-correlation matrix. In particular, the correlation between $\log_{10}(\text{Vs30})$ and the numerical proxies is systematically stronger when the proxies are expressed in $\log_{10}$ scale. This trend is summarized in Table 10, which reports the correlation indices obtained using both the non-log and the log transformed proxies. For example, the correlation index between $\log_{10}(\text{Vs30})$ and TRI at 150 m increases from 0.48 (non-log scale) to 0.61 ($\log_{10}$ scale), and for VRM at 150 m it increases even more markedly,

from 0.30 to 0.58. Similar improvements are observed for elevation, roughness, and slope-based proxies.

This systematic increase in correlation supports the use of log transformed proxies in the regression models, as they provide a stronger linear relationship with $\log_{10}(\text{Vs30})$.

Table 10. Shows the correlation indices between $\log_{10}(\text{Vs30})$ and the numerical proxies described in Chapter 2, calculated for each proxy in its natural scale and in $\log_{10}$ scale. The third column reports the difference between the two correlation values ($\log_{10} - \text{natural}$).

Proxy	Correlation index of the proxy with $\log_{10}(\text{Vs30})$	Correlation index of the logarithm of the proxy with $\log_{10}(\text{Vs30})$
TRI 150m	0.48	0.61
TRI 250m	0.48	0.62
TRI 500m	0.48	0.62
TRI 1000m	0.47	0.62
VRM 150m	0.3	0.58
VRM 250m	0.31	0.58
VRM 500m	0.32	0.59
VRM 1000m	0.29	0.59
Elevation 150m	0.49	0.61
Elevation 250m	0.49	0.61
Elevation 500m	0.49	0.61
Elevation 1000m	0.49	0.61
Roughness 150m	0.47	0.61
Roughness 250m	0.48	0.62
Roughness 500m	0.48	0.62
Roughness 1000m	0.47	0.63
Tan slope 150m	0.44	0.6
Tan slope 250m	0.43	0.6
Tan slope 500m	0.42	0.6
Tan slope 1000m	0.41	0.6

The following indices were used to measure correlations between different types of variables:

- For numerical variables, we used Pearson's correlation coefficient (Pearson, 1895), which measures the linear correlation between two numerical variables. It ranges from -1 to 1, where: 1 indicates a perfect positive correlation, -1 indicates a perfect negative correlation, 0 indicates no linear correlation.

- For categorical variables, we applied Cramér's V (Cramér, 1999), based on the chi-square test. It ranges from 0 to 1, where: 0 indicates no association between the two categorical variables and 1 indicates a perfectly deterministic relationship between the categories.
- For numerical and categorical variables, we used the eta-square index (η^2) (Fisher, 1925), which measures the proportion of variance in a numerical variable explained by a categorical variable. It also ranges from 0 to 1, where: 0 means the categorical variable does not explain any variation in the numerical variable, and 1 means the categorical variable fully explains the variance in the numerical variable.

Figure 37, is a reduced version of the cross-correlation matrix, where numerical proxies with a correlation index below a threshold of ± 0.15 have been removed. In the cross-correlation matrix shown in Figure 37, the lithological classifications by ISPRA and Bucci et al., 2022, referred to as "Code Lito Bucci" and "Code Lito ISPRA", respectively, and the geomorphological classes ("Geom class"), are categorical variables. The topographic indices TRI, VRM, Slope, Roughness, and Elevation are numerical variables. Additionally, slope was converted to its tangent value following Wald and Allen (2007), and the correlation analysis was performed at four different spatial resolutions

The analysis shows that $\log_{10}(\text{Vs}_{30})$ is negatively correlated with sediment thickness, (e.g. as sediment thickness increases, Vs_{30} values tend to decrease). Moreover, $\log_{10}(\text{Vs}_{30})$ is strongly correlated with the $\log_{10}$ of TRI, Elevation, Roughness, and tangent of slope across all four spatial resolutions analyzed.

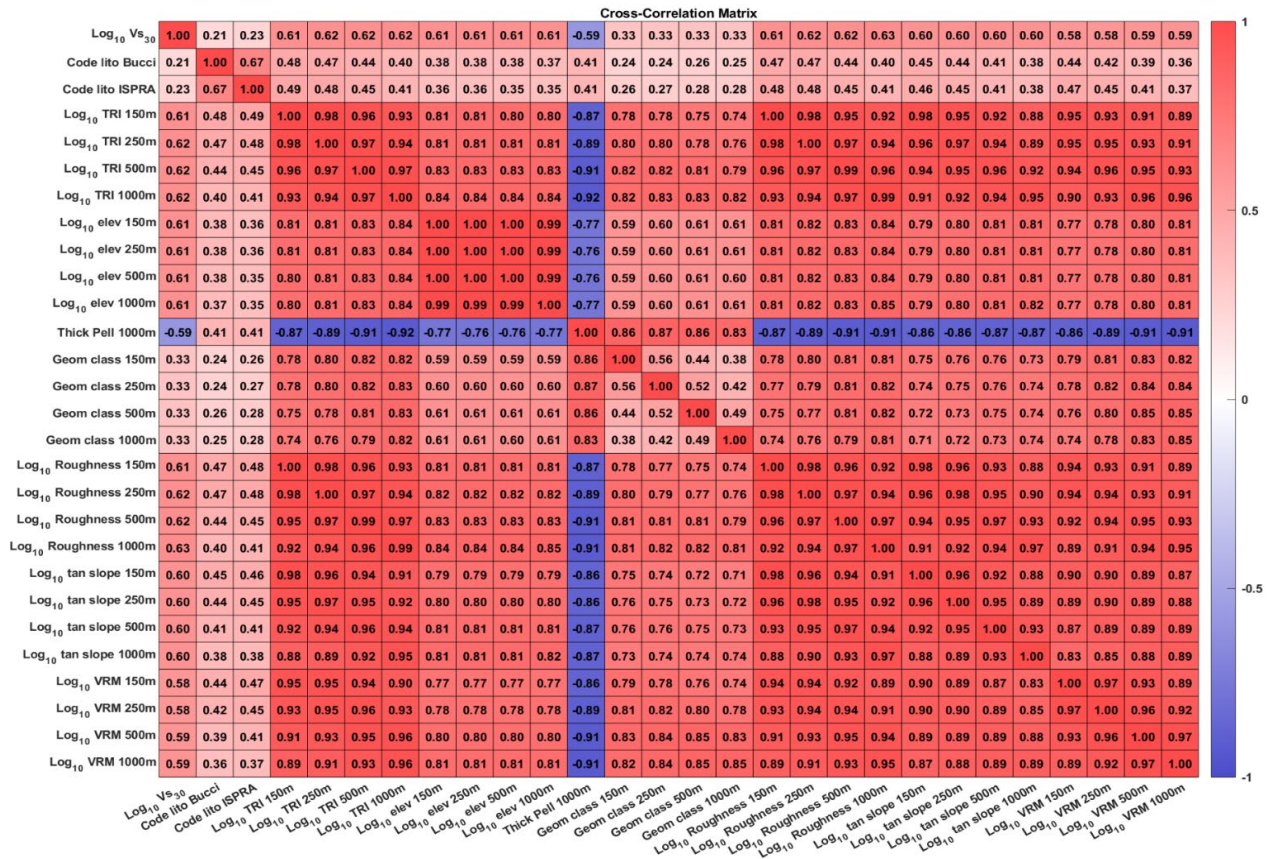


Figure 37. Cross-correlation matrix between $\log_{10}(Vs_{30})$ and the numerical and categorical proxies obtained from the literature and those specifically produced for this study

Scatter plots in Figures 38, 39, 40, and 41 relate the Vs_{30} and four topographic proxies with the highest correlation across the four spatial resolutions; the mean and standard deviation per each Vs_{30} bin is also shown.

Figures 38 to 41 display: the slope (Figure 38), the TRI (Figure 39), the roughness (Figure 40), and elevation (Figure 41). All four topographic proxies exhibit a similar trend with Vs_{30} , but only the slope shows a stable trend independently from the spatial resolution and can explain the largest values of Vs_{30} . Therefore, $\log_{10}(\tan(\text{slope}))$ with a spatial resolution of 150 m was selected as the best numerical proxy.

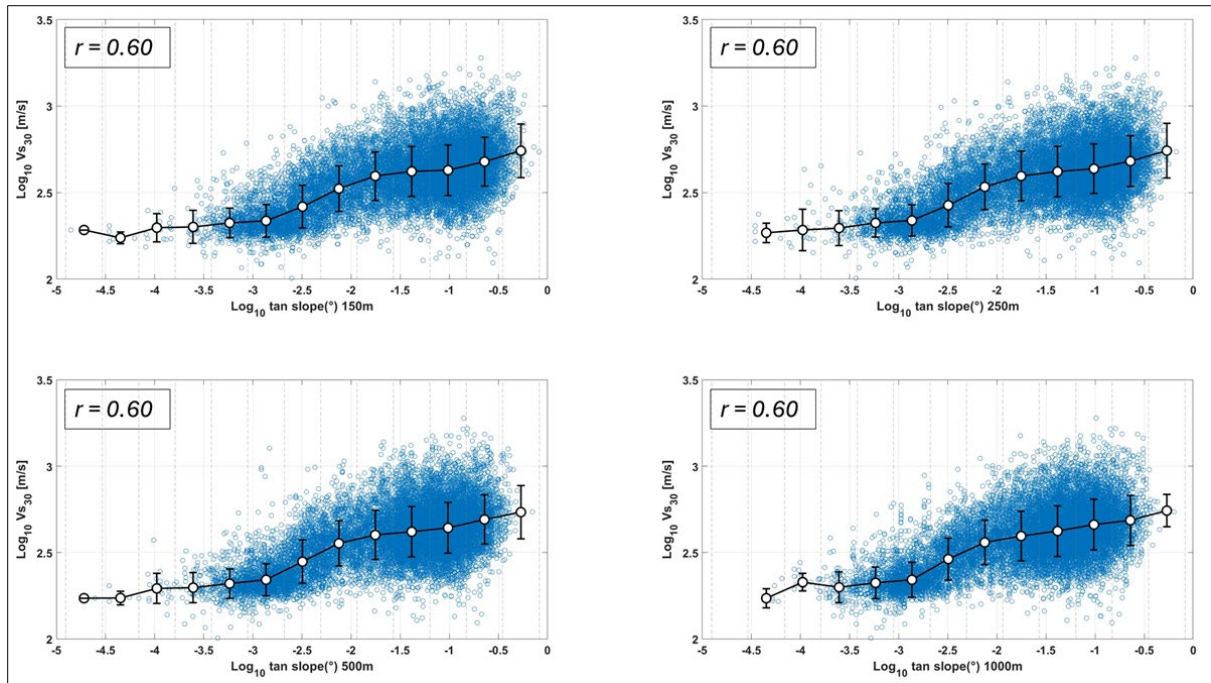


Figure 38. Scatter plot between $\log_{10} V_{s30}$ [m/s] and $\log_{10} (\tan(\text{slope}^\circ))$. The white markers and error bars indicate the mean and standard deviation within each bin (grey vertical dashed lines)

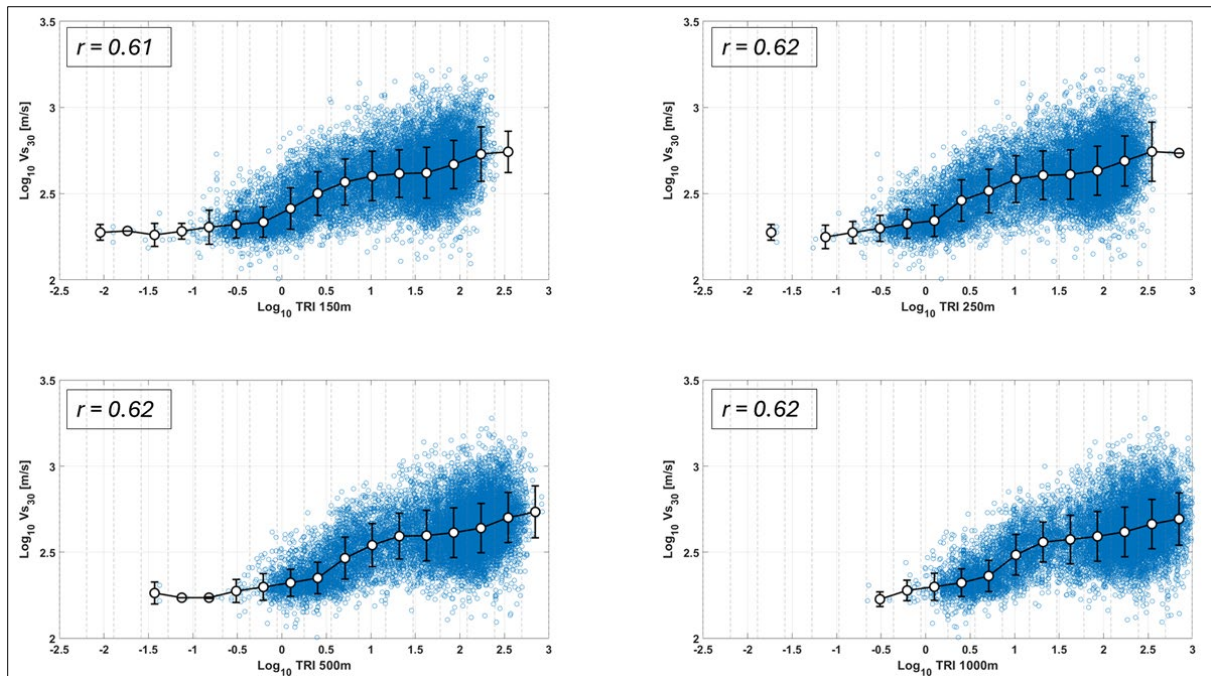


Figure 39. Scatter plot between $\log_{10} V_{s30}$ [m/s] and $\log_{10} (\text{TRI})$. The white markers and error bars indicate the mean and standard deviation within each bin (grey vertical dashed lines)

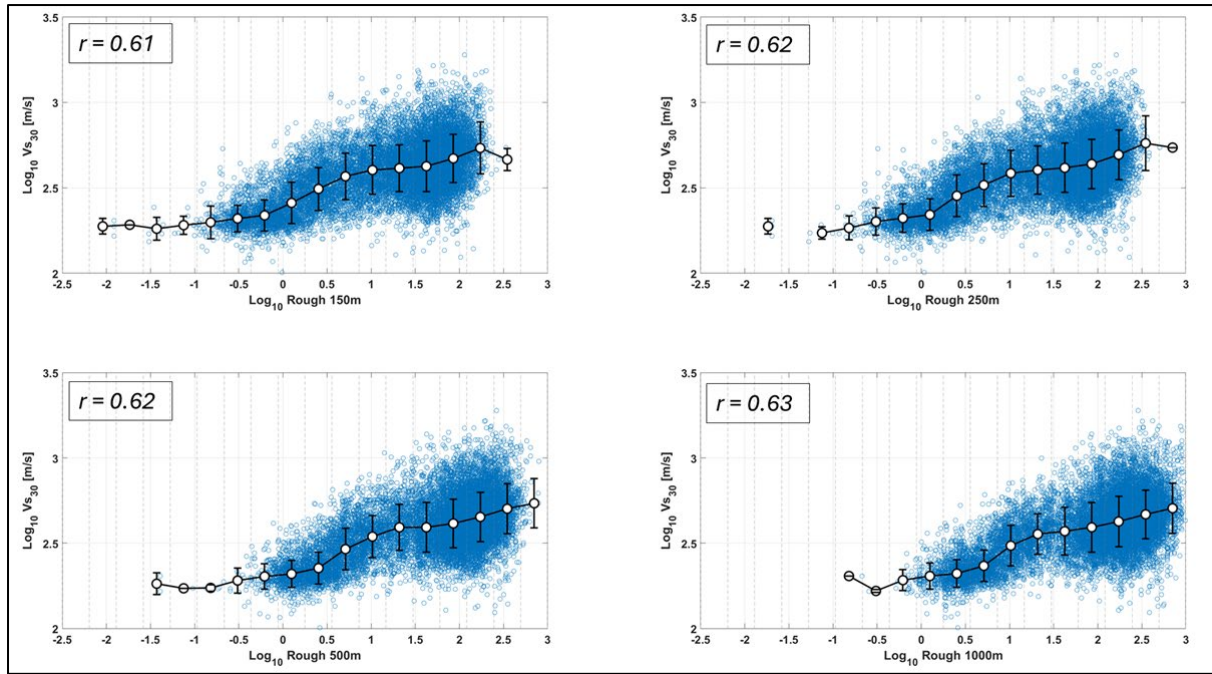


Figure 40. Scatter plot between $\log_{10}Vs_{30}$ [m/s] and $\log_{10}(\text{Roughness})$. The white markers and error bars indicate the mean and standard deviation within each bin (grey vertical dashed lines)

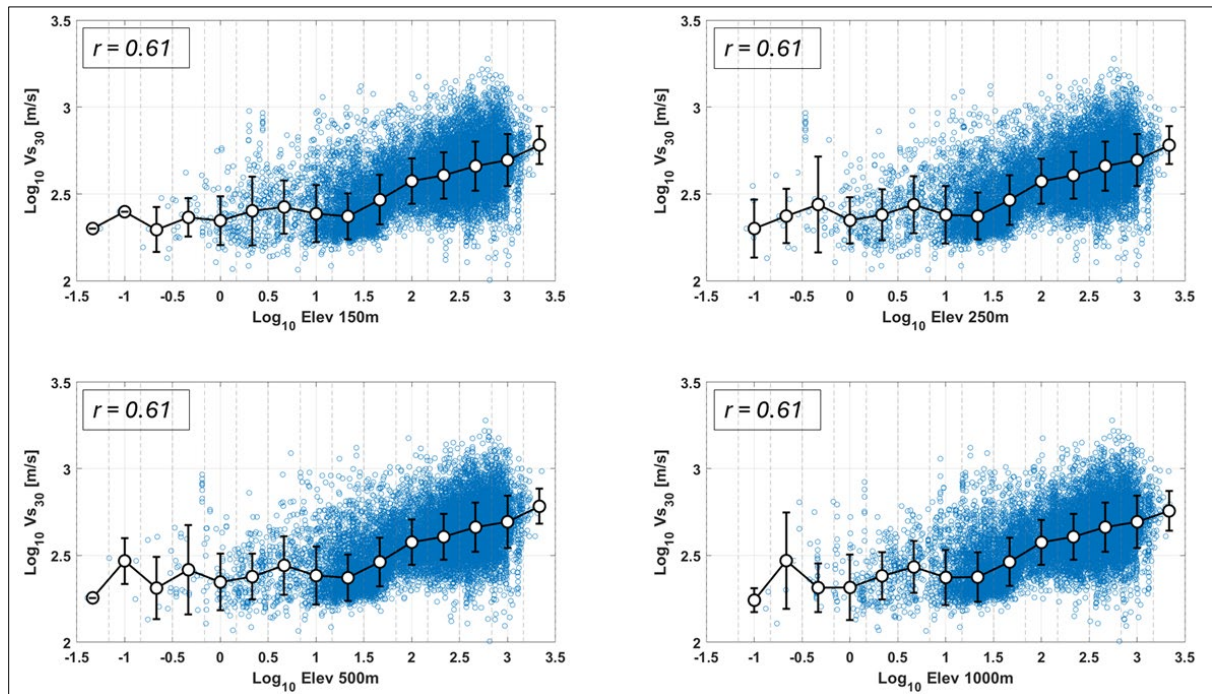


Figure 41. Scatter plot between $\log_{10}Vs_{30}$ [m/s] and $\log_{10}(\text{Elevation})$. The white markers and error bars indicate the mean and standard deviation within each bin (grey vertical dashed lines)

The deposit thickness was not considered appropriate as a proxy for Vs30 although the correlation coefficient ($r = -0.59$) is not negligible (Figure 42) because of the poor resolution in depth. Rather than providing a continuous variation, this proxy tends to separate the data into two distinct clusters: thickness within the first 5–10 meters or thickness equal to or greater than 50 meters, limiting its effectiveness as a predictive variable.

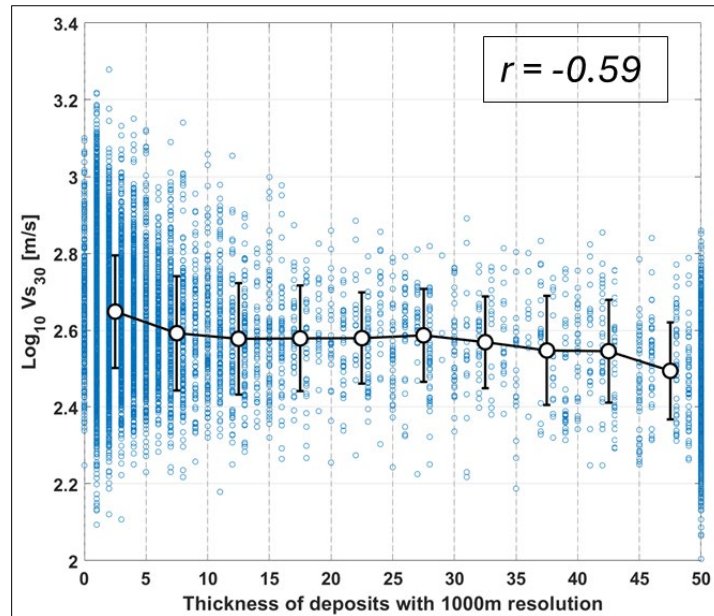
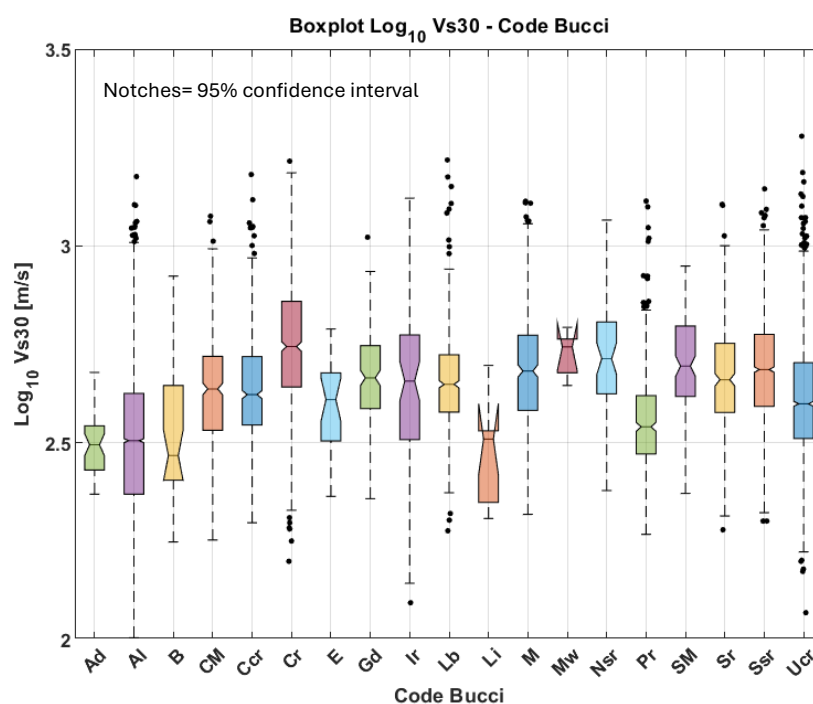


Figure 42. Scatter plot of $\log_{10}(Vs_{30})$ [m/s] versus the deposit thickness in meters derived from Pelletier et al. (2016). The white markers and error bars indicate the mean and standard deviation within each bin (grey vertical dashed lines)

To better assess how the lithological variables capture the variability in terms of Vs30, one-way ANOVA (Analysis of Variance) tests were performed between Vs30 and the lithological classes proposed by ISPRA (see chapter 2) and Bucci et al., 2022 (see chapter 2).

Figure 43 presents the results of a one-way ANOVA analysis to evaluate the variability of $\log_{10}(Vs_{30})$ across lithological classes as defined by Bucci et al. (2022).



- **Prob>F less than 0.05**
- **F=229.53** Between-group variance greater than within-group variance

Lithologies	Code	n° of Vs profiles	means (log10Vs30[m/s])
Anthropogenic deposits	Ad	40	2.50
Alluvial and marine deposits	Al	6944	2.51
Beach deposits	B	34	2.52
Chaotic – mélange	CM	408	2.63
Consolidated clastic rocks	Ccr	701	2.64
Carbonate rocks	Cr	849	2.74
Evaporite	E	25	2.60
Glacial drift	Gd	168	2.67
Intrusive rocks	Ir	70	2.65
Lavas and basalts	Lb	631	2.65
Lakes and Ice	Li	10	2.48
Marlstone	M	366	2.69
Mass wasting material	Mw	5	2.72
Non-schistose metam. rocks	Nsr	51	2.71
Pyroclastic rocks	Pr	424	2.56
Mixed sedimentary rocks	SM	117	2.70
Schistose metamorphic rocks	Sr	304	2.67
Siliciclastic sedimentary rocks	Ssr	1791	2.68
Unconsolidated clastic rocks	Ucr	2362	2.61
Anthropogenic deposits	Ad	40	2.50

Figure 43. One-way ANOVA evaluates the variability of log10(Vs30) across lithological classes as defined by Bucci et al. (2022)

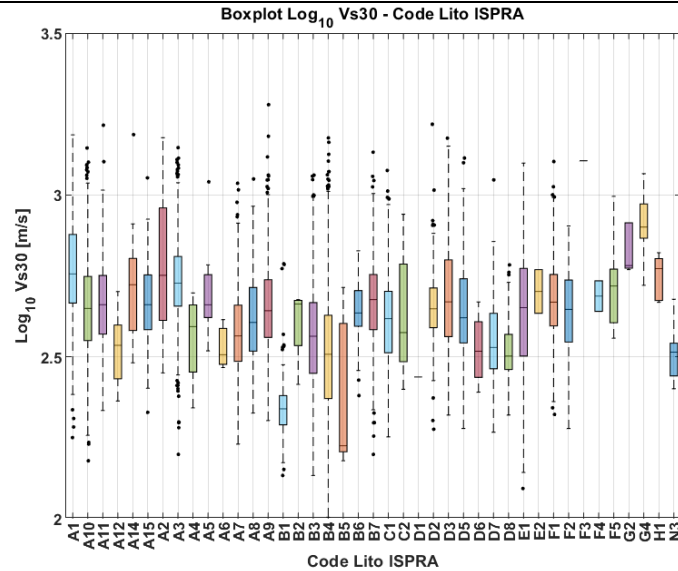
ANOVA is a statistical test used to compare the means of multiple groups and determine whether there are statistically significant differences within and between the groups. In particular, the F-statistic is calculated as the ratio of the variance between groups (MSB, Mean Square Between) to the variance within groups (MSW, Mean Square Within). A high F-value indicates that the variability between groups is significantly greater than the variability within groups, suggesting that the observed differences in means are unlikely to be due to random chance.

As shown in Figure 43, the F-value of 229.53 is remarkably high, and the p-value is lower than 0.05, meaning that the probability of the differences between groups occurring by chance is virtually zero. This confirms that the lithology significantly influences Vs30.

The boxplot in Figure 43 illustrates the distribution of $\log_{10}(\text{Vs}_{30})$ for each lithological class. Each box represents the interquartile range (IQR, between the 25th and 75th percentiles), with the line inside indicating the median, while the notches approximate the confidence interval of the median. If the notches of the two groups do not overlap, their medians are significantly different. The whiskers extend up to 1.5 times the IQR, while the black dots represent outliers.

The table on the right in Figure 43 lists the mean $\log_{10}(\text{Vs}_{30})$ values for each lithology, showing that the highest values are for carbonate rocks ($\text{Cr} = 2.74$) and consolidated clastic rocks ($\text{Ccr} = 2.64$), whereas the lowest values are found for lacustrine and glacial deposits ($\text{Li} = 2.48$) and anthropogenic deposits ($\text{Ad} = 2.50$).

The ANOVA results confirm that lithology is a significant predictor of Vs30, although some categories have overlapping distributions, suggesting that additional factors may contribute to Vs30 variability or that a refined lithological reclassification is necessary to enhance the between-group variance, leading to a better distinction in Vs30 values.



p-value less than 0.05 F=114.41 Between-group variance greater than within-group variance

Lithologies	Code	n° of Vs profiles	means (log10Vs30 [m/s])
limestones	A1	577	2.76
pelitic-arenaceous complexes	A10	2044	2.65
calcareous-arenaceous complexes	A11	601	2.66
evaporites	A12	8	2.52
cemented debris	A14	43	2.70
travertines	A15	43	2.67
dolomites	A2	56	2.77
marly limestones	A3	395	2.73
radiolarites	A4	4	2.56
quartzarenites	A5	8	2.70
diatomites	A6	3	2.53
marls	A7	511	2.58
conglomerates	A8	203	2.62
sandstones	A9	682	2.66
predominantly clayey soils	B1	178	2.35
predominantly sandy soils	B2	4	2.60
predominantly gravelly soils	B3	1159	2.56
mixed-grain soils	B4	6154	2.51
peats	B5	7	2.39
residual soils	B6	17	2.64
indefinite grain-size soils	B7	571	2.67
chaotic clays	C1	469	2.61
marine volcanic-sedimentary complexes	C2	8	2.63
acidic lavas	D1	2	2.44
neutral lavas	D2	467	2.65
basic lavas	D3	148	2.68
pyroclastics, tuffs, ignimbrites	D5	112	2.65
pyroclastics + lavas	D6	7	2.52
pedogenized tuffs	D7	123	2.55
scoriae, lapilli, pumices, bombs	D8	195	2.51
granitoid rocks	E1	66	2.64
intermediate plutonites	E2	2	2.70
phyllites and micaschists	F1	220	2.67
gneisses	F2	107	2.64
prasinities	F3	1	3.11
ophiolites	F4	2	2.69
serpentinites	F5	10	2.72
marbles	G2	5	2.83
quartzites	G4	11	2.91
low-grade metamorphics	H1	5	2.75
landfills	N3	11	2.51

Figure 44. One-way ANOVA evaluating the variability of log10(Vs30) across lithological classes as defined by ISPRA

Figure 44 presents the results of a one-way ANOVA using the lithological classification from ISPRA (see Chapter 2). The F-value of 114.41 is significantly lower than that obtained with the classification proposed by Bucci et al. 2022, indicating that the between-group variance is smaller in this classification. The p-value remains below 0.05, confirming that lithology still has a significant effect on Vs30. The boxplot shows overlapping distributions for many lithological classes, suggesting that some categories may not be well differentiated in terms of Vs30. As with the previous classification, these results indicate that a refined lithological aggregation could improve the distinction between groups and increase the between-group variance.

5.4. Geological reclassification

Following the one-way ANOVA tests conducted on the two lithological maps from the literature (ISPRA, Bucci et al., 2022), it became evident that a new lithological classification was necessary to capture greater variability in terms of Vs30. This classification was derived from the lithological classes proposed by the ISPRA geological map.

Before the aggregation, an important step was the identification of the missing areas of the Italian territory in the ISPRA lithological map, often corresponding to landslide areas. These areas were then assigned the lithological code of the adjacent unit with the largest perimeter.

By following primarily lithological similarity criteria and being guided by the results of statistical tests performed on the 41 ISPRA lithological classes (see Figure 45), these were grouped into 11 new lithological categories, as shown in Figure 45. In particular, Welch's t-test (Welch, 1947) was applied to compare the mean of $\log_{10}(\text{Vs30})$ between lithological classes. Before performing the t-tests, a Kolmogorov-Smirnov test (Kolmogorov, 1933; Smirnov, 1948) was conducted. The results showed that after transforming the data into $\log_{10}(\text{Vs30})$, the normality assumption was reasonably satisfied, making the application of Welch's t-test appropriate. Additionally, Levene's test (Levene, 1960) was performed to check for homogeneity of variances, revealing that variances were not homogeneous across lithologies ($p \leq 0.05$). Based on this result, Welch's t-test was selected due to its robustness against unequal variances. Classes with non-significant differences ($p > 0.05$) were considered for grouping.

In Figure 45, the p-value was transformed into a binary matrix where $p \leq 0.05$ is represented by 1 (red), indicating that the two lithologies are statistically different, while $p > 0.05$ is represented by 0 (blue), indicating that the two lithologies can be grouped.

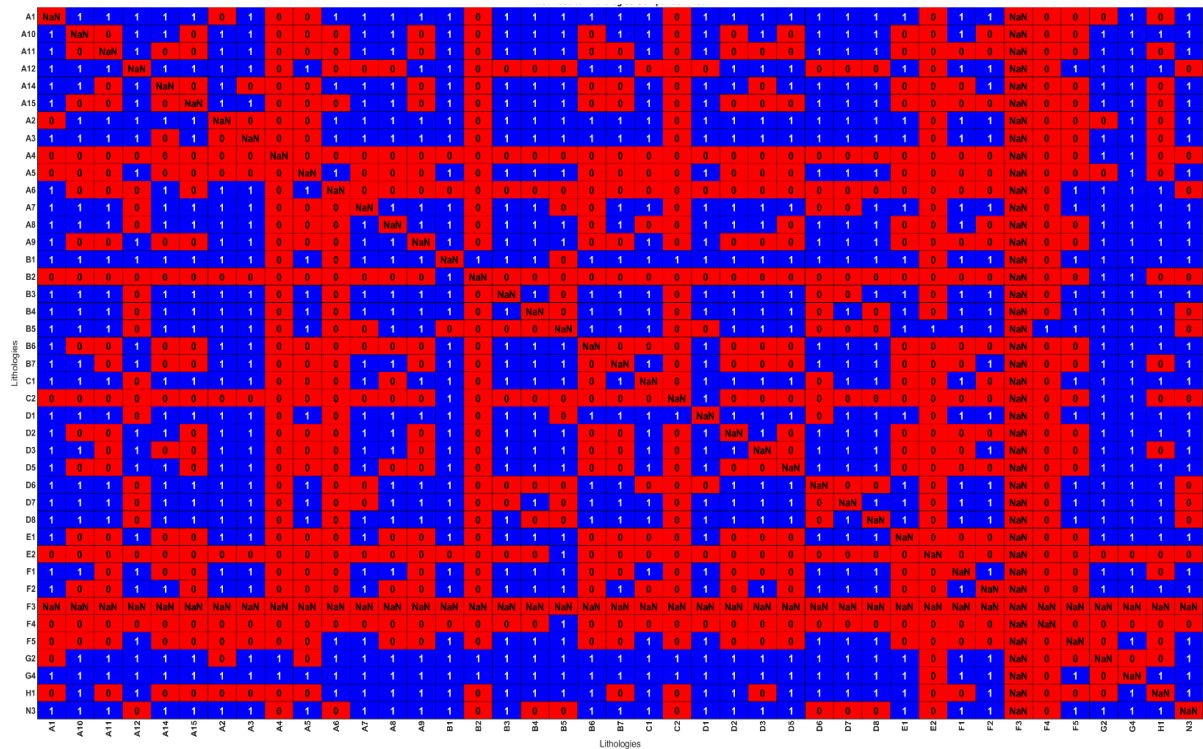


Figure 45. The binary matrix represents the results of pairwise Welch's t-tests comparing $\log_{10}(\text{Vs30})$ values between lithological classes

Only one lithological class (F3 contains only one profile, making it impossible to perform the t-test, while some other ISPRA lithological classes have less than 10 observations, see Figure 44. Nevertheless, the grouping was primarily based on lithological similarity criteria.

The new lithological grouping, see Figure 46, consists of:

- 3 classes for stiff rocks, where no correlation between Vs30 and slope is observed.
- 5 classes for soft rocks, which exhibit a correlation between Vs30 and slope.
- 1 class for volcanic pyroclastic deposits, which also show a correlation between Vs30 and slope but behave differently compared to soils. In fact, pyroclastic deposits exhibit a less pronounced trend between Vs30 and slope compared to soils.
- 2 classes for soils (coarse-to-medium-grained and fine-grained soils), both showing a correlation between Vs30 and slope.

For polygons corresponding to islands where the ISPRA lithological code was unavailable, a bibliographic study on the geology of these areas was conducted, and the most appropriate new lithological codes were assigned. For instance, Pantelleria Island was assigned the EMR code.

In particular, Table 11 shows the aggregations in detail. This reclassification aims to increase the between-group variance and improve the predictability of V_{s30} by distinguishing materials with different mechanical behaviors.

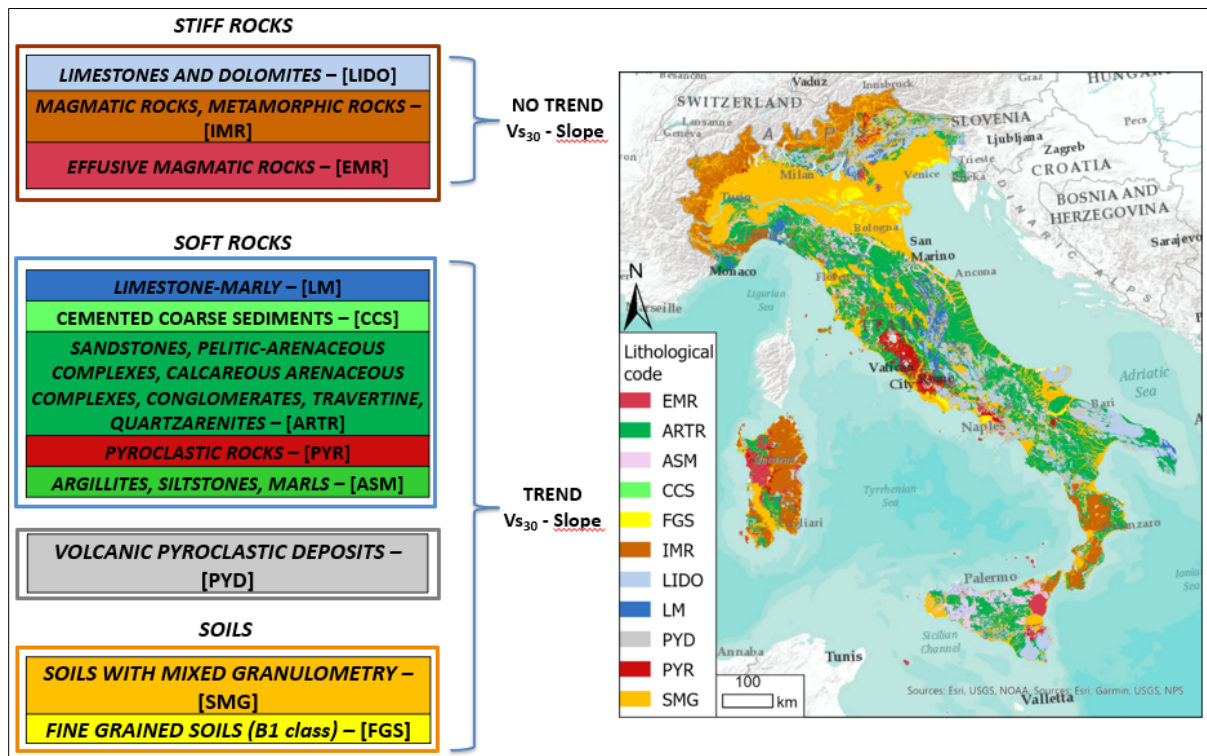


Figure 46. New lithological classification was derived by grouping the 41 original classes from the ISPRA lithological map

Lithological Code	ISPRA Code
CCS	A14
ARTR	A5, A8, A4, A9, A10, A15, A11
EMR	D1, D2, D3, D4
PYR	C2, D5
FGS	B1, B5
SMG	B4, B7, N1, N2, N3, A13, B2, B6, B3
PYD	D6, D8, D7
LIDO	A1, A2, G2, G3
IMR	E1, E2, E3, F1, F2, F3, F4, F5, G1, G4, G5, H1, H2
ASM	A7, C1, A12, A6
LM	A3

Table 11. Lithological codes from the new classification scheme showing the correspondence between each aggregated group and the original ISPRA lithological classes.

5.5. Parametric Linear Regression

Extensive testing was conducted to identify the most suitable models. Several approaches were explored to analyze the relationship between Vs30 and slope for each lithological group, incorporating different constraints, additional predictors, and random effects. The final approach consisted in applying a regression model between Vs30 and slope for soils and soft rocks, which exhibit a correlation between Vs30 and slope while assigning an average Vs30 value to stiff rocks, for which no trend was observed.

Below is the description of the models for the new lithological classes:

- SOFT ROCKS Figure 47 and volcanic PYROCLASTIC DEPOSITS Figure 48: A linear regression model was chosen, with a threshold applied to the minimum predicted Vs30 value to prevent unrealistically low estimates.
- SOILS Figure 49: A two-step linear regression model was used between Vs30 and slope, incorporating a random effect in the second segment based on soil class (Soils with medium/coarse granulometry, soils with fine granulometry). The breakpoint, where the slope of the second segment changes, corresponds to a slope value of 0.018° and was chosen to minimize the standard deviation of the regression model.
- STIFF ROCKS Figure 50: The geometric mean of Vs30 was assumed for each lithology. No regression was applied in this case, as these classes even showed an inverse trend between Vs30 and slope.

This final methodology ensured a more geologically consistent approach while addressing the limitations observed in previous tests.

SOFT ROCKS

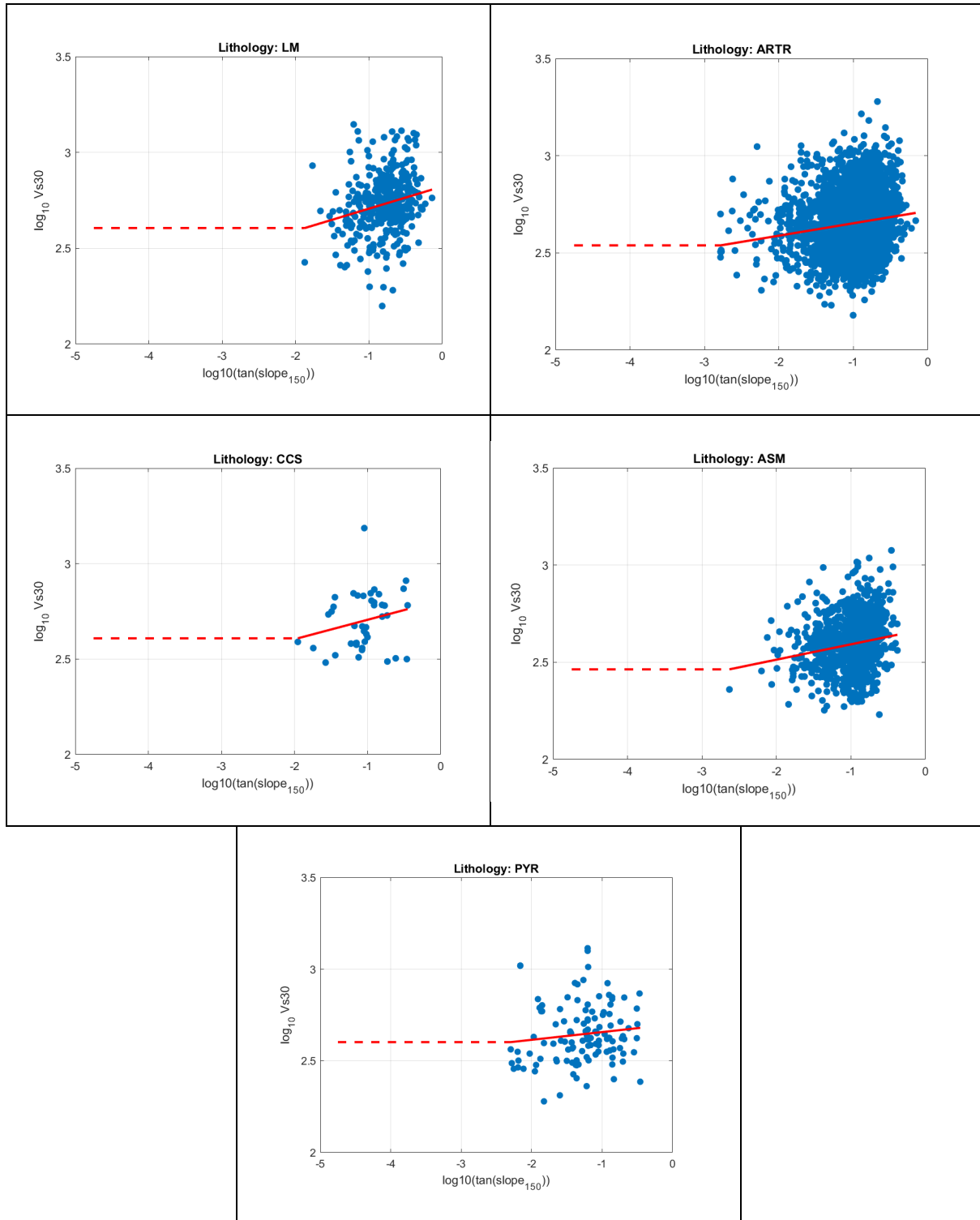


Figure 47. Linear regression models between $\log_{10}(V_{s30})$ and $\log_{10}(\tan(\text{slope}_{150}))$ for the lithological classes grouped under soft rocks, with scatter plots of the observed data. A minimum threshold was applied to the predicted V_{s30} values to avoid unrealistically

PYROCLASTIC DEPOSITS

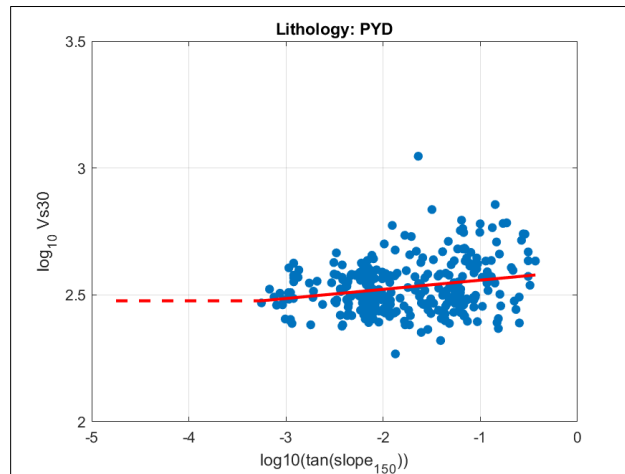


Figure 48. Linear regression model between $\log_{10}(Vs_{30})$ and $\log_{10}(\tan(\text{slope}_{150}))$ for the lithological class of pyroclastic deposits, with a scatter plot of the observed data

SOILS

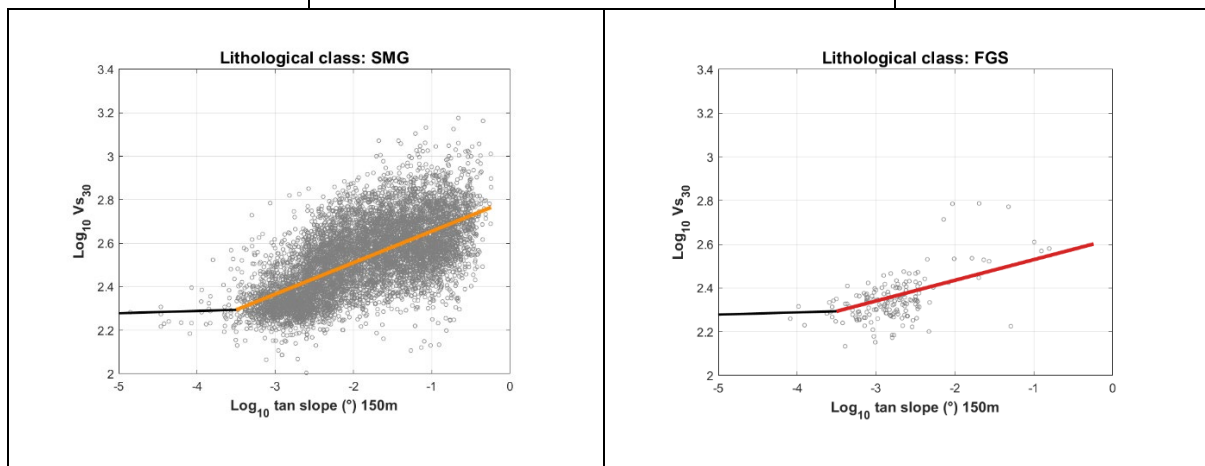
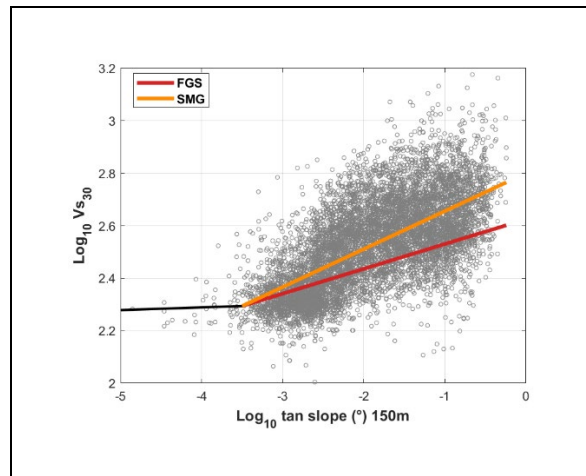


Figure 49. Piecewise linear regression models between $\log_{10}(Vs_{30})$ and $\log_{10}(\tan(\text{slope}_{150}))$ for soils. A random effect based on soil class (SMG and FGS) was introduced in the second segment of the model. Scatter plots of the observed data are shown along with the fitted regression lines.

STIFF ROCKS

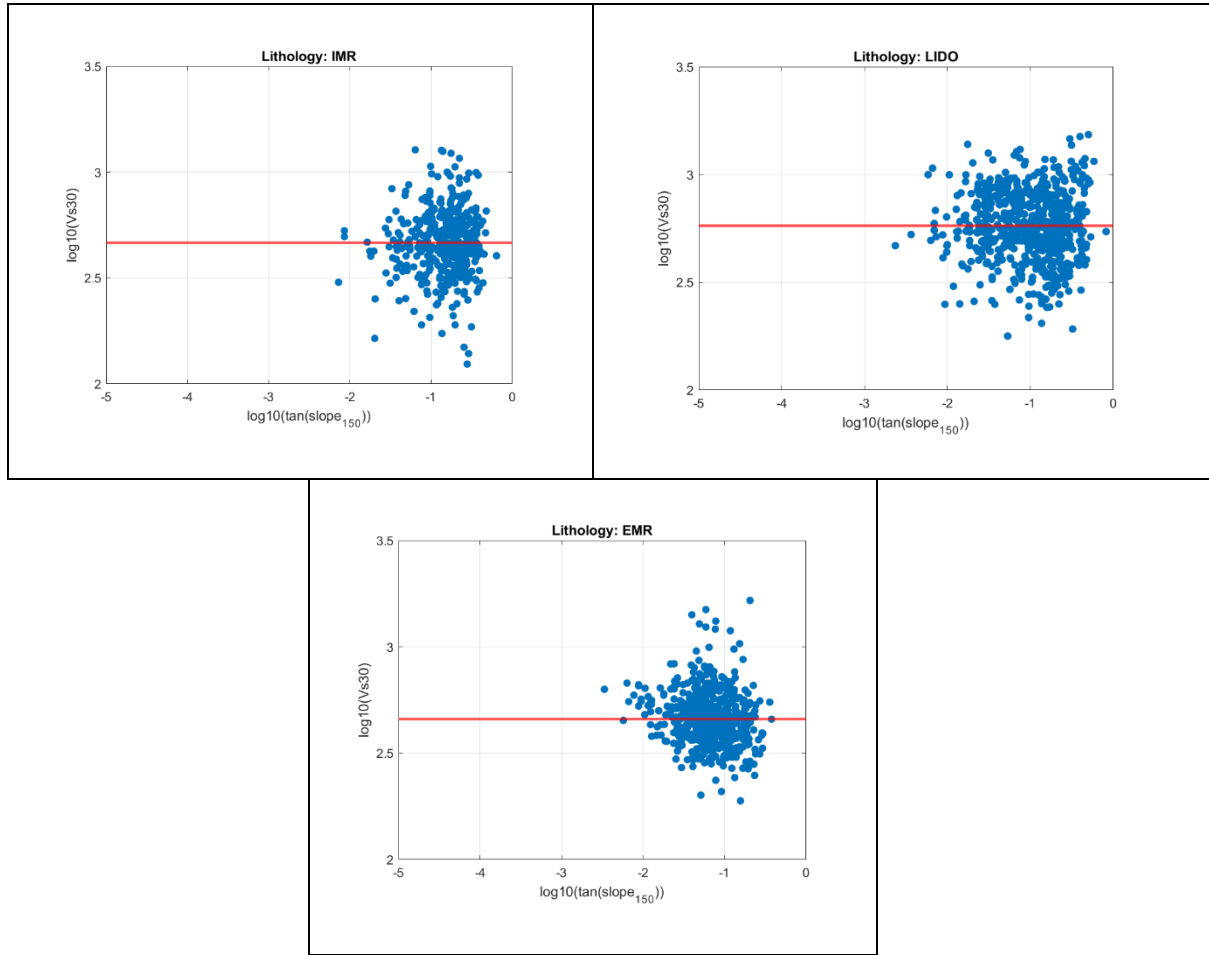


Figure 50. $\log_{10}(\text{Vs}_{30})$ values for stiff rock lithologies (IMR, LIDO, EMR) plotted against $\log_{10}(\tan(\text{slope}_{150}))$. The horizontal line representing the geometric mean of Vs_{30} was assigned to each lithology

A Vs_{30} map for the Italian territory was generated using the previously described regression models (Figure 51), accompanied by a standard deviation map of the models (Figure 52) and a residuals map (Figure 53). The residuals were calculated as the difference between $\log_{10}\text{Vs}_{30}$ (observed) and $\log_{10}\text{Vs}_{30}$ (predicted).

It is important to highlight that the lowest standard deviation values are observed for soils (SGM), see Figure 51. This is because the stepwise regression model effectively captures their variability. Conversely, for stiff rocks, such as limestones (LIDO), where the model predicts an average value, the highest standard deviation values are observed. In general, where stiff rocks (LIDO, IMR, EMR) are present (see Figure 51), such as in western Sardinia, along the Alpine chain, in southwestern Sicily, and the central-southern Apennines, the models show the highest standard deviation values, as expected.

Examining the residuals map (Figure 53), it can be seen that in the Apulian Platform (southeastern Italy), where limestones (LIDO) are predominant, residuals are generally positive. This further confirms that the mean Vs30 does not fully capture the variability of stiff rocks. Negative residual values are primarily observed along the coast and in the Po Plain, where deposits (SGM) are present.

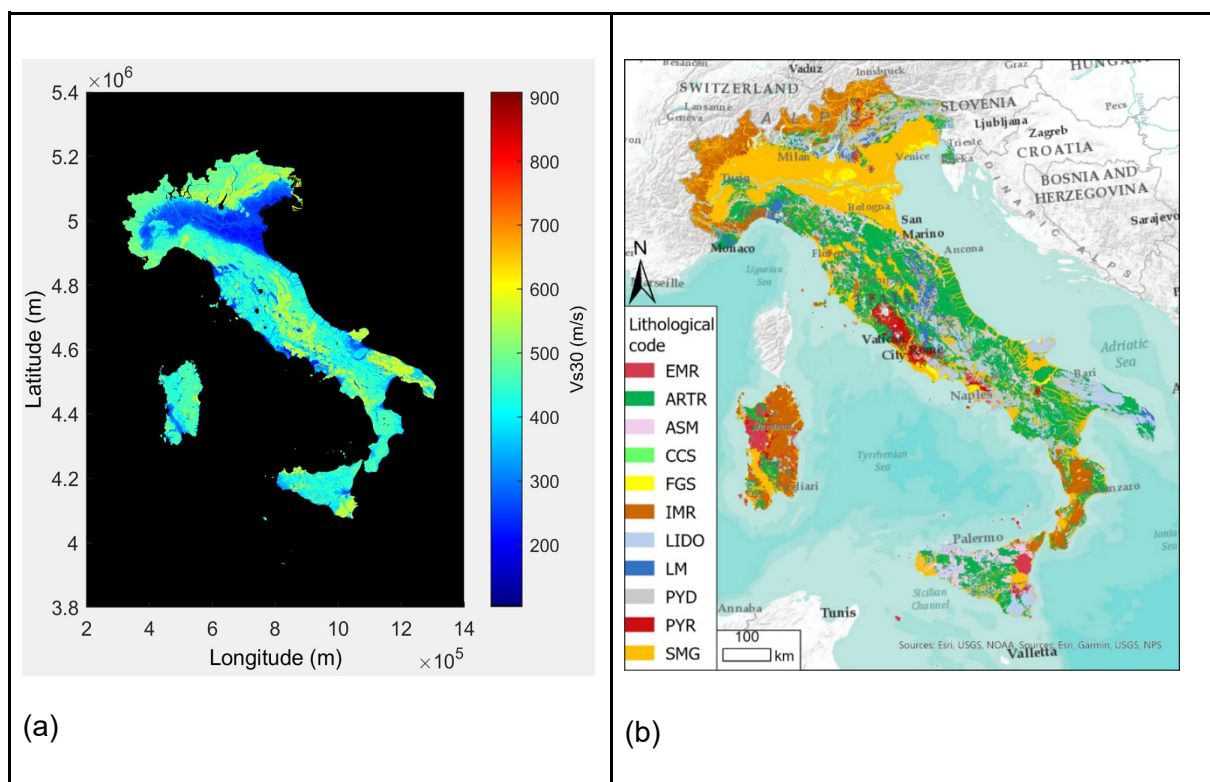


Figure 51. (a) Vs30 map for the Italian territory generated using either the regression models or the mean values per lithology, as described in the previous sections. (b) Map showing the spatial distribution of the new lithological groups used to estimate Vs30

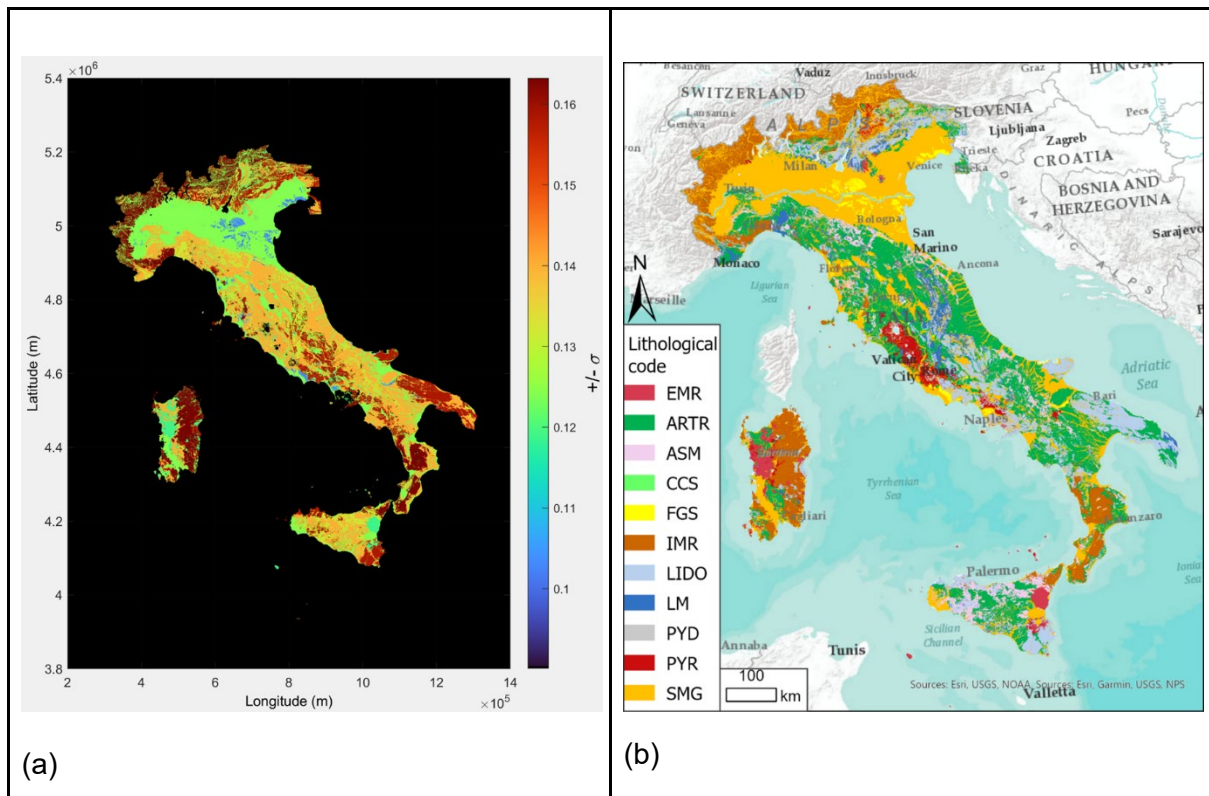


Figure 52. (a) Standard deviation map (σ) of the residuals for each new lithological class. (b) Map showing the spatial distribution of the new lithological groups used to estimate $Vs30$

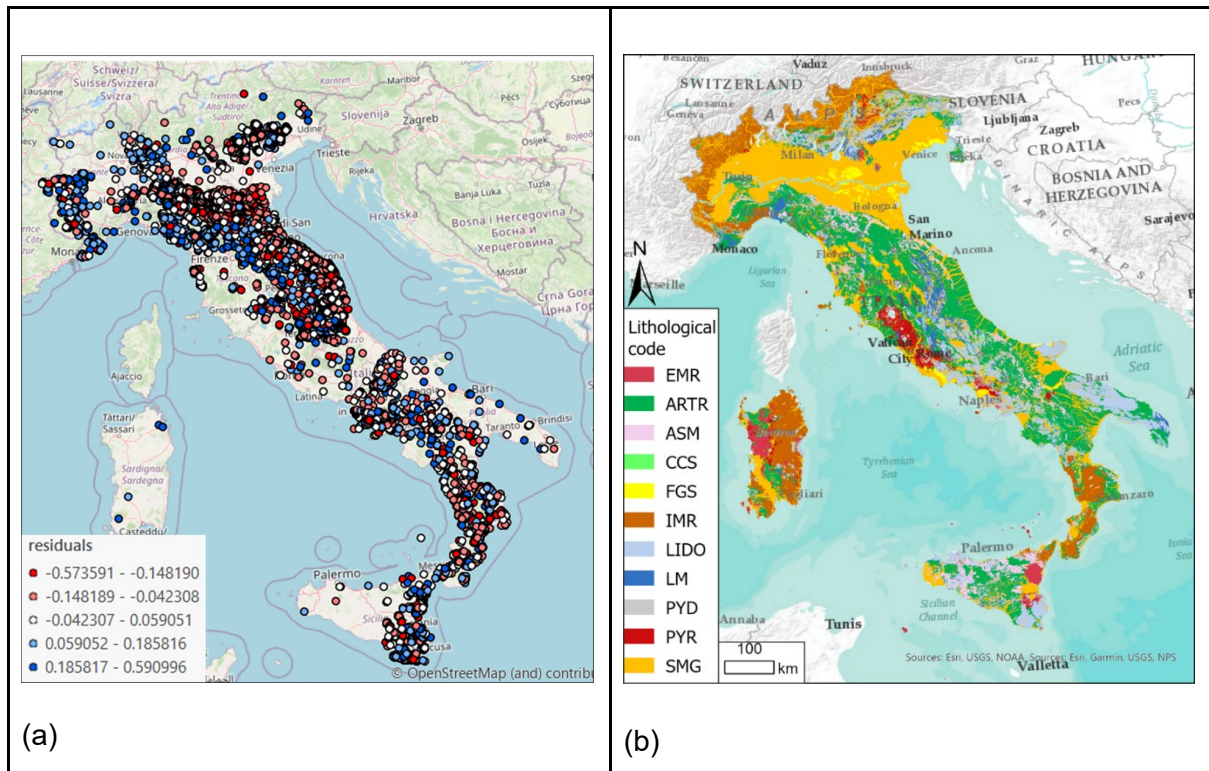


Figure 53. Spatial distribution of residuals, calculated as the difference between observed and predicted $\log_{10}(Vs30)$. (b) Map showing the spatial distribution of the new lithological groups used to estimate $Vs30$

5.6. Geostatistical Analysis

To improve the prediction and reduce the uncertainty of the Vs30 map, especially for the stiff rocks, the residuals were interpolated, and the resulting map (Figure 59) was added to the Vs30 map produced by the regression models described in Chapter 5.5.

For each lithology, synthetic semivariograms of the residuals were computed and subsequently used for spatial interpolation via ordinary kriging; for this purpose, a custom script was developed in the language R (R Core Team, 2022; RStudio Team, 2021; Pebesma E.J., 2004; Robert J. Hijmans 2023; Pebesma E.J. and R.S. Bivand, 2005; Bivand et al., 2013).

The process of kriging begins with calculating the empirical semivariogram, which evaluates the semivariance of point pairs at various separation distances, up to a predefined, a maximum distance (cutoff) beyond which data are not considered for semivariance calculation. In this study, the empirical semivariogram was computed under the assumption of a constant mean, meaning no additional covariates were considered. This approach isolates the spatial correlation structure of the residuals without incorporating potential influences from external factors such as slope.

The cutoff distance and the bin width (distance intervals for grouping point pairs) were empirically selected after multiple tests to clearly represent the spatial variability of the residuals. The chosen values for each lithology are shown in Table 12.

Table 12. Maximum distance and bin width for each lithology considered in this study

Lithology	Max Distance (m)	Bin Width (m)
SMG	110000	2200
FGS	10000	500
IMR	25000	1500
ARTR	100000	2000
LIDO	30000	1500
EMR	15000	1000
ASM	25000	625
PYR	6000	600
LM	60000	2400
PYD	38000	2666
CCS	20000	1000

However, since the empirical semivariogram is directly based on data, it can exhibit noise and irregularities. To provide a more continuous and regular representation of spatial variability, a theoretical semivariogram model was fitted. A Matérn model was chosen (Matérn, 1986), as it offers flexibility in describing spatial variance through key parameters:

- Range: the distance, in meters, beyond which observations are no longer correlated.
- Nugget: the variability observed at very small distances.
- Sill: the maximum level of semivariance reached by the model.

Additionally, the kappa parameter, which controls the smoothness of the Matérn function, was set to 0.5, making it similar to an exponential semivariogram.

Finally, to optimally fit the theoretical model to the empirical data, a weighted least squares method was applied, which attributes more weight to distance bins with a higher number of point pairs. This process minimizes the difference between the empirical and theoretical semivariograms, resulting in a synthetic semivariogram that effectively captures the spatial structure of residuals and serves as the basis for kriging interpolation.

The residuals of soils (SGM and FGS) (Figure 56) and cemented arenites and conglomerates (ARTR and CCS) (Figure 55) were merged, and for these two groups, only two synthetic semivariograms were produced instead of four. Meanwhile, no semivariogram is built from the residuals of the PYR lithology (pyroclastic rocks), as they exhibited no spatial correlation (see Figure 58).

For each lithology, the sill, range, and nugget of the synthetic semivariograms are shown in Table 13.

Table 13. Sill, range, and nugget values of the synthetic semivariograms for each lithological class considered in this study

Lithologies	Sill	Range (m)	Nugget
IMR	0.0215	3244	0.0162
LIDO	0.0186	1776	0.0078
EMR	0.0106	1678	0.0050
ASM	0.0138	1474	0.0047
LM	0.0187	13446	0.0087
PYD	0.0086	8252	0.0020
SMG+FGS	0.0145	40386	0.0075
ARTR+CCS	0.0237	50078	0.0132

As shown in Figure 56, soil residuals exhibit spatial correlation extending up to 40 km, while limestone residuals [LIDO] (Figure 54) show a much more limited spatial correlation, around 2 km. This is geologically consistent, as soils generally have homogeneous velocity characteristics, whereas rocks exhibit greater heterogeneity. Additionally, since a 1:100,000-scale lithological map was used, some profiles acquired on deposits might be included within the rock classes, due to possible errors in the map digitizing and profile coordinates. This could bias the mean Vs30 values assigned to the rock classes. Figure 57 shows the synthetic semivariograms for pyroclastic deposits [PYD].

STIFF ROCKS

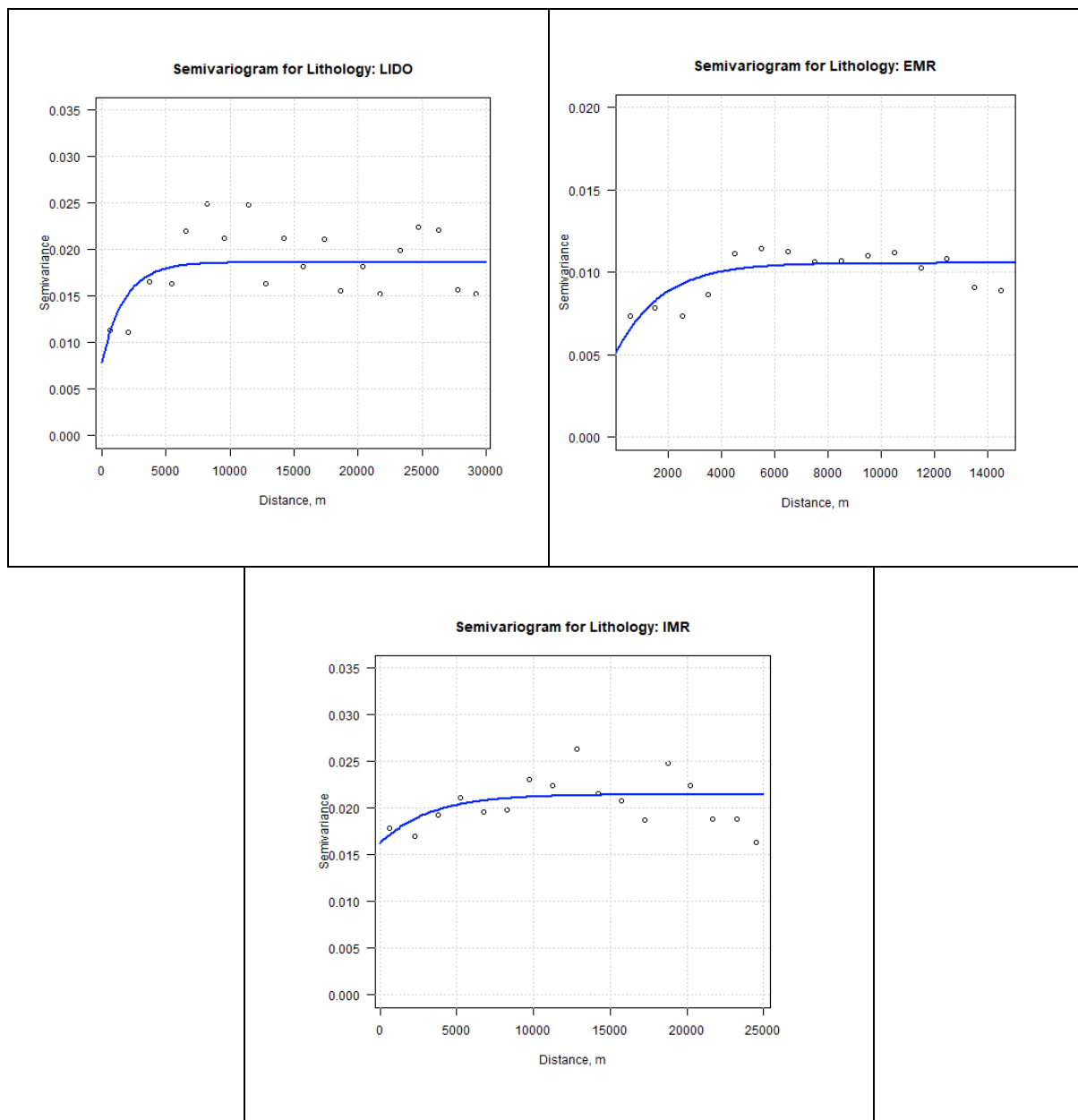


Figure 54. Synthetic semivariograms for stiff rocks

SOFT ROCKS

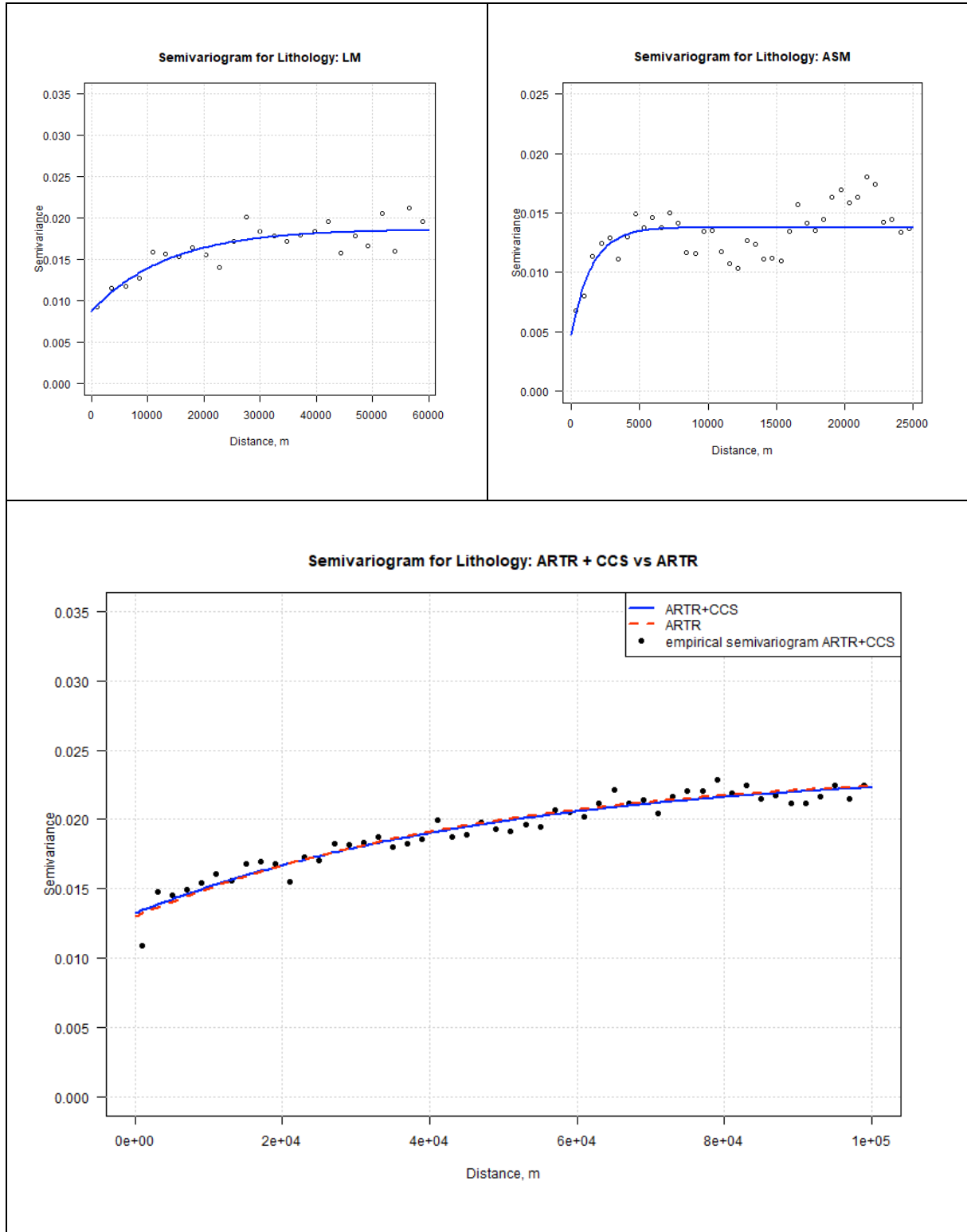


Figure 55. Synthetic semivariograms for soft rocks

SOILS

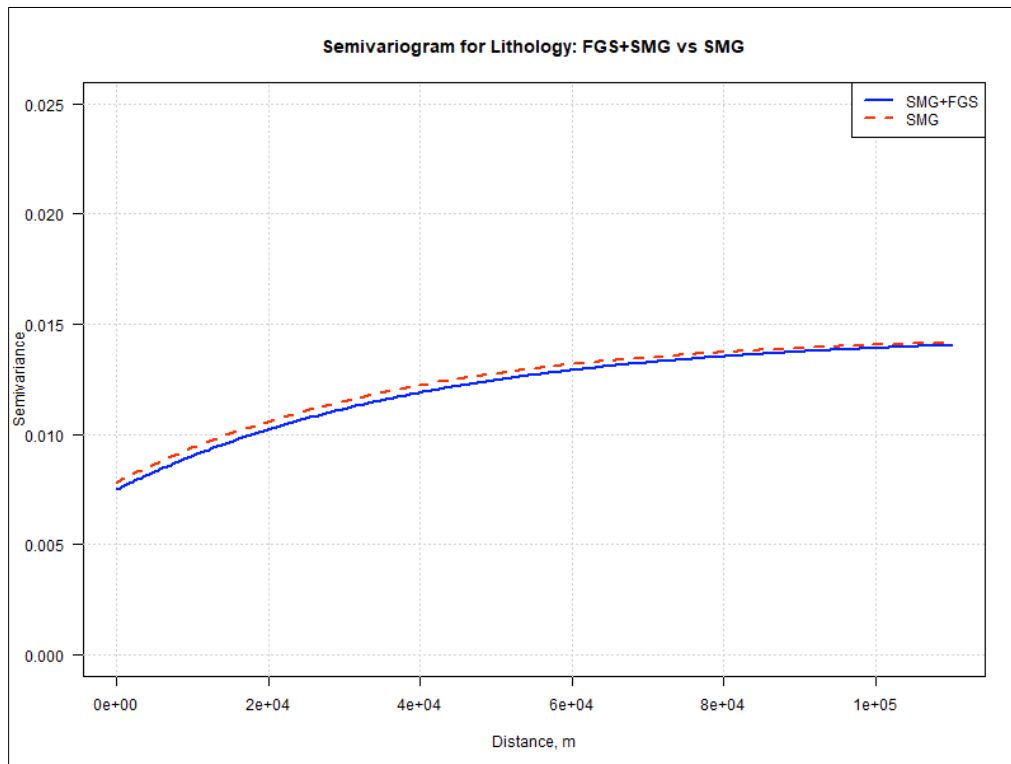


Figure 56. Synthetic semivariogram for soils

PYROCLASTIC DEPOSITS

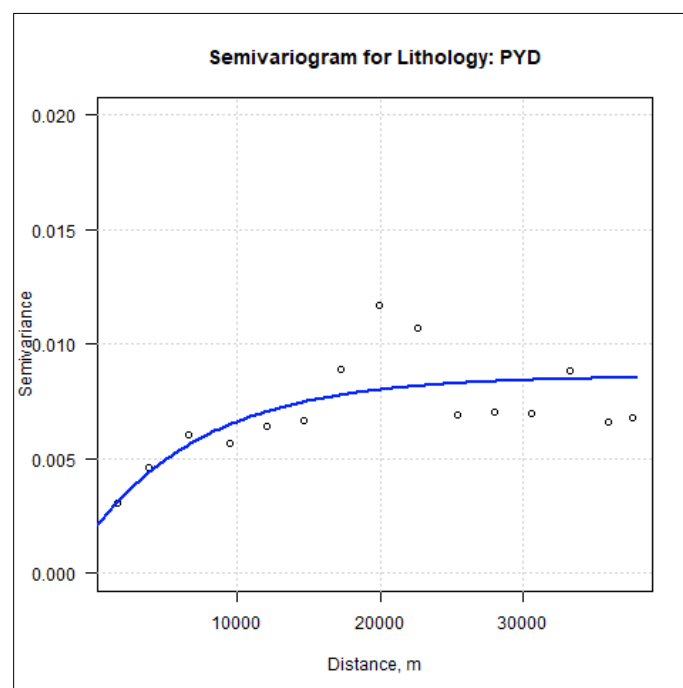


Figure 57. Synthetic semivariogram for pyroclastic deposits

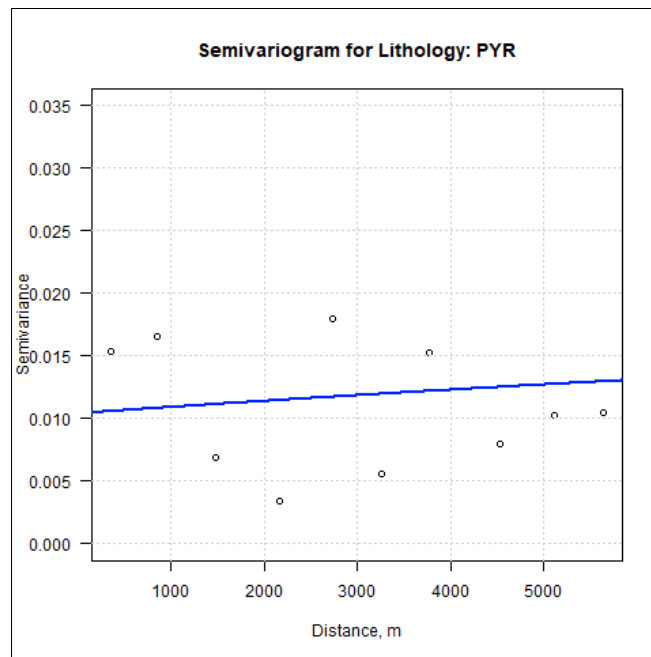


Figure 58. No valid semivariogram could be fitted for the PYR lithology due to insufficient spatial correlation

The interpolation of residuals was performed using ordinary kriging (Matheron, 1963) and only within the lithological polygons that contain residuals. Ordinary kriging is a geostatistical interpolation method that assumes the mean of the variable is unknown but constant within the local neighborhood of each estimation point. It uses a weighted linear combination of observed values, where weights are determined based on the spatial autocorrelation structure described by the semivariogram. The result of the interpolation is shown in Figure 59.

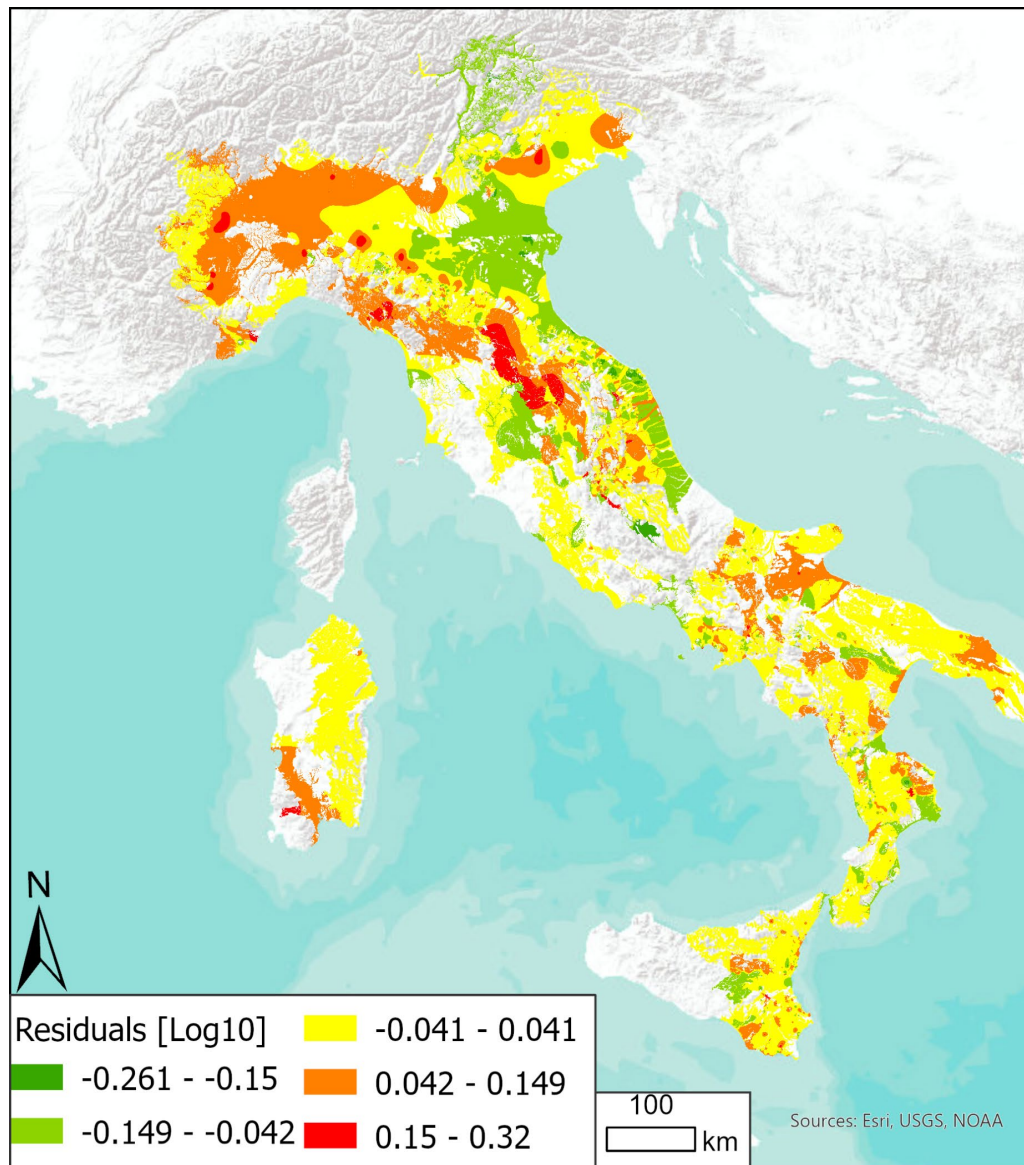


Figure 59. Interpolated map of residuals based on the semivariogram models described above. The interpolation was performed only within the polygons where residual data were available

Figure 59 shows that residuals are predominantly negative in the Po Plain and along the Adriatic coast, where soils and soft rocks are dominant, while markedly positive values are observed in the Prealps, along the Apennine chain, and in the regions of Calabria, Apulia, and south-eastern Sicily, where rock formations prevail.

5.7. Model Construction

Finally, the following maps were generated:

- Mean Vs30, obtained by summing the median predictions of the statistical models for each lithology and the interpolated residuals Figure 60;
- Vs30 +1 σ Figure 61;
- Vs30 -1 σ Figure 62.

The standard deviation (σ) accounts for both the uncertainty of the models and that of the spatial interpolation of the residuals. The total uncertainty (σ_{tot}) was computed using the formula for error propagation, assuming no correlation between the two sources of uncertainty:

$$\sigma_{\text{tot}} = \sqrt{\sigma_{\text{reg}}^2 + \sigma_{\text{interp}}^2} \quad [22]$$

where:

- σ_{reg} represents the uncertainty of the regression models, derived from the standard deviation of the residuals in the statistical models;
- σ_{interp} represents the uncertainty associated with the kriging interpolation of residuals, obtained from the kriging variance.
-

$$\sigma_{\text{interp}} = \sqrt{\text{Var}_{\text{kriging}}} \quad [23]$$

Finally, the measured Vs30 was assigned to the corresponding pixels. If multiple surveys were present, their average was used. Consequently, the uncertainty is zero in these pixels, though the uncertainty of the Vs survey itself was not considered, because it was not available.

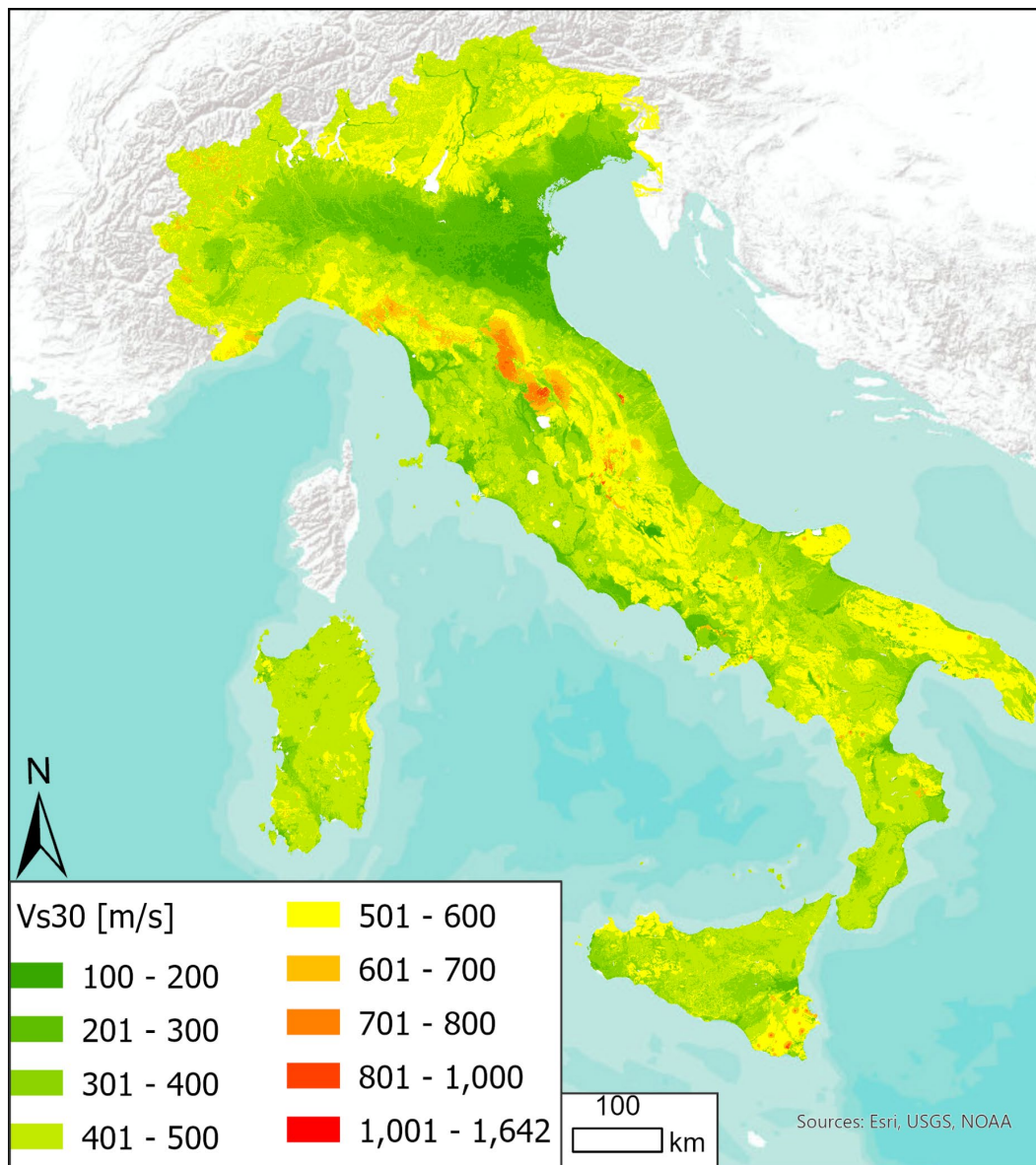


Figure 60. Final Vs30 map was obtained by combining either the median or regression-based values assigned to each lithological class with the interpolated residuals

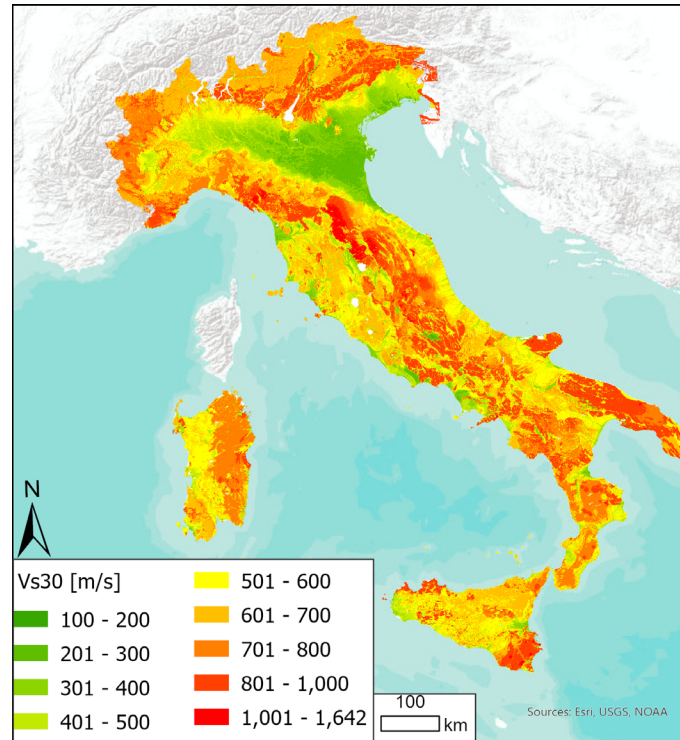


Figure 61. Final Vs30 map with the addition of $+1\sigma$. The standard deviation combines the uncertainty from both the regression models and the spatial interpolation of residuals

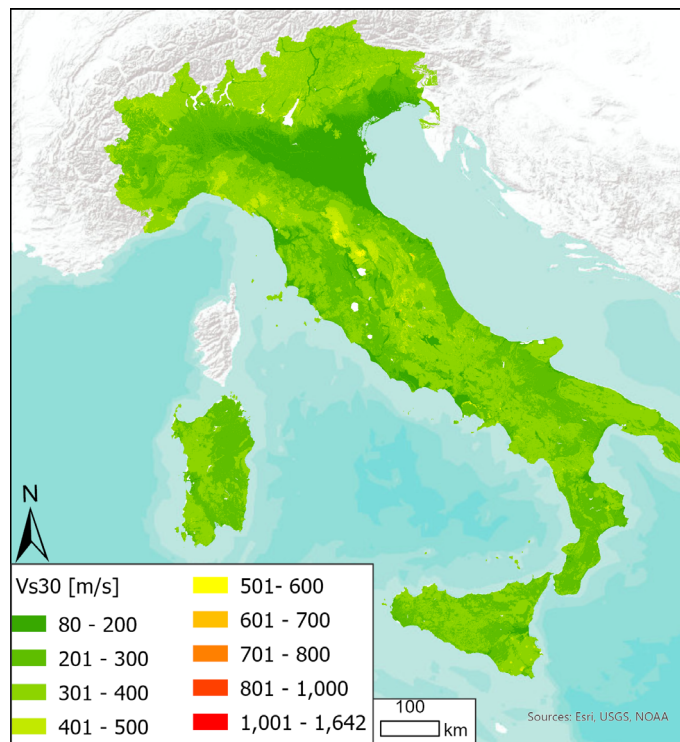


Figure 62. Final Vs30 map with the addition of -1σ . The standard deviation combines the uncertainty from both the regression models and the spatial interpolation of residuals

By observing Figure 61, the highest values of $V_{s30} + \sigma$ tot occur in areas dominated by stiff rocks (e.g., the Alps, Apennines, south-eastern Sicily, eastern Sardinia, and the Apulia region), where a mean V_{s30} value was assigned. In contrast, lower $V_{s30} + \sigma$ values are found in areas dominated by soils, such as the Po Plain.

5.9. Comparison with V_{s30} Maps from literature

Figure 63 shows the histogram of the relative frequency of V_{s30} values for the three maps available for the Italian territory by: i) Mori et al. (2020), hereinafter *Mori*; ii) estimates based on Wald and Allen (2007), produced for the Italian territory by Claudia Mascandola (personal communication), hereinafter *WA*; iii) this map, hereinafter *Brunelli*.

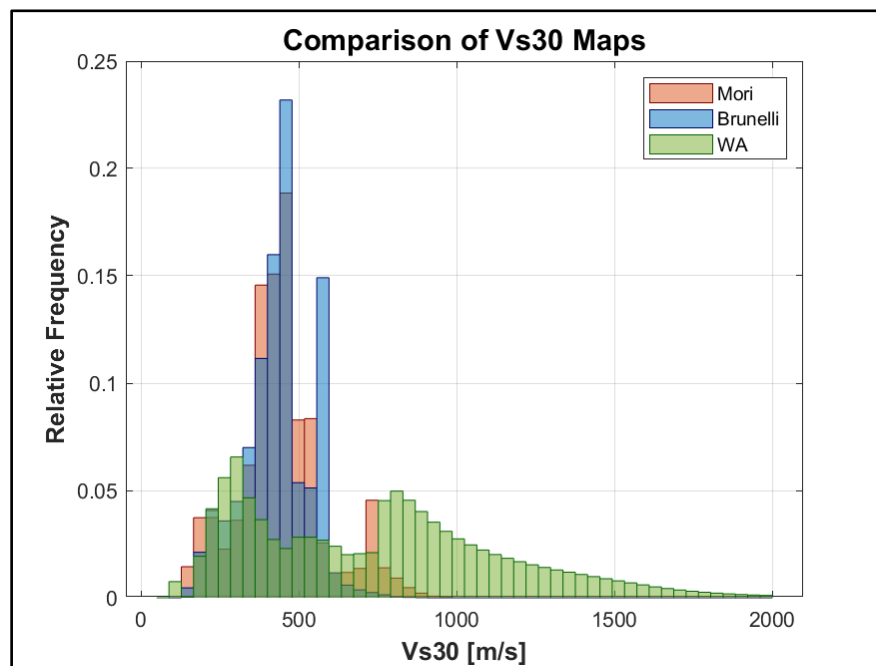


Figure 63. Histogram of the relative frequency of V_{s30} values from the Mori et al. 2020, Brunelli, and Wald and Allen (2007) maps

Table 14. Summary statistics of V_{s30} values from the Mori et al. 2020, Brunelli, and Wald and Allen (2007) maps, including minimum, maximum, mean, and standard deviation

Map	Min [m/s]	Max [m/s]	Mean [m/s]	Std [m/s]
Mori	150	1060	445	139
Brunelli	101	1642	434	105
WA	100	3500	761	425

.The Vs30 maps Mori and Brunelli exhibit similar distributions, with comparable mean values (445 m/s for Mori and 434 m/s for Brunelli) and relatively low variability compared to the WA map, as indicated by their respective standard deviations. In particular, the histogram of Brunelli shows a bimodal distribution peaked at about 450 m/s and between 550 and 600 m/s, consistent with the nature of the input data (see Figure 9). The latter interval includes the average Vs30 value assigned to the carbonate lithological units [LIDO], which represent a significant portion of the model.

In contrast, WA shows a significantly broader distribution of Vs30 values, ranging from 100 to 3500 m/s, with a higher mean (761 m/s) and a notably large standard deviation (425 m/s), strongly reflecting the distribution of slope angles in Italy. The distribution is asymmetric, and the model tends to predict higher Vs30 values more frequently, typically associated with rigid terrains or rocks with high slope angles. Despite adopting entirely different approaches, *Mori* and *Brunelli* show overall consistency, likely due to the use of a very similar input dataset.

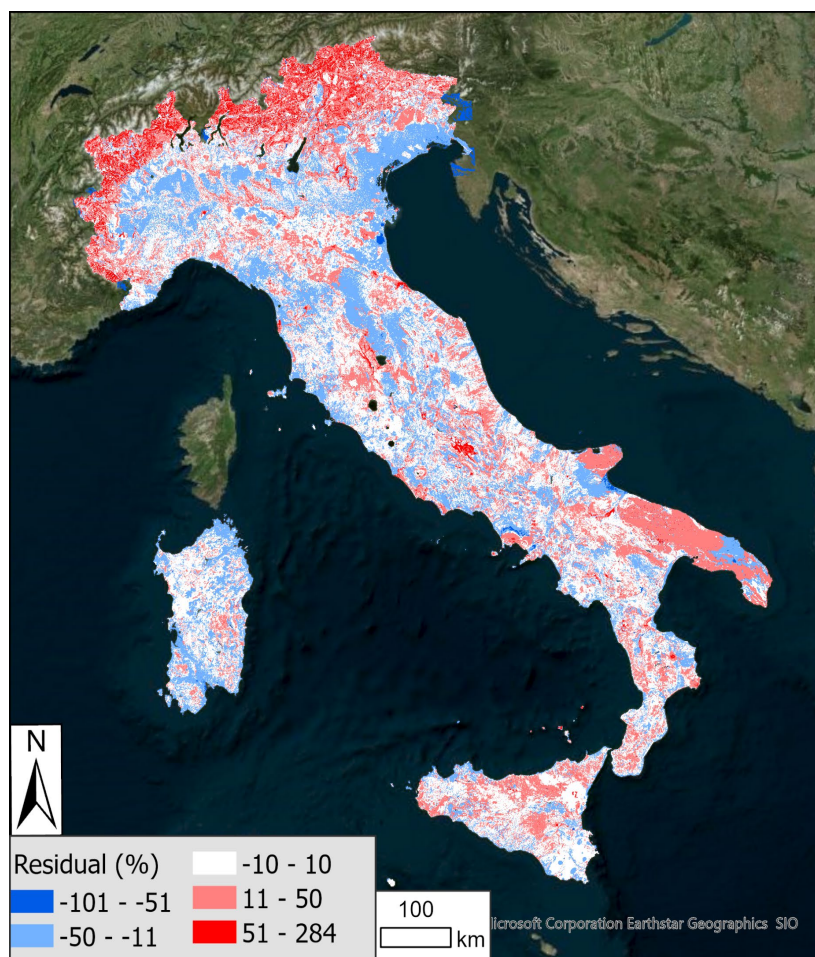


Figure 64. Map of Vs30 residuals (in percentage) between Mori and Brunelli

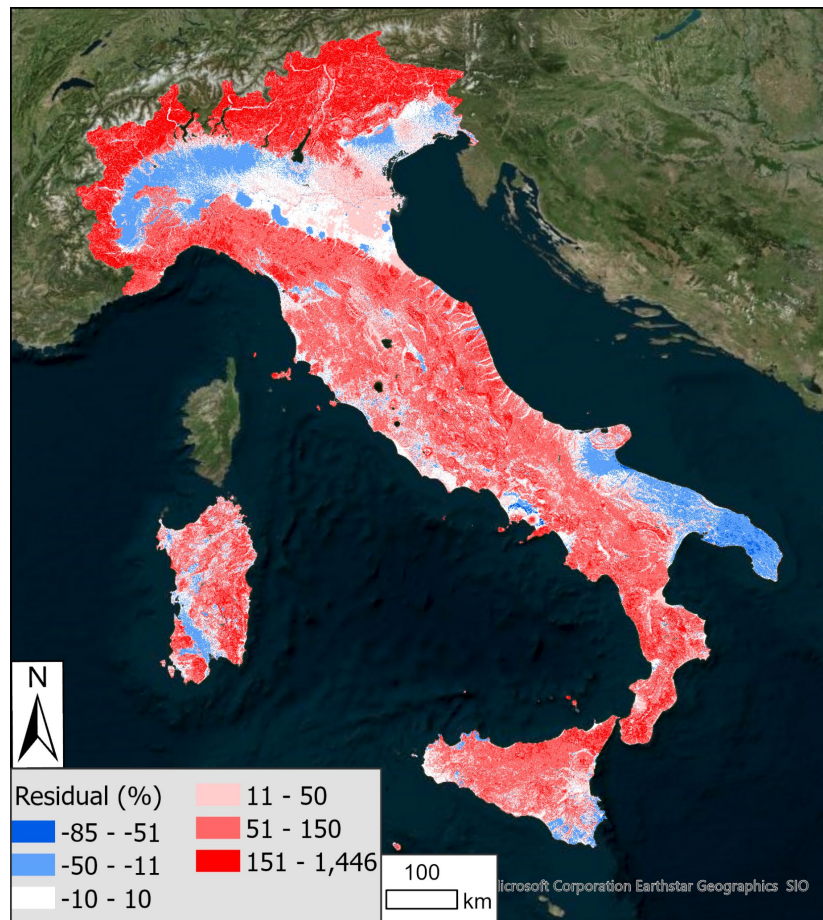


Figure 65. Map of Vs30 residuals between WA and Brunelli

Figures 64 and 65 show the percentage difference between Mori and Brunelli and WA and Brunelli, respectively. Mori estimates higher Vs30 values, ranging between 10% and 50% and sometimes even higher, along the Alpine arc and the Apulian platform, particularly in correspondence with rock sites, where Brunelli predicts an average value corrected by the interpolated residuals, when available.

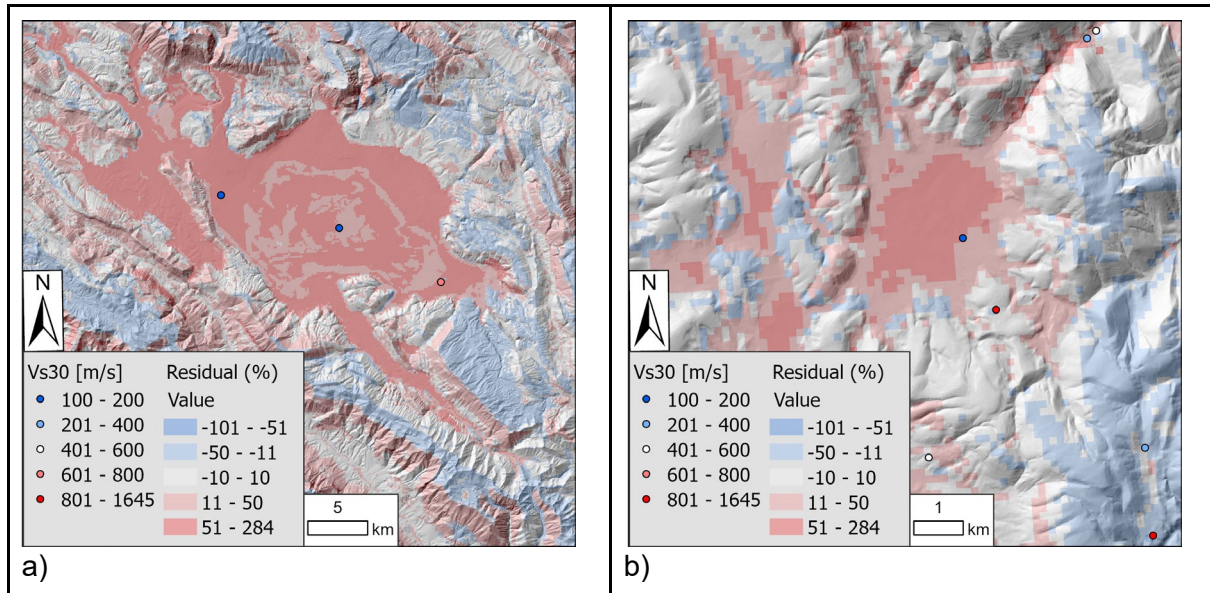


Figure 66. Difference between Mori and this map in two intramontain basins: a) Fucino and b) Colfiorito. Dots represent the values of the observed Vs30

Mori also predicts significantly higher values, exceeding 50%, in intramontane basins such as the Fucino and Colfiorito basins (Figure 66). In contrast, Mori estimates lower Vs30 values, ranging from -10% to -50% , in soils such as those found in many areas of the Po Plain. However, it can be observed that across the rest of the peninsula, the difference between the two estimates remains limited.

A key difference between the proposed map and those by Mori (2020) map and WA lies in the role of observations; in fact, this maps integrates interpolated residuals and assigns the exact Vs30 measured at 15,000 sites, including 358 accelerometric stations.

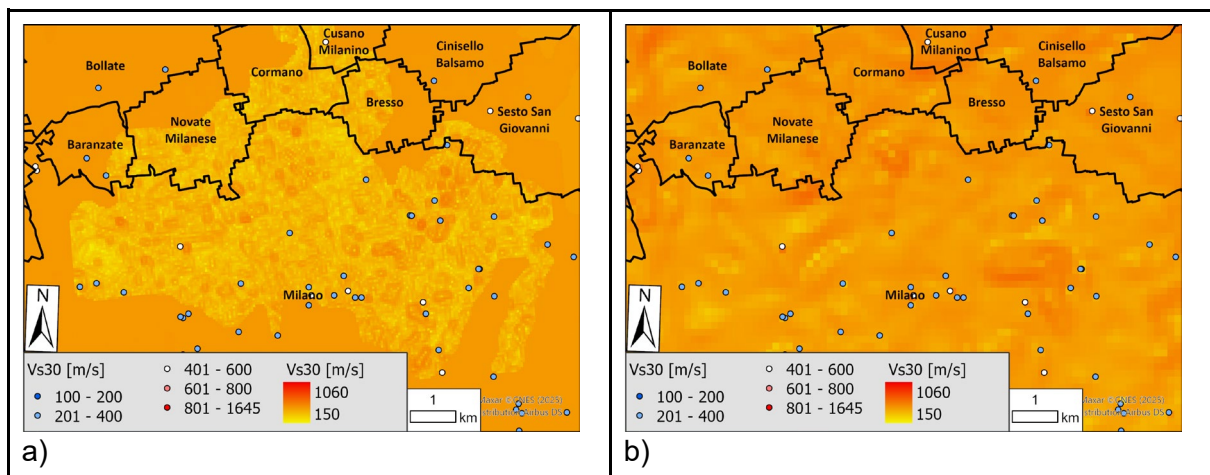


Figure 67. Comparison between a) Mori and b) Brunelli in the metropolitan area of Milano, where the influence of the slope calculated from the DSM is visible

Moreover, Mori et al (2020) used a Digital Surface Model (DSM) to derive the slope angle from ALOS World 3D at 30 m resolution (AW3D30 Version 2.2), which includes the height of buildings in urban areas. The use of a DSM can lead to slope overestimations in urban settings due to the presence of vertical obstacles, directly affecting the Vs30 estimation. The proposed map instead is based on the slope derived from a Digital Terrain Model (DTM), which represents the elevation of the terrain and thus, more suitable for representing the natural morphology of the terrain, excluding artificial structures (see Figure 67).

The comparison with WA (see Figure 65) shows higher values of Vs30 for 90% of the territory except for areas where rock sites have flat settings (e.g. Puglia region or Iblean plateau in Southern Italy). This occurs because Wald and Allen (2007) assume a positive correlation between Vs30 and slope for both soils and rock, resulting in underestimated values in flat rock regions. Conversely, WA predicts lower Vs30 values than this map in the western sector of the Po Plain because of low slope angles, although the observed Vs30 results in higher values than the eastern Po plain.

These observations lead to the conclusion that the map calibrated in this study reflects more closely the measured Vs30 values and the Vs30 estimations are neither strongly influenced by the slope angle such as WA nor altered by a base map that includes man-made infrastructures, such as Mori.

6. Shaking scenarios

Two example ground shaking scenarios were generated for earthquakes that occurred in central Italy: the Mw 6.0 Colfiorito on 26 September 1997 and the Mw 6.5 Norcia on 30 October 2016. These scenarios were specifically developed to highlight the role of empirical correction factors and the Vs30 map in improving the spatial prediction of ground shaking. The Colfiorito event represents a case with limited instrumental coverage, whereas the Norcia event benefits from a dense network of strong-motion recordings, thus offering a useful comparison for assessing the impact of the applied corrections. The scenarios were computed for two spectral acceleration periods representative of short and intermediate intervals, i.e. 0.01 s and 1 s, using the ITA18 ground motion model.

To demonstrate that the method is robust, the scenarios were calculated assuming the simplest source representation, expressed by the event magnitude and the fault surface projection (2D). The ground motion is calculated within a 50 km buffer from each epicenter and the correction for attenuation is not applied as it is deemed negligible at such distances. Similarly, no event-specific correction was applied, since both earthquakes fall within the most sampled region of Italy, therefore characterized by a negligible value of the correction factor. Since both earthquakes are real, the use of the actual δBe (between-event residual) could have been considered. However, the regional source correction is more appropriate in predictive scenarios or when modeling historical events without instrumental data. Based on these considerations, the Vs30 map plays a fundamental role.

The ground motion predictions were computed over a regular national grid of nodes, spaced at 500 m intervals and to account for site effects, two different Vs30 maps were tested: the first proposed by Mori et al. (2020) and the second developed as part of the present research.

A comparison was carried out between the predicted response spectra, derived using the two different Vs30 maps, and the observed spectra at accelerometric stations located within the 50 km buffer zone. The comparison was performed through the calculation of the Root Mean Square Error (RMSE), considering three spectral period ranges: i) [0.01 s – 0.4 s]; ii) [0.4 s – 2 s]; iii) [2 s – 10 s].

Each recording station is uniquely identified by a Network Code (e.g., IT for the Italian Strong Motion Network managed by the Department of Civil Protection – DPC, or IV for the National Seismic Network – RSN managed by INGV) and a Station Code (e.g., CLF or CSA, which are unique identifiers for each seismic station).

6.1 Simulation of the Mw 5.7 Colfiorito 1997 earthquake scenario

The first simulated event is the earthquake that occurred on 26 September 1997 at 00:33:11 (UTC), with a magnitude Mw 5.7 (Pondrelli et al., 2002), a normal faulting focal mechanism (Pondrelli et al., 2002), a hypocentral depth of 5.7 km and epicentral coordinates of 43.023° N and 12.892° E (Chiaraluce et al., 2004).

15 accelerometric recordings are available for this event in the ITACA database (Felicetta et al., 2023), 11 of which fall within a 50 km buffer from the epicenter. All recordings were manually processed following the procedure described by Paolucci et al. (2011).

The event is located near the intermontane basin of Colfiorito Province of Perugia. This earthquake was selected for the analysis because site effects play a fundamental role in the ground motion, as the Colfiorito basin is a tectonic depression filled with soft sediments, characterized by low shear-wave velocities (< 200 m/s) (Di Giulio et al., 2006).

The results of the scenario calculation are shown in Figure 68. Panels a) and b) display the results for 0.01 s and 1 s periods, respectively, obtained using the Vs30 map developed in this study. Panels c) and d) show the corresponding results for the same periods, computed using the Vs30 map proposed by Mori et al. (2020).

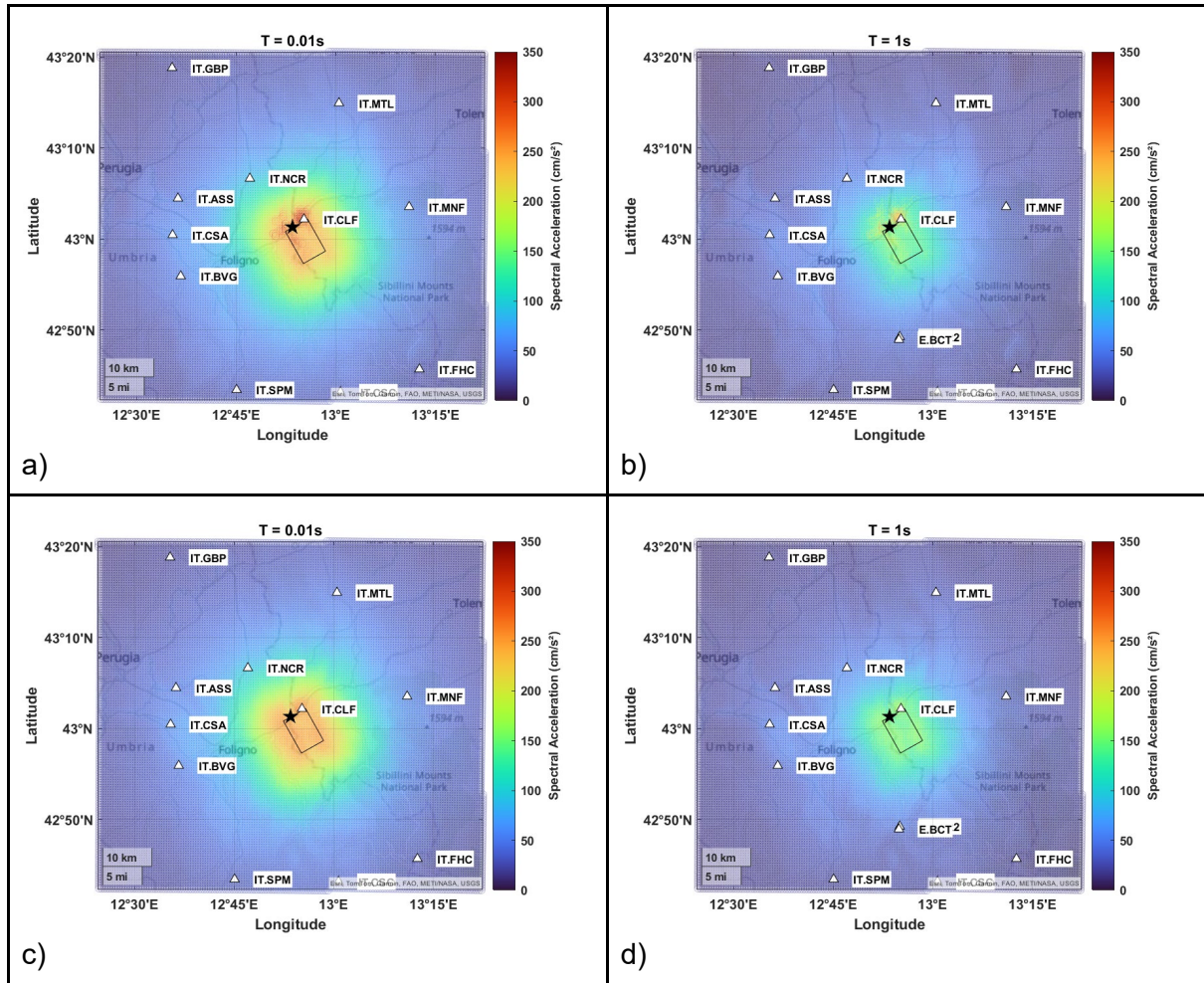


Figure 68. – Ground shaking scenarios in terms of spectral acceleration for the Mw 5.7 Colfiorito earthquake (ITACA event ID IT-1997-0004). Panels a) and b) display the results for 0.01 s and 1 s periods, respectively, obtained using the Vs30 map developed in this study. Panels c) and d) show the corresponding results for the same periods, computed using the Vs30 map proposed by Mori et al. (2020). The location of each recording station is shown using white triangles, along with the corresponding network and station codes

Figure 68 shows that the ground-shaking scenarios produced using the two different Vs30 maps are, overall, very similar. However, despite the general agreement between the two scenarios, some local differences are observed. In particular, the scenario generated using the Vs30 map developed in this study shows, at both short and long periods, higher spectral acceleration values within the Colfiorito basin, resulting in a closer match to the observed data recorded at station IT.CLF (see Figure 69, panels (a) and (b), showing the predictions using the Brunelli and Mori Vs30 maps, respectively), which is located within the basin.

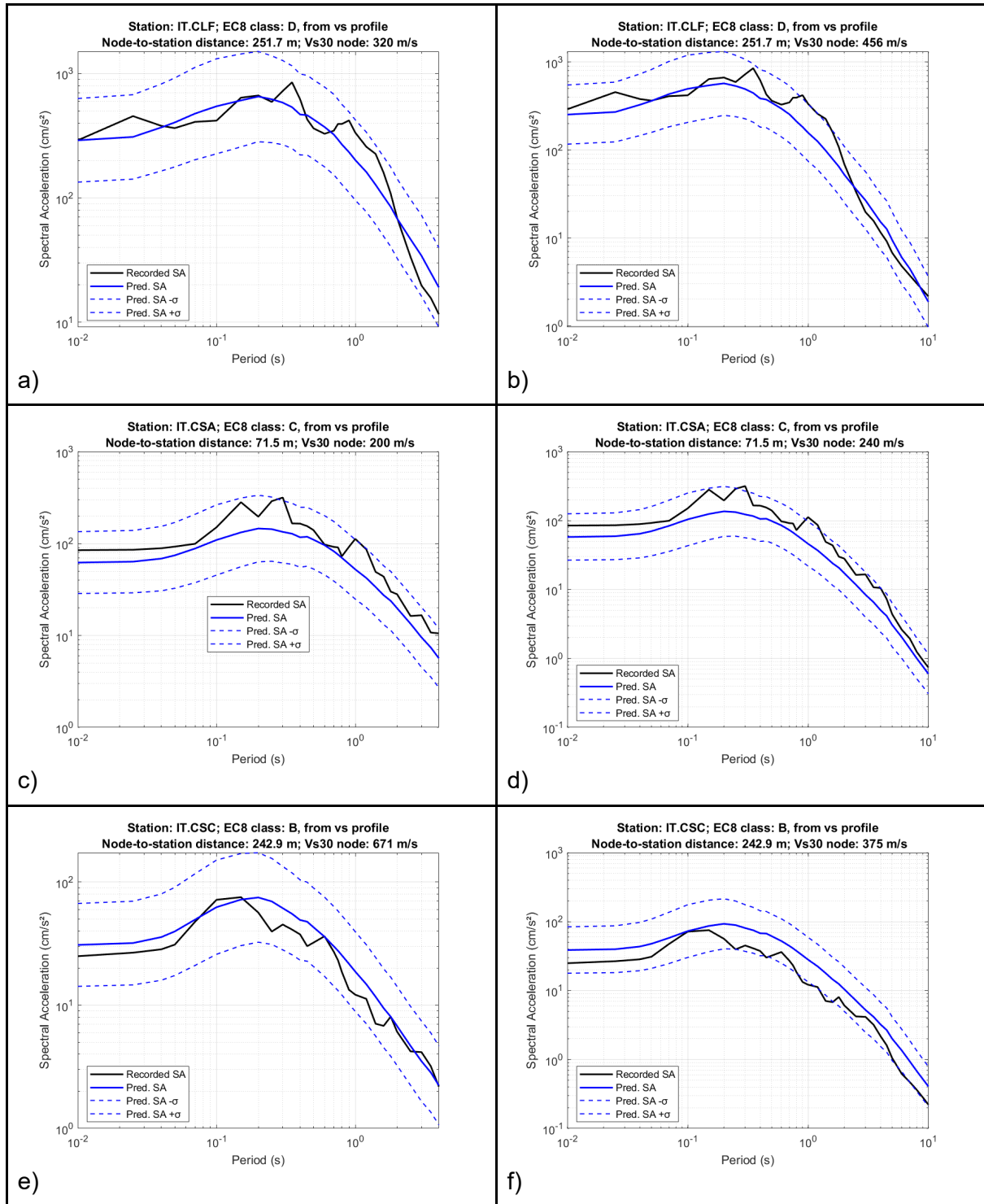


Figure 69. Predicted acceleration response spectra (in blue) overlaid with the observed spectra (in black) for three recording stations: Colfiorito (IT.CLF), panels a) and b); Castelnovo Assisi (IT.CSA), panels c) and d); and Cascia (IT.CSC), panels e) and f). Panels a), c), and e) show the predictions based on the Vs30 map developed in this study, while panels b), d), and f) use the Vs30 map proposed by Mori et al. (2020).

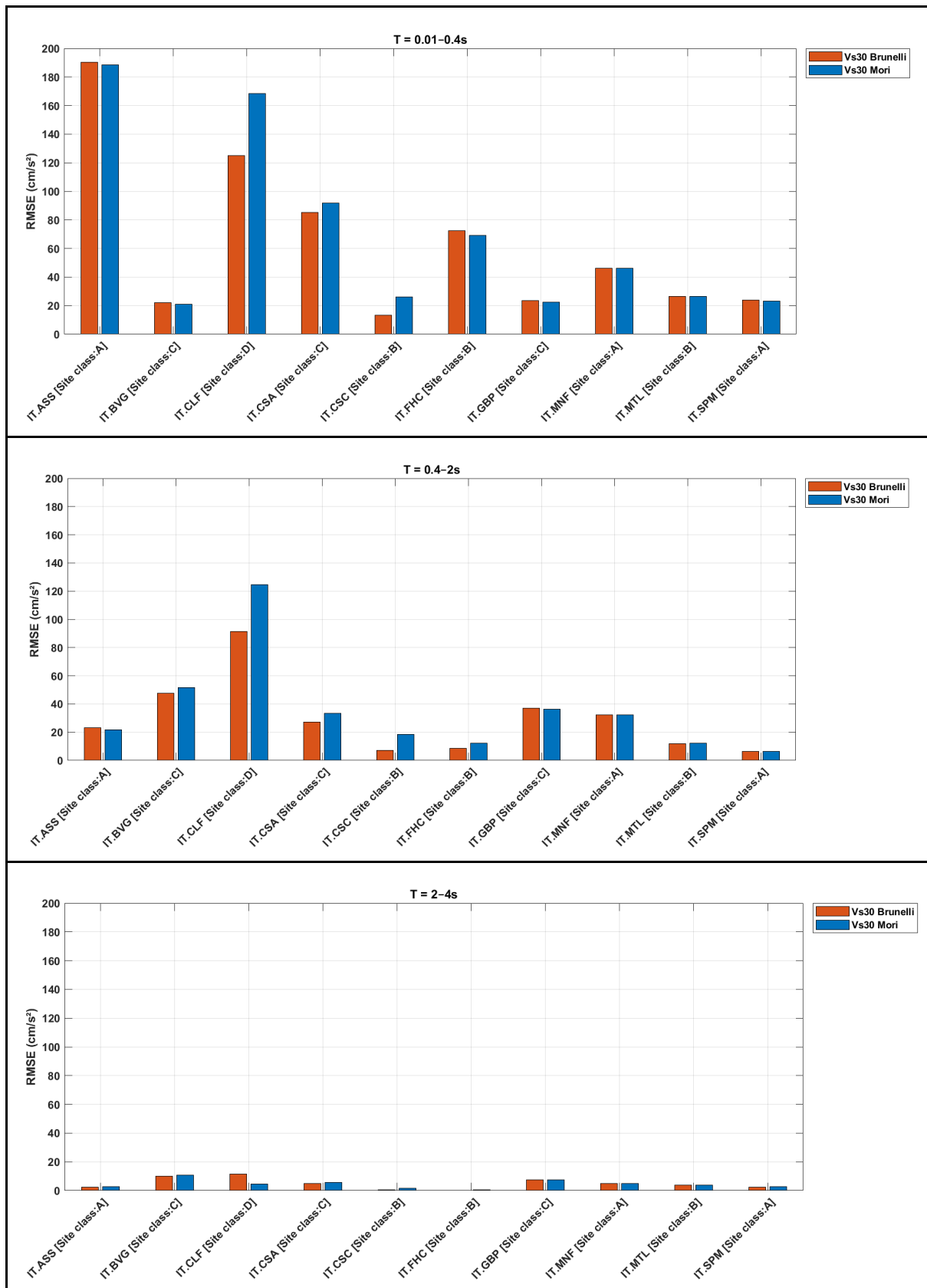


Figure 70. RMSE between observed and predicted response spectra using the two Vs30 maps, calculated for three distinct period ranges ($[0.01-0.4\text{ s}]$, $[0.4-2\text{ s}]$, $[2-10\text{ s}]$)

Table 15. Root Mean Square Error (RMSE) in cm/s^2 between predicted and observed spectral acceleration values for three-period ranges ($[0.01-0.4 \text{ s}]$, $[0.4-2 \text{ s}]$, and $[2-4 \text{ s}]$) across 10 seismic stations. The table compares the performance of two Vs30 maps: one developed in this study (Brunelli) and the other by Mori et al. (2020) (Mori). The EC8 site class and the method used to determine Vs30 (from geological information or measured profiles) are also reported for each station

Network Code	Station Code	EC8 Class	Vs30 Method	RMSE $[0.01-0.4 \text{ s}]$ (cm/s^2) Brunelli	RMSE $[0.4-2 \text{ s}]$ (cm/s^2) Brunelli	RMSE $[2-4 \text{ s}]$ (cm/s^2) Brunelli	RMSE $[0.01-0.4 \text{ s}]$ (cm/s^2) Mori	RMSE $[0.4-2 \text{ s}]$ (cm/s^2) Mori	RMSE $[2-4 \text{ s}]$ (cm/s^2) Mori
IT	ASS	A	geology	190	23	2	188	22	3
IT	BVG	C	vs profile	22	48	10	21	52	11
IT	CLF	D	vs profile	125	91	11	169	125	5
IT	CSA	C	vs profile	85	27	5	92	34	6
IT	CSC	B	vs profile	13	7	0	12	6	0
IT	FHC	C	geology	73	9	2	69	18	4
IT	GBP	C	vs profile	46	37	8	36	36	7
IT	MNF	A	vs profile	23	32	4	22	17	5
IT	SPM	A	geology	24	7	2	23	6	3

Supporting this observation, at station IT.CLF, the RMSE between the observed and predicted spectra is lower when using the newly developed Vs30 map, compared to the result obtained with the Vs30 map proposed by Mori et al. (2020) (see Figure 70 and Table 15).

Similar behavior is observed at stations IT.CSA and IT.CSC (Figure 70 and Table 15), for which the new Vs30 map assigns the measured Vs30 values, available in the ITACA database (Felicetta et al., 2023), to the grid cells corresponding to the station locations.

Lastly, it is noted that, for the long-period scenario, the RMSE was calculated up to a maximum period of 4 seconds, as data for longer periods are affected by significant noise, mainly due to the use of early analog instruments at the time of recording.

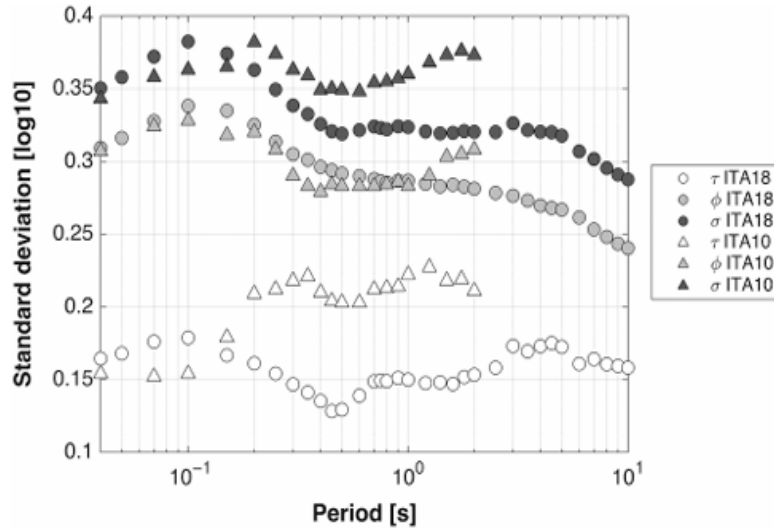


Figure 71. – Standard deviations of two ground motion models (GMMs) developed for Italy: ITA18 (Lanzano et al., 2019) and ITA10 (Bindi et al., 2011). σ (total sigma, dark gray), τ (between-event sigma, white), and ϕ (within-event sigma, light gray). Figure reproduced from Lanzano et al. (2019), with permission from the authors.

It can be noted that RMSE values at short periods (0.01-0.4s) are higher compared to long periods (Figure 70); this trend reflects the general findings observed by Lanzano et al (2019), Figure 71, who highlighted that the standard deviation of the within-event term decrease with increasing period.

In summary, at short periods, the RMSE highlights the limitations of Vs30 as a predictive parameter, as it becomes less effective in capturing the detailed complexity of local site effects.

Appendices B and C present, for each of the 11 recording stations, the predicted acceleration response spectra overlaid with the observed spectra: Appendix B refers to the predictions based on the Vs30 map developed in this study, while Appendix C uses the Vs30 map by Mori et al. (2020).

Each figure in Appendices B and C also reports: i) the corresponding EC8 site class of the recording stations, specifying whether it was derived from a measured Vs profile or inferred from geological information; ii) the distance between the station and the nearest prediction grid point; iii) and the Vs30 value at that grid point.

It is worth highlighting that the accelerometric station IT.NCR shows a marked site effect at short periods. In contrast, in Gubbio, a strong site effect is observed at long periods, likely due to the generation of surface waves within the basin (see Appendix B-C).

6.2 Simulation of the Mw 6.6 Norcia 2016 earthquake scenario

The second simulated event is the earthquake that occurred on October 30, 2016, at 06:40:18 UTC, with a magnitude of Mw 6.6 (Pondrelli et al., 2002). The focal mechanism is normal (Pondrelli et al., 2002), with a hypocentral depth of 6.2 km and epicentral coordinates of 42.838°N, 13.123°E (Spallarossa et al., 2021).

For this event, 298 strong-motion recordings are available in the ITACA database (Felicetta et al., 2023), of which 84 fall within a 50 km buffer from the epicenter. All recordings were manually processed following the approach described by Paolucci et al. (2011).

This event was selected due to its good instrumental coverage and its geographical setting near the Norcia tectonic basin (Province of Perugia), which is characterized by a carbonate bedrock buried beneath Quaternary fluvial and lacustrine sediments, overlain by coarse deposits associated with alluvial fans extending from the surrounding mountain ranges (Blumetti et al., 1990).

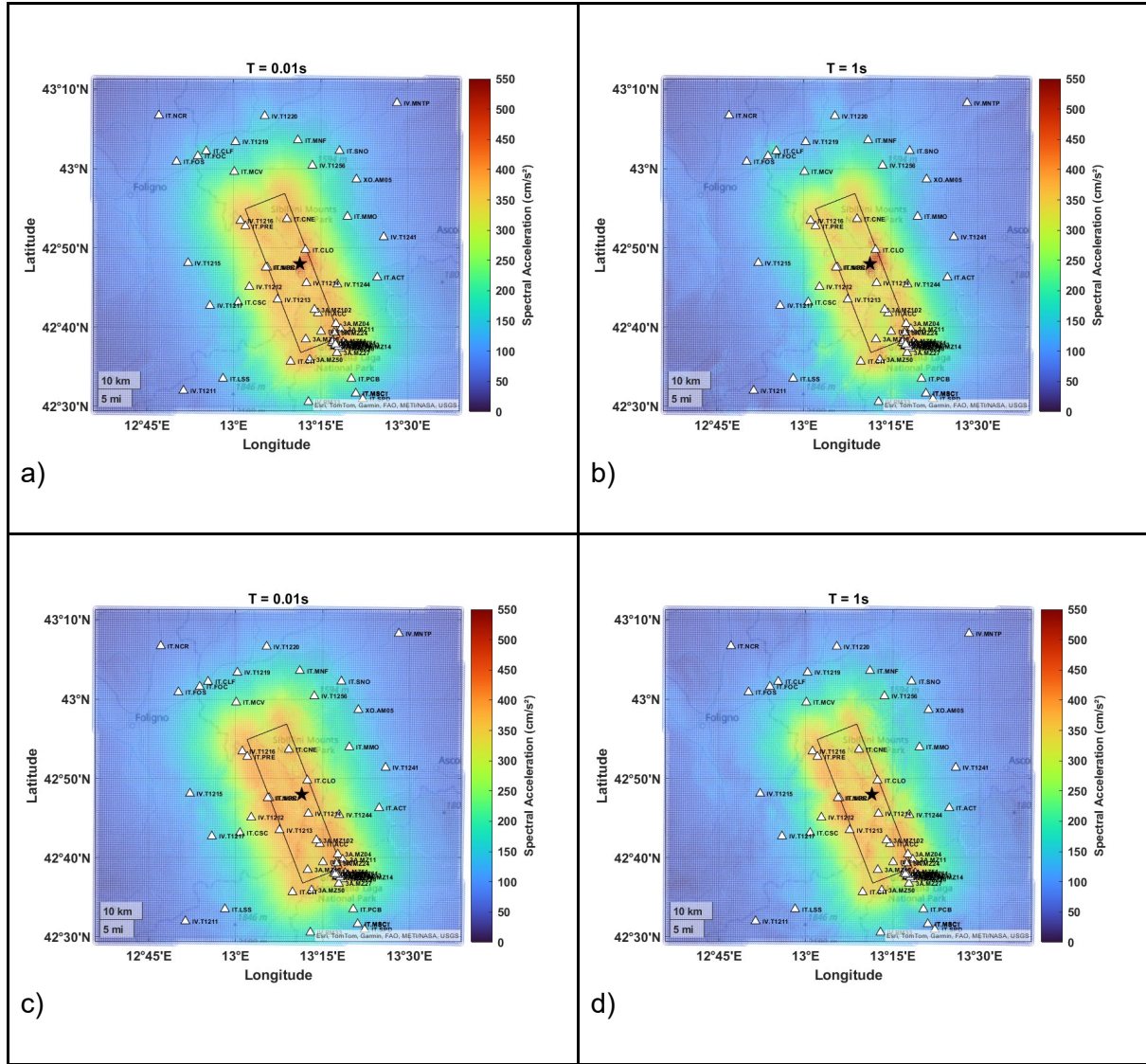
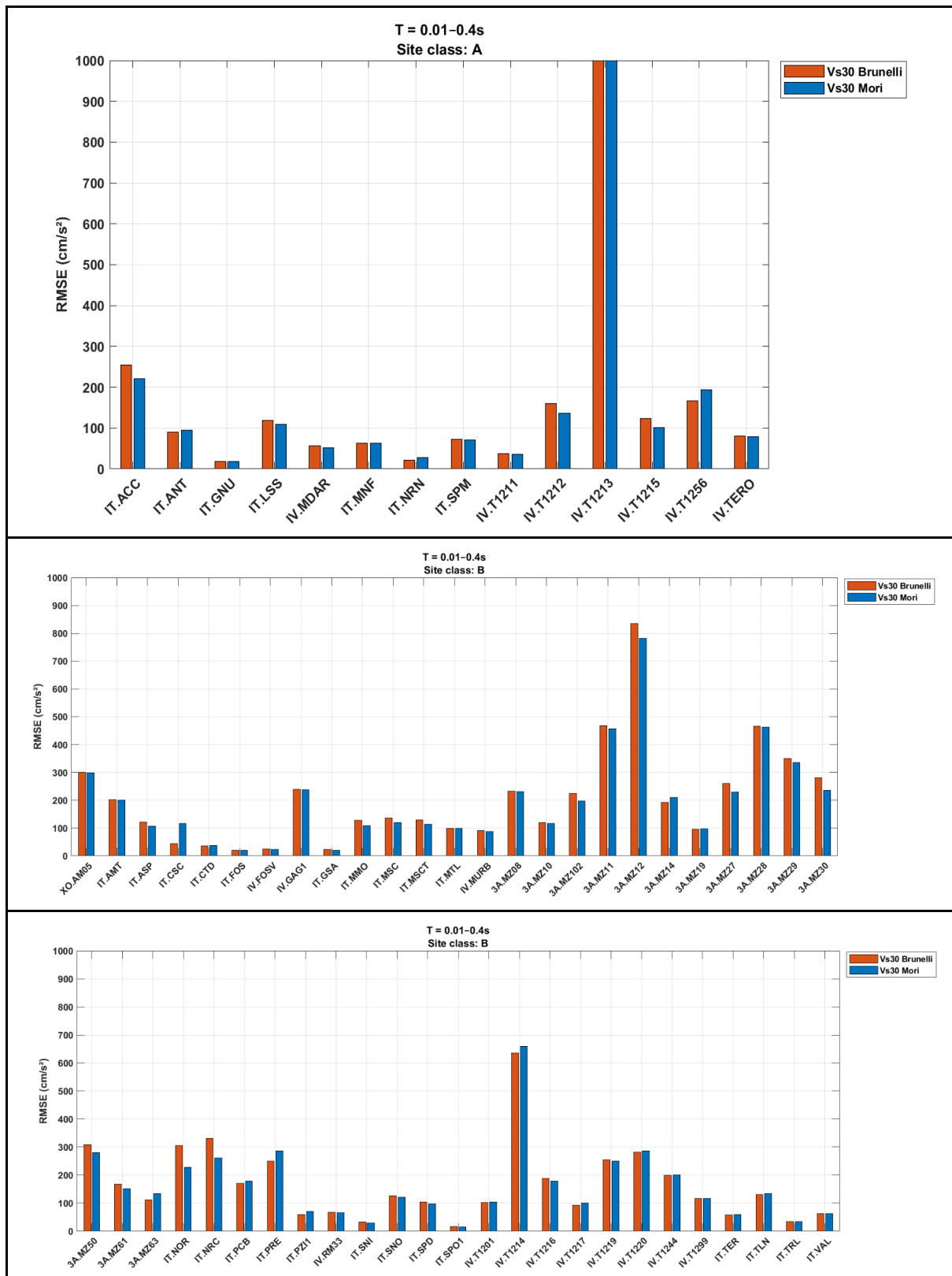


Figure 72. – Ground shaking scenarios in terms of spectral acceleration for the Mw 6.6 Norcia earthquake (ITACA event ID: EMSC-20161030_000029). Panels a) and b) display the results for 0.01 s and 1 s periods, respectively, obtained using the Vs30 map developed in this study. Panels c) and d) show the corresponding results for the same periods, computed using the Vs30 map proposed by Mori et al. (2020). The location of each recording station is shown using white triangles, along with the corresponding network and station codes.

Figure 72 shows that also in this case, the two Vs30 maps produce a very similar ground-shaking scenario at both short and long periods. Figure 73 reports the RMSE between the predicted spectra (using both Vs30 maps) and the observed spectra for short periods (0.01–0.4 s), Figure 74 for intermediate periods (0.4–2 s), and Figure 75 for long periods (2–10 s). All figures are categorized by EC8 site class.



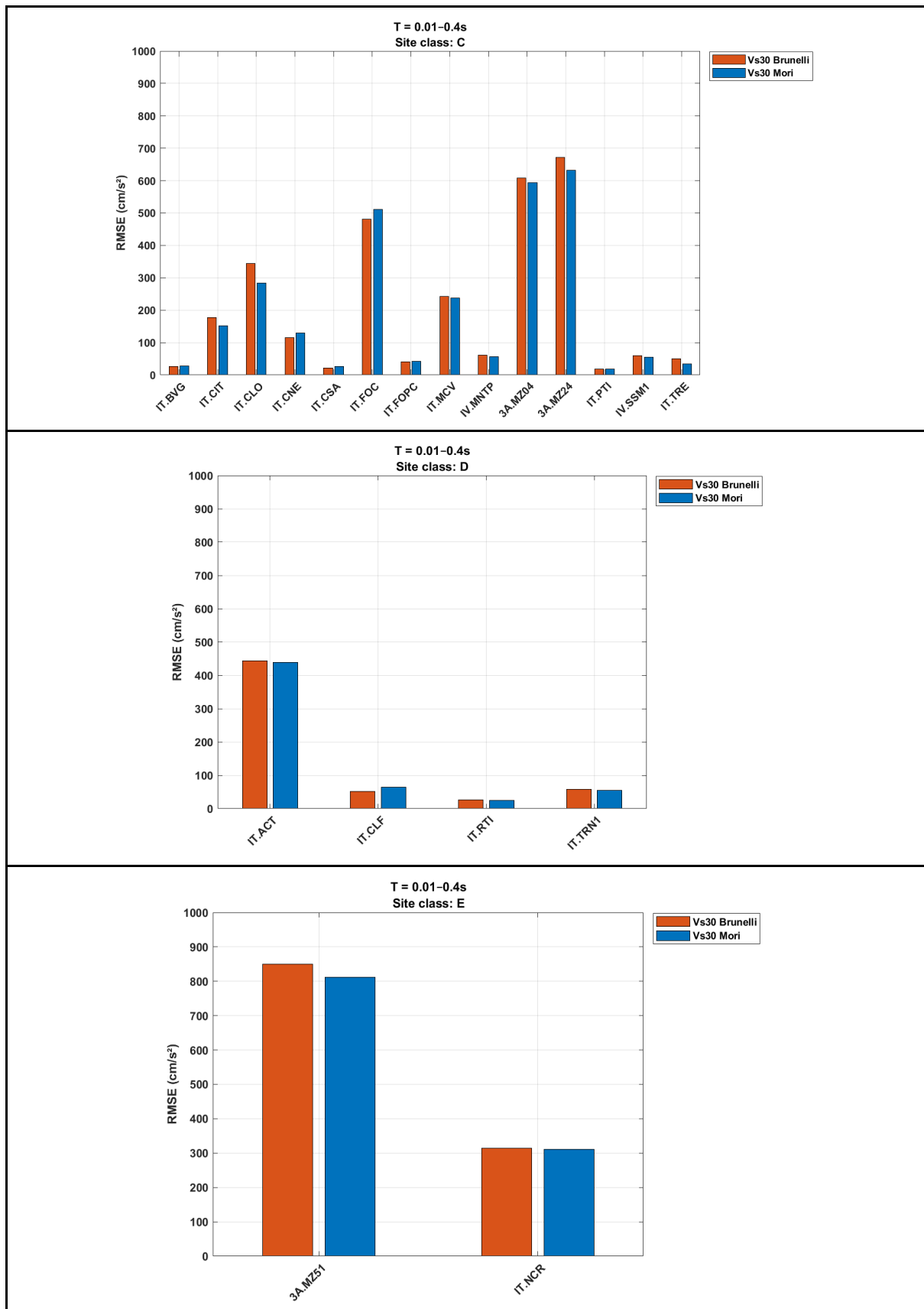
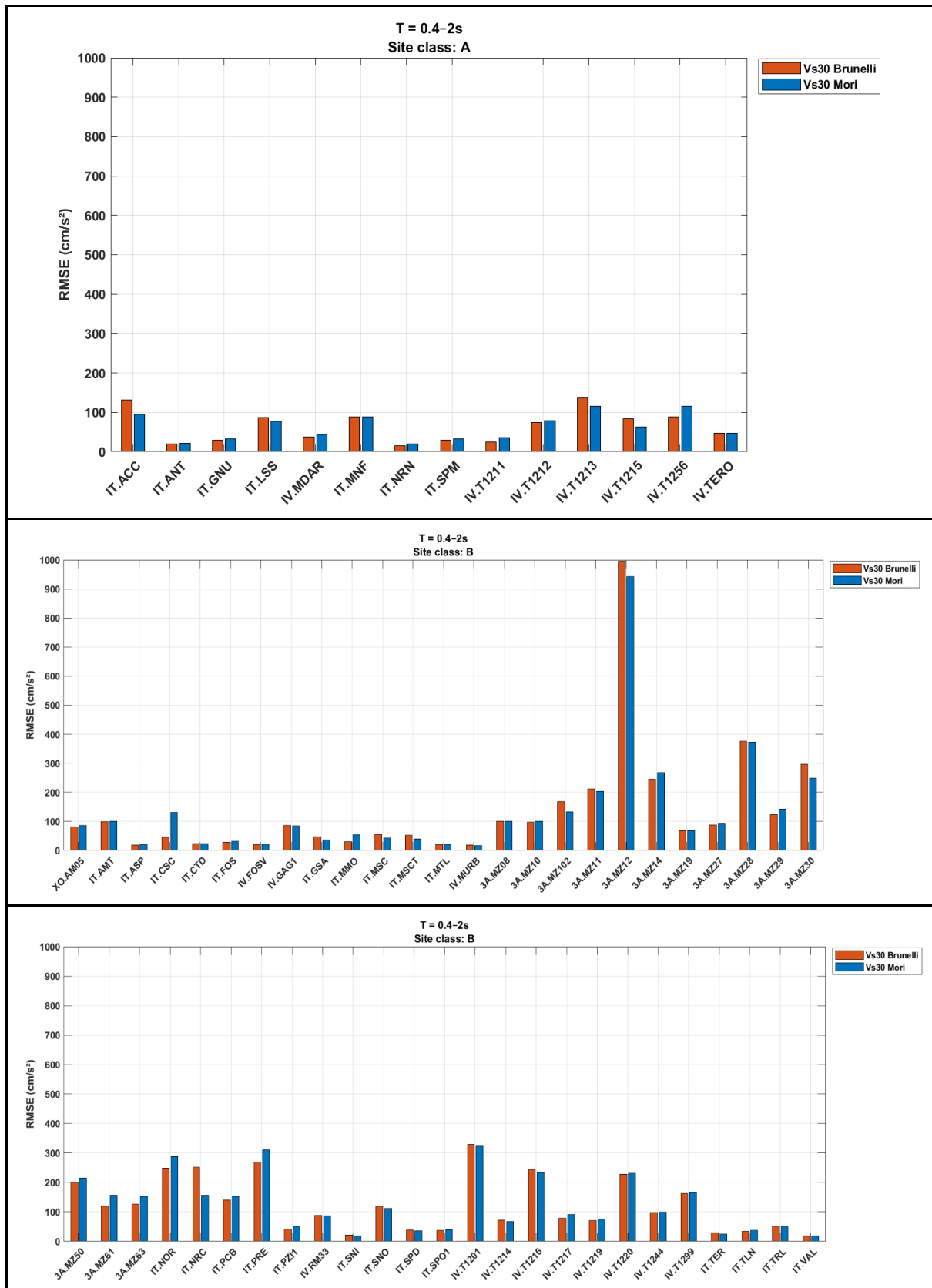


Figure 73. RMSE between observed and predicted response spectra using the two Vs30 maps, calculated for short periods [0.01–0.4 s]



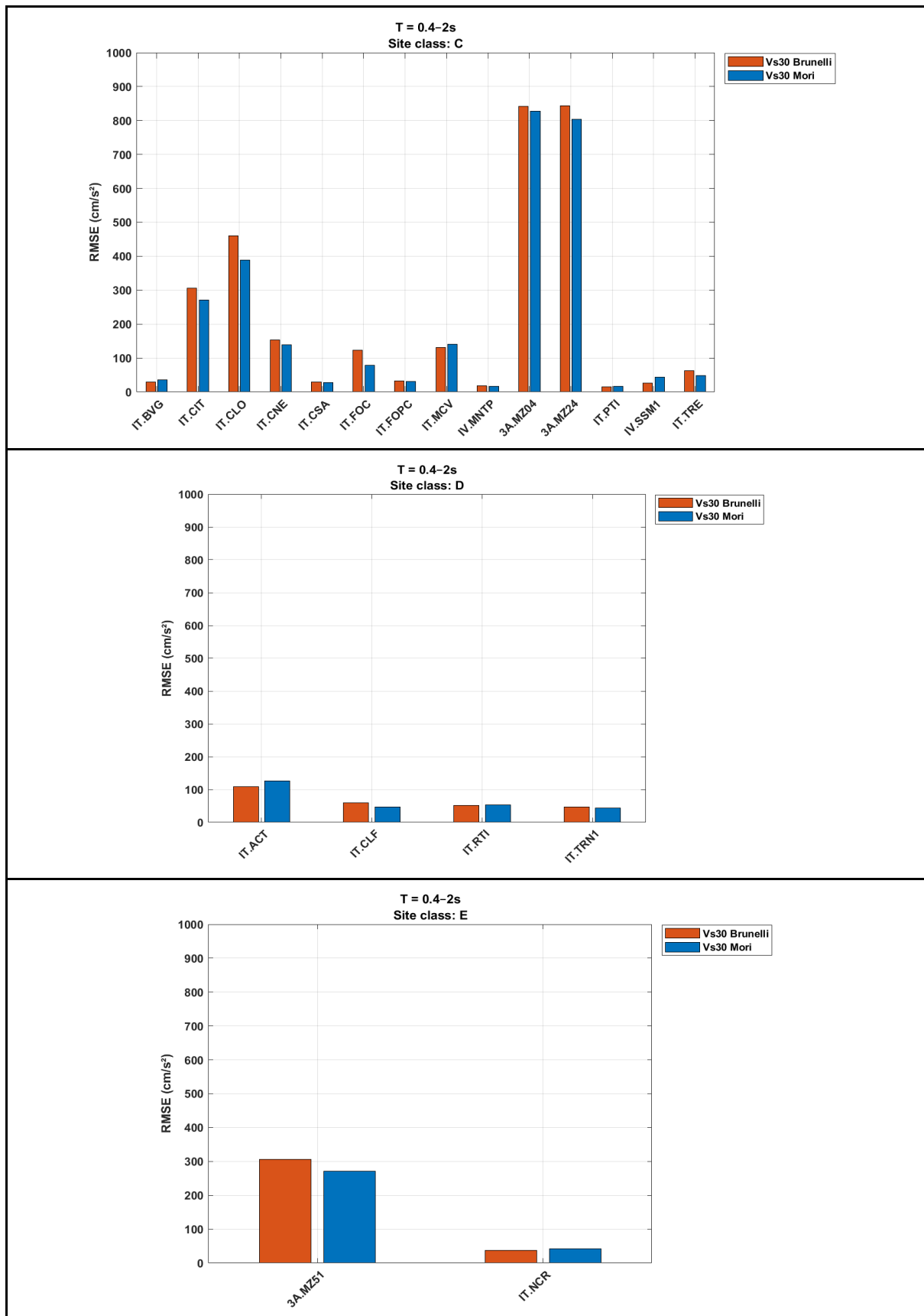
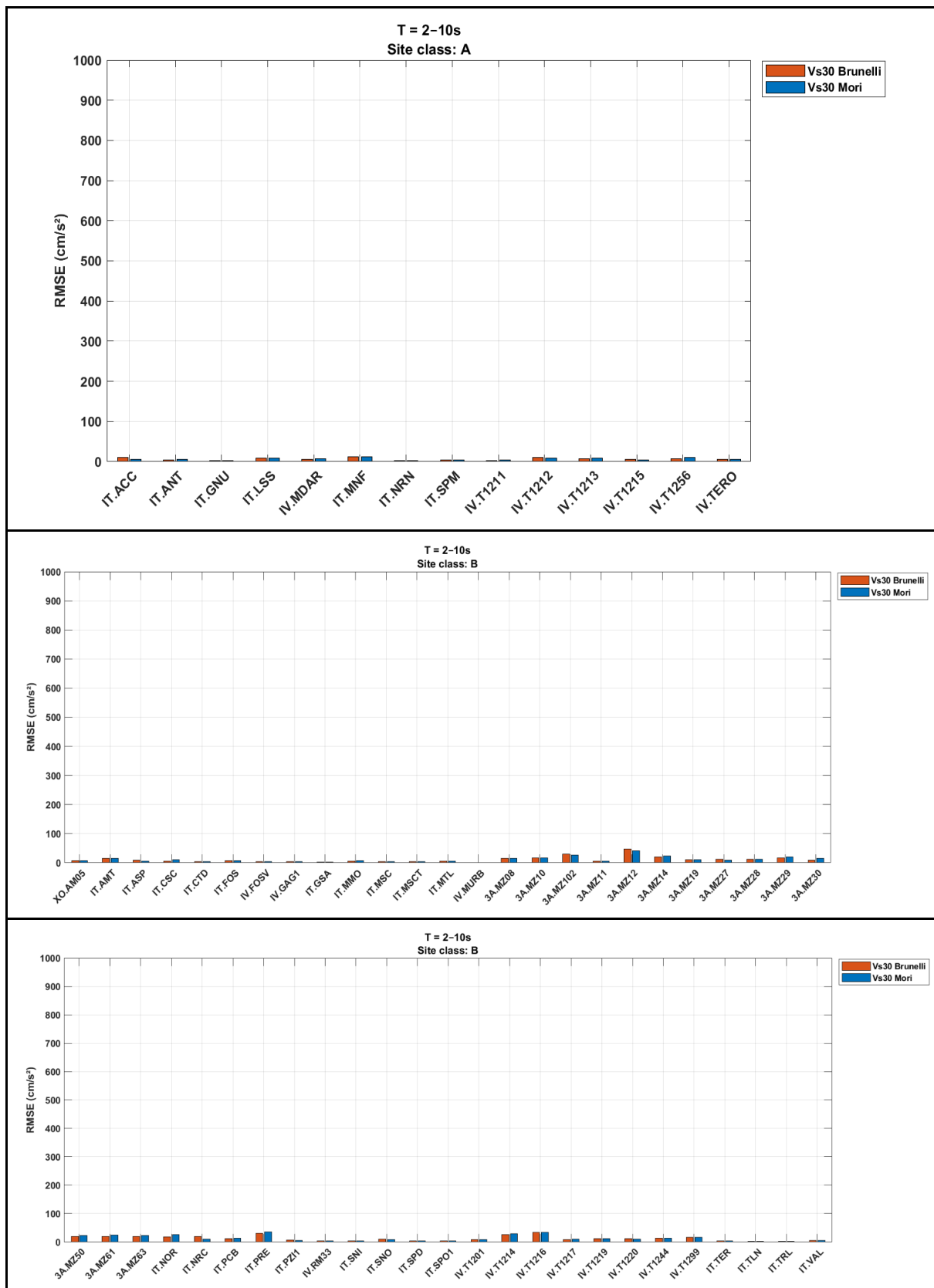


Figure 74. RMSE between observed and predicted response spectra using the two Vs30 maps, calculated for short periods [0.4–2 s].



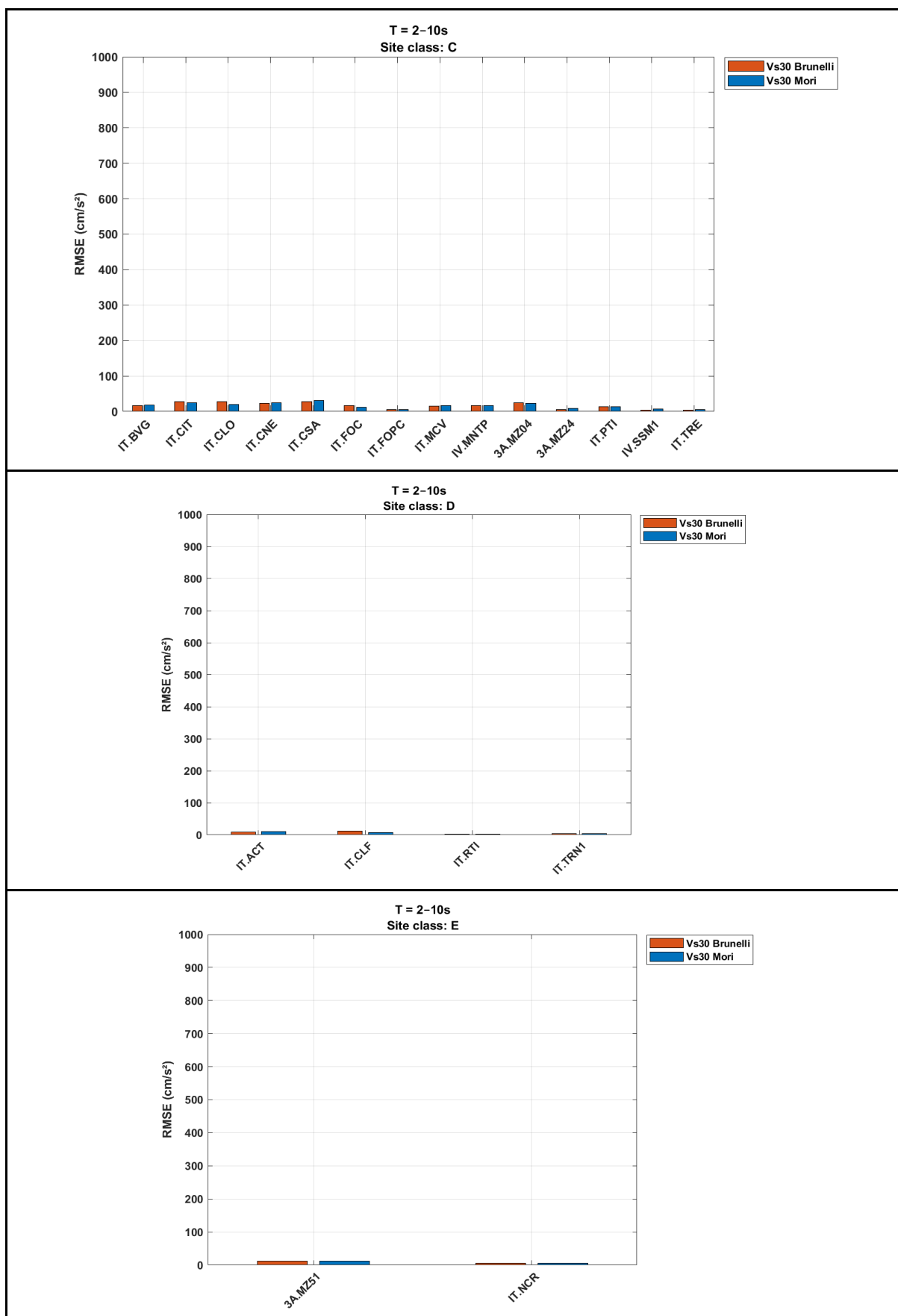


Figure 75. RMSE between observed and predicted response spectra using the two Vs30 maps, calculated for short periods [2–10 s]

A general reduction in RMSE is observed for all stations as the period increases from short to long. No significant differences are found between the predictions obtained using the two Vs30 maps across different site classes. The highest RMSE values, particularly at short periods, are observed for stations located near the source, where source effects or soil non-linearity could be present.

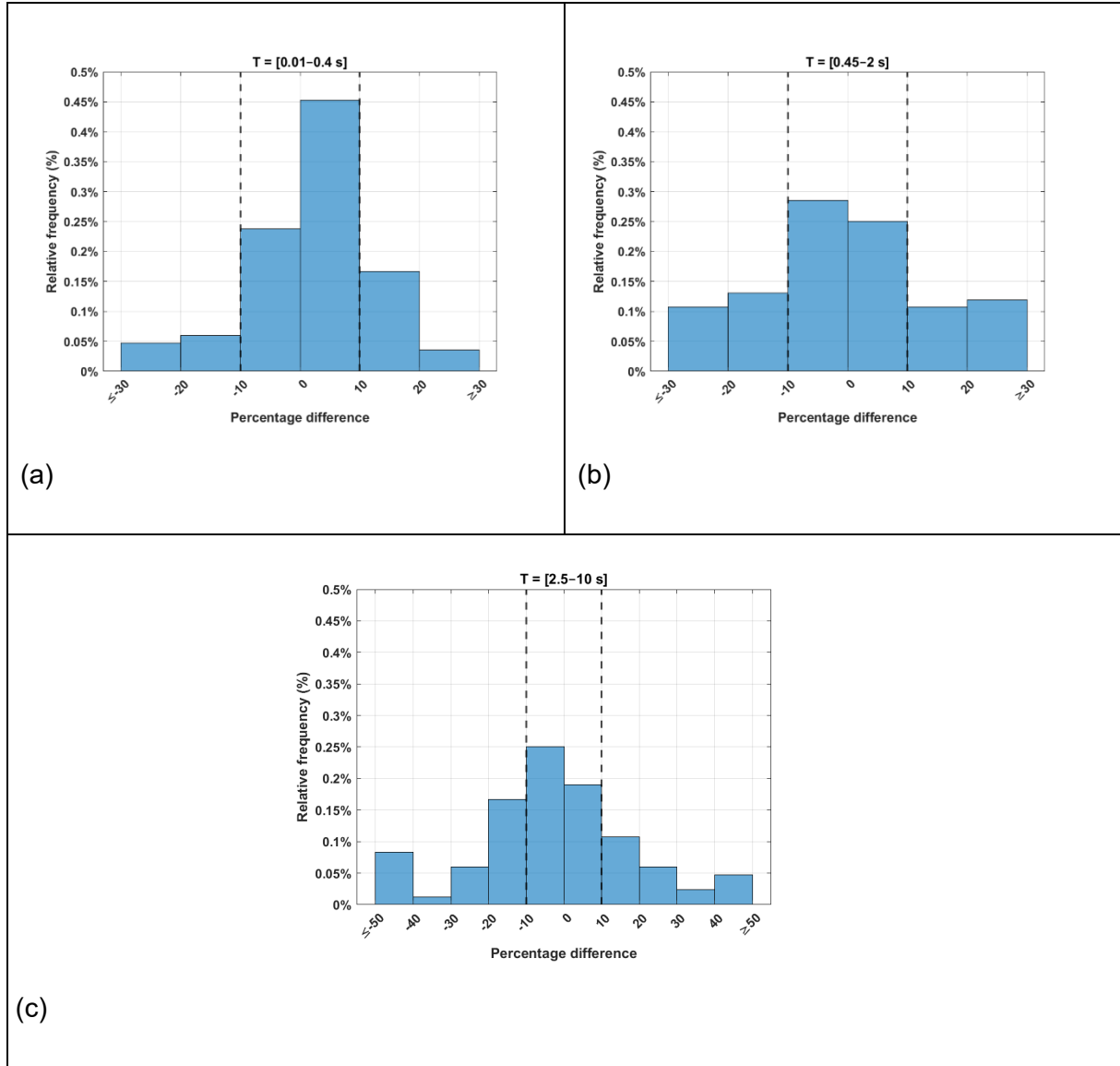


Figure 76. Histograms of difference (in percentage) between RMSE values obtained with the Brunelli and Mori Vs30 maps, shown for three spectral period intervals: [0.01–0.4 s] (a), [0.45–2 s] (b), and [2.5–10 s] (c)

Figure 76 (a), (b), and (c) shows histograms of the percentage differences between RMSE values obtained using the Brunelli Vs30 map and those calculated with the Mori Vs30 map, for three distinct period intervals: [0.01–0.4 s], [0.45–2 s], and [2.5–10 s], respectively.

These differences were computed for a total of 84 seismic stations, of which 51 do not have a measured Vs profile, while 33 are associated with direct Vs measurements. For these 33 stations, the Brunelli Vs30 map incorporates the Vs30 values derived from the measured Vs profiles, whereas the Mori map does not use this information and relies instead on a proxy-based classification.

Overall, the differences are generally small and centered around zero, indicating a good average agreement between the two maps in representing the site amplification.

At intermediate periods ([0.45–2 s], Figure 76b), the two maps are equivalent in terms of RMSE: approximately 60% of the stations show percentage differences within $\pm 10\%$, reflecting the stability of the predictions in this period range.

At long periods ([2.5–10 s], Figure 76c), the distribution shifts slightly towards negative values: about 10% of the stations exhibit percentage differences $\leq -30\%$, suggesting that the Brunelli Vs30 map yields significantly lower RMSE values than the Mori map, and thus shows better agreement with the observed data in this period range.

At short periods ([0.01–0.4 s], Figure 76a), the distribution is more asymmetric. Approximately 45% of the predictions show positive percentage differences, but these are mostly contained within 10%, indicating that, in these cases, the Brunelli Vs30 map tends to produce slightly higher RMSE values compared to the Mori map. This trend is also observed at stations with measured Vs profiles, suggesting that local differences in site characterization can influence predictive performance at high frequencies.

In summary, the two maps exhibit comparable overall performance, with a slight advantage for the Brunelli Vs30 map at long periods, and limited but more variable differences at short periods.

These differences are, however, clearly linked to this specific ground-motion scenario and dataset, and may vary under different seismic events and in other geographic areas where the differences between the two Vs30 maps are more significant

Appendices D and E present, for each of the 84 recording stations, the predicted acceleration response spectra overlaid with the observed spectra: Appendix D refers to the predictions based on the Vs30 map developed in this study, while Appendix E uses the Vs30 map by Mori et al. (2020). Each figure also reports the corresponding EC8 site class of the recording stations, specifying whether it was derived from a measured Vs profile or inferred from geological information; the distance between the station and the nearest prediction grid point; and the Vs30 value at that grid point.

7. Discussion and General Conclusions

This doctoral thesis developed an innovative framework for the generation of empirical correction factors for source, attenuation, and site effects with the aim of improving the spatial prediction of ground shaking across Italy.

The work is situated within the context of ground motion modeling at the national scale, with the goal of overcoming the ergodicity assumption that characterizes most traditional GMMs. The regionalization of source, path, and site effects was carried out through a detailed geospatial analysis and a robust statistical approach relying on a high-quality accelerometric dataset consisting of approximately 37,000 waveforms (ITACAext, Brunelli et al., 2022) and a database of around 15,000 shear-wave velocity (V_s) profiles collected throughout the Italian territory. The first step for the regionalization of source and attenuation effects involved calculating the residuals, defined as the logarithmic difference between recorded observations and model predictions, in this case using the ITA18 GMMs (Lanzano et al. 2019) as the reference model. These residuals were subsequently decomposed into repeatable components, following the methodology of Al Atik et al. (2010), into three main contributions: i) between-event term (δBe), representing source-specific effects not captured by the ground motion model, such as variations in stress drop or rupture kinematics; ii) site-to-site term ($\delta S2S$), describing the systematic local site response not accounted for by the model, often associated with specific lithostratigraphic or geomorphological conditions; iii) residual term (δWes), the residual corrected for both source and site effects, which may include additional contributions not modeled explicitly, such as regional propagation phenomena or source directivity effects.

Assuming that the event- and site-corrected residuals (δWes) reflect regional variations in anelastic attenuation not accounted for by the ground motion model, a correction factor $\delta c3$ was introduced. This parameter quantifies the adjustment needed to the anelastic attenuation coefficient $c3$ to compensate for the observed δWes on a record-by-record basis.

To explore the spatial variability of $\delta c3$, the analysis was limited to recordings with source-to-site distances exceeding 80 km, where the influence of anelastic attenuation is expected to be more pronounced (Sedaghati & Pezeshk, 2017). Each $\delta c3$ value was assigned to the midpoint of the corresponding source-to-site path, considered representative of the region most affected by wave propagation.

Subsequently, both δBe and $\delta c3$ were interpolated separately, averaging their values over two distinct period ranges: short periods ($T = 0.07\text{--}0.4$ s) and long periods ($T = 0.7\text{--}4$ s). This

procedure enabled the identification of homogeneous zones with distinct source and attenuation characteristics, with particularly evident spatial patterns at short periods.

In particular, markedly positive event residuals (δBe) are observed along the entire Alpine chain, in the Apulian platform, in south-eastern Sicily, and in the western lobe of the Ionian Fault (see Figure 14a). Conversely, strongly negative event residuals are found in northern Sicily, in the volcanic district between Tuscany and Lazio, in the Molise region, in the central Adriatic Sea, and in the eastern lobe of the Ionian Fault (see Figure 14a).

The $\delta c3$ coefficient reveals a distinct contrast in attenuation between the Adriatic side (slower) and the Tyrrhenian side (faster) along the Apennine arc, as shown in Figure 20a. In the Alpine region, attenuation patterns differ noticeably, with slower attenuation observed in the eastern sector and faster attenuation in the western portion of the arc. In Sicily, the Hyblean foreland area (southwestern part of the island) stands out for its significantly reduced attenuation compared to surrounding regions. Additionally, elevated $\delta c3$ values are found in the Po Valley.

Starting from the seismogenic zonations ZS16 (Visini et al., 2022) and the zonation developed by (Basili et al., 2018) within the framework of the TSUMAPS-NEAM project, the zones were suitably modified to capture the spatial patterns observed at short period in the event residuals (δBe) and in the attenuation coefficient ($\delta c3$), respectively, and for each modified zone, average correctives were then computed to enable a more accurate regional modeling of source and attenuation effects.

Subsequently, site effects were investigated in greater detail through the development of a new national map of shear-wave velocity in the uppermost 30 meters ($Vs30$), based on a dataset of approximately 15,000 Vs profiles collected across the Italian territory. The map was generated using parametric linear regressions and geostatistical analysis, after carefully analyzing the correlation between $Vs30$ and several site proxies available at different spatial resolutions. A new lithological classification was also developed and, together with terrain slope, was used as a key predictor in the final $Vs30$ model.

During the construction process, some challenges were encountered, particularly regarding rock lithologies. In several cases, the estimated mean $Vs30$ values for these units were lower than expected, likely due to the inclusion of velocity profiles from soft sediments mistakenly associated with rock polygons and, in general, very few measurements were available. Future refinements may involve the use of more detailed regional geological maps than the 1:100,000 scale, to represent soil units with finer detail and correctly assign the $Vs30$ measurements.

In general, compared to the maps by Mori et al. (2020) and Wald and Allen (2007), the map presented here preserves the measured Vs30 and aligns better with observations, due to the integration of interpolated residuals and a reduced dependence on slope, differently from the WA model.

Subsequently, as a practical example, two ground shaking scenarios were generated for two different periods of the acceleration response spectrum (0.01 s and 1 s), related to two seismic events that occurred in central Italy: the earthquake of September 29, 1997, near the Colfiorito basin, and the one of October 30, 2016, near the Norcia basin. The scenarios were produced using two different Vs30 maps—one proposed by Mori et al. (2020) and the one developed in this study—in order to test the influence of the site effects. To demonstrate the robustness of the model, for both events, no information on the 3D fault geometry was assumed considering the Joyner-Boore distance. The simulations were restricted to a 50 km buffer around the epicenter, and no correction was applied for either path attenuation or source effects.

While this approach is effective under limited information conditions, it presents several *limitations*:

1. Inability to account for near-source effect and directivity.
2. Site effects modeled only through Vs30: Although Vs30 is a widely adopted proxy, it does not always capture the actual local site response, especially in cases involving deep bedrock or resonant effects not linked to the average velocity in the upper 30 meters.
3. Inability to generate ground motion time series: The model provides only response spectra and does not simulate synthetic accelerograms, which are often required in advanced engineering applications.

Despite these limitations, the approach offers several clear *advantages*:

1. Computational efficiency: The method enables rapid generation of ground-shaking scenarios, making it suitable for both real-time applications and large-scale simulations.
2. Compatibility with 3D fault geometry: When available, detailed information on the fault geometry can be incorporated into the model, improving near-source predictions.

3. Flexibility for future developments: The framework is designed to accommodate additional empirical corrections, such as $\delta S2S$ terms, near-source effects, and other refinements as new data or methods become available.

Future developments

To improve the representation of *site effects*, several strategies for future development can be considered:

1. Although $Vs30$ is widely used to represent site effects, its performance varies with spectral period. RMSE plots (see Figure 75) show lower values at long periods for rock sites (EC8 class A) compared to sites with thick sedimentary layers (class C), suggesting that $Vs30$ alone may be insufficient to fully capture site response under deep-soil conditions. Conversely, $Vs30$ appears to better describe site effects at short periods, although the higher RMSE observed in this range (see Figure 70) may be more attributable to the greater standard deviation (σ) of ground-motion models at short periods (see Figure 71) than to limitations in the $Vs30$ proxy itself. To improve the modeling of site response, future developments could explore alternative proxies such as $Vs5$, $Vs10$, or the site fundamental frequency (f_0), which may enhance model performance, particularly at short periods.
2. Increasing the number of observed $Vs30$ profiles is also crucial. Since the current regression-based method is highly influenced by available measurements, a denser distribution of profiles would enhance the reliability of the resulting $Vs30$ map.
3. Alternative modeling strategies, such as machine learning techniques, could be explored to model $Vs30$ more effectively, allowing the integration of multiple site proxies and the ability to capture complex, non-linear relationships among variables.
4. Finally, the use of empirical site terms such as $\delta S2S$ could be explored in place of $Vs30$. A comparative study between $\delta S2S$ and transfer functions derived from 1D numerical modeling (e.g., response spectrum ratios) may help assess whether these approaches can outperform $Vs30$ in capturing local site effects.

As for *source* modeling, future improvements could involve:

1. Incorporating 3D fault geometries into the simulation process—when available—and adopting more physical finite-fault distance metrics, such as those proposed in ITA18 by Lanzano et al. (2019), including the rupture distance (R_{rup}), would allow for a more realistic reproduction of near-fault ground motion effects.

Several future developments and considerations can be envisaged to further enhance the *framework*:

1. Since the two V_{s30} maps (this study and Mori et al., 2020) yield very similar results across most of central Italy—except in basin areas—it would be valuable to apply the methodology to other regions where the two maps show more pronounced differences. This would allow for a more meaningful assessment of the impact of different V_{s30} models on ground motion predictions.
2. Including source directivity through empirical correction such as the one proposed by Colavitti et al. (2022) and near-source global empirical corrections, like those proposed by Sgobba et al. (2021).
3. The empirical corrective factors developed in this study for source effects (δB_e) and anelastic attenuation (δc_3), together with the updated V_{s30} map for site effects, could be applied to simulate historical earthquakes for which no strong-motion recordings are available. This would allow for the reconstruction of past ground-shaking scenarios in a physically grounded and data-consistent way.
4. A promising direction involves shifting from response spectra to Fourier amplitude spectra, recalculating the empirical corrections accordingly. This would allow the use of stochastic methods to generate synthetic time series (e.g., multiple realizations per grid point), which are often required for engineering applications and risk analyses.
5. Finally, the development of a custom shakemap-like system for real earthquakes could be explored. This would involve adding the between-event correction of the actual

event, the δc_3 correction for anelastic attenuation from this study, V_{s30} to account for site effects and interpolated ground motion residuals via kriging to the baseline ground motion prediction.

In summary, this method offers an efficient and flexible framework for generating ground-shaking scenarios of real or virtual earthquakes, with the potential for significant improvement through the integration of empirical corrections and more detailed site characterization.

References

- Abrahamson NA, Kuehn NM, Walling M, Landwehr N. Probabilistic seismic hazard analysis in California using nonergodic ground-motion models. *Bull Seismol Soc Am* 2019;109(4):1235–49.
- Ahdi, S.K., Stewart, J.P., Ancheta, T.D., Kwak, D.Y., Mitra, D., 2017. Development of Vs profile database and proxy-based models for Vs30 prediction in the pacific northwest region of North America. *Bull. Seismol. Soc. Am.* <https://doi.org/10.1785/0120160335>.
- Aki K. Scaling law of seismic spectrum. *Journal of geophysical research*. 1967 Feb 15;72(4):1217-31.
- Aki, K., & Richards, P. G., 2002. *Quantitative Seismology*. University Science Books.
- Akinci A, Mueller C, Malagnini L, Lombardi A. Seismic hazard estimate in the Alps and Apennines (Italy) using smoothed historical seismicity and regionalized predictive ground-motion relationships. *Boll Geofis Teor Appl* 2004; 45:285–304.
- Akinci A, Munafò I, Malagnini L. S-Wave attenuation variation and its impact on ground motion amplitudes during 2016-2017 Central Italy earthquake sequence. *Front Earth Sci Solid Earth Geophys* 2022. <https://doi.org/10.3389/feart.2022.903955>.
- Akkar, S., Bommer, J.J., 2010. Empirical equations for the prediction of PGA, PGV, and spectral accelerations in Europe, the mediterranean region, and the Middle East. *Seismol. Res. Lett.* <https://doi.org/10.1785/gssrl.81.2.195>.
- Al-Atik L, Abrahamson NA, Bommer JJ, Scherbaum F, Cotton F, Kuehn N. The variability of ground-motion prediction models and its components. *Seismol Res Lett* 2010;81(5):794–801. <https://doi.org/10.1785/gssrl.81.5.794>.
- Allen, T.I., Wald, D.J., 2009. On the use of high-resolution topographic data as a proxy for seismic site conditions (VS30). *Bull. Seismol. Soc. Am.* <https://doi.org/10.1785/0120080255>.
- Amanti, M., Battaglini, L., Campo, V., Cipolloni, C., Congi, M. P., Conte, G., Delogu, D., Ventura, R., and Zonetti, C.: La carta litologica d'italia alla scala 1:100000, in: *Atti della*

- 11a Conferenza Nazionale ASITA, Turin, Italy, 6–9 November 2007, <http://atti.asita.it/Asita2007/Pdf/119.pdf> (last access: 14 January 2022), 2007.
- Amanti, M., Battaglini, L., Campo, V., Cipolloni, C., Congi, M. P., Conte, G., Delogu, D., Ventura, R., and Zonetti, C.: The Lithological map of Italy at 1:100000 scale: An example of re-use of an existing paper geological map, in: 33rd International Geological Conference, Oslo, Norway, 6–14 August 2008, IE102310L, <http://www.33igc.org> (last access: 14 January 2022), 2008.
- Ameri G, et al. The 6 april 2009, Mw 6.3, L'Aquila (Central Italy) earthquake: strong-motion observations. *Seismol Res Lett* 2009;80(6):951–66.
- Anderson H, Jackson J. Active tectonics of the adriatic region. *Geophys J Int* 1987; 91(3):937–83.
- Anderson JG. Seismic energy and stress-drop parameters for a composite source model. *Bull Seismol Soc Am* 1997;87(1):85–96.
- Anderson JG, Brune JN. Probabilistic seismic hazard analysis without the ergodic assumption. *Seismol Res Lett* 1999;70(1):19–28. <https://doi.org/10.1785/gssrl.70.1.19>.
- Anderson J, Uchiyama Y. A methodology to improve ground-motion prediction equations by including path corrections. *Bull Seismol Soc Am* 2011;101(4): 1822–46. <https://doi.org/10.1785/0120090359>.
- Baltay AS, Hanks TC, Abrahamson NA. Uncertainty, variability, and earthquake physics in ground-motion prediction equations. *Bull Seismol Soc Am* 2017;107(4): 1754–72. <https://doi.org/10.1785/0120160164>.
- Baltay AS, Hanks TC, Abrahamson NA. Earthquake stress drop and arias intensity. *J Geophys Res Solid Earth* 2019;124(4):3838–52.
- Bard PY. Microtremor measurements: a tool for site effect estimation. The effects of surface geology on seismic motion. 1999 Dec; 3:1251-79.
- Bard, P.Y. and Bouchon, M., 1985. The two-dimensional resonance of sediment-filled valleys. *Bulletin of the Seismological Society of America*, 75(2), pp.519-541.
- Basili R, Brizuela B, Herrero A, Iqbal S, Lorito S, Maesano FE, Murphy S, Perfetti P, Romano F, Scala A, Selva J, Taroni M, Thio HK, Tiberti MM, Tonini R, Volpe M, Glimsdal S, Harbitz CB, Løvholt F, Baptista MA, Carrilho F, Matias LM, Omira R, Babeyko A,

- Hoechner A, Gurbuz M, Pekcan O, Yalçiner A, Canals M, Lastras G, Agalos A, Papadopoulos G, Triantafyllou I, Benchechrout S, Agrebi Jaouadi H, Attafi K, Ben Abdallah S, Bouallegue A, Hamdi H, Oueslati F. NEAM tsunami hazard model 2018 (NEAMTHM18): online data of the probabilistic tsunami hazard model for the NEAM region from the TSUMAPS-NEAM project. Istituto Nazionale di Geofisica e Vulcanologia (INGV); 2018. <https://doi.org/10.5281/zenodo.3406625>.
- Bassi G, Sabadini R, Rebaï S. Modern tectonic regime in the Tyrrhenian area: observations and models. *Geophys J Int* 1997; 129:330–46.
- Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Software* 2015;67(1):1–48.
- Bigi, G., Castellarin, A., Catalano, R., Coli, M., Cosentino, D., Dal Piaz, G.V., Lentini, F., Parotto, M., Patacca, E., Pratlun, A. and Salvini, F., 1991. Synthetic structural-kinematic Map of Italy. Time of main alpidic deformations and of related sedimentary metamorphic and magmatic processes, Scale 1: 2.000. 000.
- Bindi D, Pacor F, Luzi L, Puglia R, Massa M, Ameri G, Paolucci R. Ground motion prediction equations derived from the Italian strong motion database. *Bulletin of Earthquake Engineering*. 2011 Dec; 9:1899-920.
- Bindi D, Spallarossa D, Pacor F. Between-event and between-station variability observed in the Fourier and response spectra domains: comparison with seismological models. *Geophys J Int* August 2017; 210(2):1092–104. <https://doi.org/10.1093/gji/ggx217>.
- Bindi D., Spallarossa D., Picozzi M., Scafidi D. and Cotton F.; 2018: Impact of magnitude selection on aleatory variability associated with Ground-Motion Prediction Equations: Part I - Local, energy, and moment magnitude calibration and stress-drop variability in central Italy. *Bull. Seismol. Soc. Am.*, 108, 1427–1442.
- Bindi D, Kotha SR. Spectral decomposition of the Engineering Strong Motion (ESM) flat file: regional attenuation, source scaling and Arias stress drop. *Bull Earthq Eng* 2020;18(6):2581–606.
- Bivand Roger S., Pebesma Edzer, Gomez-Rubio Virgilio, 2013. Applied spatial data analysis with R, Second edition. Springer, NY. <https://asdar-book.org/>

- Blumetti, A. M., Coltorti, M., & Farabollini, P. (1990). Two stratigraphic sections in the Middle Pleistocene of the Conca of Norcia: geomorphological and neotectonic implications. *BOLLETTINO DELLA SOCIETÀ GEOLOGICA ITALIANA*, 13, 17-26 (in Italian).
- Boatwright J. A spectral theory for circular seismic sources; simple estimates of source dimension, dynamic stress drop, and radiated seismic energy. *Bulletin of the Seismological Society of America*. 1980 Feb 1;70(1):1-27.
- Bonomo R., Capotorti F., D'Ambrogi C., Di Stefano R., Graziano R., Martarelli L., Pampaloni M. L., Pantaloni M., Ricci V., Compagnoni B., Galluzzo F., Tacchia D., Masell, G., Pannuti V., Ventura R. and Vitale, V.: Carta geologica d'Italia alla scala 1:1250000, Serv. Geol. d'It., APAT, Roma, 2005
- Bommer JJ, Douglas J, Strasser FO. Style-of-faulting in ground-motion prediction equations. *Bulletin of Earthquake Engineering*. 2003 May; 1:171-203.
- Boore, D. M. Simulation of ground motion using the stochastic method. *Pure and Applied Geophysics*, 160(3-4), 635-676; 2003. <https://doi.org/10.1007/PL00012553>
- Boore DM. Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. *Bulletin of the Seismological Society of America*. 2010 Aug 1;100(4):1830-5.
- Boore DM, Thompson EM, Cadet H. Regional correlations of VS 30 and velocities averaged over depths less than and greater than 30 meters. *Bulletin of the Seismological Society of America*. 2011 Dec 1;101(6):3046-59.
- Borcherdt RD, Glassmoyer G. Influences of local geology on strong and weak ground motions recorded in the San Francisco Bay region and their implications for site-specific building-code provisions. *US geological survey professional paper*. 1994:77-108.
- Bucci F; Santangelo M; Fongo L; Alvioli M; Cardinali M; Melelli L; Marchesini I (2021): A new digital lithological Map of Italy at 1:100.000 scale [dataset]. PANGAEA, <https://doi.org/10.1594/PANGAEA.935673>
- Bucci F., Santangelo M., Fongo L., Alvioli M., Cardinali M., Melelli L., Marchesini I.; A new digital lithological map of Italy at the 1: 100 000 scale for geomechanical modelling. *Earth System Science Data*. 2022 Sep 7;14(9):4129-51.

- Bradley BA, Huang J, Motha J, Tarbali K, Lee R, Bae S, Polak V, Zhu M, Schill C, Paterson J, Lagrava D. Cybershake NZ v19.5: New Zealand simulation-based probabilistic seismic hazard analysis. Wellington, New Zealand: New Zealand Society for Earthquake Engineering Conference; 2020.
- Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research*, 75(26), 4997-5009. <https://doi.org/10.1029/JB075i026p04997>
- Brune, J.N., 1971. Correction. *J. Geophys. Res.*, 76, 5002.
- Brunelli G, Lanzano G, D'Amico MC, Felicetta C, Luzi L, Mascandola C, Pacor F, Russo E, Sgobba S. V.3.2 ITACAext flatfile. 2022. https://doi.org/10.13127/itaca32/itacaext_flatfile.1.0.
- Brunelli G, Lanzano G, Luzi L, Sgobba S. Data-driven zonations for modelling the regional source and propagation effects into a Ground Motion Models in Italy. *Soil Dynamics and Earthquake Engineering*. 2023 Mar 1; 166:107775.
- Brunelli G., Lanzano G., Sgobba S. and Luzi L.; 2023: Spatial variability of non-ergodic GMM residuals related to source and path effects in Italy. *Bull. Geoph. Ocean.*, 64, 417-432, doi: 10.4430/bgo00425.
- Carletti F, Gasperini P. Lateral variations of seismic intensity attenuation in Italy. *Geophys J Int* 2003; 155:839–56.
- Castro RR, Mucciarelli M, Monachesi G, Pacor F, Berardi R. A review of nonparametric attenuation functions computed for different regions of Italy. *Ann Geophys* 1999;42(4). <https://doi.org/10.4401/ag-3753>.
- CEN, 2004. EN 1998-1, 2004 (English): Eurocode 8: Design of structures for earthquake resistance– Part 1: General rules, seismic actions and rules for buildings [Authority: The European Union Per Regulation 305/2011]. (Directive 98/34/EC, Directive 2004/18/EC).
- Chiaraluce L., Amato A., Cocco M., Chiarabba C., Selvaggi G., Di Bona M., Piccinini D., Deschamps A., Margheriti L., Couboulex F., Ripepe M. (2004). Complex Normal Faulting in the Apennines Thrust-and-Fold Belt: The 1997 Seismic Sequence in Central Italy, *Bull. Seism. Soc. Am.*, 94, 99–116.

- Chiaraluce L., Michele M., Waldhauser F., Tan Y.J., Herrmann M., Spallarossa D., Beroza G.C., Cattaneo M., Chiarabba C., De Gori P., Di Stefano R., Ellsworth W., Main I., Mancini S., Margheriti L., Marzocchi W., Meier M.A., Scafidi D., Schaff D. and Segou M.; 2022: A comprehensive suite of earthquake catalogues for the 2016-2017 central Italy seismic sequence. *Sci. Data*, 9, 710, 12 pp., doi: 10.1038/s41597-022-01827-z.
- Cipolloni, C., Pantaloni, M., Ventura, R., Vitale, V., and Tacchia, D.: The GEO1MDB: the database of the 1:1000000 scale Geological Map of Italy, I, 6th EUREGEO, European Congress on Regional Geoscientific Cartography and Information Systems, Munich, Bavaria, Germany, 9–12 June 2009, 215–217, 2009.
- Colavitti L, Lanzano G, Sgobba S, Pacor F, Gallović F. Empirical evidence of directivity patterns for small-to-moderate earthquakes on normal faults: the study case of Central Italy. *J Geophys Res Solid Earth* 2022: e2021JB023498. <https://doi.org/10.1029/2021JB023498>.
- Compagnoni, B.: La Carta geologica d'Italia, alla scala 1:1000000, Mem. Descr. Carta Geol. It., 71, 207–212, 2004.
- Compagnoni, B., Damiani, A. V., and Valletta, M.: Carta geologica d'Italia alla scala 1:500000. In 5 fogli e note illustrative, Servizio Geologico d'Italia, Roma, 1976–1983.
- Commissione tecnica per la Microzonazione sismica (2016). Standard di rappresentazione e archiviazione informatica. Versione 4.0b. Dipartimento delle Protezione Civile Nazionale. Roma, ottobre 2015.
- Cramér H. Mathematical methods of statistics. Princeton university press; 1999 Apr 12.
- Cressie N. Fitting variogram models by weighted least squares. *J Int Assoc Math Geol* 1985; 17:563–86.
- Crowley H, Bommer JJ, Stafford PJ. Recent developments in the treatment of ground-motion variability in earthquake loss models. *Journal of Earthquake Engineering*. 2008 May 14;12(S2):71-80.
- Danciu, L., Giardini, D., Weatherill, G., Basili, R., Nandan, S., Rovida, A., Beauval, C., Bard, P.-Y., Pagani, M., Reyes, C. G., Sesetyan, K., Vilanova, S., Cotton, F., and Wiemer, S.: The 2020 European Seismic Hazard Model: overview and results, *Nat. Hazards Earth Syst. Sci.*, 24, 3049–3073, <https://doi.org/10.5194/nhess-24-3049-2024>, 2024.

- Dawood HM, Rodriguez-Marek A. A method for including path effects in ground-motion prediction equations: an example using the Mw 9.0 Tohoku earthquake aftershocks. *Bull Seismol Soc Am* 2013;103(2B):1360–72.
- Della Vedova B, Mongelli F, Pellis G, Squarci P, Taffi L, Ziti G. Heat flow map of Italy. CNR-Pisa: International Institute for Geothermal Research; 1991.
- De Natale G, Faccioli E, Zollo A. Scaling of peak ground motions from digital recordings of small earthquakes at Campi Flegrei, southern Italy. *Pure Appl Geophys* 1988; 126(1):37–53.
- Delogu, D., Campo, V., Cipolloni, C., Congi, M. P., Falcetti, S., Moretti, P., Pampaloni, M. L., Pantaloni, M., Roma, M., and Ventura, R.: Il Portale del Servizio Geologico d'Italia: uno strumento al servizio dei geologi professionisti, *Professione Geologo*, 32, 24–27, 2012.
- Di Capua, G., Peppoloni, S., Amanti, M., Cipolloni, C., Conte, G., 2016. Site Classification Map of Italy Based on Surface Geology. *Geol. Soc. Eng. Geol. Spec. Publ* <https://doi.org/10.1144/EGSP27.13>.
- Di Giulio G., Cornou C., Ohrnberger M., Wathelet M., and Rovelli A. (2006). Deriving Wavefield Characteristics and Shear-Velocity Profiles from Two-Dimensional Small-Aperture Arrays Analysis of Ambient Vibrations in a Small-Size Alluvial Basin, Colfiorito, Italy, *Bull. Seism. Soc. Am.*, Vol. 96, No. 5, pp. 1915–1933, doi: 10.1785/0120060119
- Di Stefano R, Ciaccio MG. Seismic velocity model of P- and S-waves for the Italian lithosphere (version 1.0). 2020. Zenodo, <https://doi.org/10.13127/tomorama.1>.
- DISS Working Group. Database of individual seismogenic sources (DISS), version 3.2.1. Istituto Nazionale di Geofisica e Vulcanologia (INGV); 2018. <https://doi.org/10.6092/ingv.it-diss3.2.1>.
- Dobry, R., Borcherdt, R.D., Crouse, C.B., Idriss, I.M., Joyner, W.B., Martin, G.R., Power, M.S., Rinne, E.E. and Seed, R.B., 2000. New site coefficients and site classification system used in recent building seismic code provisions. *Earthquake spectra*, 16(1), pp.41-67.
- Douglas J. Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates. *Earth-Science Reviews*. 2003 Apr 1;61(1-2):43-104.

- Douglas J, Edwards B. Recent and future developments in earthquake ground motion estimation. *Earth-Science Reviews*. 2016 Sep 1; 160:203-19.
- Douglas J. Ground motion prediction equations 1964–2021. Department of Civil and Environmental Engineering University of Strathclyde, Glasgow, United Kingdom. 2021.
- Dreger DS, Lee WHK, Kanamori H, Jennings PC, Kisslinger C. Time-domain moment tensor INVerse code (TDMT-INVC) release 1.1. *Int Handb Earthq Eng Seismol* 2003; B:1627.
- Faccenna C, Mattel M, Funiciello R, Jolivet L. Styles of back-arc extension in the central Mediterranean. *Terra Nova* 1997; 9:126–30.
- Felicetta C., Russo E., D'Amico M., Sgobba S., Lanzano G., Mascandola C., Pacor F., Luzi L. (2023) Italian Accelerometric Archive v4.0 - Istituto Nazionale di Geofisica e Vulcanologia, Dipartimento della Protezione Civile Nazionale. doi: 10.13127/itaca.4.0
- FEMA 273, 1997. NEHRP guidelines for the seismic rehabilitation of buildings: FEMA 273, Federal Emergency Management Agency. Dev. by Appl. Technol. Council.
- Forte G., Chioccarelli E., De Falco M., Cito P., Santo A. and Iervolino I.: Seismic soil classification of Italy based on surface geology and shear-wave velocity measurements, *Soil Dynamics and Earthquake Engineering*, 122, 79–93, <https://doi.org/10.1016/j.soildyn.2019.04.002>, 2019.
- Gabrielli S, Akinci A, Ventura G, Napolitano F, Del Pezzo E, De Siena L. Fast- changes in seismic attenuation of the upper crust due to fracturing and fluid migration: the 2016-2017 Central Italy seismic sequence. *Front Earth Sci* 2022. <https://doi.org/10.3389/feart.2022.909698>.
- Gasperini P, Lolli B, Vannucci G, Boschi E. Calibration of moment magnitude estimates for the European–Mediterranean and Italian regions. *Geophys J Int* 2012; 190(3):1733–45.
- Gasperini P, Lolli B, Vannucci G. Empirical calibration of local magnitude data sets versus moment magnitude in Italy. *Bull Seismol Soc Am* 2013;103(4):2227–46. <https://doi.org/10.1785/0120120356>.
- Gironelli V, Volatili T, Luzi L, Brunelli G, Tondi E. Dual-proxy estimation of Vs30: the case study of the Marche Region (central Italy). *Journal of Maps*. 2024 Dec 31;20(1):2349787.

- Graves Robert, Jordan Thomas, Callaghan Scott, Deelman Ewa, Field Edward, Juve Gideon, Kesselman Carl, Maechling Philip, Mehta Gaurang, Milner Kevin, Okaya David, Small Patrick, Vahi Karan. CyberShake: a physics-based seismic hazard model for Southern California. *Pure Appl Geophys* 2010;168:367–81. <https://doi.org/10.1007/s00024-010-0161-6>.
- Haskell NA. Total energy and energy spectral density of elastic wave radiation from propagating faults. *Bulletin of the Seismological Society of America*. 1964 Dec 1;54(6A):1811-41.
- Hobson, R.D.; Surface Roughness in Topography: A Quantitative Approach. In: Chorley, R.J., Ed., *Spatial Analysis in Geomorphology*, Methuen & Co., London, 1972; 221-245.
- Ishihara K. Simple method of analysis for liquefaction of sand deposits during earthquakes. *Soils and Foundations*. 1977 Sep 1;17(3):1-7.
- Iwahashi, J., Kamiya, I., Matsuoka, M., Yamazaki, D., 2018. Global terrain classification using 280 m DEMs: segmentation, clustering, and reclassification. *Prog. Earth Planet. Sci.* <https://doi.org/10.1186/s40645-017-0157-2>.
- Jayaram N, Baker JW. Considering spatial correlation in mixed-effects regression and the impact on ground-motion models. *Bull Seismol Soc Am*. 2010; 100:3295-3303.
- Jasiewicz, J and Stepinski, T. J., *Geomorphons - a pattern recognition approach to classification and mapping of landforms*, *Geomorphology*, 182, January 15, 2013: 147-56. <https://doi.org/10.1016/j.geomorph.2012.11.005>
- Kanamori H, Anderson DL. Theoretical basis of some empirical relations in seismology. *Bulletin of the seismological society of America*. 1975 Oct 1;65(5):1073-95.
- Knopoff, L. Q, Review of the attenuation of seismic waves in the Earth. *Reviews of Geophysics*, 2(4), 625-660; 1964. <https://doi.org/10.1029/RG002i004p00625>
- Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, 4, 83–91.
- Komatitsch D, Tromp J. Spectral-element simulations of global seismic wave propagation-I. Validation. *Geophys J Int* 2002;149(2):390–412. <https://doi.org/10.1046/j.1365-246X.2002.01653.x>.

- Komatitsch D, Tromp J. Spectral-element simulations of global seismic wave propagation-II. Three-dimensional models, oceans, rotation and self-gravitation. *Geophys J Int* 2002;150(1):303–18.
- Kotha SR, Bindi D, Cotton F. Partially non-ergodic region specific GMPE for Europe and Middle-East. *Bulletin of Earthquake Engineering*. 2016 Apr; 14:1245-63.
- Kotha SR, Weatherill G, Bindi D, Cotton F. A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe. *Bull Earthq Eng* 2020;18: 4091–125.
- Kramer SL, Stewart JP. *Geotechnical earthquake engineering*. CRC Press; 2024 Nov 29.
- Kubo H, Kunugi T, Suzuki W, Suzuki S, Aoi S. Hybrid predictor for ground-motion intensity with machine learning and conventional ground motion prediction equation. *Scientific reports*. 2020 Jul 27;10(1):11871.
- Kuehn NM, Abrahamson NA. Spatial correlations of ground motion for non-ergodic seismic hazard analysis. *Earthq Eng Struct Dyn*. 2019;49:4-23.
- Landwehr N, Kuehn NM, Scheffer T, Abrahamson N. A nonergodic ground-motion model for California with spatially varying coefficients. *Bull Seismol Soc Am* 2016; 106(6):2574–83.
- Lanzano G, D'Amico M, Felicetta C, Puglia R, Luzi L, Pacor F, Bindi D. Ground-motion prediction equations for region-specific probabilistic seismic-hazard analysis. *Bull Seismol Soc Am* 2016;106(1):73–92.
- Lanzano G, Pacor F, Luzi L, D'Amico M, Puglia R, Felicetta C. Systematic source, path and site effects on ground motion variability: the case study of Northern Italy. *Bull Earthq Eng*. 2017;15(11):4563-4583.
- Lanzano G, Sgobba S, Pacor F, Luzi L, Puglia R, D'Amico M, Felicetta C. Spatial correlation of the systematic site-and pathspecific corrections of a GMPE calibrated in Northern Italy. In 16th European Conference Earthquake Engineering 2018 Jun 18.
- Lanzano G, Luzi L, Pacor F, Felicetta C, Puglia R, Sgobba S, D'Amico M. A revised ground motion prediction model for shallow crustal earthquakes in Italy. *Bull Seismol Soc Am* 2019a;109(2):525–54. <https://doi.org/10.1785/0120180210>.
- Lanzano G, Sgobba S, Luzi L, Puglia R, Pacor F, Felicetta C, D'Amico M, Cotton F, Bindi D. The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and

- data statistics. *Bull Earthq Eng* 2019b;17(2):561–82. [https://doi.org/ 10.1007/s10518-018-0480-z](https://doi.org/10.1007/s10518-018-0480-z).
- Lanzano G, Luzi L. A ground motion model for volcanic areas in Italy. *Bull Earthq Eng* 2020;18(1):57–76.
- Lanzano G, Sgobba S, Caramenti L, Menafoglio A. Ground-motion model for crustal events in Italy by applying the multisource geographically weighted regression (MS-GWR) method. *Bull Seismol Soc Am* 2021;111(6):3297–313.
- Lanzano G, Felicetta C, Pacor F, Spallarossa D, Traversa P. Generic-To-Reference Rock Scaling Factors for Seismic Ground Motion in Italy. *Bulletin of the Seismological Society of America* 2022;112(3):1583–606.
- Lekshmy PR, Raghukanth ST. A hybrid genetic algorithm-neural network model for power spectral density compatible ground motion prediction. *Soil dynamics and earthquake engineering*. 2021 Mar 1; 142:106528.
- Levene, H. (1960). Robust tests for equality of variances. In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (pp. 278–292). Stanford University Press.
- Li C, Ji D, Zhai C, Ma Y, Xie L. Vertical ground motion model for the NGA-West2 database using deep learning method. *Soil Dynamics and Earthquake Engineering*. 2023 Feb 1;165:107713.
- Lin PS, Chiou B, Abrahamson N, Walling M, Lee CT, Cheng CT. Repeatable source, site, and path effects on the standard deviation for empirical ground-motion prediction models. *Bull Seismol Soc Am*. 2011;101:2281-2295.
- Locardi E. Identification of seismogenic structures from the examination of the volcano-tectonic evolution of the Apennines and the Tyrrhenian Sea. *Mem Soc Geol Ital* 1982; 24:569–96 [In Italian].
- Lombardi AM, Akinci A, Malagnini L, Mueller CS. Uncertainty analysis for seismic hazard in Northern and Central Italy. *Ann Geophys* 2005;48(6):853–65.
- Luzi L, et al. Overview on the strong-motion data recorded during the May–June 2012 Emilia seismic sequence. *Seismol Res Lett* 2013;84(4):629–44.

- Luzi L, et al. The central Italy seismic sequence between August and December 2016: analysis of strong-motion observations. *Seismol Res Lett* 2017;88(5): 1219–31.
- Manu-Marfo D, Aoudia A, Pachhai S, Kherchouche R. 3D shear wave velocity model of the crust and uppermost mantle beneath the Tyrrhenian basin and margins. *Scientific reports* 2019;9(1):1–10
- Malagnini L, Herrmann RB, Di Bona M. Ground-motion scaling in the Apennines (Italy). *Bull Seismol Soc Am* 2000;90(4):1062–81.
- Malagnini L, Akinci A, Herrmann RB, Pino NA, Scognamiglio L. Characteristics of the ground motion in northeastern Italy. *Bull Seismol Soc Am* 2002;92(6): 2186–204.
- Malagnini L, Scognamiglio L, Mercuri A, Akinci A, Mayeda K. Strong evidence for non-similar earthquake source scaling in central Italy. *Geophys Res Lett* 2008;35 (17):1–6.
- Malagnini L, Akinci A, Mayeda K, Munafo' I, Herrmann RB, Mercuri A. Characterization of earthquake-induced ground motion from the L'Aquila seismic sequence of 2009, Italy. *Geophys J Int* 2011;325–37. <https://doi.org/10.1111/j.1365-246X.2010.04837>
- Malinverno A, Ryan WBF. Extension in the Tyrrhenian Sea and shortening in the Apennines as result of arc migration driven by sinking of the lithosphere. *Tectonics* 1986;5 2:227–45.
- Mascandola C, Lanzano G, Pacor F. Consistency check of ITACAext, the flatfile of the Italian accelerometric archive. *Geosciences* 2022;12(9):334.
- Matérn, B. (1986). *Spatial Variation*. Springer. Matérn, B. (1986). *Spatial Variation*. Springer.
- Matheron, G. (1963). Principles of Geostatistics. *Economic Geology*, **58**(8), 1246-1266.
- Matsuoka, M., Wakamatsu, K., 2006. Development of the 7.5-ARC-second engineering geomorphologic classification database and average shear-wave velocity map for Japan's metropolitan areas. In: 8th US National Conference on Earthquake Engineering, pp. 2006.
- Matheron G. Les variables régionalisées et leur estimation. Paris: Masson; 1965.
- Mazzieri I, Stupazzini M, Guidotti R, Smerzini C. SPEED: spectral elements in elastodynamics with discontinuous galerkin: a non-conforming approach for 3D multi-scale problems. *Int J Numer Methods Eng* 2013;95:991–1010.

- McCallen D, Petersson A, Rodgers A, Pitarka A, Miah M, Petrone F, et al. EQSIM - a multidisciplinary framework for fault-to-structure earthquake simulations on exascale computers part I: computational models and workflow. *Earthq Spectra* 2021;37(2):707–35. <https://doi.org/10.1177/8755293020970982>.
- McCallen D, Petrone F, Miah M, Pitarka A, Rodgers A, Abrahamson N. EQSIM - a multidisciplinary framework for fault-to-structure earthquake simulations on exascale computers, part II: regional simulations of building response. *Earthq Spectra* 2021;37(2):736–61. <https://doi.org/10.1177/8755293020970980>.
- Meletti C, Patacca E, Scandone P. Construction of a seismotectonic model: the case of Italy. *Pure Appl Geophys* 2000;157:11–35.
- Meletti C, et al. A seismic source zone model for the seismic hazard assessment of the Italian territory. *Tectonophysics* 2008;450(1–4):85–108.
- Meletti C, Marzocchi W, D’Amico V, Lanzano G, Luzi L, Martinelli F, Pace B, Rovida A, Taroni M, Visini F. The new Italian seismic hazard model (MPS19). *Ann Geophys* 2021;64(1).
- Menafoglio A, Sgobba S, Lanzano G, et al. Simulation of seismic ground motion fields via object-oriented spatial statistics with an application in Northern Italy. *Stoch Environ Res Risk Assess.* 2020;34:1607-1627.
- Michellini, A., Faenza, L., Lauciani, V., Malagnini, L., 2008. ShakeMap implementation in Italy. *Seismol. Res. Lett.* <https://doi.org/10.1785/gssrl.79.5.688>.
- Michellini A, Faenza L, Lanzano G, Lauciani V, Jozinović D, Puglia R, Luzi L. The new ShakeMap in Italy: progress and advances in the last 10 yr. *Seismol Res Lett* 2020;91(1):317–33.
- Mohammadi A, Karimzadeh S, Banimahd SA, Ozsarac V, Lourenço PB. The potential of region-specific machine-learning-based ground motion models: application to Turkey. *Soil Dynamics and Earthquake Engineering.* 2023 Sep 1;172:108008.
- Molinari I, Verbeke J, Boschi L, Kissling E, Morelli A. Italian and Alpine three dimensional crustal structure imaged by ambient noise surface wave dispersion. *Geochem Geophys Geosys* 2015;16(12):4405–21.

- Montaldo V, Faccioli E, Zonno G, Akinci A, Malagnini L. Treatment of ground motion predictive relationships for the reference seismic hazard map of Italy. *J Seismol* 2005; 9:295–316.
- Moore, I. D., R. B. Grayson, and A. R. Landson. 1991. Digital Terrain Modelling: A Review of Hydrological, Geomorphological, and Biological Applications. *Hydrological Processes* 5: 3–30.
- Morasca P, Malagnini L, Akinci A, Spallarossa D, Herrmann RB. Ground-motion scaling in the western Alps. *J Seismol* 2006;10(3):315–33.
- Morasca P, D’Amico M, Sgobba S, Lanzano G, Colavitti L, Pacor F, Spallarossa D. Empirical correlations between an FAS non-ergodic ground motion model and a GIT derived model for Central Italy. *Geophys J Int* 2023; 233(1):51–68. <https://doi.org/10.1093/gji/ggac445>.
- Mori F., Mendicelli A., Moscatelli M., Romagnoli G., Peronace E. and Naso G.: A new Vs30 map for Italy based on the seismic microzonation dataset, *Eng. Geol.*, 275, 105745, <https://doi.org/10.1016/j.enggeo.2020.105745>, 2020.
- Nardoni C, De Siena L, Cammarano F, Magrini F, Mattei E. Modelling regional-scale attenuation across Italy and the Tyrrhenian Sea. *Physics of the Earth and Planetary Interiors* 2021;318:106764.
- Neri G, Marotta AM, Orecchio B, Presti D, Totaro C, Barzaghi R, Borghi A. How lithospheric subduction changes along the Calabrian Arc in southern Italy: geophysical evidences. *Int J Earth Sci* 2012;101(7):1949–69. <https://doi.org/10.1007/s00531-012-0762-7>.
- NTC-Italian Building Code. Decree 17-1-2018, official bulletin 42, 20-02-2018. 2018. Ordinary Suppl. n. 8 (in Italian).
- Pantaloni M.: La Carta geologica d’Italia alla scala 1 1000000: una pietra miliare nel percorso della conoscenza geologica, *Geologia Tecnica & Ambientale*, 11, 88–99, 2011.
- Paolucci R, Pacor F, Puglia R, Ameri G, Cauzzi C, Massa M. Record processing in ITACA, the new Italian strong-motion database. In: *Earthquake data in engineering seismology*. Dordrecht: Springer; 2011. p. 99–113.

- Paolucci R, Gatti F, Infantino M, Smerzini C, Ozcebe AG, Stupazzini M. Broadband ground motions from 3D physics-based numerical simulations using artificial neural networks. *Bull Seismol Soc Am* 2018;108:1272–86.
- Paolucci R, Aimar M, Ciancimino A, Dotti M, Foti S, Lanzano G, Mattevi P, Pacor F, Vanini M. Checking the site categorization criteria and amplification factors of the 2020 draft of Eurocode 8 Part 1-1. *Bull Earthq Eng* 2021; 19(11):4199–234.
- Park J, Bazzurro P, Baker JW. Modeling spatial correlation of ground motion intensity measures for regional seismic hazard and portfolio loss estimation. *Applications of statistics and probability in civil engineering*. 2007 Jul 31;2:1-8.
- Parker, G.A., Harmon, J.A., Stewart, J.P., Hashash, Y.M.A., Kottke, A.R., Rathje, E.M., Silva, W.J., Campbell, K.W., 2017. Proxy-based Vs30 estimation in central and Eastern North America. *Bull. Seismol. Soc. Am.* <https://doi.org/10.1785/0120160101>.
- Parker GA, Baltay AS, Thompson EM. Repeatable source, path, and site effects from the 2019 M 7.1 Ridgecrest earthquake sequence. *Bull Seismol Soc Am*. 2020;110:1530-1548.
- Parker GA, Stewart JP, Boore DM, Atkinson GM, Hassani B. NGA-subduction global ground motion models with regional adjustment factors. *Earthq Spectra* 2022;38 (1). <https://doi.org/10.1177/87552930211034889>.
- Patacca E, Scandone P. Post-Tortonian mountain building in the Apennines. The role of the passive sinking of a relic lithospheric slab. In: Boriani A, et al., editors. *The lithosphere in Italy*, vol. 80. Acc. Naz. Lincei; 1989. p. 157–76.
- Patacca E, Sartori R, Scandone P. Tyrrhenian basin and Apenninic arcs. Kinematic relations since late Tortonian times. *Mem Soc Geol It* 1990; 45:425–51.
- Patanè D, Ferrucci F, Giampiccolo E, Scaramuzzino L. Source scaling of microearthquakes at Mt. Etna volcano and in the Calabrian Arc (southern Italy). *Geophys Res Lett* 1997;24:1879–82.
- Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, 30: 683-691. Benedikt Gräler, Edzer Pebesma and Gerard Heuvelink, 2016. Spatio-Temporal Interpolation using gstat. *The R Journal* 8(1), 204-218
- Pebesma, E.J., R.S. Bivand, 2005. Classes and methods for spatial data in R. *R News* 5 (2), <https://cran.r-project.org/doc/Rnews/>.

- Pelletier JD, Broxton PD, Hazenberg P, Zeng X, Troch PA, Niu GY, Williams Z, Brunke MA, Gochis D. A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. *Journal of Advances in Modeling Earth Systems*. 2016 Mar;8(1):41-65.
- Polonia A, Torelli L, Mussoni P, Gasperini L, Artoni A, Klaeschen D. The Calabrian Arc subduction complex in the Ionian Sea: regional architecture, active deformation, and seismic hazard. *Tectonics* 2011;30:TC5018. [https://doi.org/ 10.1029/2010TC002821](https://doi.org/10.1029/2010TC002821).
- Pondrelli S, Morelli A, Ekström G, Mazza S, Boschi E, Dziewonski AM. European–Mediterranean regional centroid-moment tensors: 1997–2000. *Phys Earth Planet In* 2002;130(1–2):71–101.
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Riley S J, DeGloria S D, Elliot R; 1999: Index that quantifies topographic heterogeneity. *intermountain Journal of sciences*, 5(1-4), 23-27.
- Robert J. Hijmans (2023). raster: Geographic Data Analysis and Modeling. R package version 3.6-20. <https://CRAN.R-project.org/package=raster>
- Rodriguez-Marek A, Cotton F, Abrahamson NA, et al. A model for single-station standard deviation using data from various tectonic regions. *Bull Seismol Soc Am*. 2013;103:3149-3163.
- Rosenberg CL. Shear zones and magma ascent: a model based on a review of the Tertiary magmatism in the Alps. *Tectonics* 2004;23:TC3002. <https://doi.org/10.1029/2003TC001526>.
- Rovida A, Locati M, Camassi R, Lolli B, Gasperini P. Parametric Catalogue of Italian Earthquakes (CPTI15). Istituto Nazionale di Geofisica e Vulcanologia (INGV). 2016. <https://doi.org/10.6092/INGV.IT-CPTI15>.
- Russo E, Felicetta C, D'Amico MC, Sgobba S, Lanzano G, Mascandola C, Pacor F, Luzi L. Italian ACcelerometric Archive (ITACA), version 3.2. Istituto Nazionale di Geofisica e Vulcanologia (INGV); 2022. <https://doi.org/10.13127/itaca.3.2>.
- RStudio Team (2021). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.

- Sabetta F, Pugliese A. Estimation of response spectra and simulation of nonstationary earthquake ground motions. *Bull Seismol Soc Am* 1996;86(2): 337–52.
- Sahakian VJ, Baltay A, Hanks TC, et al. Ground motion residuals, path effects, and crustal properties: a pilot study in Southern California. *J Geophys Res Solid Earth*. 2019;124:5738-5753.
- Sappington J M, Longshore K M, Thompson D B; 2007: Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave Desert. *The Journal of wildlife management*, 71(5), 1419-1426.
- Sartori R. Stratigraphic and tectonic notes on the Messina rise (northwestern Ionian Sea). *Giorn Geol* 1991;53:49–64 [In Italian)].
- Sato, H., Fehler, M. C., & Maeda, T. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*. Springer; 2012. <https://doi.org/10.1007/978-3-642-23029-5>
- Schiappapietra E, Douglas J. Modelling the spatial correlation of earthquake ground motion: insights from the literature, data from the 2016–2017 central Italy earthquake sequence and ground-motion simulations. *Earth Sci Rev.* 2020; 203. <https://doi.org/10.1016/j.earscirev.2020.103139>
- Schmid SM, Pfiffner OA, Froitzheim N, Schoenborn G, Kissling E. Geophysical- geological transect and tectonic evolution of the Swiss-Italian Alps. *Tectonics* 1996; 15(5):1036–64. <https://doi.org/10.1029/96TC00433>.
- Scognamiglio L, Malagnini L, Akinci A. Ground-motion scaling in eastern Sicily, Italy. *Bull Seismol Soc Am* 2005;95(2):568–78.
- Sedaghati F, Pezeshk S. Partially nonergodic empirical ground-motion models for predicting horizontal and vertical PGV, PGA, and 5% damped linear acceleration response spectra using data from the Iranian Plateau. *Bull Seismol Soc Am* 2017; 107:934–48.
- Servizio geologico d'Italia: Carta geologica d'Italia interattiva. In *teractive geological map of Italy: 1 100000*. 3CD., Etruria innovazione, 2004.
- Spallarossa D., Cattaneo M., Scafidi D., Michele M., Chiaraluce L., Segou M., Main I.G. (2021) An automatically generated high-resolution earthquake catalogue for the 2016–2017 Central Italy seismic sequence, including P and S phase arrival times. *Geophysical*

Journal International 225-1, pp 555–571, <https://doi.org/10.1093/gji/ggaa604>.
Catalogue DOI: <https://doi.org/10.5281/zenodo.4306165>

Sgobba S, Lanzano G, Pacor F, et al. Spatial correlation models of systematic site and path effects for ground-motion fields in Northern Italy. *Bull Seismol Soc Am*. 2019;109(4):1419-1434.

Sgobba S, Lanzano G, Pacor F. Empirical non-ergodic shaking scenarios based on spatial correlation models: an application to central Italy. *Earthq Eng Struct Dynam* 2021;50(1):60–80.

Sgobba SA, Pacor F. An application of the NonErgodic ground ShaKing (NESK) approach to an historical earthquake scenario: The case-study of the 1915 Fucino earthquake (central Italy). *Soil Dynamics and Earthquake Engineering*. 2023 Jan 1;164:107622.

Shearer PM. *Introduction to seismology*. Cambridge university press; 2019 May 30.

Smirnov, N. V. (1948). Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, 19(2), 279–281.

Stafford PJ. Crossed and nested mixed-effects approaches for enhanced model development and removal of the ergodic assumption in empirical ground-motion models. *Bull Seismol Soc Am* 2014;104(2):702–19.

Stein, S., & Wysession, M.; 2003. *An Introduction to Seismology, Earthquakes, and Earth Structure*. Blackwell Publishing.

Stewart JP, Kwok AO. Nonlinear seismic ground response analysis: Code usage protocols and verification against vertical array data. In *Geotechnical earthquake engineering and soil dynamics IV 2008* May 14 (pp. 1-24).

Stewart, J.P., Klimis, N., Savvaidis, A., Theodoulidis, N., Zargli, E., Athanasopoulos, G., Pelekis, P., Mylonakis, G., Margaris, B., 2014. Compilation of a local Vs profile database and its application for inference of Vs30 from geologic- and terrain-based proxies. *Bull. Seismol. Soc. Am.* <https://doi.org/10.1785/0120130331>.

Stucchi M, Meletti C, Montaldo V, Crowley H, Calvi GM, Boschi E. Seismic hazard assessment (2003–2009) for the Italian building code. *Bull Seismol Soc Am* 2011; 101(4):1885–911.

- Tarquini S., I. Isola, M. Favalli, F. Mazzarini, M. Bisson, M.T. Pareschi, E. Boschi (2007). TINITALY/01: a new triangular irregular network of Italy. *Annals of Geophysics*. <https://doi.org/10.4401/ag-4424>
- Tarquini S., I. Isola, M. Favalli, A. Battistini, G. Dotta (2023). TINITALY, a digital elevation model of Italy with a 10 meters cell size (Version 1.1). Istituto Nazionale di Geofisica e Vulcanologia (INGV). <https://doi.org/10.13127/tinitaly/1.1>.
- Villani M, Abrahamson NA. Repeatable site and path effects on the ground-motion sigma based on empirical data from Southern California and simulated waveforms from the CyberShake platform. *Bull Seismol Soc Am*. 2015;105(5):2681-2695.
- Visini F, Meletti C, Rovida A, D'Amico V, Pace B, Pondrelli S. An updated area- source seismogenic model (MA4) for seismic hazard of Italy. *Nat Hazards Earth Syst Sci* 2022; 22:2807–27.
- Vucetic, M. and Dobry, R., 1991. Effect of soil plasticity on cyclic response. *Journal of geotechnical engineering*, 117(1), pp.89-107.
- Wald, D.J., Allen, T.I., 2007. Topographic slope as a proxy for seismic site conditions and amplification. *Bull. Seismol. Soc. Am.* <https://doi.org/10.1785/0120060267>.
- Walling MA. Non-Ergodic Probabilistic Seismic Hazard Analysis and Spatial Simulation of Variation in Ground Motion [Ph.D. Thesis]. University of California, Berkeley; 2009.
- Weatherill G, Cotton F. A ground motion logic tree for seismic hazard analysis in the stable cratonic region of Europe: regionalisation, model selection and development of a scaled backbone approach. *Bull Earthq Eng* 2020;18(14): 6119–48.
- Weiss A; 2001: Topographic position and landforms analysis. In Poster presentation, ESRI user conference, San Diego, CA (Vol. 200).
- Welch, B. L. (1947). The generalization of 'Student's' problem when several different population variances are involved. *Biometrika*, 34(1-2), 28–35.
- Worden CB, Thompson EM, Baker JW, Bradley BA, Luco N, Wald DJ. Spatial and spectral interpolation of ground- motion intensity measure observations. *Bull Seismol Soc Am* 2018;108(2):866–75. <https://doi.org/10.1785/0120170201>

Data and Resources

The digital lithological map of Italy proposed by Bucci et al. (2021) is available at <https://doi.org/10.1594/PANGAEA.935673> (last accessed June 2025).

Elevation data were obtained from the TINITALY digital elevation model (DEM) with a 10-meter cell size (Tarquini et al., 2023), available at <https://doi.org/10.13127/tinitaly/1.1> (last accessed June 2025).

Strong-motion recordings and metadata were retrieved from the extended flat-file ITACAext v3.2 (Brunelli et al., 2022), available at https://doi.org/10.13127/itaca32/itacaext_flatfile.1.0 (last accessed June 2025). Information on seismogenic sources was taken from the Database of Individual Seismogenic Sources (DISS), version 3.2.1 (DISS Working Group, 2018), available at <https://doi.org/10.6092/ingv.it-diss3.2.1> (last accessed June 2025).

APPENDICES

Appendix A

Lithological codes from the ISPRA map

NOME_ULF	ETA_GEOL	LITOLOGIA	CLASSI_LIT	COD_LITO
calcari bianchi, dolomitici	triassico e giurassico	calcari	calcilutiti, calcareniti, calciruditi; calcari dolomitici, calcari selciferi,	A1
arenarie quarzose	permiano medio	complessi pelitico-arenacei	complesso pelitico- arenaceo, pelitico- arenaceo- conglomeratico, calcareo-marnoso- arenaceo	A10
marne	scitico	complessi calcareo- arenacei	complesso calcareo (dolomitico) - arenaceo, calcareo (dolomitico)-marnoso- arenaceo	A11
gessi micacei	triassico	evaporiti	sedimenti evaporitici	A12
sacche di bauxite	senoniano	rocce residuali	sedimenti residuali (bauxiti; lateriti)	A13

detrito cementato e crostoni di falda	neozoico	detriti cementati	sedimenti grossolani recenti-attuali cementati, detriti e falde di detrito cementati	A14
travertini	N.D.	travertini	travertini	A15
dolomia principale	norico	dolomie	dolomie, dolomie calcaree	A2
calcari marnosi e marne	carnico	calcari marnosi	calcari marnosi, marne calcaree, alternanze calcari- marne, calcari-argille, dolomie-marne, dolomie-argille	A3
rosso ad aptici	dogger e malm	diaspri	radiolariti e diaspri	A4
arenarie quarzose	carnico	quarzareniti	quarzareniti, quarzolutiti, quarzoruditi (prevalenza di cemento siliceo)	A5
marne biancastre	aquitano	diatomiti	depositi diatomitici (tripoli)	A6
strati argillosi	N.D.	marne	marne, argilliti, siltiti	A7

verrucano	permico	conglomerati	conglomerati, brecce, puddinghe	A8
arenarie calcaree gialle scistose ed arenarie quarzifere chiare	buntsandstein e vernicino	arenarie	arenarie, arenarie + conglomerati, arenarie + sabbie	A9
alluvioni torbose	quaternario	terreni prevalentemente argillosi	terreni a granulometria fine (f)	B1
sabbie fluviali	wurmiano	terreni prevalentemente sabbiosi	terreni a granulometria media (m)	B2
alluvioni recenti ed attuali	quaternario	terreni prevalentemente ghiaiosi	terreni a granulometria grossolana (g), media-grossolana (m+g)	B3
detriti di falda	quaternario	terreni a granulometria mista	terreni a granulometria mista (f+m+g)	B4
terreni torbosi	N.D.	torbe	terreni torbosi	B5
scisti carboniosi	N.D.	terreni residuali	terreni residuali	B6

coni di deiezione e coni detritici	quaternario	terreni granulometria indefinita	terreni granulometricamente indefiniti (conoidi, frane)	B7
complesso caotico	N.D.	argille caotiche	complesso indifferenziato; argille caotiche; argille varicolori; argille scagliose ecc.	C1
alla base tufiti sedimentarie verdi	carnico	Complessi vulcanico sedimentari marini	Tufiti	C2
rioliti in filoni	permiano	lave acide	quarzoandesiti; daciti; riodaciti; rioliti (lipariti); trachidaciti	D1
porfidi	N.D.	Lave neutre	trachiti; andesiti; trachianandesiti; latiti; trachianandesiti basaltiche; andesiti basaltiche	D2
colate laviche andesitico- basaltiche	ladinico e carnico	Lave basiche	basalti; trachibasalti; picrobasalti; picrite; tefriti; basaniti; fonotefriti; tefrifonoliti; fonoliti; foiditi fonolitiche, foiditi	D3

			tefritiche, foiditi (leucititi)	
diabasi	N.D.	lave a chimismo non noto	lave a chimismo non noto	D4
porfidi quarziferi normali	N.D.	piroclastiti, tufi, ignimbriti	piroclastiti, tufi, ignimbriti	D5
lave porfiritiche e metafiriche	ladinico superiore	piroclastiti+lave	Piroclastiti+lave	D6
tufi e pomici bianche, riodacitici	miocene superiore	Tufi pedogenizzati	Tufi pedogenizzati	D7
prodotti di disfacimento delle formazioni piroclastiche	quaternario	Scorie, lapilli, pomici, bombe	Scorie, lapilli, pomici, bombe	D8
rocce filoniane	N.D.	rocce granitoidi	Quarzoliti, graniti, granito a feldspati alcalini, tonaliti, granodioriti	E1

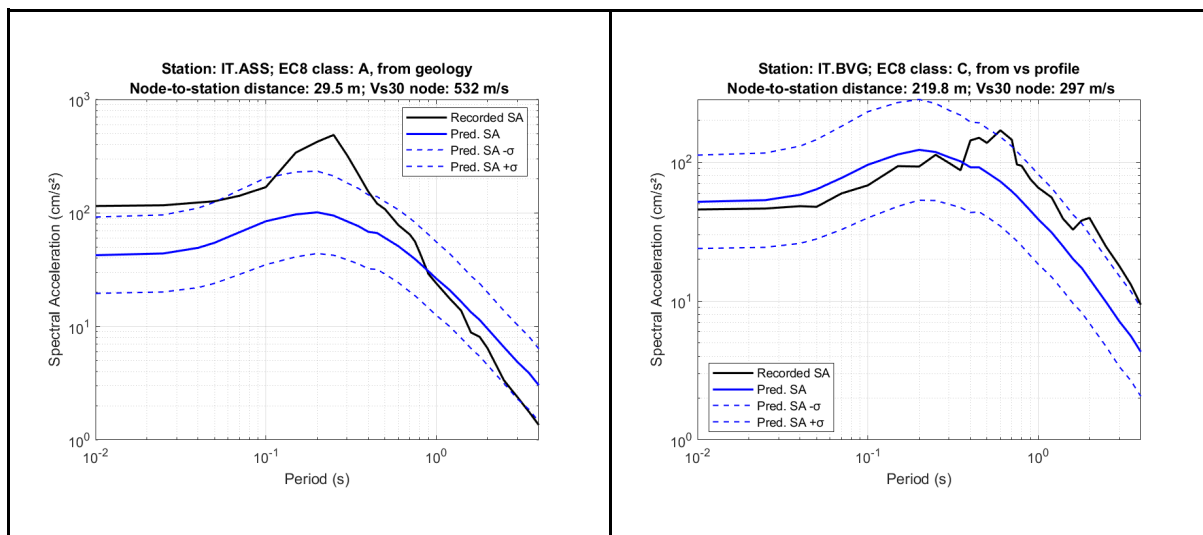
gabbro di sondalo	cenozoico	plutoniti intermedie	Monzoniti, sieniti a feldspati alcalini, sieniti, dioriti, anortositi, monzodiorite, monzogabbro, diorite, gabbro, anortosite	E2
olivinita	N.D.	plutoniti foidiche	Sienite a feldspatoidi, monzonite a feldspatoidi, essexite, theralite, foidite, ultramafiche: pirosseniti, peridotiti, duniti	E3
filladi scure a porfiroblasti biotitici	raibliano	filladi e micascisti	Filladi, anfiboliti, argilloscisti, micascisti, parascisti, quarziti scistose, calcescisti, talcoscisti	F1
gneiss granitici a muscovite	pre-permiano	gneiss	Gneiss	F2
prasiniti prevalentemente epidotiche	mesozoico	prasiniti	Prasiniti	F3
complesso c. s. con ftaniti	ladinico	ofioliti	Ofioliti	F4

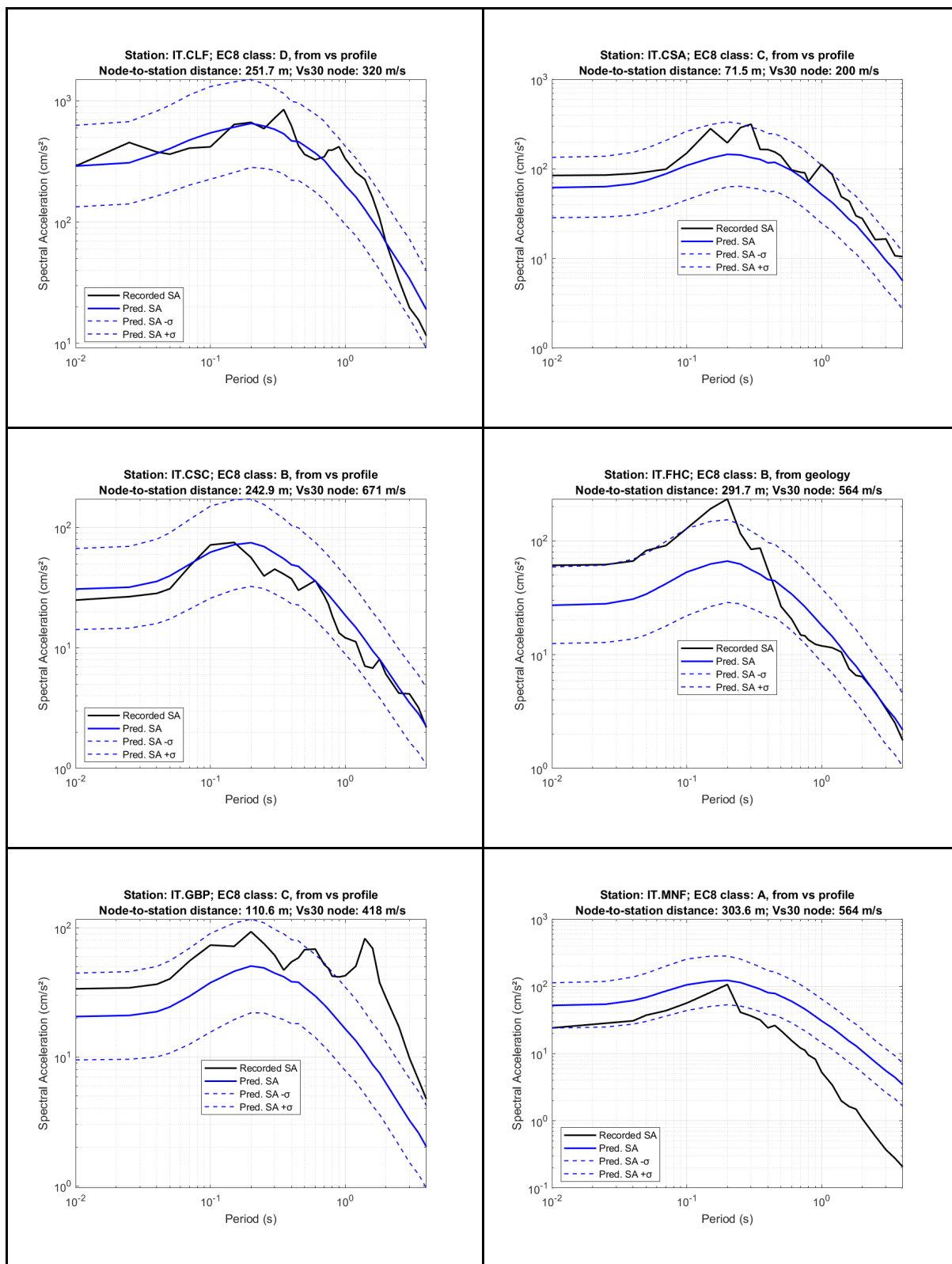
serpentini	mesozoico recente	serpentiniti	Serpentiniti	F5
oficalci e calcefiri associati ai serpentini	mesozoico	cornubianiti	Hornfelse (cornubianiti);calcefiri	G1
dolomia del wetterstein	anisico-ladinico	marmi	Marmo	G2
brecce di seravezza	norico-retico	dolomie metamorfiche	Dolomia metamorfica	G3
quarziti, quarziti nere, micascisti gneissici conglomeratici	N.D.	quarziti	Quarziti	G4
stronaliti	precarbonifero	granuliti	Granuliti, eclogiti	G5
arenarie scistose	permiano inferiore	metamorfiti di basso grado	Meta-sedimentarie, meta-magmatiche (metamorfiti di basso grado)	H1
giacimento cuprifero	N.D.	Mineralizzazioni di particolare interesse	Mineralizzazioni e affioramenti di minerali di particolare interesse	H2

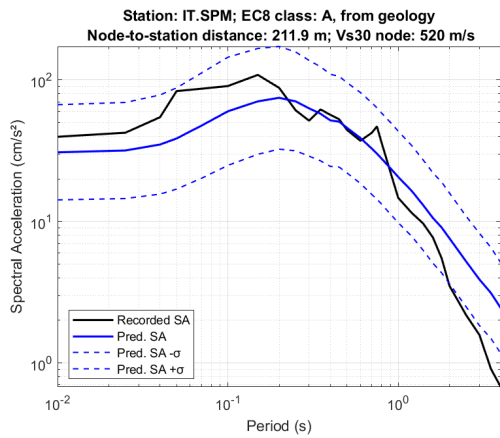
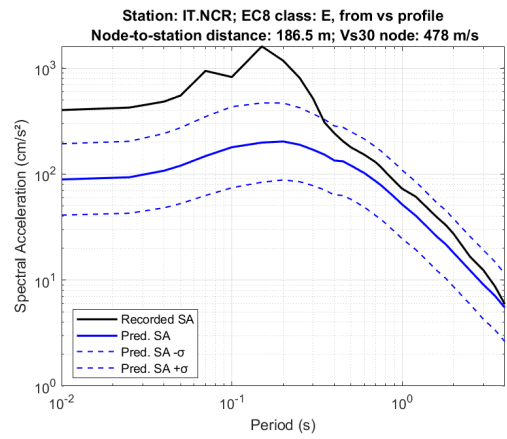
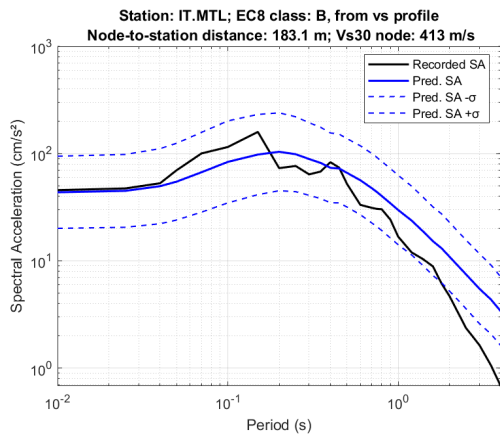
ghiacciai e nevati.	N.D.	acque ghiacciai e	Corsi d'acqua, acquitrini, laghi, ghiacciai, ecc.	N1
zona archeologica di aquileia	postglaciale recente	insediamenti antropici	Insedimenti antropici	N2
materiali argillo-sabbiosi	olocene	discariche	Discariche	N3

Appendix B

For the Mw 5.7 Colfiorito 1997 earthquake scenario, the acceleration response spectra predicted using the Vs30 map developed in this study are compared with the observed spectra recorded at 11 seismic stations.

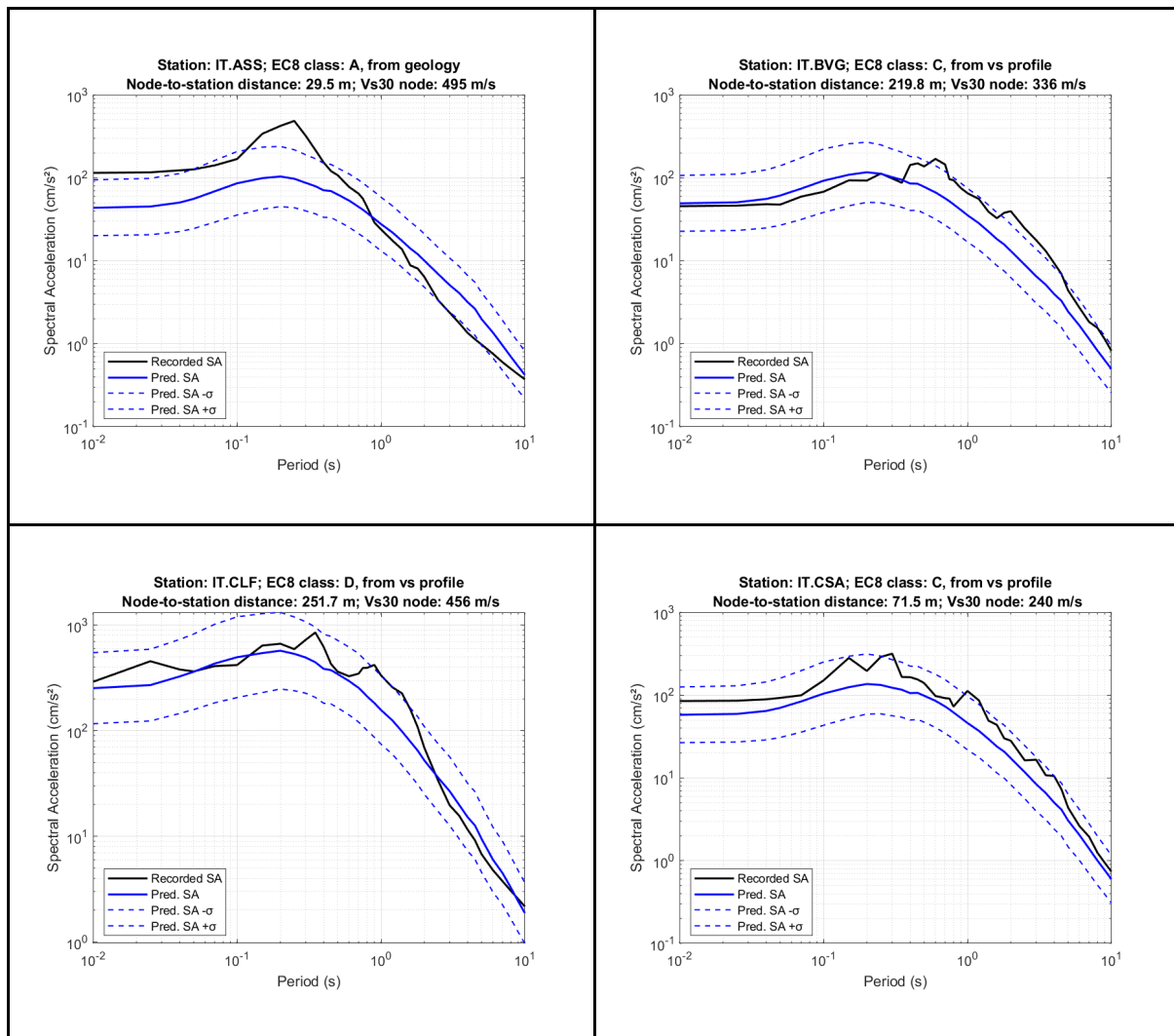


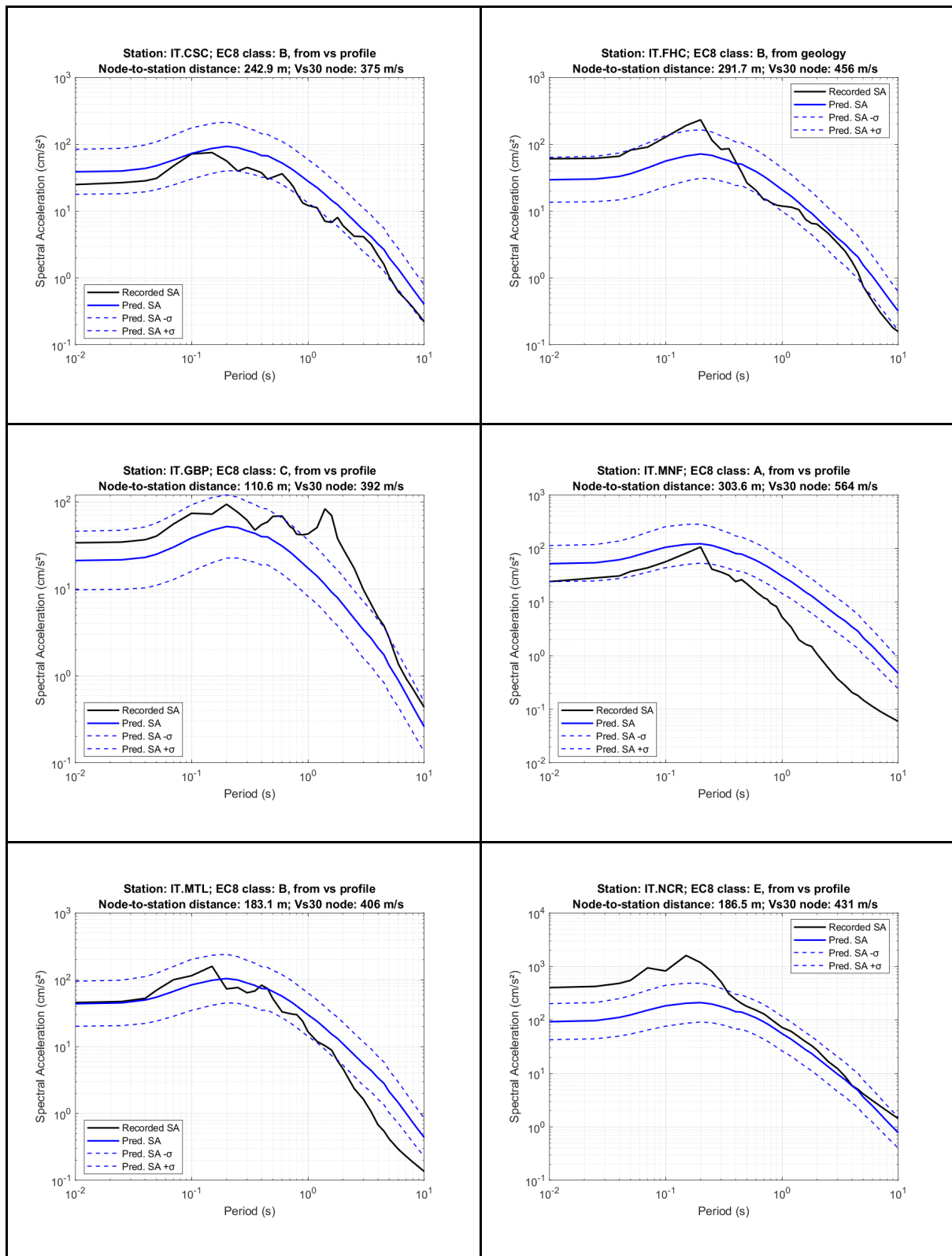


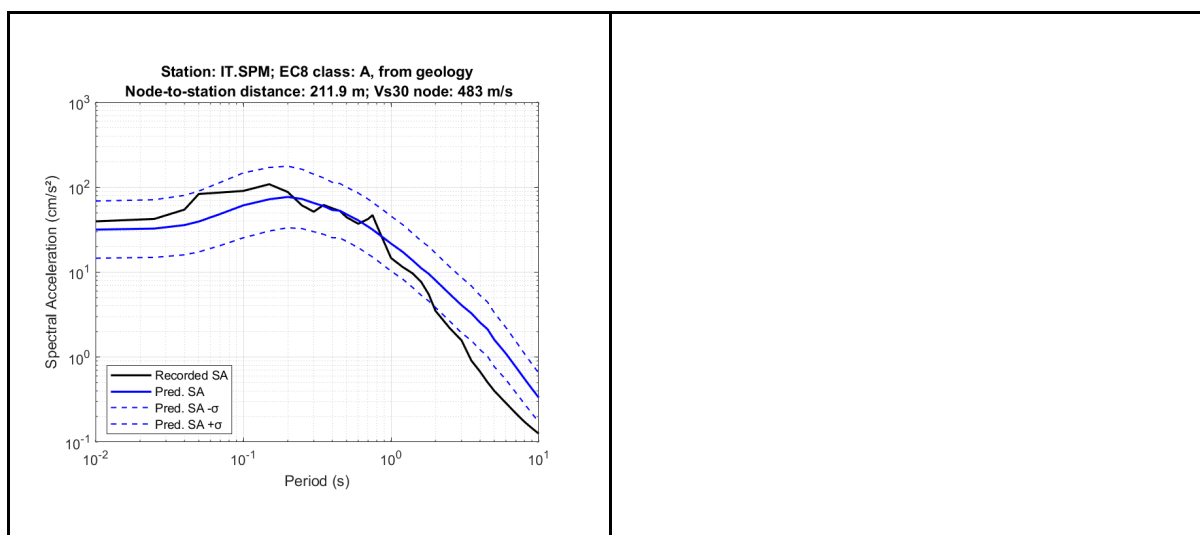


Appendix C

For the Mw 5.7 Colfiorito 1997 earthquake scenario, the acceleration response spectra predicted using the Vs30 map developed by Mori et al. 2020 are compared with the observed spectra recorded at 11 seismic stations.

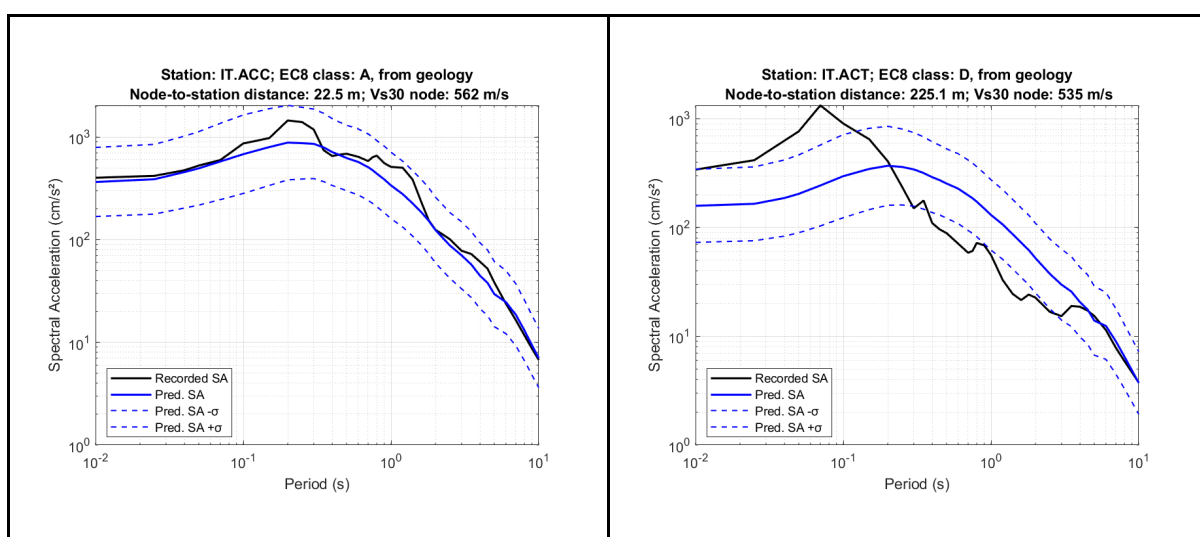


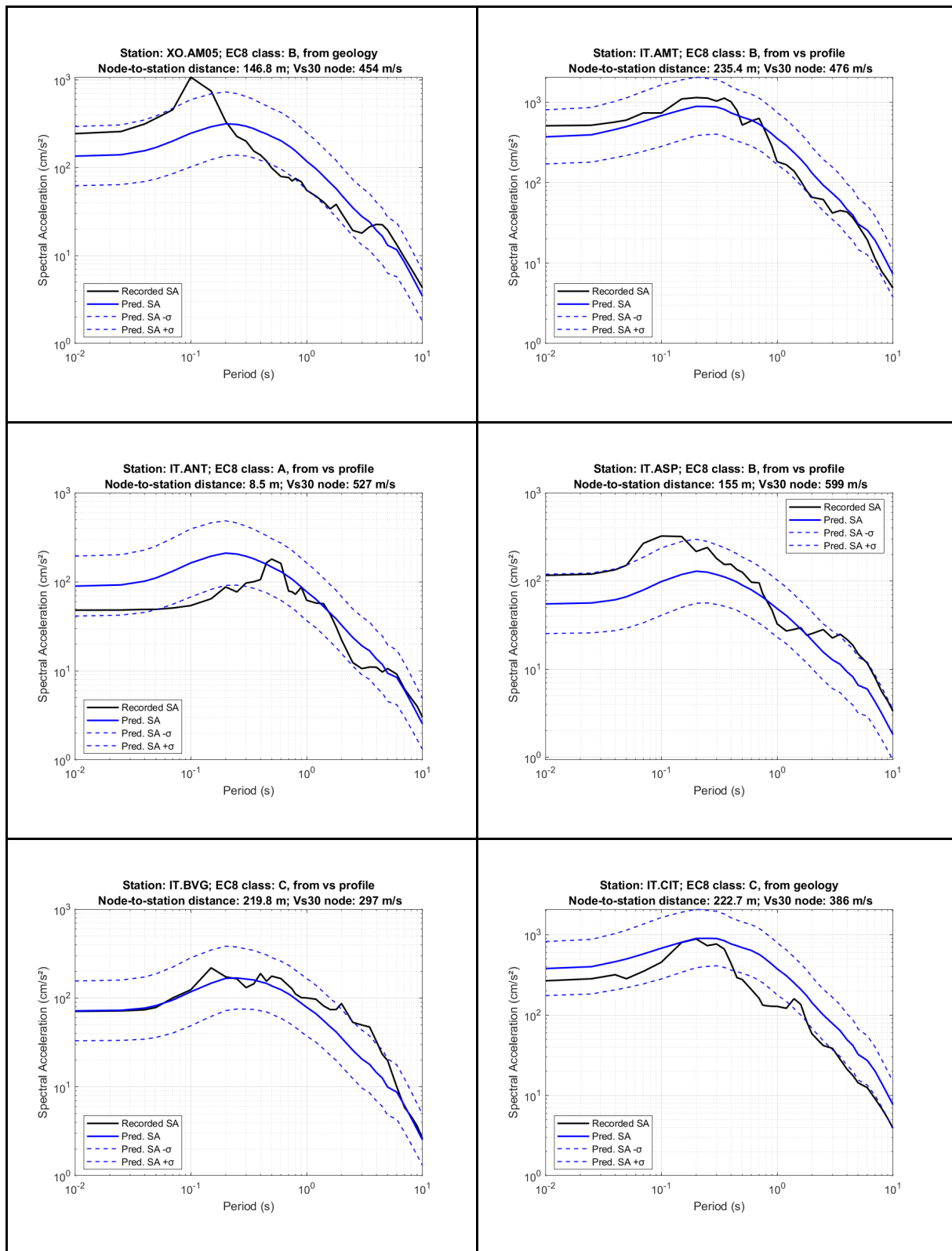


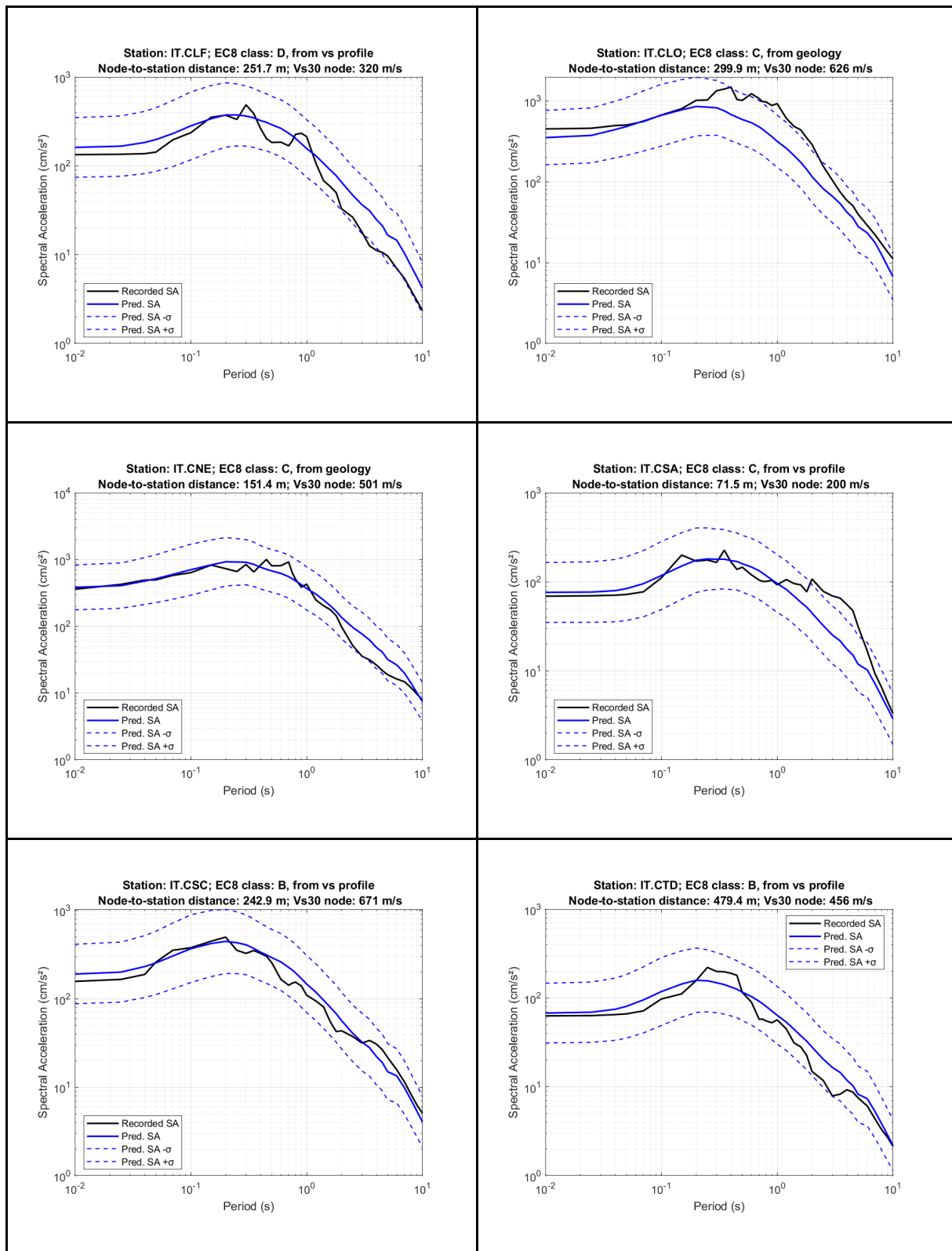


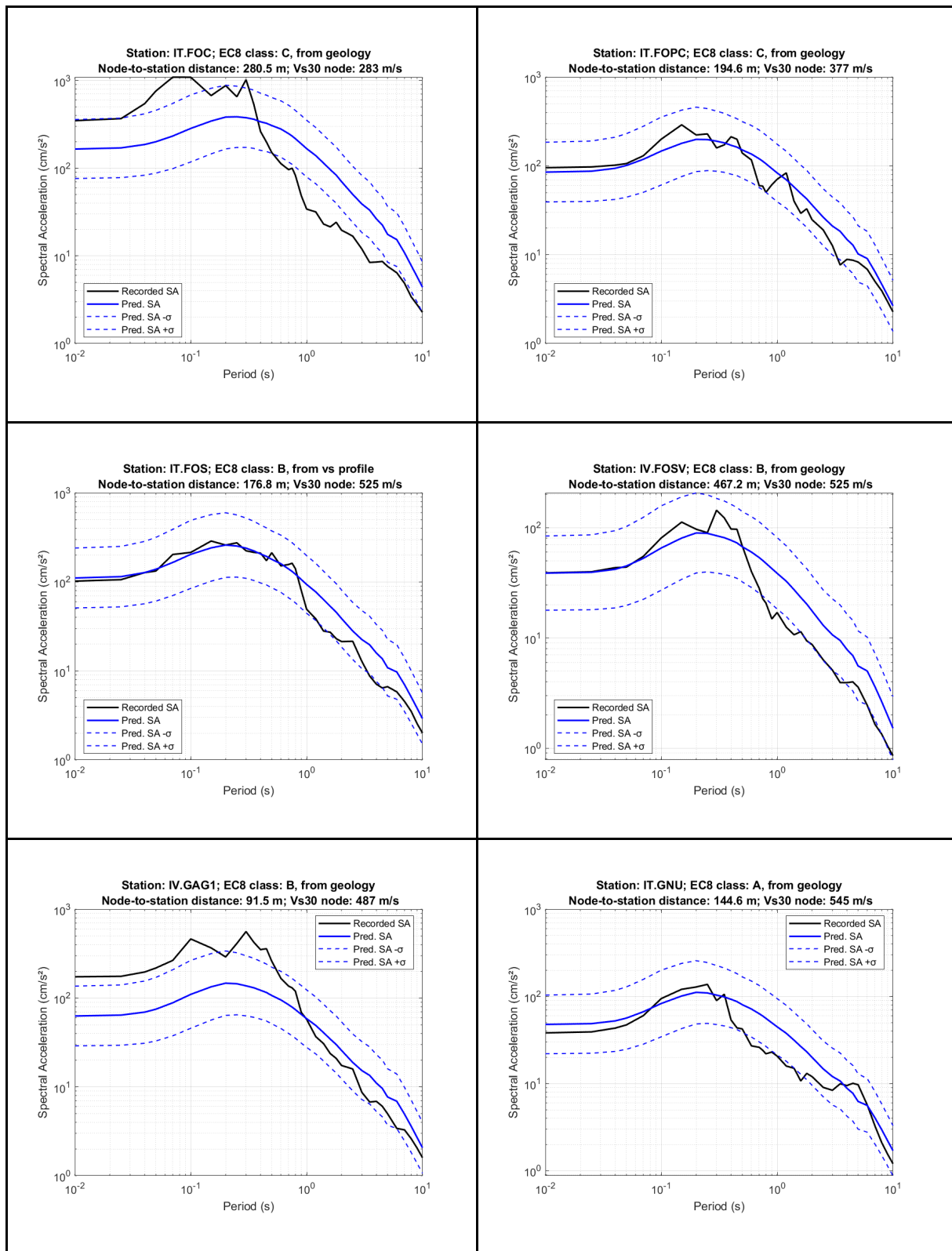
Appendix D

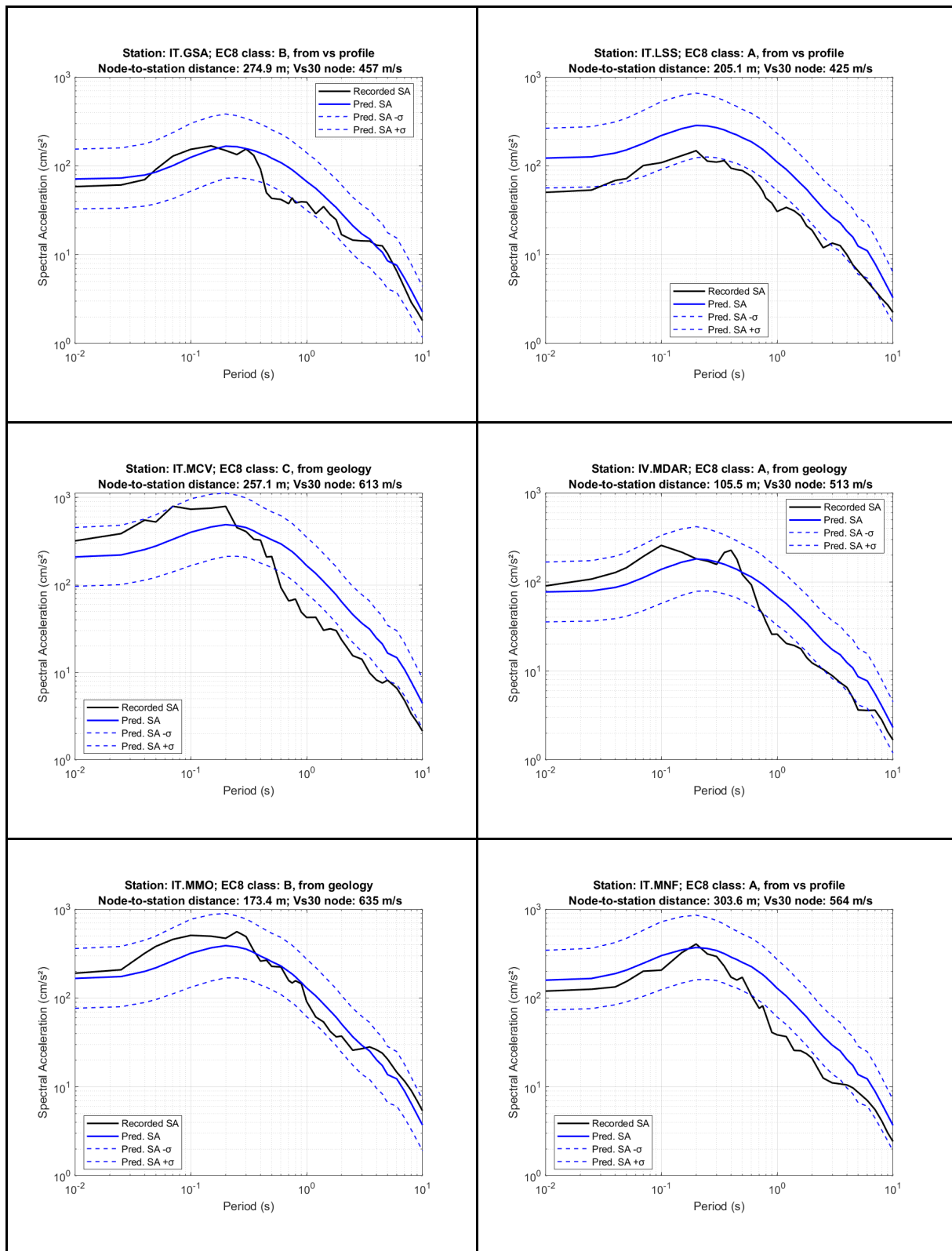
For the Mw 6.6 Norcia 2016 earthquake scenario, the acceleration response spectra predicted using the Vs30 map developed in this study are compared with the observed spectra recorded at 84 seismic stations.

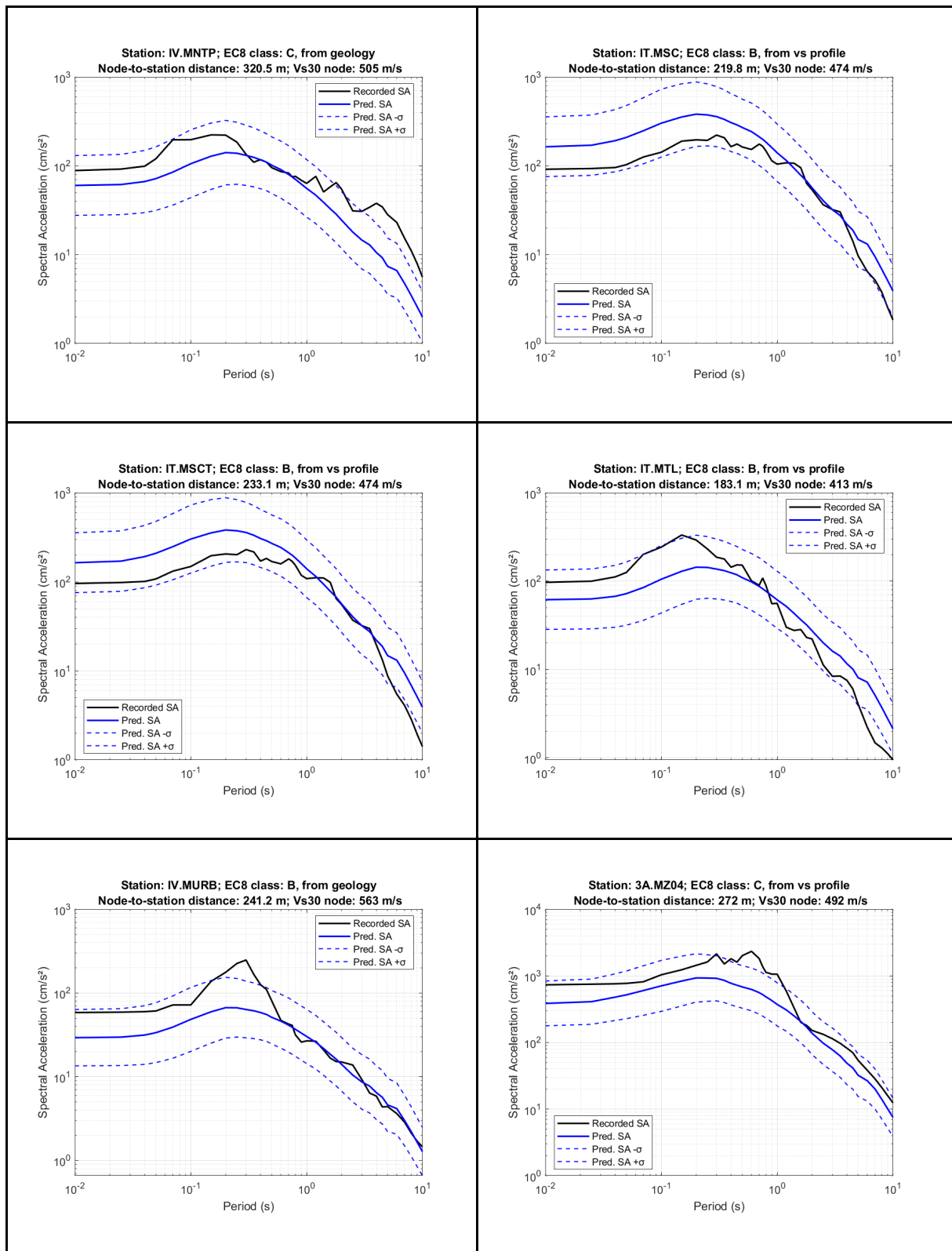


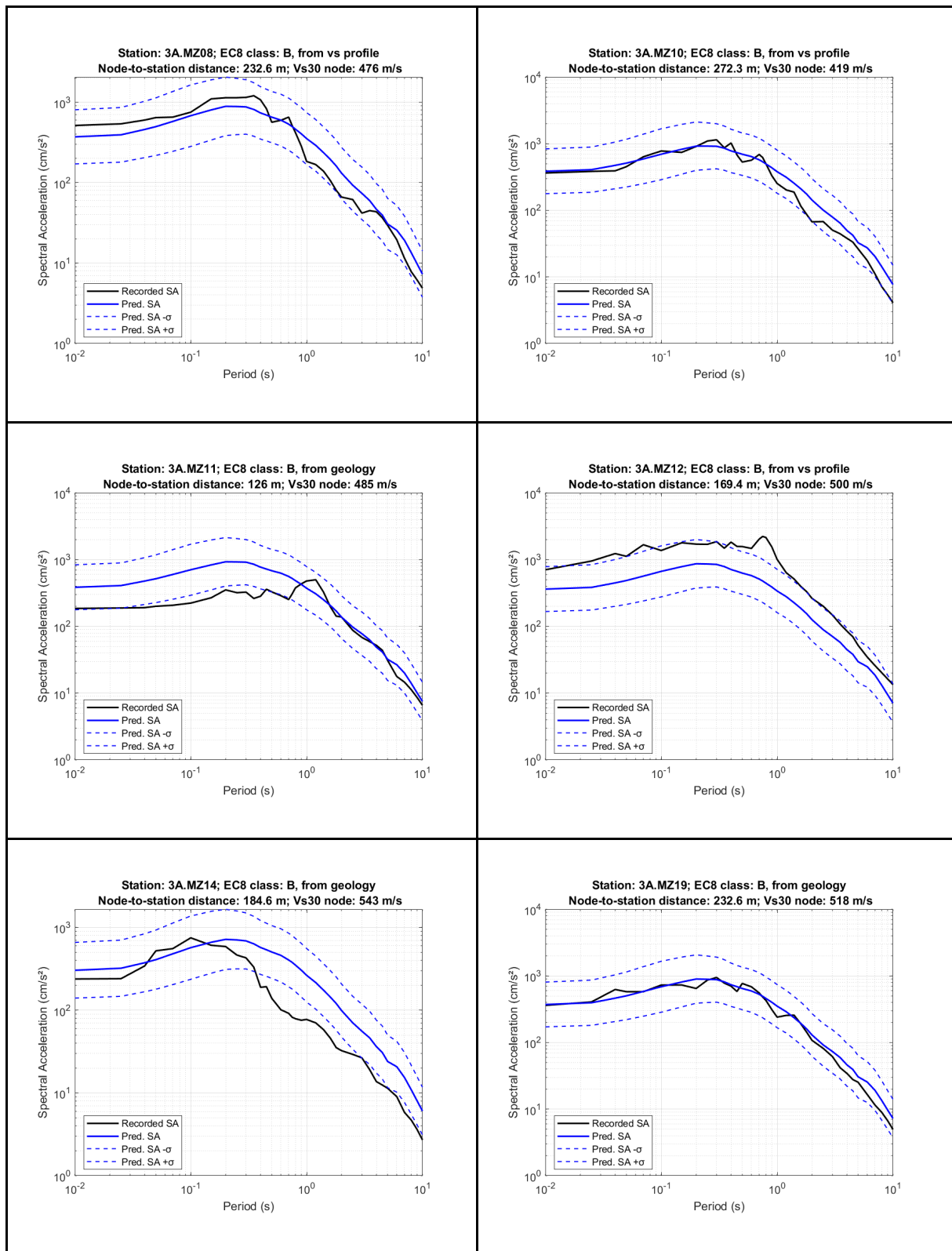


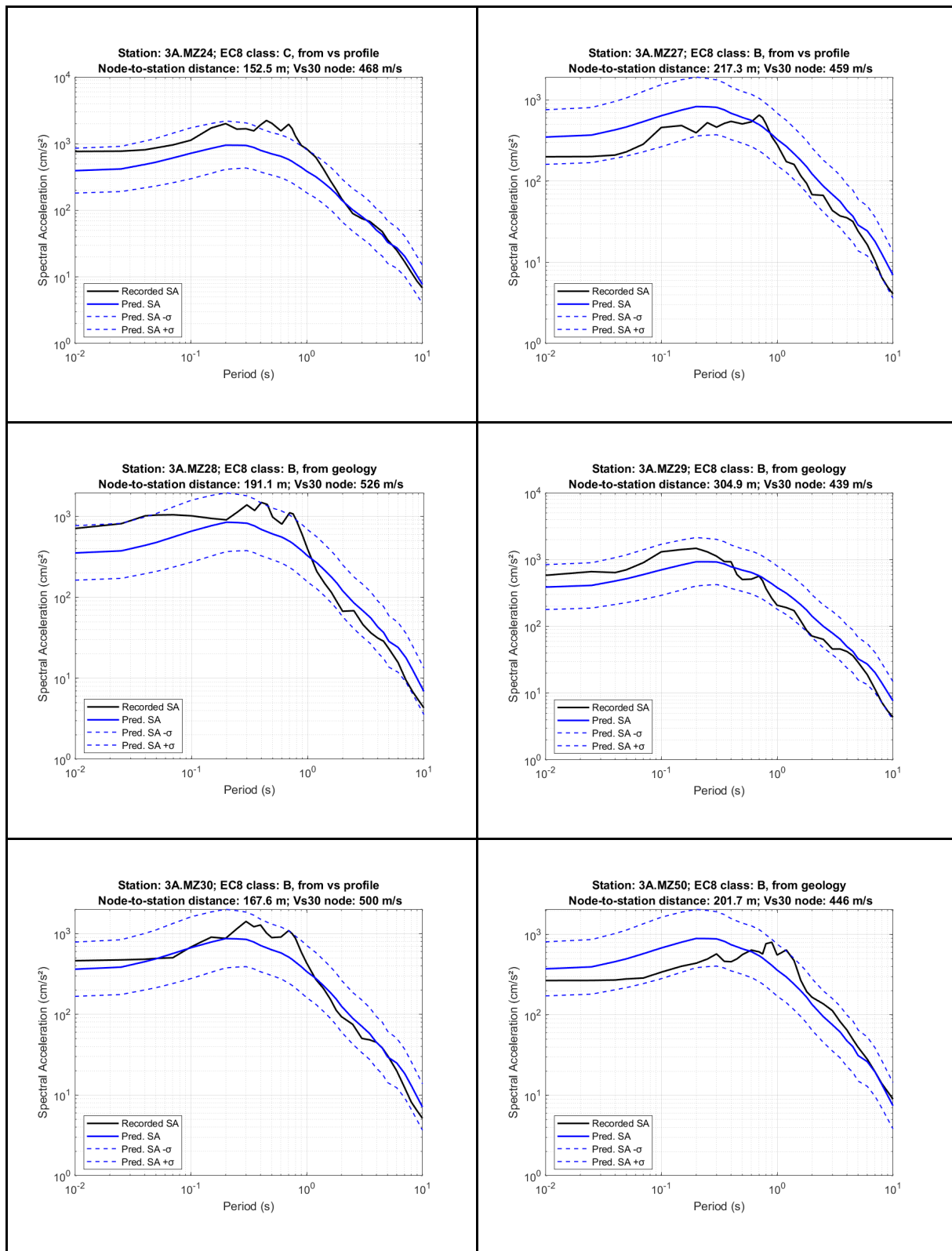


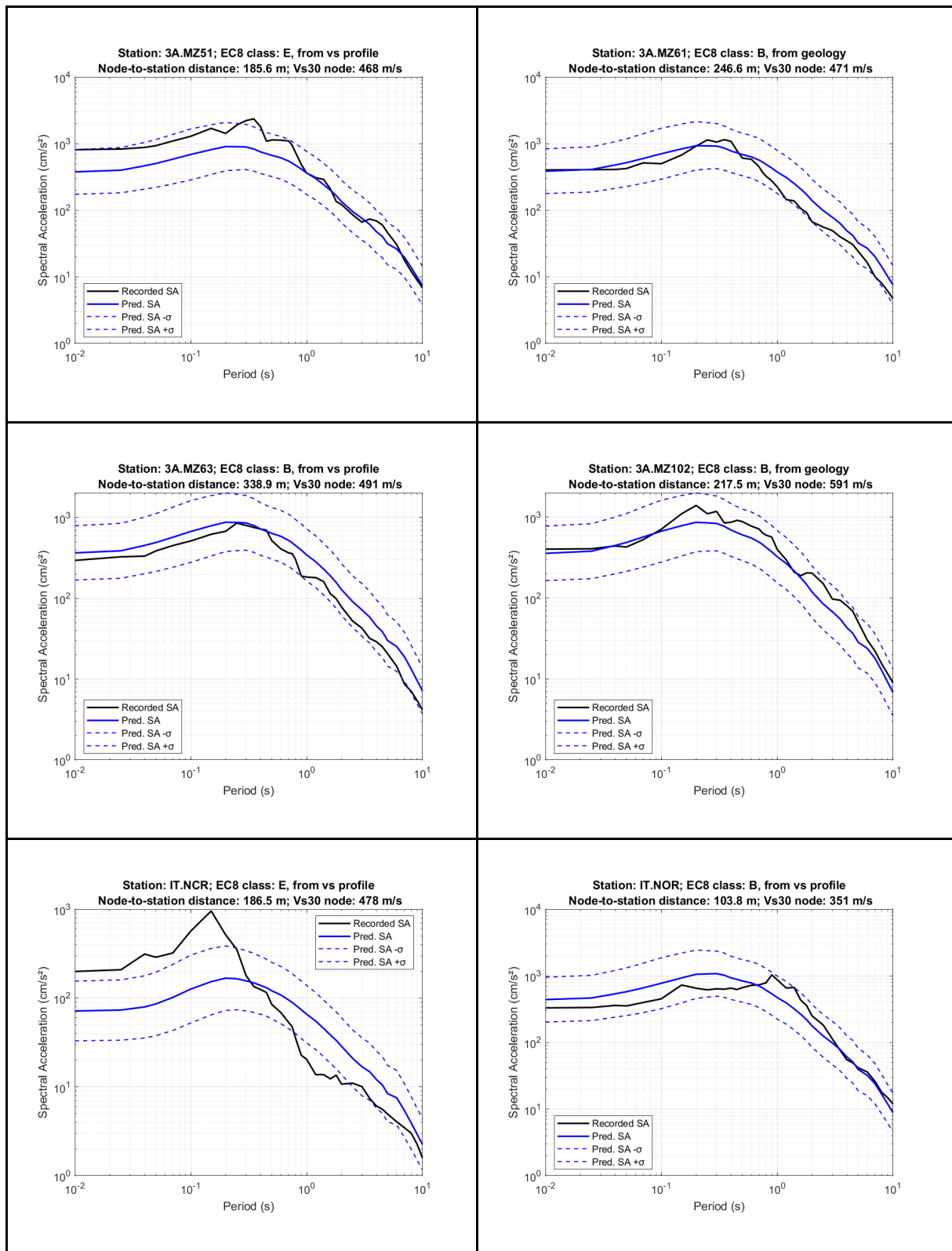


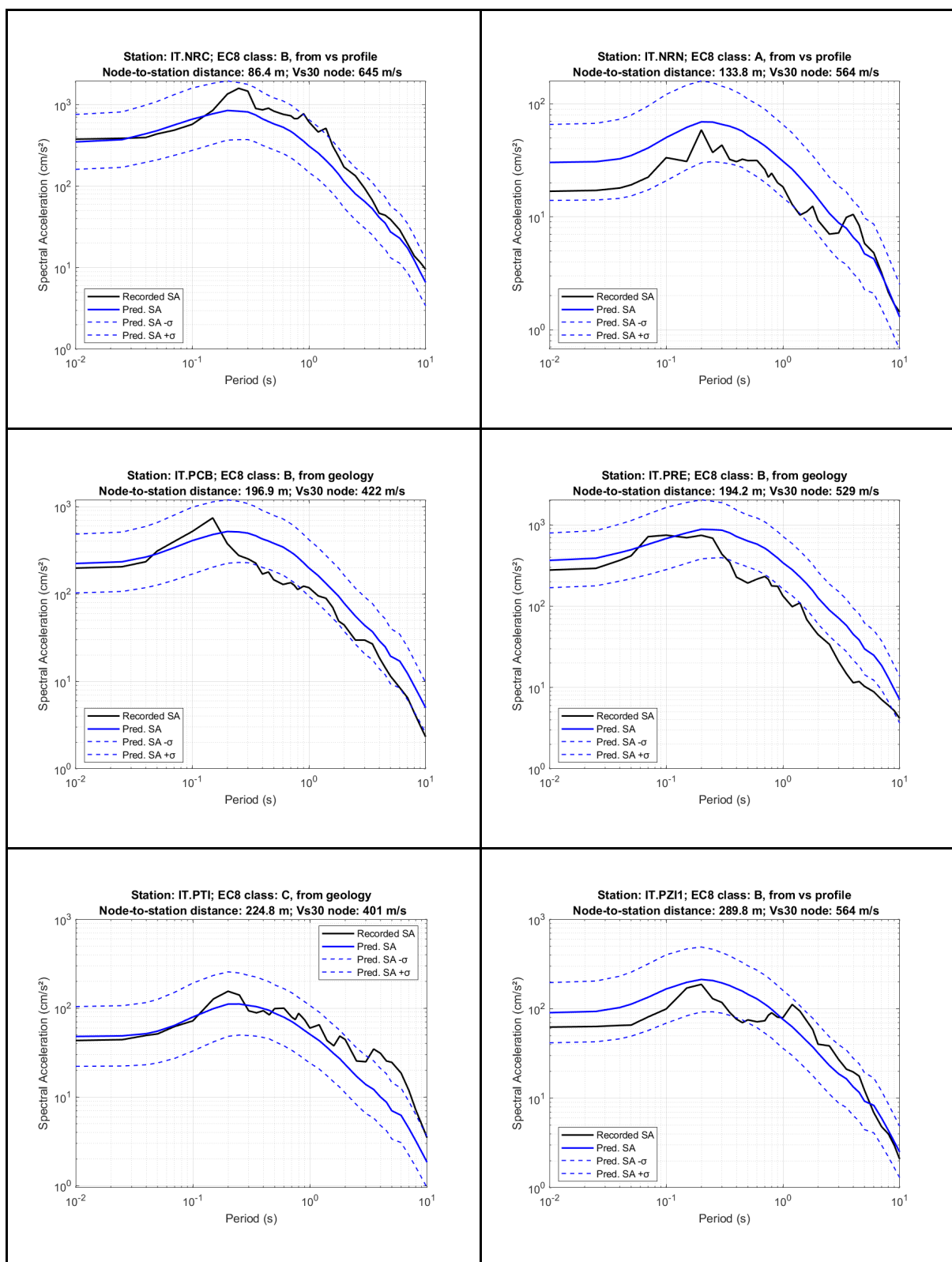


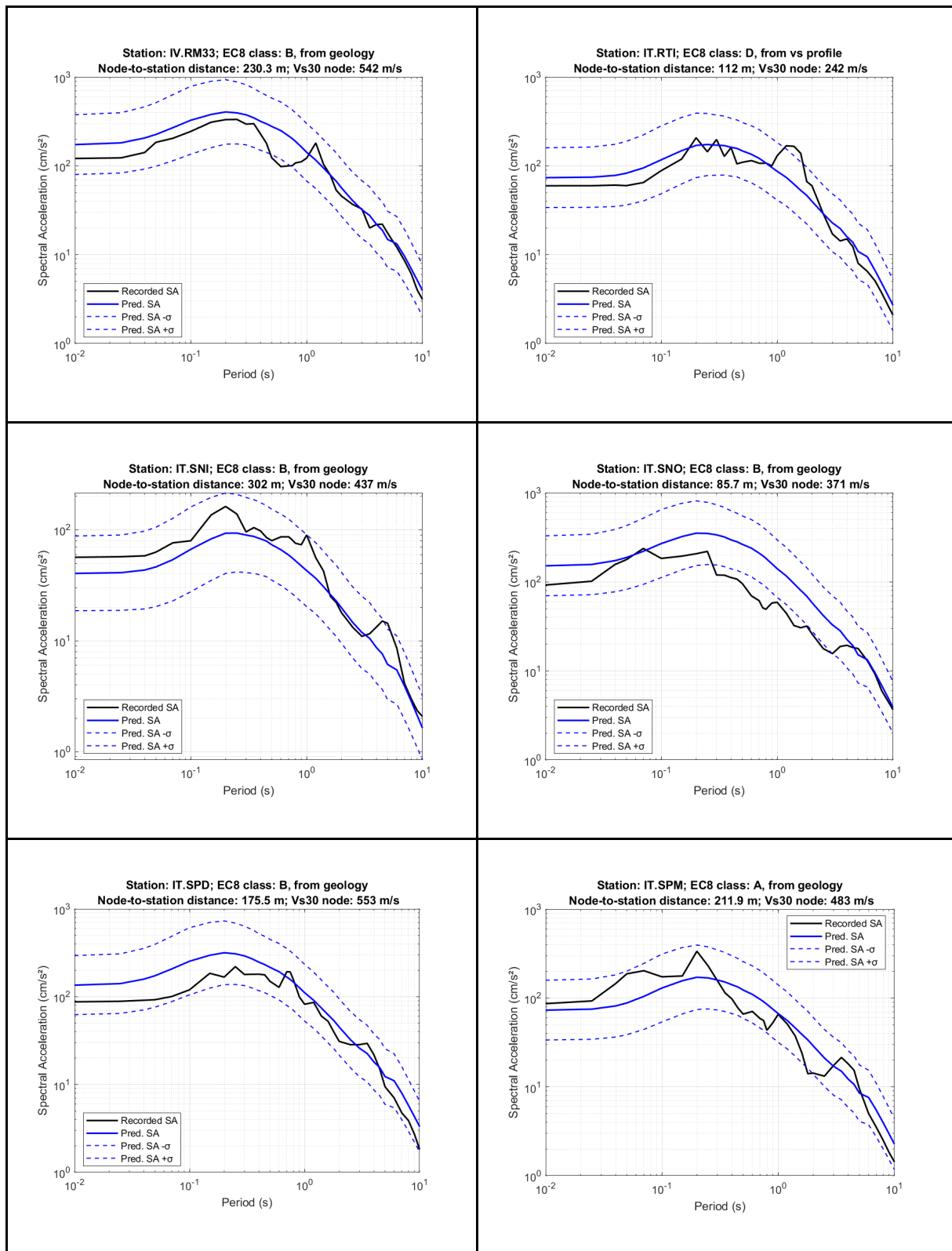


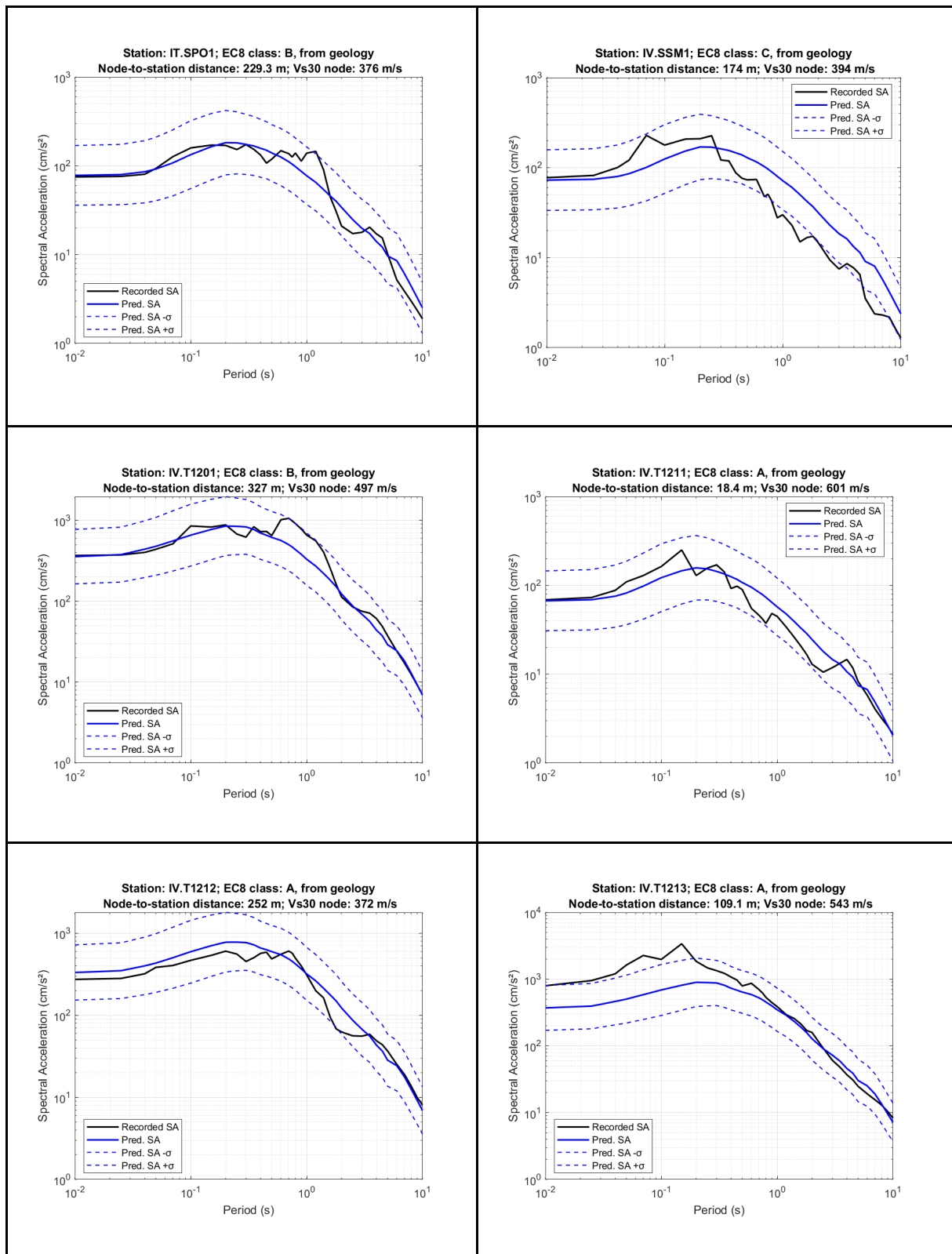


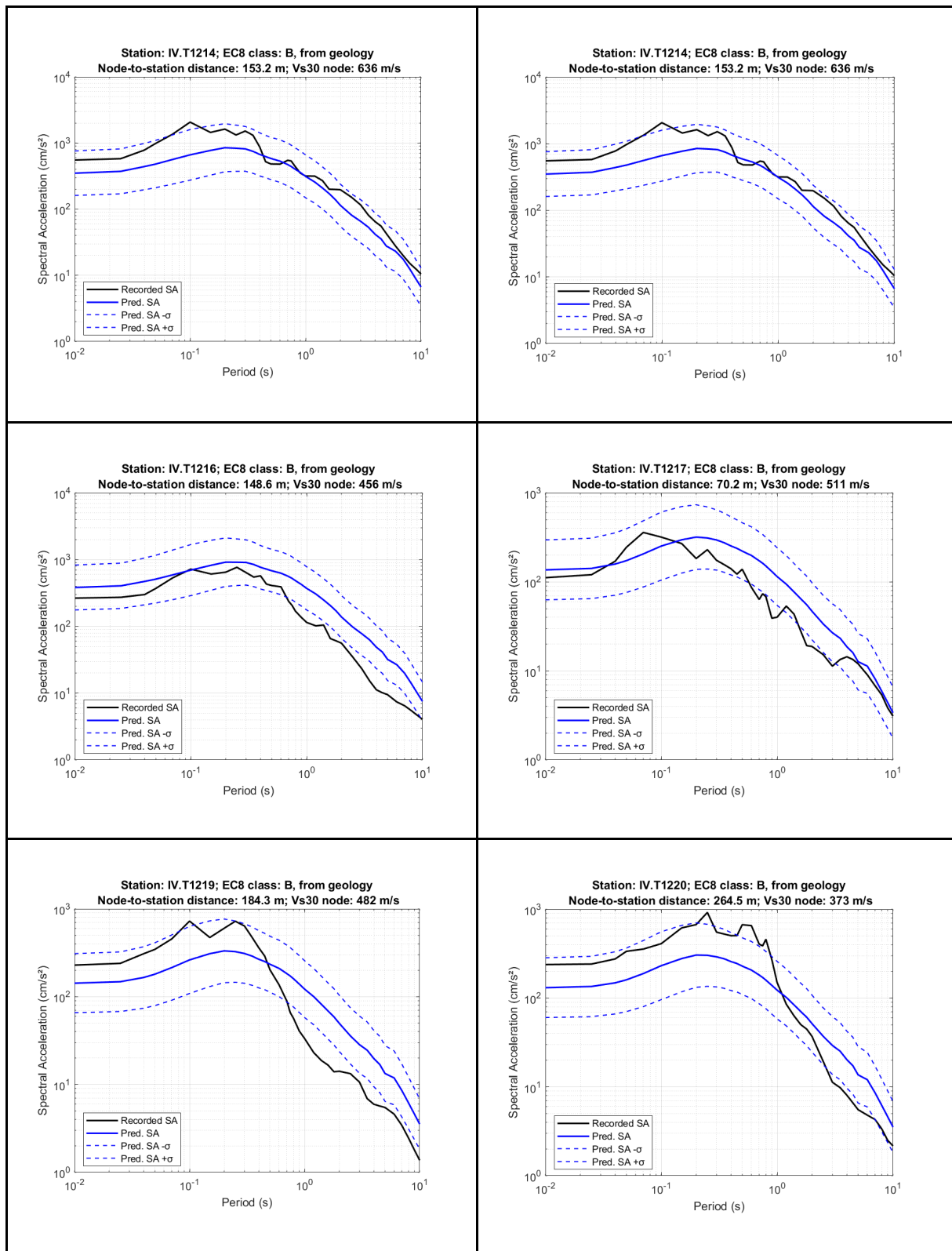


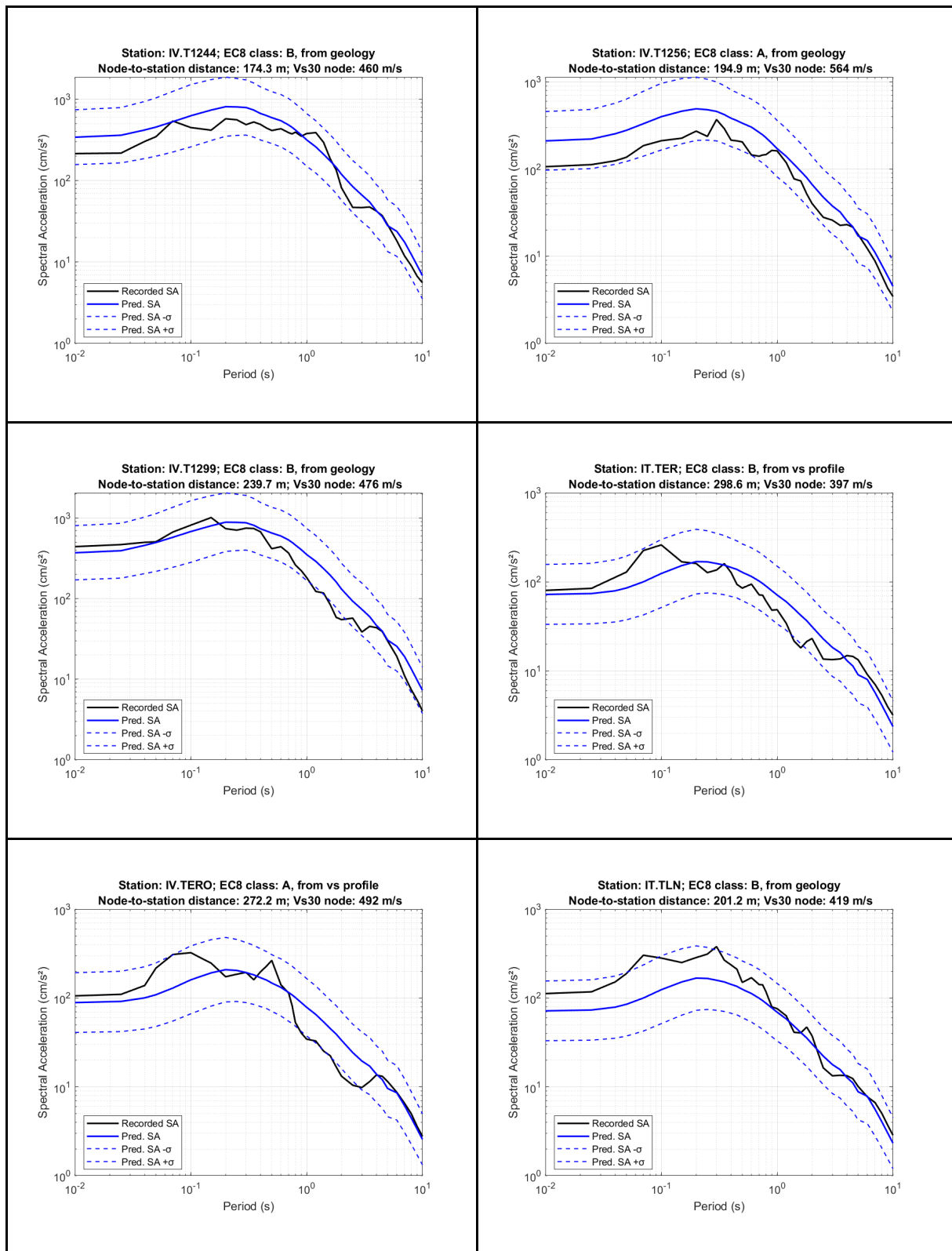


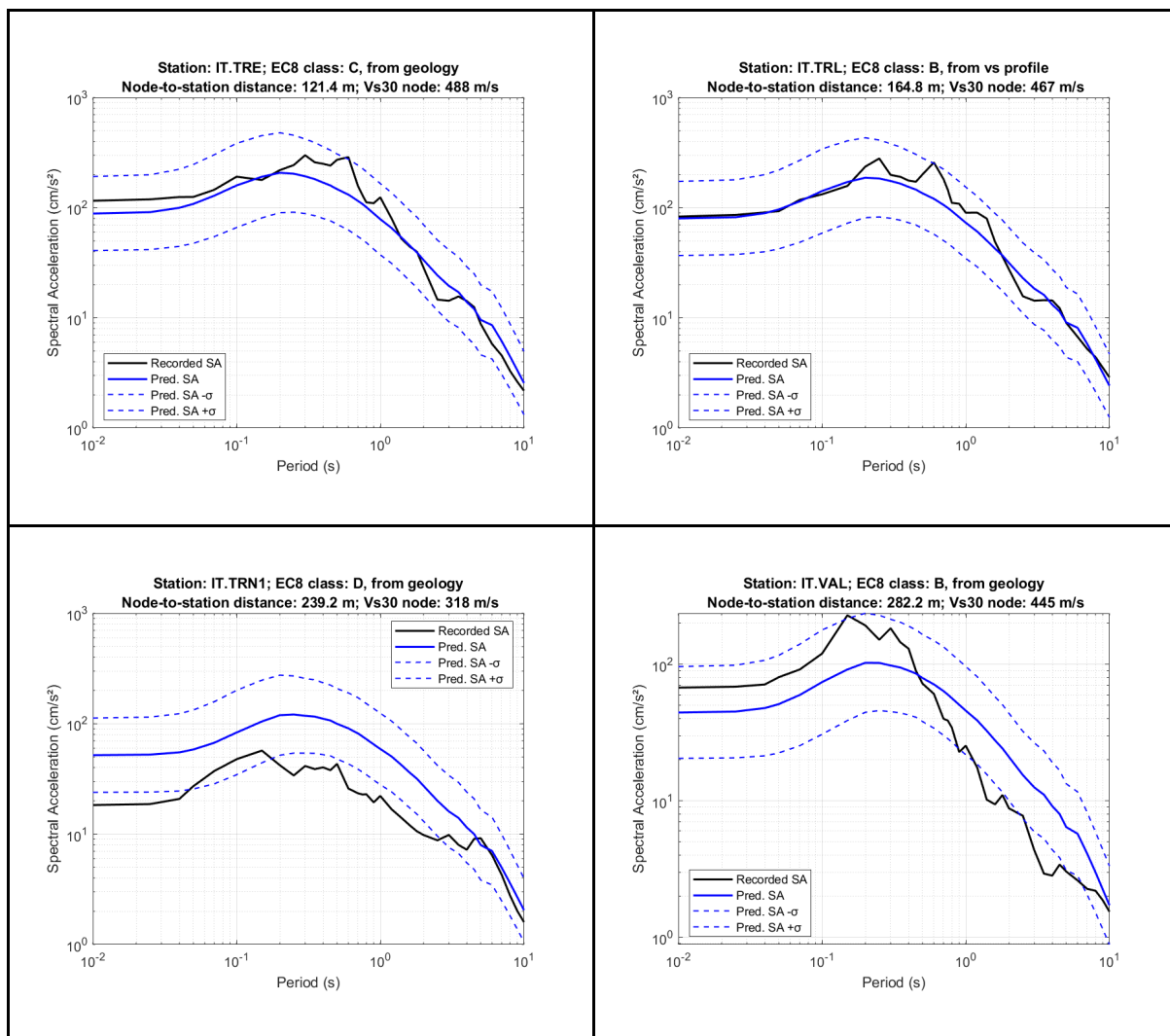












Appendix E

For the Mw 6.6 Norcia 2016 earthquake scenario, the acceleration response spectra predicted using the Vs30 map developed by Mori et al. 2020 are compared with the observed spectra recorded at 84 seismic stations.

