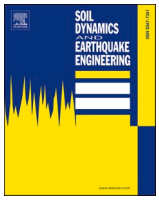
Soil Dynamics and Earthquake Engineering 179 (2024) 108496

Contents lists available at ScienceDirect

Soil Dynamics and Earthquake Engineering

journal homepage: www.elsevier.com/locate/soildyn



A geostatistical modelling of empirical amplification functions and related site proxies for shaking scenarios in central Italy

Sara Sgobba [a,*], Chiara Felicetta [a], Teresa Bortolotti [b], Alessandra Menafoglio [b], Giovanni Lanzano [a], Francesca Pacor [a]

[a] Istituto Nazionale di Geofisica e Vulcanologia (INGV), Via A. Corti 12, 20133, Milan, Italy
[b] MOX - Department of Mathematics, Politecnico di Milano, p.zza Leonardo da Vinci 32, 20133, Milan, Italy

A B S T R A C T

This work aims at identifying and modelling statistical dependencies between empirical amplification functions of sites in central Italy and the main geological and geophysical characteristics of the region, within a geostatistical analysis framework. The empirical functions, named δS2S, are estimated by decomposing the residuals of the median predictions of a non-ergodic ground motion model of elastic acceleration response spectra developed for the reference region. To select the model that best describes the spatial variability of the data, the performance of stationary and non-stationary spatial models is compared, the latter being able to constrain the prediction of the empirical functions to physical quantities available in the region and descriptive of the geology, topography and geographical location of the site. Finally, we obtain optimal models of δS2S, for each spectral ordinate, parameterised as a function of geographical coordinates and an input map of shear wave velocity in the upper 30 m (Vs30) constructed *ad hoc* by combining information gathered from two high-resolution maps available for the region. The methodology allows the development of a new practice-oriented framework for the empirical estimation of site amplification, which can be adopted for the generation of shaking scenarios in the context of regional hazard and seismic risk assessment.

1. Introduction

The estimate of ground shaking scenarios is a key-issue for evaluating damages and losses, for risk analysis, as well as for engineering purposes. Among the techniques for predicting ground shaking scenarios, the empirical methods based on the use of a ground motion model (GMM) are probably the most expeditious and widespread. GMMs are predictive equations used worldwide to estimate shaking intensity measures as a function of several parameters dependent on the reference earthquake scenario (magnitude, distance, etc.) together with variables representative of site conditions.

These empirical tools need a key-proxy to describe the effect of local site response, which is usually represented by the shear wave velocity in the first 30 m depth (Vs30) from in-situ geotechnical/geophysical measurements or inferred by other proxies (geology, topography, etc.). Vs30 is the most common site proxy, so researches focused on the application of techniques to spatialise Vs30 with the aim to get estimates of this parameter even at unsampled points. An example is the Vs30-map proposed by Thompson et al. [1] and used within the Shakemaps [2–4] or other similar maps that have been linked to different available spatial information, such as surface geology [5,6], topographic slope [7] or digital terrain [8].

Although Vs30 has the advantage of being a synthetic predictor, on the other hand it has been shown that it is not truly representative of site response [9–14].

Therefore, in parallel with the increasing availability of data from the development of new seismic networks in the world, the tendency is moving towards the use of alternative and more explanatory site proxies, such as those estimated from the residuals of empirical ground motion modelling, named site-to-site terms δS2S ([15–17], among the others). These terms represent systematic deviations of the observed amplification at the site from the median values empirically predicted by the GMM that can be quantified as random-effect terms of the model regression; they can be robustly calculated when multiple recordings are available at multiple sites. δS2S is arguably the most reliable representation of empirical site response if estimated from a sufficient number of site-specific ground-motion observations [16,18] and would help to physically explain the variability between sites [18–20]. Because of these features, δS2S is increasingly used, especially to adjust empirical predictions of partially or fully non-ergodic GMMs [21] to obtain ground motion predictions in a non-ergodic Probabilistic Seismic Hazard Assessment (PSHA) [22] or even to capture site-specific amplification with respect to reference motion in scenario predictions.

For the above applications, and thus in a similar way as for Vs30,

* Corresponding author.
E-mail address: sara.sgobba@ingv.it (S. Sgobba).

https://doi.org/10.1016/j.soildyn.2024.108496
Received 11 September 2023; Received in revised form 15 December 2023; Accepted 15 January 2024
Available online 6 February 2024

there is a need to adopt continuous spatial models of δS2S, in order to capture these effects within regional GMM predictions. As a consequence, the reliability of the techniques adopted to estimate and map the site parameters becomes relevant for generating robust shaking scenarios.

In this context, geostatistical methods have often been applied in the literature to infer the Vs30 proxy; for example, they have been used to produce region-specific correlations between seismic properties and geological units [23]. The most widespread techniques are based on the classical methods for extending a variable in space, i.e. the Kriging technique, which is a well-known geostatistical method used in various fields to estimate the value of a variable in non-sampled locations using observables of the same type in other locations [24]. Kriging is often applied in its simplest form (Ordinary Kriging, OK), which however has limited prediction capability in less densely sampled areas as it is constrained only to observations; it follows that the resulting maps ignore geological parameters or information that might be available in the region. To overcome such a limit, the authors explore statistical correlations with other mappable site proxies to develop parametric models of δS2S, and then spatially correlate the model residuals. As an example, Wheatherill et al. (2020) [25] provide a comprehensive attempt to explore these correlations by focusing on slope and geology. In their work, the authors mapped the site-to-site amplification term as a function of the other investigated proxies starting from the dense dataset of recording stations of the Japanese strong-motion networks (KiK-net and K-NET). More recently, Loviknes et al. [26] derived site amplification models from the relation between δS2S for sites over large areas (Europe and Turkey) and other proxies including the geomorphological sediment thickness. Fewer studies however are focused on the spatial predictions of the empirical site terms δS2S, considering also possible non-stationarities in their correlations (Chao et al., 2020; [27]). In some cases more advanced techniques such as the co-kriging methods have been used to extend the application of Kriging to the multivariate setting in a Bayesian framework, such as Gilder et al. [28] and De Risi et al. [29].

In the present study, we apply spatial non-stationary techniques such as Universal Kriging [24] or Kriging with-a-trend to go beyond Gaussianity of the random field and try to better capture local variability of δS2S. In such a way, the estimates are not only constrained on the observations but also on a proper set of site proxies progressively made available in the area and more readily accessible, such as Vs30 maps, lithostratigraphic maps, etc. In doing this, we also go through the development of a new map of Vs30, instrumental for mapping δS2S, which is obtained by combining the information collected by two high-resolution maps available for the region. Note that although the objective of constraining empirical site amplification to geological/geophysical parameters may be more appropriately achieved in terms of Fourier Amplitude Spectrum (FAS), as it is more closely related to physics (e.g. Refs. [30,31]), in the present study we aim to model δS2S terms of an elastic spectral acceleration (SA) model as our main goal is to predict ground motion for shaking scenario studies or hazard applications.

The area of Central Italy is chosen as the target because of the great availability of information. Indeed, it is characterised by the presence of geologic and tectonic structures, alluvial basins and mountain chains (e.g. the central Apennines), as well as a complex surface geology that motivates the need to develop high-resolution soil amplification maps. In addition, it represents a highly seismic area, affected in the last 30 years by important sequences (Umbria-Marche 1997–1998, Mw 6.0; L'Aquila 2009, Mw 6.1; Amatrice-Visso-Norcia 2016–2017, Mw 6.2-Mw 5.9-Mw 6.5), which have been recorded by a large number of seismic stations and that has led to the calibration of regional GMMs in the non-ergodic framework [31,32], as well as to studies dedicated to the local seismic response and microzonation studies [33].

The paper is outlined in the following steps: i) description of the dataset of the empirical amplification functions δS2S extracted from a regional GMM, as well as the set of site proxies considered in the explorative analysis; ii) definition of the methodological approach; iii) analysis of the spatial dependences among the different proxies and identification of predictors candidates; iv) selection of the optimal spatial model via leave-one-out-cross-validation (LOO-CV) and building of the final optimal maps.

2. Dataset

The data we aim to model are empirical site amplification functions (also defined as site-to-site terms δS2S) quantified via a mixed-effects non ergodic ground motion modelling (GMM) of 5 % damped elastic response spectral acceleration (SA) (the flatfile of the input δS2S and site metadata is provided in **ESUPP1_flatfile.zip**).

The calibration dataset consists of more than 30,000 waveforms relative to about 450 earthquakes, occurred in Central Italy in the time interval 2009–2018, in the magnitude range 3.2–6.5 (local magnitude for $M < 4.5$ and moment magnitude for $M \geq 4.5$) recorded by about 460 stations within 250 km from the epicentres (see also detail on the webpage https://shake.mi.ingv.it/central-italy/; last accessed August 1, 2023). The PGA values are distributed in the range [0.003–1000] cm/s^2. On this dataset, it appears that a marginal fraction of the sites can potentially be affected by a non-linear response (0.06 %), assuming that non-linearity may occur above a PGA threshold of 100 cm/s^2 and with a shear stress ratio PGV/Vs30 > 0.1 % [34]; thus, we can assume that the assumption of linearity holds.

The GMM calibrated for this study is a slightly modified version of that proposed by Sgobba et al. [32] calibrated for shallow active crustal earthquakes in Central Italy. This model is *ad-hoc* developed for the present study (details are provided in the **APPENDIX**) and is calibrated on the geometric mean of the horizontal components of the PGA and 69 spectral ordinates of the acceleration response spectra (5 % damping) SA in the period range T = 0.04–2 s; however, for the sake of simplicity, in the following we will only show results and maps referred to two parameters, representative of short and long period behavior: PGA and SA (1s). Coefficients and parameters of the reference GMM are provided in **ESUPP2_GMM.zip**.

Below, we focus on δS2Sref terms (herein simply denoted by δS2S) that represent the average misfit of ground-motion at the site with respect to the event-corrected median value predicted by the GMM. They include site-related features due to missing covariates that are not fully explained by the regression model. In more detail, the δS2S terms define the systematic bias of ground motions recorded at each station *s* with respect to the reference rock motion predicted by the fixed-effect part of the GMM, which is calibrated on 6 reference rock sites (see **APPENDIX** for major details). The reference rock differs from the generic rock conditions typically defined by the time-averaged shear wave velocity at the top 30 m of a 1-D soil column, (Vs30 = 800 m/s); it is defined as out-cropping rock or stiff soils that show a flat, unamplified response over a frequency range of engineering interest (Felicetta et al. [35]) - usually f = 0.5–25 Hz - and identified by a multi-criteria weighted decision matrix method (Reference Rock Identification Method, RRIM, [15]) used to rank alternative criteria. In this way, the definition of the δS2S term is slightly different from that of the typical systematic site-to-site residual that takes into account effects not captured by the median prediction of the reference GMM. In fact, the term here is derived from a GMM that lacks an explicit linear site amplification response term in the fixed coefficient model, so the resulting δS2S represents the soil amplification w.r.t. reference rock condition, as also shown by Sgobba et al. [32] and Kotha et al. [19].

Hence, the so-computed empirical site functions are able to capture the effects of soil response including amplification due to shallow soft sediments, contrast of impedances, resonances and 2D/3D effects, which are not described from synthetic proxies as the Vs30 [33,36–38]. Recently δS2S, was also used as a reference proxy to validate the amplification coefficients suggested in the new Eurocode proposal [39],

as well as to compare and validate the amplification factors obtained from microzonation studies at local and national scales in Italy [40]; they also were found to contribute more heavily to the ground shaking variability [32,41].

In the following, all the analyses are performed in a scalar framework, meaning that δS2S extracted from SA models are treated separately at every period T, and that no correlation between different ordinates is considered. Each record (i.e. δS2S vs T) is associated with the latitude and longitude identifying the geographical location of the corresponding station (Fig. 1a–b), allowing us to exploit geostatistical techniques for modelling the site proxies.

2.1. Geological and geophysical site proxies

We use a proxy-set for investigating correlation with the empirical amplification functions to select the covariates useful for geospatial modelling (Table 1), which include:

i) The shear-wave velocity Vs30 [42]: it is inferred from different source data and accounting for litho-stratigraphic effects. It provides a synthetic description of the subsurface structure (first 30 m depth) that is more accessible to geotechnical investigations. It is also the most commonly included explanatory variable in ground motion models, as well as an easy and low-cost parameter for evaluating site response classification (e.g., Ref. [43,44]);
ii) the high-frequency attenuation parameter κ0 (or "site kappa"): it measures the attenuation of the acceleration spectrum in log-linear space at high-frequencies, according to definition provided by Anderson and Hough [45] and the taxonomy by Ktenidou et al. [46];
iii) the topographic slope;
iv) the local descriptions of the litho-stratigraphic units.

Note that we do not consider here additional geological/geophysical data potentially relevant for this study, such as f0 (fundamental frequency of the soil deposit) and H800 (the depth of the engineering bedrock) because these data are presently scarce in our dataset, due to the large number of temporary stations, and maps of them at national scale are not currently available.

Table 1
List of the parameters and site proxies considered in the analysis.

Site proxy	Notation and unit of measure	Type of estimation	Reference
Empirical site amplification function	δS2S [log10]	Residual decomposition with respect to a regional GMM	Sgobba et al. [31,32]
Shear-wave velocity in the uppermost 30 m	Vs30 [m/s]	Geophysical measurements	ESM - Luzi et al. [47]
	Vs30-WA [m/s]	Inferred from empirical correlation with topographic slope from the function-Tinitaly DEM 10 m resampled at 40 m	Wald and Allen [7]
	Vs30-Mori [m/s]	Inferred from site-specific microzonation dataset and one-dimensional numerical approach	Mori et al. [13]
Slope	Slope [deg]	Topographic slope from the function-Tinitaly DEM 10 m resampled at 40 m	Mascandola et al. [48]
High-frequency decay parameter	κ0	Empirical procedure from Generalised Inversion Technique (GIT) analysis	Morasca et al. [49]
Geology	Litho-stratigraphic units	Large-scale lithological maps	i. ISPRA 1:100.000; ii. Forte et al [50]

The maps showing the distribution of the input data in the study region are provided in separate plots in **ESUPP3_plot_of_input_data.docx**.

Data Sources.

Herein, the data of Vs30 (m/s) are taken from different estimates.

1) *direct in-situ measurements* of the S-wave velocity profile at stations provided by the Engineering Strong-Motion database, ESM https://esm-db.eu/, [47]. Less than 25 % of the dataset is associated with this

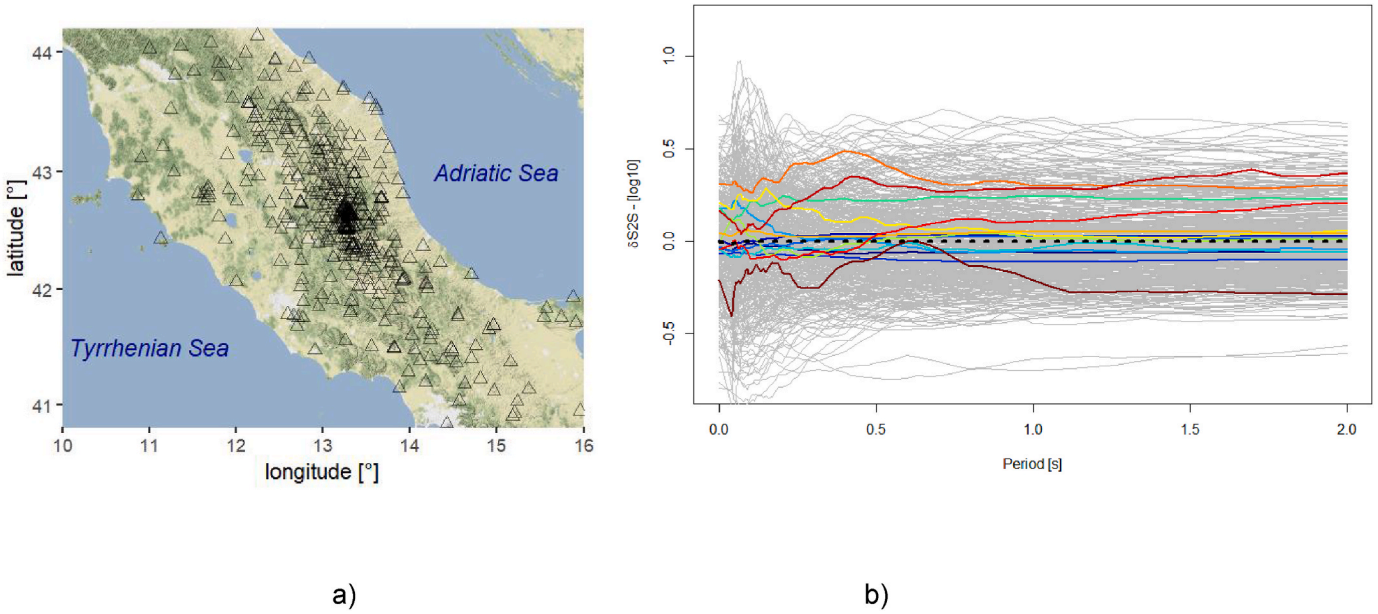


Fig. 1. Geographical distribution of the set of stations located in the region of Central Italy (a); amplification curves δS2Ss vs period T derived from the reference GMM (the dashed black line is the mean curve over all sites; the colored lines are the median of each lithological class color-coded according to the grouping shown in Fig. 3) (b).

type of measurement, while the information is missing in the rest of the dataset;

2) *inferred from empirical correlation with the topographic slope* as prescribed in Wald and Allen [7] on the basis of slope measurements. Herein, slopes come from high-resolution digital elevation models (DEM) of Italy sampled at 40 m [48]. In the following, the corresponding inferred proxy is denoted as Vs30-WA;
3) *inferred from site-specific microzonation* dataset (Mori et al., [13]). It is here referred to as Vs30-Mori. This map (50 m × 50 m raster) covers the whole of Italy and integrates data from the Italian seismic microzonation dataset ([51], www.webms.it), consisting of about 11,000 shear wave velocity profiles from geophysical surveys and 35,000 borehole logs.

The reason why we chose to introduce these last 2 proxies of Vs30 (Vs30-WA and Vs30-Mori) lies in the fact that dataset of Vs30 observations present a lot of missing values (75.8 % of the sample) so its sampling is not enough to condition the amplification estimates in a robust manner. The availability of high-resolution maps for Vs30-WA (Fig. 2a) and Vs30-Mori (Fig. 2b) in the Italian territory allows Vs30 information to be obtained almost continuously in space, thus motivating the evaluation of a non-stationary spatial model for Vs30 as dependent on Vs30-Mori and Vs30-WA. This will allow us not only to predict Vs30 at locations where its direct measurement is missing, but also to construct a high-resolution shear-wave velocity map for Central Italy.

The distribution of these two inferred maps (Fig. 2c) reveals the presence of a sampling bias, as both Vs30-WA and Vs30-Mori are unable to capture the full range of variability in the Vs30 observations. This is particularly evident in the case of the Mori' s map, whose statistical distribution appears particularly narrow and with a smaller standard deviation than that observed in the actual data. On the other hand, the WA map, although showing a more realistic variance, tends to overestimate the Vs30 observations.

Regarding κ0, we consider the values extracted by Morasca et al. [49] on the same dataset of Central Italy from a Generalised Inversion Technique (GIT), which is a non-parametric tool for evaluating empirical source, path and site contributions from earthquake recordings in the frequency domain [52]. About 17 % of observations (82 vs. 460 stations) are missing for this variable. The average κ0 for the 6 reference rock sites on which the GMM was calibrated is 0.012s [31], while the average value over all stations in this dataset is equal to 0.026 s.

Another source of information continues in space is related to shallow geology and more specifically to large-scale lito-stratigraphic

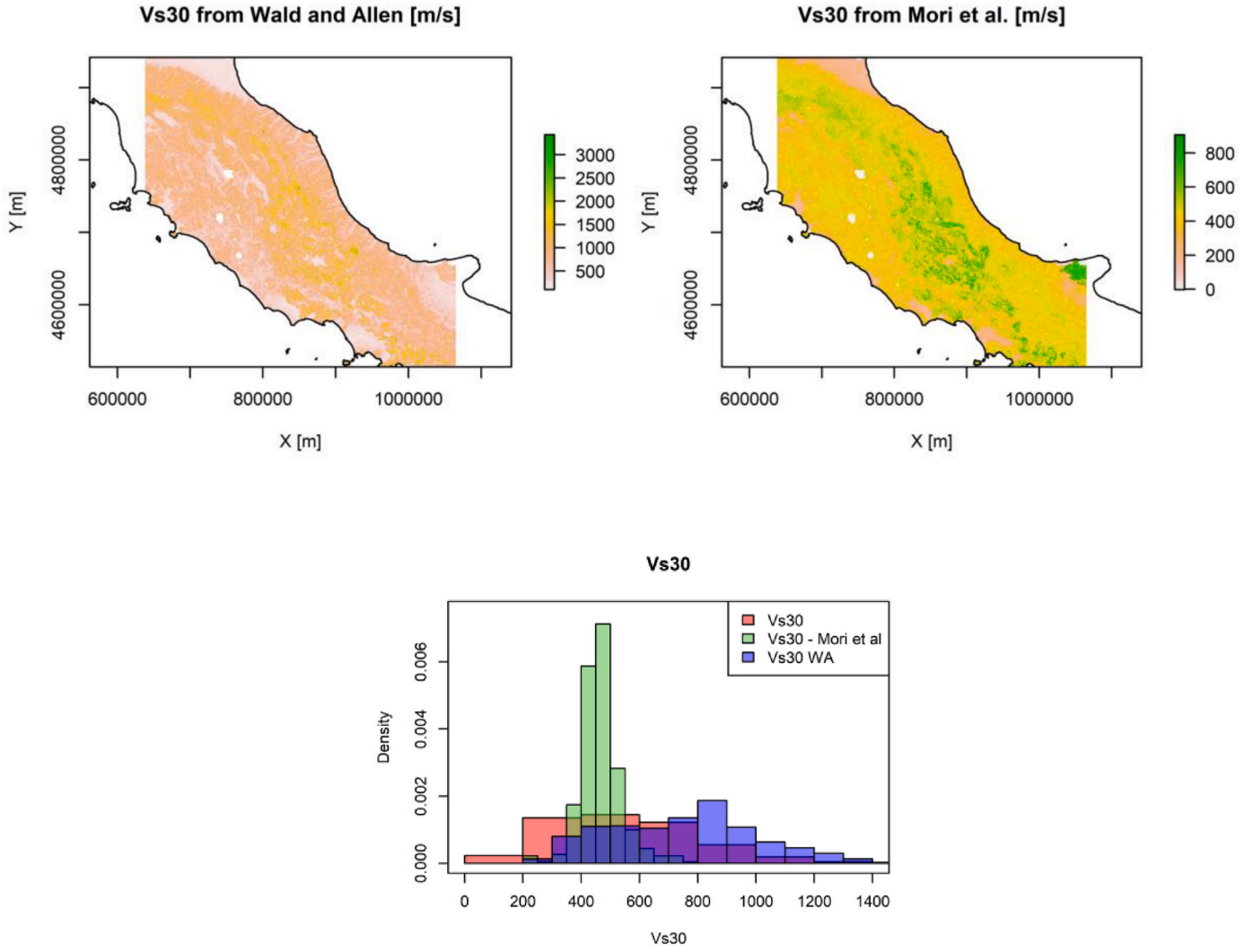


Fig. 2. Raster maps of Vs30 inferred from Wald and Allen relationship [7] (a); map by Mori et al. [13] (b) and their statistical distributions (c). The original coordinate system in a) and b) was re-projected to match that of the dataset in projected coordinate system UTM (zone N32). The prediction grid of the maps is built within the bounding box of the data locations. Note that the maps are plotted on a different full scale to highlight patterns within maps; hence, comparisons across figures should be made carefully.

units made available for Italy, conforming to two classifications. One is in accordance with the ISPRA classification (scale 1:100.000), that partitions sites into 50 lithological categories, 19 of which are represented in our dataset. The other is the seismic shallow soil classification for Italy provided by Forte et al. [50], that groups data into 18 more generic categories, 14 of which appear in the dataset. The choice of considering two alternative classification criteria is made to assess the impact that a more detailed or a broader grouping of data may have on the significance of lithology both in the linear and in the spatial modelling of site proxies. The link between the stations of the dataset and the ISPRA and Forte et al. classifications are provided in the flatfile of **ESUPP1_flatfile.csv**, whereas a map showing the spatial distribution of the lithological classes in the target region can be found in **ESUPP3_plot_of_input_data.docx**.

The lithologic classes are not uniformly represented in our dataset. As an example, in Fig. 3, the pie chart shows the distribution of the geo-lithological complexes according to Forte et al. While some classes have a high numerosity, others are underrepresented; for instance we can note a dominance of arenaceous and marly flysch, marly limestones, (the modal category "AFB") along with marly calcareous (McB) and carbonates (CB).

3. Methodology

The adopted methodology follows the scheme below (Fig. 4):

STATISTICAL DEPENDENCIES (STEP 1): we explore the statistical dependencies and correlations of δS2S with the site proxies described above in order to select the set of candidate predictors, used in the subsequent analyses to constrain the spatial model. To evaluate the dependencies on the lithostratigraphic classes, we adopt the analysis of variance (ANOVA) test, as implemented in the statistical computing language R (R Core Team, 2021); this is a technique used to compare the means of two or more populations and to assess whether these differences are statistically significant. In cases where the dependent variables are more than one, we apply the Multivariate ANOVA (MANOVA) test.

GEOSTATISTICAL ANALYSIS AND MODEL BUILDING (STEP 2): we estimate the spatial dependencies of the stochastic field generating the data and build the spatial prediction models by using the R package *gstat*. Under the assumption of isotropy and stationarity of the spatial process underlying the δS2S measurements, we first characterise the spatial structure of the data by using the semivariogram, which is a classic geostatistic tool to quantify the covariance structure of the field. First, an experimental semivariogram γ(h) of the random field is computed to measure the average dissimilarity between spatial data separated by a lag *h*. To this end, a classic estimator based on the method of moments [53] is used, which is defined as follows:

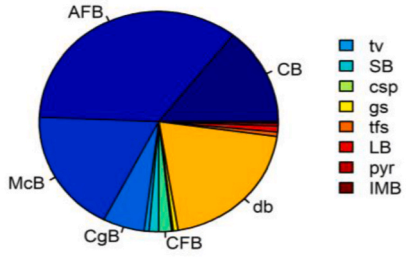
$$\gamma(h) = \frac{1}{|N(h)|} \sum_{(s_i,s_j)\in N(h)} \left(\delta S2S(s_i) - \delta S2S(s_j)\right)^2 \quad [1]$$

where $N(h)$ denotes the set of locations whose distance is (close to) h, $|N(h)|$ indicates the numerosity of the set, and $\delta S2S(s_i)$ denotes the target parameter measured at location s_i. In practical terms, the semivariogram provides the degree of similarity of couples of data as function of the separating distance h. In the presence of spatial stationarity, an increasing trend with respect to h is typical of the experimental semivariogram up to a certain correlation length, after which no correlation is observed; then the sample points of the semivariogram tend asymptotically to the variance of the data.

The experimental semivariogram can be modeled with parametric functions. Amongst the commonly used models, in the following we rely on the exponential model with nugget, that is defined by a positive, exponentially increasing function associated with a non-zero limit to zero (named *nugget,* τ^2), an asymptote (named *sill,* $\tau^2 + \sigma^2$), approximately reached at a distance named *range (a):*

$$\gamma(h;\ \tau,\sigma,a) = \tau^2\ + \sigma^2\left(1 - e^{-\frac{h}{a}}\right), h > 0 \quad [2]$$

The semivariogram model is then used within a linear unbiased spatial predictor named Kriging. An Ordinary Kriging algorithm (OK) is considered if the assumption of spatial stationarity holds true, and a Universal Kriging (UK) is used under non-stationary conditions. In the stationary case, the predictor assumes the form $\delta S2S^*(s_0) = \sum_{i=1}^{n}\lambda_i^*\delta S2S(s_i)$ where $\lambda_1, \ldots, \lambda_n$ are optimal weights, found by solving the so-called kriging system, which is a linear system that depends on the



Geologic Bedrock	ID - Forte	ID - ISPRA	Cover deposits	ID - Forte	ID - ISPRA
Lava	LB	D3	Pyroclastic soil	pyr	D3/D8
Sand	SB	A9	Tuff and scoriae	tfs	D5
Conglomerate	CgB	A8/A14	Clay silt and peat	csp	B6
Clay and Clay flysch	CFB	A7/C1	Gravel and sand	gs	B3
Arenaceous flysch	AFB	A10	Shallow debris	db	B4/B7
Marly calcareous	McB	A3/A11	Travertine	tv	A15
Carbonate	CB	A1/A2			
Igneous metamorphic	IMB	F1			

Fig. 3. Pie chart of the stations in the dataset distributed according to the geo-lithological classes in the map of Forte et al. [50]. The legend also shows the correspondence with the classification according to ISPRA.

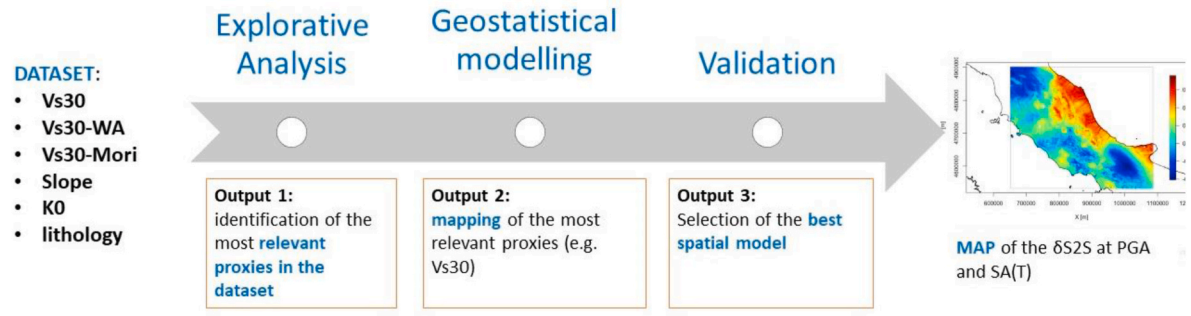


Fig. 4. Workflow of the methodology adopted in this study.

(estimated) variogram model [54]:

$$\begin{pmatrix} \gamma(0) & \cdots & \gamma(|s_1 - s_n|) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(s_n - s_1) & \cdots & \gamma(0) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \zeta \end{pmatrix} = \begin{pmatrix} \gamma(s_0 - s_1) \\ \vdots \\ \gamma(s_0 - s_n) \\ 1 \end{pmatrix} \quad [3]$$

Here, ζ is a Lagrange multiplier which accounts for the unbiasedness constraint, and $|s_i - s_j|$ denotes the distance between the locations s_i and s_j.

In the non-stationary case, we consider that a significant spatial trend is underlying the mean of the data, due to unmodeled effects. Note that we expect that the spatial modelling of δS2S will depend on the presence of regional geological structures and litho-stratigraphic characteristics and thus might turn out to be non-stationary. In this case, the expected value of the random variable E[δS2S(SA(T))] is not constant across all sites, but may vary depending on location [55], i.e. the spatial distribution of the data shows a trend called *drift*. Non-stationarity can be handled by modelling the observed spatial variability from each data point, i.e. the *drift*, with a regression model that captures the effect of dependence on the main variables [32,56].

Intuitively, if such modelling captures effectively the main dependencies, then the residual component (ε) should be (or closer to being) stationary and associated with a more stable variogram model. In mathematical terms, the variable $\delta S2S(s_i)$ is modeled through a linear regression model

$$\delta S2S(s_i) = \sum_{l=0}^{L} a_l f_l(s_i) + \epsilon(s_i), \quad [4]$$

where a_l are the (unknown) regression coefficients, $f_l(s_i)$ are the (known) regressors, and $\epsilon(s_i)$ is assumed to be a stationary residual. In this case, the UK predictor $X^*(s_0) = \sum_{i=1}^{n} \lambda_i^* X(s_i)$ is found by solving the following linear system

$$\begin{pmatrix} \gamma(0) & \cdots & \gamma(|s_1 - s_n|) & f_0(s_1) & \cdots & f_L(s_1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \gamma(|s_n - s_1|) & \cdots & \gamma(0) & f_0(s_n) & \cdots & f_L(s_n) \\ f_0(s_1) & \cdots & f_0(s_n) & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f_L(s_1) & \cdots & f_L(s_n) & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \zeta_0 \\ \vdots \\ \zeta_L \end{pmatrix} = \begin{pmatrix} \gamma(s_0 - s_1) \\ \vdots \\ \gamma(s_0 - s_n) \\ \zeta_0 \\ \vdots \\ \zeta_L \end{pmatrix} \quad [5]$$

where $\zeta_0, \ldots, \zeta_L$ are the $L + 1$ Lagrange multipliers needed to account for the unbiasedness constraint under the linear model [4]. Note that UK predicts the value of the site response as the sum of the regional non-stationary trend as optimally estimated by least squares, and the optimal linear prediction of the stationary component (i.e., the model residuals).

MODEL VALIDATION (STEP 3): Finally, we perform validation of the selected stationary and non-stationary models via leave-one-out cross-validation (LOO-CV) analysis [57,58]. The analysis quantifies the accuracy of the model in predicting site terms by applying the following procedure: i) remove a sample station from the dataset to use it as a test site; ii) with the remaining stations (training set), estimate the target variable at the location of the removed test station; iii) use a diagnostic index of the quality of spatial prediction that is the mean square error (MSE) between the station prediction and the actual value for each selected parameter.

Then, the final model is selected on the basis of the minimisation of the MSE assessed via LOO cross-validation. In case of equivalent predictive performances, the choice falls on the simplest model and for which the covariates are better known at the target locations.

4. Results

4.1. Statistical dependencies (Step 1)

To explore the correlation among the proxies, we use the matrix in Fig. 5a, which shows the Pearson correlation coefficient of δS2S at PGA and various exemplificative spectral ordinates SA (uniformly selected in log-scale) against the investigated proxies. δS2S is linearly correlated with Vs30 measurements (about −0.6 in the long-period range above 0.3s) and less with the inferred estimates (Vs30-WA and Vs30-Mori). The negative correlation is expected and agrees with the findings of Zhu et al. [59], even if Derras et al. [60] and Bergamo et al. [14] identified Vs30 as an optimal proxy for site amplification only for short periods (T < 0.6 s). Also Kamai et al. [61] showed that at short periods (e.g. PGA), Vs30 captures the site amplification for both soil and soft rock sites; instead at long periods, while capturing the overall amplification for soil sites, the linear trend starts to break down at Vs30 of approximately 600 m/s, due to the loss of correlation with the deeper structure, as a result of surface weathering. Our results are closer to those provided by Kotha et al. [19] who showed that at T = 0.02 s and T = 0.2 s correlation of δS2S versus Vs30 is close to zero, implying that high frequency soil response is weakly correlated to Vs30; such correlation tend to increase with period, which is consistent with other past studies (e.g., Ref. [62]). For instance, Thompson and Wald [37] show that measured Vs30 reduce the standard deviation of δS2S by 3 %–9 % for periods less than 0.2 s and by 26 %–40 % for periods longer than 0.5 s. However, it should be noted that a systematic correlation between δS2S and Vs30 has not been observed at all frequency bands in the literature and that the variability of the results depends on the training dataset and on the geographical context of analysis, as the scaling of amplification with Vs30 can be different between regions due to different correlations between Vs30 and the full Vs profile [61]. Furthermore, it should be stressed that in the present study we refer to a not typical estimate of the empirical soil amplification function, which is based on response spectra, while more rigorous, comparable yet physically-based interpretations should be made on the estimates from Fourier amplitude spectra.

In Fig. 5a we note that Vs30-correlates equally well with Vs30-WA and slope (−0.53 vs −0.55, respectively); therefore, from now on, we will only consider Vs30-WA instead of the slope as a proxy for building

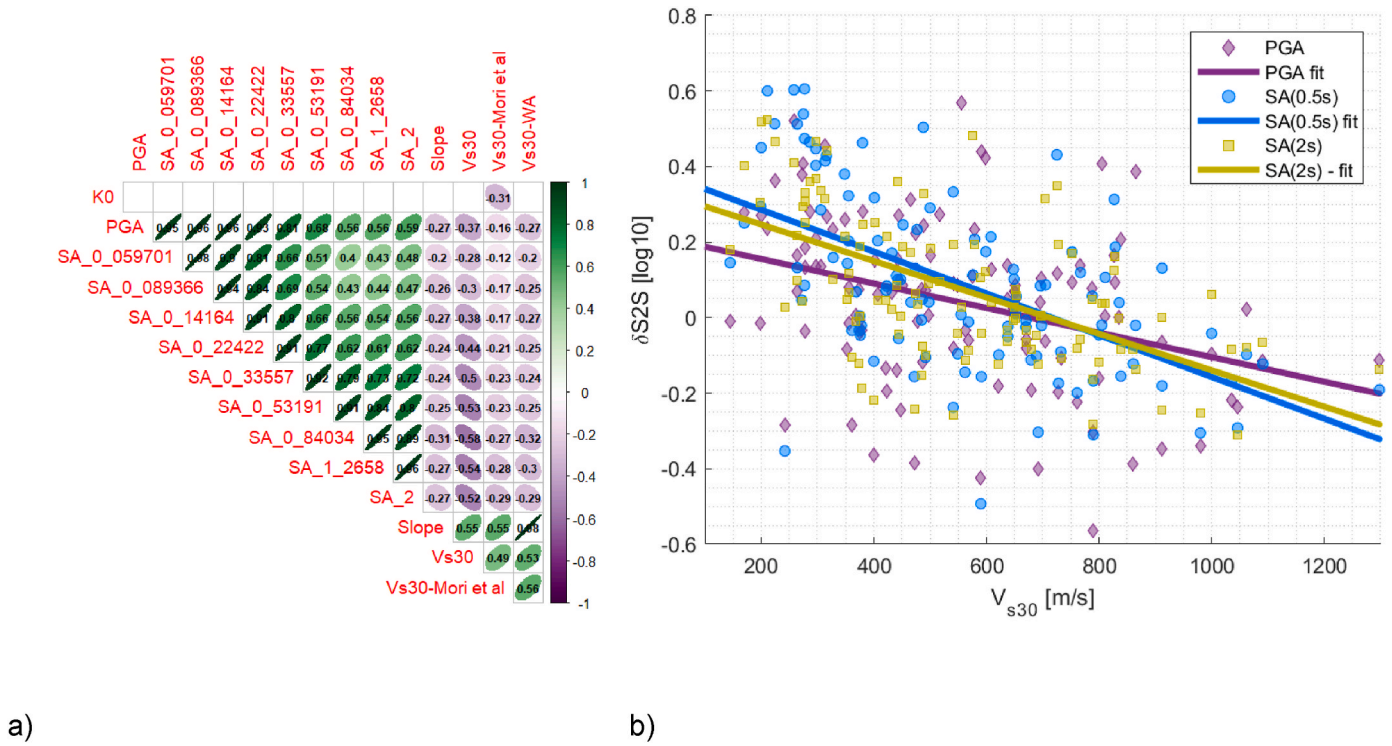


Fig. 5. a) Correlogram of investigated proxies with non-missing values (i.e. data with missings in κ0 or Vs30 were removed for the computation). Blank cells relate to non-significant coefficients based on the resulting p-values (i.e. p-values>0.05); b) δS2S against Vs30 for PGA, SA(0.5s) and SA(2s) along with the corresponding fit lines.

the spatial model of δS2S. The reason why we prefer to use only the inferred estimate of Vs30-WA instead of the slope is that it is more convenient to use it in modelling Vs30 for reasons of dimensional homogeneity. Finally, we see that the dependence on κ0 is not statistically significant for any period (p-values>0.05).

In Fig. 5b, we observe that the δS2S terms show a trend with Vs30: this is more evident (steeper slope) for intermediate and long-period spectral ordinates (T = 0.5 s and T = 2 s), implying that the PGA soil response is less correlated with Vs30, as also observed by Kotha et al. [19]. In any case, the observed trend is in accordance with the expected physics since lower values of Vs30 (soft soils) correspond to positive δS2S (amplifications), while higher values of Vs30, particularly from around 700–800 m/s, correspond to negative δS2S (de-amplifications).

Concerning the assessment of the impact of lithology, this is performed both with reference to the detailed classification of the ISPRA map 1:100.000 (50 classes) and to the broader groupings of 18 lithological classes as proposed by Forte et al [50]. Indeed, as greater geologic precision in site classification does not necessarily correspond to greater significance of the lithology variable, the impact of this term on spatial proxy predictions is assessed at the two levels of classification detail. Fig. 6 shows that all proxies present statistically significant dependence on lithology when they are tested jointly over the periods (p-value = 0 in MANOVA one-way) and singularly within each period (p-value = 0 in ANOVA one-way). Since the transition from the classification of Forte et al. to the 50-class ISPRA lithological map (Fig. 6a) does not lead to any particular improvement in terms of significance in linear models, the spatial analysis for δS2S in the subsequent section will only consider the simpler Forte et al. classification (Fig. 6b).

Furthermore, to select the best possible spatial model for δS2S, we need to reconstruct the Vs30 data in order to elaborate a non-stationary prediction that is constrained to this proxy. To this aim, we also check for possible dependencies of Vs30 from lithology as well (Fig. 6c). It can be noted that Vs30 from measurements and Vs30-WA show a dependence on the litho-classes whereas Vs30-Mori is less sensitive to this classification (p-value ANOVA: 0.077). Regarding κ0 (Fig. 6d) we observe moderate dependence on lithology (p-value ANOVA: 0.036).

4.2. Geostatistical Building models and validation (Step 2–3)

In this section, we focus on studying the spatial correlation structure of the data to infer the best spatial model of δS2S on a regular grid with cell size 1500 m, which is a compromise value for the scope of this work between resolution and computational speed. In fact, the interpolation algorithm, as before discussed, requires the solution of a system of linear equations that has the size of the number of observed data covering the region of interest. The system must be solved for each target location in the prediction grid, so the computational cost of building predictions over the grid increases linearly with the number of locations within the grid.

To this end, we need to study and map, with the same resolution, the identified proxies that have been found to be most appropriate for predicting empirical site amplification at each spectral period. Based on the results of STEP 1, the main predictor is the parameter Vs30 from all sources considered (a. Mori et al.; [13] b. WA [7]; c. measurements). Therefore, here, the spatial analysis is performed first for Vs30 and then for δS2S. In this way, we are able to obtain (input) proxy maps useful for constraining the predictions of the empirical amplification functions.

4.2.1. Spatial model of Vs30

A preliminary spatial study of Vs30 data is provided by the variography analysis and in particular on the trend of the experimental semivariogram. The analysis of the experimental semivariogram of Vs30 (Fig. 7a) reveals a degree of non-stationarity identified by unstable plateau (i.e. tendency to increase with increasing separation distance leading to overestimated correlations), which implies the need to go beyond the stationarity hypothesis.

In the following, we compute a first-order trend surface to model the *drift* component and specify the spatial dependence with the below

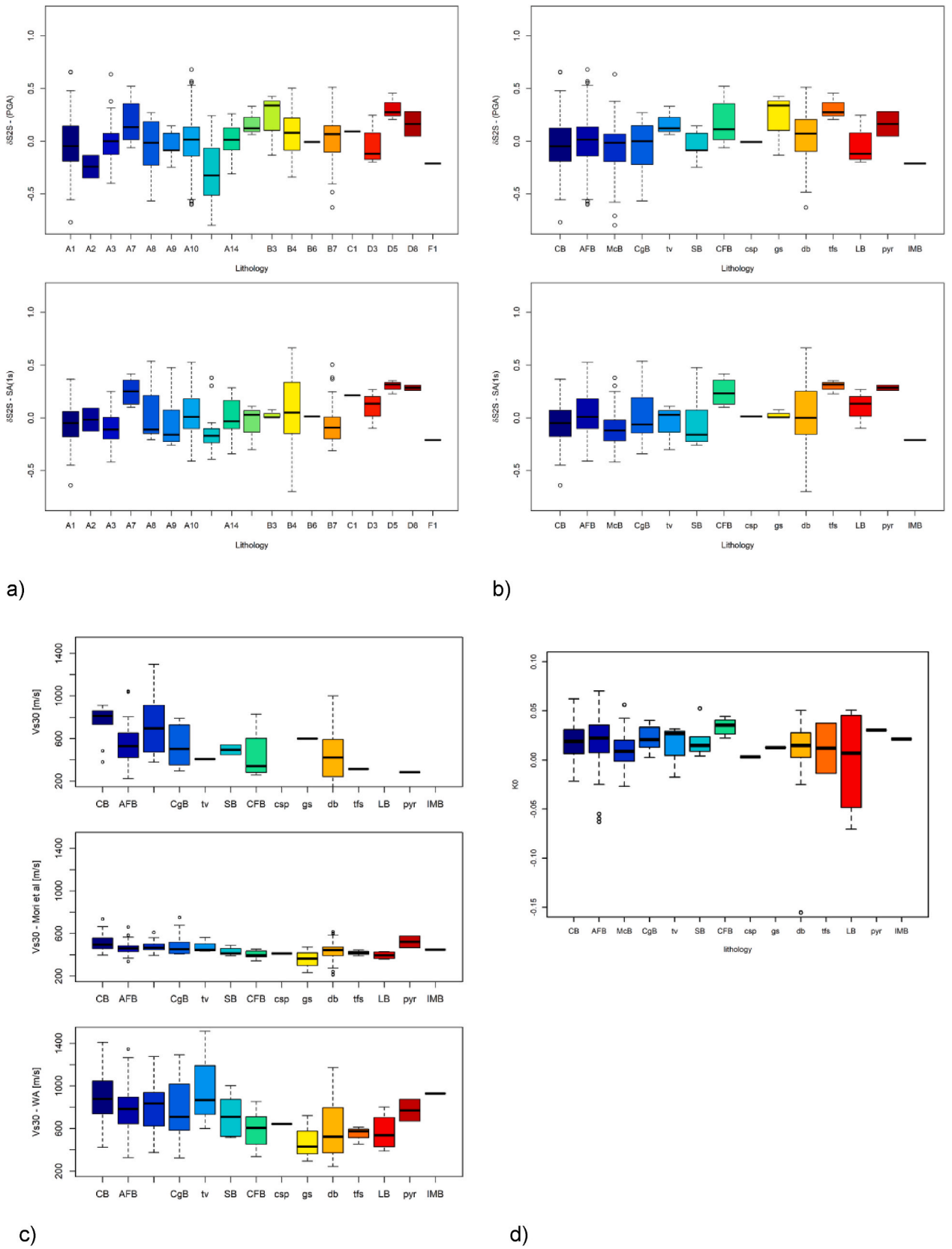


Fig. 6. Dependence on lithostratigraphic classification: a) δS2S at PGA, SA(1s) and SA(2s) versus lithology according to ISPRA; b) δS2S at PGA, SA(1s) and SA(2s) versus lithology according to Forte et al. [50] classification; c) Vs30 versus lithology according to Forte et al. classification; d) κ0 versus lithology according to Forte et al. classification.

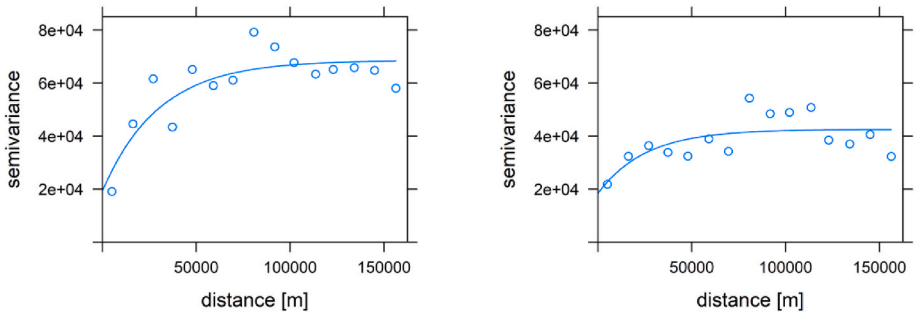


Fig. 7. a) Variograms of Vs30 under the assumption of stationarity (exponential fit - parameters: psill 6.0e4, range 3.0e4, nugget 1.0e-4); b) sample semi-variogram of the residuals of model iii) i.e. Vs30 dependent on Vs30-Mori + Vs30-WA (exponential fit - parameters: psill 4.0e4, range 3.0e4, nugget 1.0e-4).

relationships (this has the same meaning as linear regression model specification).

i) Non-stationary model Vs30 $\sim \beta_0$ (litho) $+ \varepsilon$
ii) Non-stationary model Vs30 $\sim \beta_0 + \beta_1 x + \beta_2$ y $+ \beta_3$ (x*y) $+ \beta_4$ litho $+ \varepsilon$
iii) Non-stationary model Vs30 $\sim \beta_0$ (Vs30-Mori) $+ \beta_1$ (Vs30-WA) $+ \varepsilon$
iv) Stationary model Vs30

In addition to the stationary model (iv), we implement the non-stationary model (ii) based on linear regression of the lithology classes by Forte et al. [50] and the geographical coordinates (x, y) of the data points. We also implement and test the performance of a model dependent on Vs30-Mori and Vs30-WA (iii) that aims to reconstruct the data at unsampled locations based on the other Vs30 information available in the region. Basically, in model iii) the distribution of Vs30 is interpolated on the missing values based on Vs30–Mori or Vs30–WA, whose statistical distributions have been shown in Fig. 2c. In this sense, using the two maps jointly appears advantageous in order to reconstruct the distribution of Vs30 at unobserved points.

When analyzing the correlations of the Vs30 data with the other estimates in detail, it can be seen that this quantity is positively correlated with Vs30-Mori and Vs30-WA (Fig. 8). Assuming a linear model as a function of Vs30-Mori and Vs30-WA, we find that it is significant in both regression coefficients (p-value 0.0058 and 7.29*e−5, respectively) and thus is the two inferred maps are promising candidate for modelling the observed non-stationarity. We remind that the assumed linear dependency is considered to be not statistically significant when its p-value is greater than the level of 0.05.

To confirm this, we observe that the semivariogram of the residuals improves in terms of stability at the level of sill (Fig. 7b), and provide the best performance (i.e. lowest MSE) measured via LOO-CV and compared to the other candidate models (Table II).

Additional rationale for choosing the non-stationary model (iii) is the availability of high-resolution maps of Vs30-Mori and Vs30-WA, as they allow us to obtain stable predictions of Vs30 over a high-resolution grid.

We report below the resulting optimal maps of Vs30 (Fig. 9a) anchored to the values of Vs30-Mori and Vs30-WA. The map shows a global trend with higher values in correspondence of the Central Apennines and mountain chains while lower values near the coasts and alluvial basins. The associated prediction variance is depicted in Fig. 9b and shows the highest values occurring at unsampled locations and lower variance near the sampling locations.

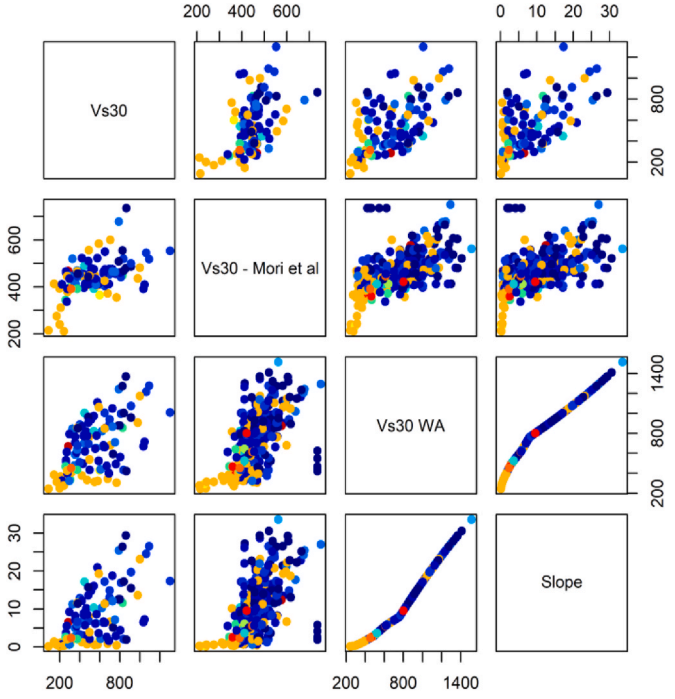


Fig. 8. Correlation matrix of measured Vs30 against other Vs30-proxies (color-coded according to lithology classes of Fig. 4).

Table 2
Results of LOO-CV and p-values of linear regressions. The best performing models for each parameter are denoted in bold.

Variable to model	Covariates	p-value of regression	LOO Mean Squared Error	LOO Standard Deviation
Vs30	i. Litho	0.006548	$43.54*10^{3}$	$73.24*10^{3}$
	ii. Litho + geo coordinates	1.588e-10	$44.88*10^{3}$	$71.21*10^{3}$
	iii. Vs30-Mori + Vs30-WA	9.543e-11	$\mathbf{36.89*10^{3}}$	$\mathbf{57.46*10^{3}}$
	iv. Stationary	–	$47.11*10^{3}$	$83.29*10^{3}$
δS2S (PGA)	i. Geo-coordinates	<2.2e-16	$48.88*10^{-3}$	$67.63*10^{-3}$
	ii. Geo-coordinates + Vs30	<2.2e-16	$\mathbf{46.47*10^{-3}}$	$\mathbf{66.14*10^{-3}}$
	iii. Geo-coordinates + Vs30-map + Litho	<2.2e-16	$45.13*10^{-3}$	$64.32*10^{-3}$
	v. Stationary	–	$49.22*10^{-3}$	$68.00*10^{-3}$
δS2S (SA (1s))	i. Geo-coordinates	0.04104	$44.61*10^{-3}$	$73.16*10^{-3}$
	ii. Geo-coordinates + Vs30	<2.2e-16	$\mathbf{36.98*10^{-3}}$	$\mathbf{58.04*10^{-3}}$
	iii. Geo-coordinates + Vs30-map + Litho	<2.2e-16	$34.13*10^{-3}$	$55.29*10^{-3}$
	v. Stationary	–	$44.64*10^{-3}$	$71.90*10^{-3}$

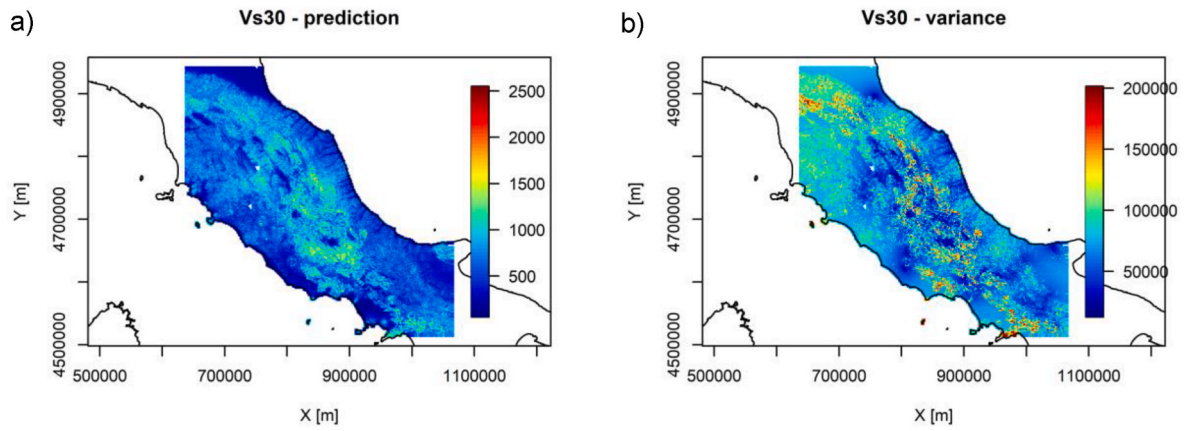


Fig. 9. Maps (projected coordinate system UTM zone N32) of Vs30 predictions (a) and associated variance (b). The prediction grid of the maps is built within the bounding box of the data locations.

4.2.2. Spatial model of δS2S

Here we aim to construct and test both stationary and non-stationary spatial models of the empirical site response, including the dependence on the Vs30-map.

The empirical variograms suggest the presence of an exponential trend both for PGA (Fig. 10a) and SA(T = 1s) (Fig. 10b).

Therefore, it is necessary to model spatial non-stationarity to capture the trend observed in the data by a combination of the covariates according to the following.

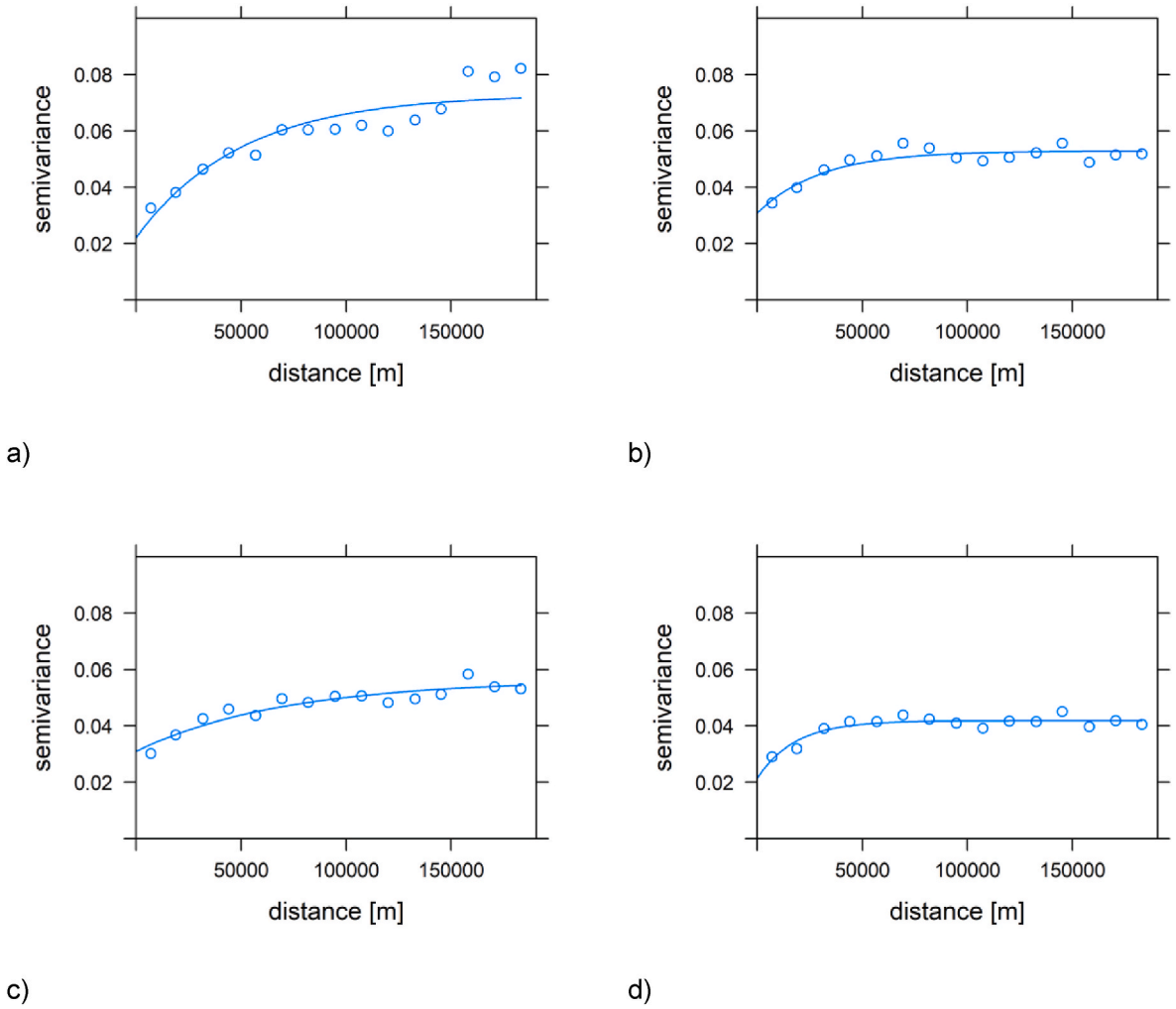


Fig. 10. Experimental and model semi-variograms of δS2S for a) PGA stationary (exponential fit - parameters: psill 0.051, range 5.0e5, nugget 0.022); b) SA(1s) stationary (exponential fit - parameters: psill 0.0234, range 3.0e4, nugget 0.028); c) PGA non-stationary, model iii) (exponential fit - parameters: psill 0.035, range 7.0e4, nugget 0.035); d) SA(1s) non-stationary, model iii) (exponential fit - parameters: psill 0.028, range 3.0e4, nugget 0.02).

i) Non-stationary model: $\delta S2S \sim \beta_0 + \beta_1 x + \beta_2 y + \beta_3 (x*y) + \varepsilon$
ii) Non-stationary model: $\delta S2S \sim \beta_0 + \beta_1 x + \beta_2 y + \beta_3 (x*y) + \beta_4$ (Vs30-map) $+ \varepsilon$
iii) Non-stationary model: $\delta S2S \sim \beta_0 + \beta_1 x + \beta_2 y + \beta_3 (x*y) + \beta_4$ (Vs30-map) $+ \beta_5$ litho $+ \varepsilon$
iv) Stationary model $\delta S2S$

The LOO-CV results in Table 2 show that the trend of PGA and SA(T = 1s) is captured by different combinations of these covariates. For example, in Fig. 10c–d, the spatial trend adjusts slightly when lithology is included (model iii). Given similar performance, the choice can be guided by Occam's razor principle that suggests choosing the simplest model over those involving a larger number of variables. In line with this principle, we decide not to include the dependence on lithology as the use of this map is not justified by an appreciable gain in the predictive performance of the model. On the other hand, the fact that the predicted values of Vs30-map obtained with the non-stationary model are more correlated with $\delta S2S$ than the values obtained with the stationary model, pushes us towards the use of this proxy in the construction of the model.

Based on the above, we select the non-stationary model (ii) as the best compromise between performance and simplicity for constructing the $\delta S2S$ median map both for PGA (Fig. 11 a) and SA(1s) (Fig. 11 c) with the corresponding prediction variances (Fig. 11b–d). Results of model (ii) predictions for median and related prediction variances at each node of the grid are collected in the flat-files of **ESUPP4_dS2S.zip**. Examples of raster maps for PGA and SA(1s) are provided in **ESUPP5_maps.zip**.

It can be seen that the model is able to capture some characteristics on a local scale; for example, amplification is greater in plains and basins (clearly visible for instance higher values near Avezzano at 1s in correspondence of the Fucino deep sedimentary basin), and in general in the presence of areas with lower slopes, river valleys and recent surface deposits (e.g. towards the coastal area in the east).

Standard deviation (sigma) associated to the spatial predictions is shown in Fig. 12 as a function of vibration periods. Since this quantity varies geographically, as evident from Fig. 11b-d, we plot for simplicity the maximum sigma (recorded at locations far from the data points) and

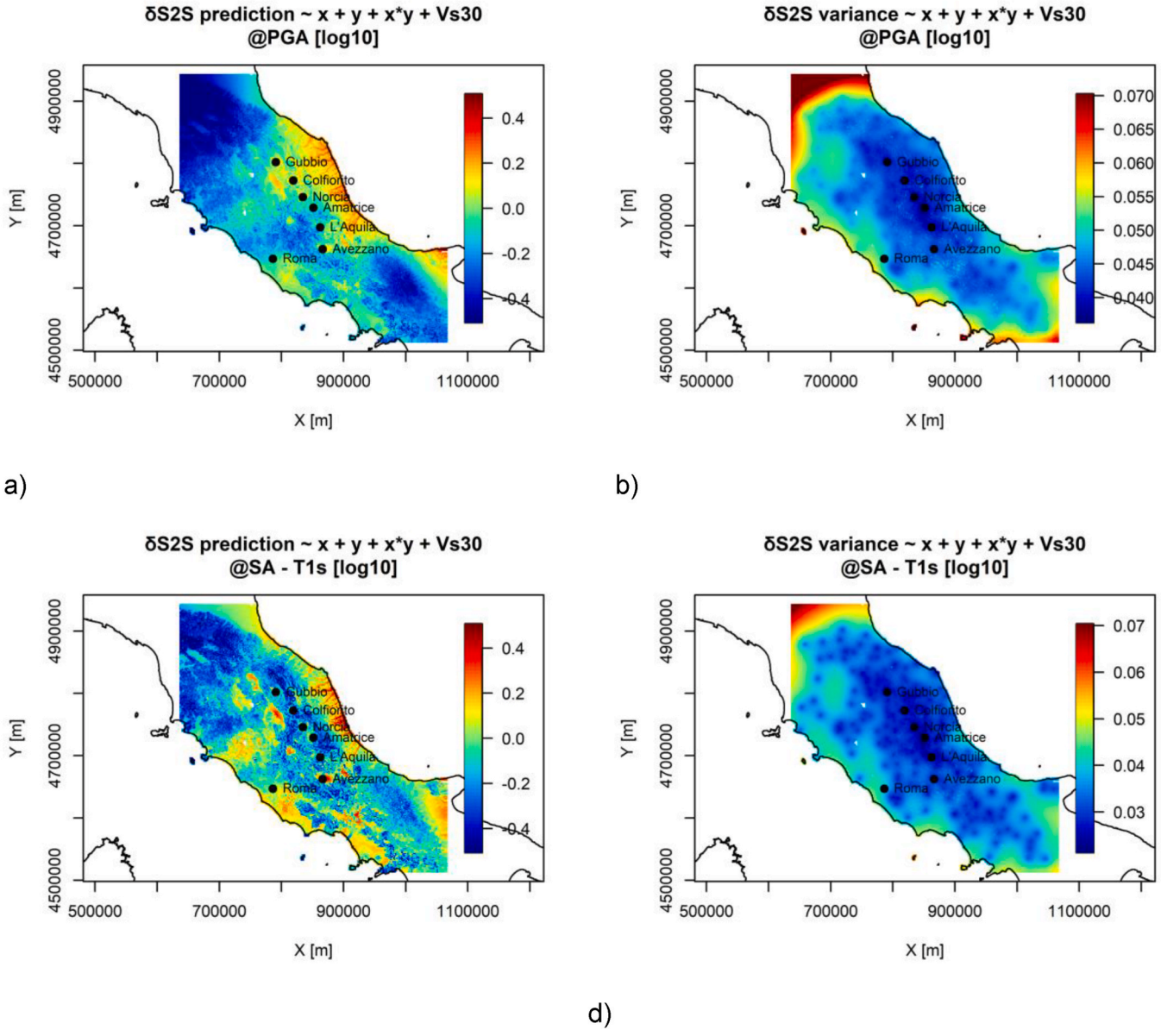


Fig. 11. Maps (projected coordinate system UTM zone N32) of the best $\delta S2S$ model predictions (a, c) and associated variance (b, d) for PGA and SA(1s), respectively. The prediction grid of the maps is built within the bounding box of the data locations.

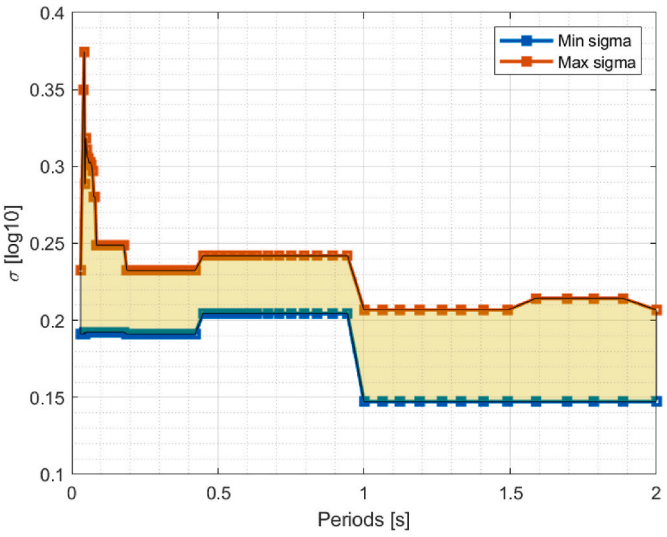


Fig. 12. Standard deviations (minimum and maximum) of the model predictions as a function of vibration periods. Shaded area indicates the range of variability of sigma values.

minimum sigma (obtained at locations near dense observation areas). Note that the maximum sigma is highly variable at very short periods (T < 0.2s) reaching up to 0.38 log10 units and revealing the model's poor constraint on predictors at the borders of the domain; while the overall trend decreases around 0.2 log10 for periods longer than T = 1s, a value that is obtained at most central points of the testing area.

Looking in more detail, we compare the prediction of the entire δS2S curve with the observations in the dataset at individual stations, used as testing sites within the LOO-CV. In other words, we consider the stations, one at a time, as independent sites to assess the goodness of the prediction. In Fig. 13, we show some illustrative cases of stations with good performance (the observed curve is within the 1σ confidence band) and stations where the observed curve is outside the band 2 or 3 times the σ for at least a number of spectral periods (30 % and 10 % over all periods, respectively). In Table III we report the list of stations out of range and the corresponding statistics. We observe that more than 7 % of the sites are affected by sub-optimal prediction, which is probably due to incomplete parameterisation of the site function. This is made more evident by the lack of constraint at the peaks of the site response on many stations, particularly when the peak is located in a short-period interval (e.g. stations IV. ATSC, IV. RM16 in Fig. 12). In these cases, the prediction fails to capture the amplitude and position of the peak, suggesting that some other predictor, such as the fundamental frequency

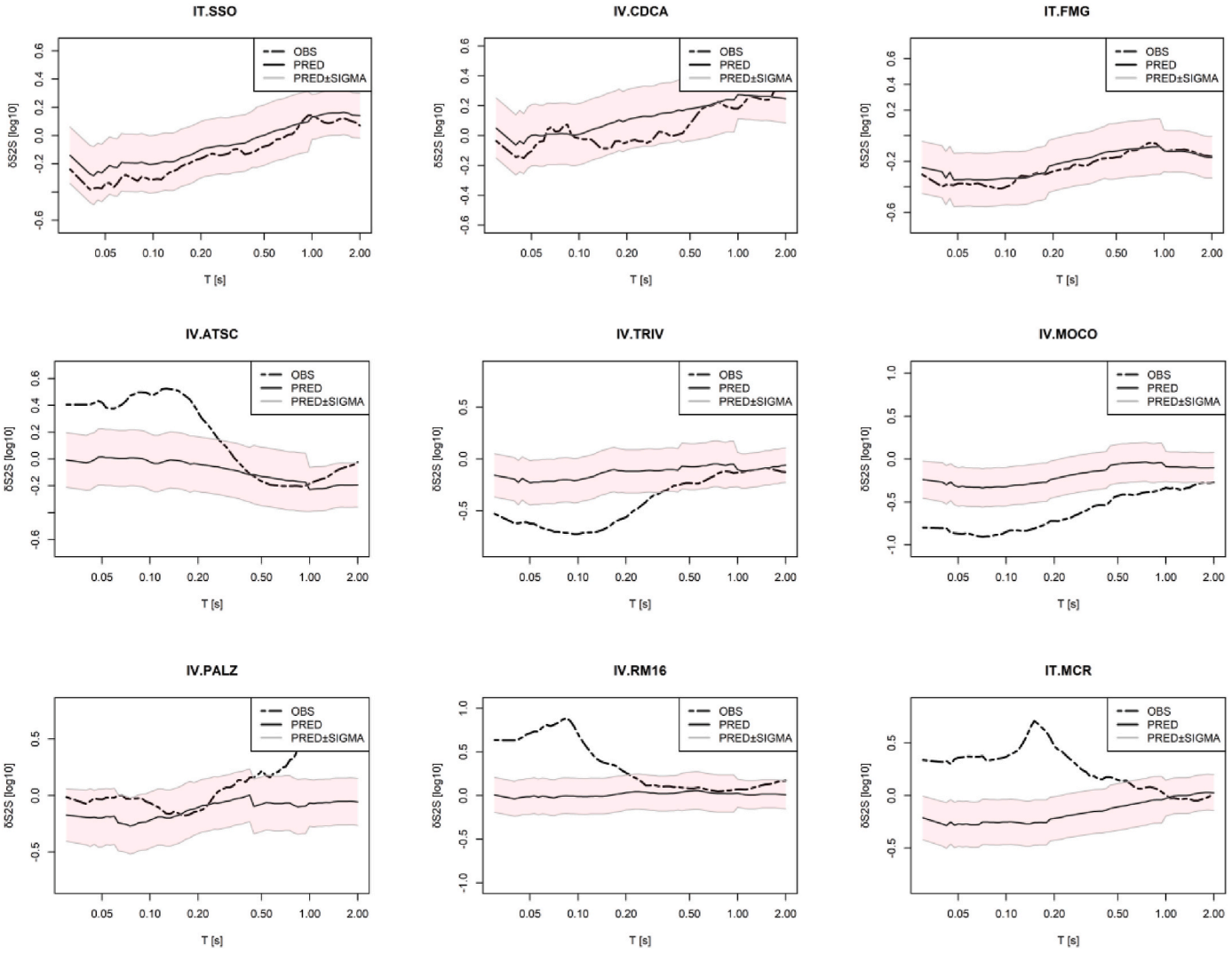


Fig. 13. δS2S predictions on some testing stations versus observations: RMSE<1σ(top panel), RMSE>2σ on 30 % of periods (middle panel), RMSE>3σ on 10 % of periods (bottom panel).

Table 3
List of stations out of the confidence range (in bold are highlighted the ones reported in Fig. 13).

Identification Criteria	Station codes	Statistics
RMSE>2σ on almost 30 % of periods	AM05 AOI **ATSC** BGR CAMP CIGN GIGS LRP MCR **MOCO** MZ10 MZ104 MZ11 NCR PGG PTQR RM03 RM16 RM22 RSM2 SCF SCO SENI SNG SPT1 SSP9 T1220 **TRIV** TVL VLN VVDG	7.5 % (31 stations) over the total
RMSE>3σ on almost 10 % periods	AM05 CAMP GIGS **MCR** MZ10 **PALZ** **RM16** RSM2 SCF SPT1 SSP9	2.7 % (11 stations) over the total

of the soil deposit, must be introduced in the modelling.

As a matter of fact, among the out-of-range stations, more than half belong to the geo-lithological complexes AFB, db and McB corresponding to arenaceous flysch (arenaceous sand, marly flysch, marly limestones, gypsums, clayey metamorphic rocks), shallow debris (infill, alluvial fan, debris, colluvium, breccia, etc.) and marly calcareous (marly, calcareous sand siliceous successions deposited in pelagic environment), respectively.

This may indicate that for this type of (soft soil) classes, an additional parameter needs to be introduced to constrain the site response (especially for stations on soil deposits, such as for the db class, e.g. the MCR station) or that the litho class is not sufficiently resolved with the current input map.

5. Conclusions

In this paper we have proposed a comprehensive methodological approach based on geostatistical analysis to investigate and model the spatial dependencies of the empirical site-to-site δS2S response terms derived from the residual decomposition of a reference GMM specific for the Central Italy region.

The method used is based on a systematic pre-analysis of the dependencies of the site-to-site terms on other site proxies to identify the main statistical correlations. Then, variographic analysis and Kriging estimations (under both stationary and non-stationary assumptions) were applied to produce different parameterised spatial models. Namely, non-stationary modelling of the δS2S prediction passed through the mapping of the predictor Vs30 and lithology, which is available in continuous form; we therefore constructed also an input map of Vs30. The results in terms of LOO-CV allowed us to select the best performing models to obtain interpolated maps of δS2S at various spectral ordinates, conditioned on the most relevant proxies, along with the associated variance.

The final proposed maps of δS2S could be used in conjunction with the predictions of the reference GMM (provided in the **Appendix**) in spectral accelerations to produce non-ergodic shaking scenarios for regional hazard and risk assessments.

Specific findings of our study are resumed in the following:

- δS2S estimates at different spectral periods correlated more with Vs30 measurements, especially at longer periods, and less on the other investigated proxies (Vs30 inferred from different proxies and lithostratigraphic classifications). We did not observe a statistically significant correlation of the δS2S terms estimated by the SA model with the κ0 parameter, probably due to the weaker and highly non-linear link between the high-frequency parts of the Fourier domain spectra and the response spectra. This implies that residual terms δS2S calibrated directly from Fourier amplitude spectra should show a clearer dependence on the high-frequency parameter than was observed in this study;
- Vs30 input map was constrained to the dataset of observations and interpolated on the missing values based on the estimates inferred from the topographic slope, according to Wald and Allen's global empirical relationship (Vs30-WA), and the Italian map provided by Mori et al. [13] (Vs30-Mori); indeed, we observed a different distribution (in terms of mean value and standard deviation) in these two maps, whereby neither Vs30-WA nor Vs30-Mori alone were able to capture the full range of variability of Vs30 observations. The combination of the Vs30-Mori map and the Vs30-WA map had a twofold advantage: first, it provided the lowest mean square error when tested with a LOO-CV; second, it exploited high-resolution raster maps, which enables widespread, non-stationary predictions over the entire region of interest;
- Validation results demonstrated that non-stationary spatial models that include lithology showed no significant improvement in the prediction of Vs30 compared to the corresponding stationary models, both using the broader geological classification proposed by Forte et al. [50] and the more detailed ISPRA map. Furthermore, the lithological classification did not provide a significant improvement in the Kriging predictions of δS2S. In general, the inclusion of lithology in the modelling did not show a significant ability of the lithological classification to capture the spatial pattern of the data. We attribute this discrepancy to the presence of underrepresented lithological classes, which did not allow us to fully capture the spatial variability of the data. These results suggest that the lithological classification in its current form (in terms of the quality and resolution of the available maps), is not useful for the construction of the spatial model of the proxies under analysis;
- A map of δS2S was obtained with a non-stationary model as a function of geographical coordinates and the combined Vs30 map. Although this model was able to capture the main trend of the empirical site response in Central Italy, it was not able to reproduce the full range of amplifications at the local level and the actual variability between periods, so the identification of new key-proxies to constrain the predictions may be necessary for future improvements of the proposed model, such as fundamental frequency of the soil deposit f0, mean shear wave velocity at bedrock, sediment thickness, alternative maps of the geo-lithological complexes, etc. Subsequent development of this research could therefore be directed towards the inclusion of these additional site proxies as predictors of δS2S modelling. Improvements could be achieved by considering, for example, a smaller target area within which a higher density of information and geophysical/morphological parameters, as well as small-scale maps could be available as input data. Further developments could also be aimed at quantifying and possibly utilizing the spatial correlation between spectral periods in a multivariate or functional framework.

CRediT authorship contribution statement

Sara Sgobba: Conceptualization, Data curation, Formal analysis, Investigation, Supervision, Validation, Writing – original draft. **Chiara Felicetta:** Data curation, Investigation, Software, Writing – review & editing. **Teresa Bortolotti:** Formal analysis, Investigation, Methodology, Software, Validation, Writing – review & editing. **Alessandra Menafoglio:** Formal analysis, Investigation, Methodology, Software, Validation, Writing – review & editing. **Giovanni Lanzano:** Conceptualization, Investigation, Supervision, Writing – review & editing. **Francesca Pacor:** Funding acquisition, Project administration, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We share data in the Electronic Supplements

Acknowledgment

This study was partially funded in the framework of INGV and Dipartimento della Protezione Civile (INGV-DPC) AGREEMENT B1 2019–2021 and by the project PRIN-SERENA 2020 (mapping Seismic site Effects at REgional and National scAle), Prot. 2020MMCPER; Decreto Direttoriale n. 223 del 18/02/2022 del Ministero dell'Università e della Ricerca, Segretariato Generale Direzione Generale della ricerca, coordinated by Prof. D. Albarello (Italy).

Alessandra Menafoglio and Teresa Bortolotti acknowledge the support by MUR, grant Dipartimento di Eccellenza 2023–2027. The authors wish to thank Giulio Brunelli (INGV) for his technical support during the development of the first stage of this work. We thank also the associate editor, Chuanbin Zhu and the anonymous reviewer for their constructive suggestions that helped to improve the manuscript.

The contents of this paper represent the authors' ideas and do not necessarily correspond to the official opinion and policies of DPC.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.soildyn.2024.108496.

Appendix. Reference GMM for δS2S estimation

The reference GMM used to estimate the empirical site-to-site functions is calibrated for the geometric mean of horizontal peak ground accelerations (PGA) and 69 SA ordinates in the frequency range logarithmically equispaced from 0.01 to 2s seconds The model is calibrated via a mixed-effect regression, providing the repeatable effects on the seismic motion (i.e. source, site and path), along with the associated aleatory variability.

The adopted model is a slightly modified version of the Sgobba et al. [32] model calibrated for Central Italy. It has the following functional form:

$$log_{10}\, Y = a + F_M(M) + F_R(M,R) + F_S + \delta B_e + \delta S2S_s + \delta L2L_{source} + \delta P2P_p + \delta W_0 \quad [1A]$$

where Y is the spectral parameter, a, $F_M(M)$, $F_R(M,R)$, F_S represent the fixed effects, and δB_e, $\delta S2S_s$, $\delta L2L_{source}$, $\delta P2P_p$ stands for zero-mean gaussian-distributed random effects.

Parameter a is the offset and F_M describes the scaling with magnitude:

$$F_M(M) = b_1(M_w - M_h) \text{ for } M \le M_h,$$

$$F_M(M) = b_2(M_w - M_h) \text{ otherwise.} \quad [2A]$$

The coefficients b_1 and b_2 are obtained from nonlinear least-square regression, while M_h is the hinge magnitude fixed at 5.0. The term $F_R(M,R)$ is the other fixed term and represents the scaling with distance:

$$F_R(M,R) = \left[c_1\left(M_w - M_{ref}\right) + c_2\right]\, log_{10}\frac{\sqrt{R^2+h^2}}{R_{ref}} + c_3\left(\sqrt{R^2+h^2} - R_{ref}\right). \quad [3A]$$

$F_R(M,R)$ is divided into a contribution due to the geometrical spreading also computed by nonlinear regression (including magnitude-dependent terms with coefficient c_1, and magnitude-independent with coefficient c_2) and the anelastic attenuation (described by c_3, which is typically regionally dependent). M_{ref} is the reference magnitude obtained from a preliminary nonlinear regression. R_{ref} is the reference distance fixed at 1 km, h is the pseudo depth fixed at 6 km, and R is the Joyner-Boore distance for events with magnitude larger than 5.5, for which the fault geometry was defined. For lower magnitude events, the Joyner-Boore distance is assumed to be equal to the epicentral distance, since the small area of the rupture surface makes the event equivalent to a point-like source.

F_S is the site term which is a dummy variable introduced only to distinguish between reference and non-reference sites, as:

$$F_S = \begin{cases} 0 & \text{for reference rock sites} \\ \mathrm{f_s} & \text{for non} - \text{reference rock sites} \end{cases}$$

The median ground motion level is scaled to that observed on 6 reference rock sites (IT.LSS, IT. MNF, IT. SLO, IT. SNO, IT. SDM, IT. NRN), which were detected in the work by Morasca et al [49], according to several proxies based on geophysical, seismological and geomorphological features. This selection updates the 36 reference sites, originally used by Sgobba et al. [32]. The term $\mathrm{f_s}$, on the contrary, quantifies the median deviation of all non-reference sites from the prediction of the reference sites.

Model coefficients, correction terms and related standard deviations are provided in **ESUPP2_GMM.zip**.

The random effects represent error terms with respect to the median prediction of the GMM in equation [1A], defined as follows.

- δB_e is the between-event term, i.e., a measure of the systematic deviation from the median of each group of recordings for the same event;
- $\delta S2S_s$ is the site-to-site term, defining the between-station residual distribution measuring the systematic deviation from the median of the recordings relevant to the same station. Note that, for the purposes of this study, it is necessary to refer to a consistent estimate of repeatable terms, that is, uniformly calculated with respect to the same reference motion, which we assume to be corresponding to the median motion of the 6 reference sites. So the site term estimated with respect to the reference ground motion, called $\delta S2S_{ref}$, is computed as:

$$\delta S2S_{ref} = \begin{cases} \delta S2S_s & \textit{for the reference rock sites} \\ \delta S2S_s + f_s & \textit{for non} - \textit{reference rock sites} \end{cases}$$

- $\delta L2L_{source}$ represents the source term, describing the systematic bias of the source regions. This term is estimated via the clustering approach that is based on geological considerations used to define the source areas, differently from what was originally published by Sgobba et al. [32] who adopted event aggregation mainly based on Reasenberg's algorithm (Reasenberg 1985), as the main spatial clustering criterion. Here instead we group the events considering the boundaries defined by the main fault structures that have been identified in the geological model RETRACE3D - centRal Italy Earthquakes integRAted Crustal modEl (https://doi.org/10.13127/retrace-3d/geomod.2021) for central Italy. Following these criteria, the events are aggregated within 4 polygons, delimited by the main thrusts and geological domains, identifying the following clusters: #1 (Lazio-Abruzzo), #2 (Laga Foredeep), #3 (Sibillini), #4 (Umbria-Marche). In the rest of the region we considered a background seismicity with a mean value of the location term equal to zero. A map of the clusters is reported in Fig. A1 a:
- $\delta P2P_p$ is the path-term, denoting the systematic deviation along each source-to-site path (from each identified source region to the sites) and is related to anisotropy in the properties of the crustal propagation medium. The path terms are estimated by dividing the whole region into squared cells (0.2°-spaced) – Fig. A1 b that allow to capture the spatial distribution of the attenuation behavior (cell-specific attenuation), as defined by Sgobba et al. [31,32].

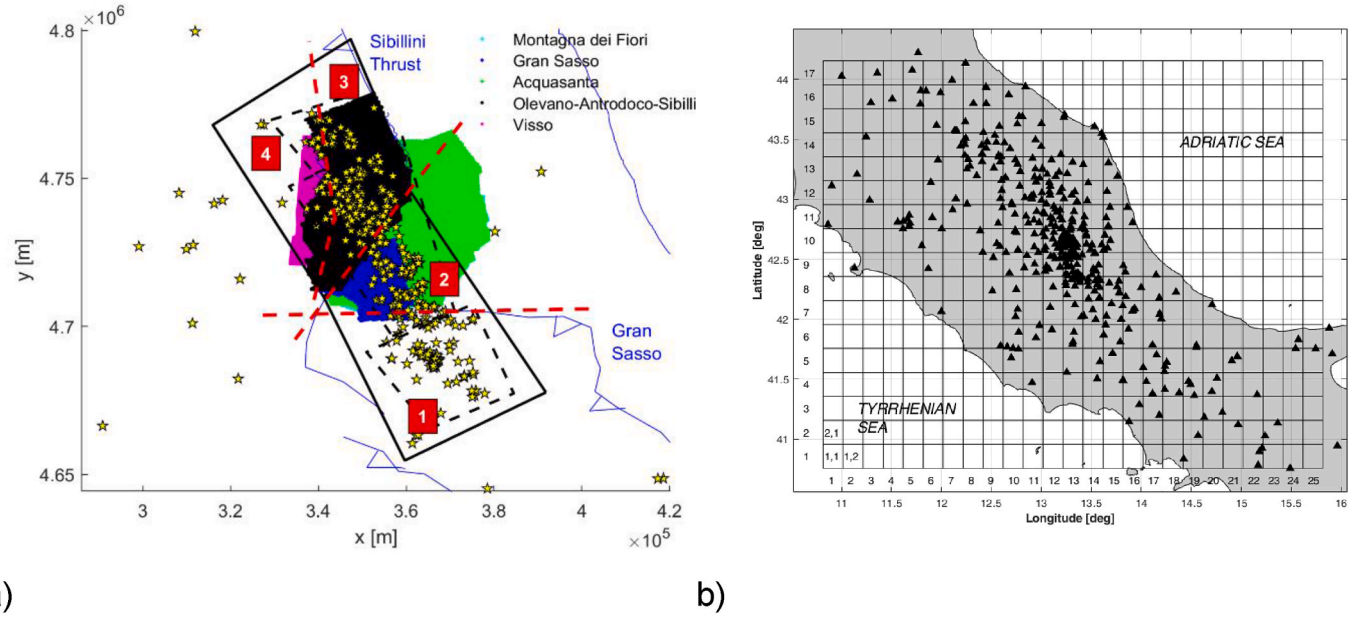


Fig. A1. a) Map of the source regions identified in Central Italy: clusters are delimited by red dashed lines and numbered from 1 to 4: #1 (Lazio-Abruzzo), #2 (Laga Foredeep), #3 (Sibillini), #4 (Umbria-Marche). The surface projections of the boundary fault systems according to the RETRACE-3D geological model are shown as colored lines and reported in the legend. Black dashed lines delimit the original clusters identified by Sgobba et al. [32]. Epicentres of events are marked by stars. The map is defined in UTM32 reference system with metric coordinates (x for Latitude and y for Longitude); b) reference grid used to estimate the path-to-path correction terms $\delta P2P_p$ (from Ref. [32])

References

[1] Thompson EM, Wald DJ, Worden C. A Vs30 map for California with geologic and topographic constraints. Bull Seismol Soc Am 2014;104:2313–21.
[2] Worden CB, Wald DJ. ShakeMap manual, version 2.0. 2016. p. 1–113.
[3] Verros SA, Wald DJ, Worden CB, Hearne M, Ganesh M. Computing spatial correlation of ground motion intensities for ShakeMap. Comput Geosci 2017;99:145–54. https://doi.org/10.1016/j.cageo.2016.11.004.
[4] Michelini 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. https://doi.org/10.1785/0220190130.
[5] Wills CJ, Clahan KB. Developing a map of geologically defined site-condition categories for California. Bull Seismol Soc Am 2006;96(4A):1483–501.
[6] Di Capua Giuseppe, Peppoloni Silvia, Amanti Marco, Cipolloni Carlo, Conte G. Site classification map of Italy based on surface geology. Geological Society, vol. 27. London: Engineering Geology Special Publications; 2016. p. 147–58. https://doi.org/10.1144/EGSP27.13.
[7] Wald DJ, Allen TI. Topographic slope as a proxy for seismic site conditions and amplification. Bull Seismol Soc Am 2007;97(5):1379–95.
[8] Yong A, Hough SE, Iwahashi J, Braverman A. A terrain-based site-conditions map of California with implications for the contiguous United States. Bull Seismol Soc Am 2012;102(1):114–28.
[9] Stewart JP, Liu AH, Choi Y. Amplification factors for spectral acceleration in tectonically active regions. Bull Seismol Soc Am 2003;93:332–52.
[10] Castellaro S, Mulargia F, Rossi PL. Vs30: proxy for seismic amplification? Seismol Res Lett 2008;79:540–3. https://doi.org/10.1785/gssrl.79.4.540.
[11] Luzi L, Puglia R, Pacor F, Gallipoli MR, Bindi D, Mucciarelli M. Proposal for a soil classification based on parameters alternative or complementary to Vs(,30). Bull Earthq Eng 2011;9(6):1877–98.
[12] Luzi L, Bindi D, Puglia R, Pacor F, Oth A. Single-station sigma for Italian strong-motion stations. Bull Seismol Soc Am 2014;104(1):467–83. https://doi.org/10.1785/0120130089. 2014.
[13] Mori F, Mendicelli A, Moscatelli M, Romagnoli G, Peronace E, Naso G. A new Vs30 map for Italy based on the seismic microzonation dataset. Eng Geol 2020;275:105745.
[14] Bergamo P, Hammer C, Fäh D. Correspondence between site amplification and topographical, geological parameters: collation of data from Swiss and Japanese stations, and neural networks-based prediction of local response. Bull Seismol Soc Am 2021;112:1008–30. https://doi.org/10.1785/0120210225.
[15] Lanzano G, Felicetta C, Pacor F, Spallarossa D, Traversa P. Generic-to-reference rocks scaling factors for seismic ground motion in Italy. Bull Seismol Soc Am 2022. https://doi.org/10.1785/0120210063.
[16] Loviknes K, Kotha SR, Cotton F, Schorlemmer D. Testing nonlinear amplification factors of ground-motion models. Bull Seismol Soc Am 2021;111(4):2121–37.
[17] Zhu C, Pilz M, Cotton F. Which is a better proxy, site period or depth to bedrock, in modelling linear site response in addition to the average shear-wave velocity? Bull Earthq Eng 2020;18(3):797–820.
[18] Bard PY, Bora SS, Hollender F, Laurendeau A, Traversa P. Are the standard Vs30-kappa host-to-target adjustments the only way to get consistent hard-rock ground motion prediction? Pure Appl Geophys 2020;177:2049–68.
[19] Kotha Sreeram Reddy, Cotton Fabrice, Bindi Dino. A new approach to site classification: mixed-effects ground motion prediction equation with spectral clustering of site amplification functions. Soil Dynam Earthq Eng 2018. https://doi.org/10.1016/j.soildyn.2018.01.051.
[20] Shible H. Development of a new approach to define reference ground motions applicable to existing strong-motion databases. Earth sciences. Université Grenoble Alpes. (PhD Thesis) NNT : 2021GRALU021. 2021.
[21] 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;103:3149–63.

[22] Kotha Sreeram Reddy, Bindi Dino, Cotton Fabrice. From ergodic to region- and site-specific probabilistic seismic hazard assessment: method development and application at European and middle eastern sites. Earthq Spectra 2017;33:1433–53. https://doi.org/10.1193/081016EQS130M.
[23] Thompson EM, Baise LG, Kayen RE, Tanaka Y, Tanaka H. A geostatistical approach to mapping site response spectral amplifications. Eng Geol 2010;114:330–42.
[24] Goovaerts P. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma 1999;89(1–2):1–45. https://doi.org/10.1016/S0016-7061(98)00078-0. ISSN 0016-7061.
[25] Weatherill G, Kotha SR, Cotton F. Re-thinking site amplification in regional seismic risk assessment. Earthq Spectra 2020;36:274–97. https://doi.org/10.1177/8755293019899956.
[26] Loviknes K, Cotton F, Weatherill G. Introducing inferred geomorphological sediment thickness as a new site proxy to predict ground-shaking amplification at regional scale. Application to Europe and Eastern Turkey 2023. https://doi.org/10.5194/egusphere-2023-1370.
[27] Kuehn Nicolas. Investigations into non-stationary correlation models for site terms of ground-motion models due to basins, vol. 10; 2022. p. 31224–2232.
[28] Gilder CE, De Risi R, De Luca F, Mohan Pokhrel R, Vardanega PJ. Geostatistical framework for estimation of VS 30 in data-scarce regions. Bull Seismol Soc Am 2022;112(6):2981–3000.
[29] De Risi R, De Luca F, Gilder CE, Pokhrel RM, Vardanega PJ. The SAFER geodatabase for the Kathmandu valley: Bayesian kriging for data-scarce regions. Earthq Spectra 2021;37(2):1108–26. https://doi.org/10.1177/8755293020970977.
[30] Bora SS, Scherbaum F, Kuehn N, Stafford P, Edwards B. Development of a response spectral ground-motion prediction equation (GMPE) for seismic-hazard analysis from empirical Fourier spectral and duration models. Bull Seismol Soc Am 2015;105(4):2192–218.
[31] Sgobba S, Lanzano G, Colavitti L, Morasca P, D'Amico MC, Spallarossa D. Physics-based parametrization of a FAS nonergodic ground motion model for Central Italy. Bull Earthq Eng 2023;21:4111–37. https://doi.org/10.1007/s10518-023-01691-1.
[32] Sgobba S, Lanzano G, Pacor F. Empirical non-ergodic shaking scenarios based on spatial correlation models: an application to Central Italy. Earthq Eng Struct Dyn; 2021. https://doi.org/10.1002/eqe.3362.
[33] Priolo E, Pacor F, Spallarossa D, et al. Seismological analyses of the seismic microzonation of 138 municipalities damaged by the 2016–2017 seismic sequence in Central Italy. Bull Earthq Eng 2020;18:5553–93. https://doi.org/10.1007/s10518-019-00652-x.
[34] Chandra J, Guéguen P, Bonilla LF. PGA-PGV/Vs considered as a stress–strain proxy for predicting nonlinear soil response. Soil Dynam Earthq Eng 2016;85:146–60. https://doi.org/10.1016/j.soildyn.2016.03.020. ISSN 0267-7261.
[35] Felicetta C, Lanzano G, D'Amico M, Puglia R, Luzi L, Pacor F. Ground motion model for reference rock sites in Italy. Soil Dyn Earthq Eng 2018;110:276–83. https://doi.org/10.1016/j.soildyn.2018.01.024.
[36] Lanzano G, Felicetta C, Pacor F, Spallarossa D, Traversa P. Methodology to identify the reference rock sites in regions of medium-to-high seismicity: an application in Central Italy. Geophys J Int 2020;222(3):2053–67. https://doi.org/10.1093/gji/ggaa261.
[37] Thompson EM, Wald DJ. Uncertainty in VS30-based site response. Bull Seismol Soc Am 2016;106(2):453–63. https://doi.org/10.1785/0120150214.
[38] Pilz M, Cotton F. Does the one-dimensional assumption hold for site response analysis? A study of seismic site responses and implication for ground motion assessment using KiK-net strong-motion data. Earthq Spectra 2019;35:883–905.
[39] Paolucci R, Aimar M, Ciancimino A, et al. Checking the site categorization criteria and amplification factors of the 2021 draft of Eurocode 8 Part 1–1. Bull Earthq Eng 2021;19:4199–234. https://doi.org/10.1007/s10518-021-01118-9.
[40] 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 Dynam Earthq Eng 2023;166:107775. https://doi.org/10.1016/j.soildyn.2023.107775. ISSN 0267-7261.
[41] 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.
[42] Borcherdt RD. Estimates of site-dependent response spectra for design (methodology and justification). Earthq Spectra 1994;10(4):617–53.
[43] Rodriguez-Marek A, Bray JD, Abrahamson NA. An empirical geotechnical seismic site response procedure. Earthq Spectra 2001;17(1):65–87.
[44] Pitilakis K, Riga E, Anastasiadis A. New code site classification, amplification factors and normalized response spectra based on a worldwide ground-motion database. Bull Earthq Eng 2013;11(4):925–66.
[45] Anderson John, Hough Susan. A model for the shape of the Fourier amplitude spectrum at high frequencies, vol. 74. Bulletin of the Seismological Society of America; 1984. p. 1969–93.
[46] Ktenidou OJ, Cotton F, Abrahamson NA, Anderson JG. Taxonomy of κ: a review of definitions and estimation approaches targeted to applications. Seismol Res Lett 2014;85(1):135–46.
[47] Luzi L, Lanzano G, Felicetta C, D'Amico MC, Russo E, Sgobba S, Pacor F, ORFEUS Working Group 5. Engineering strong motion database (ESM) (version 2.0). Istituto Nazionale di Geofisica e Vulcanologia (INGV) 2020. https://doi.org/10.13127/ESM.2.
[48] Mascandola C, Luzi L, Felicetta C, Pacor F. A GIS procedure for the topographic classification of Italy, according to the seismic code provisions. Soil Dynam Earthq Eng 2021;148:106848. https://doi.org/10.1016/j.soildyn.2021.106848.
[49] Morasca P, D'Amico M, Sgobba S, Lanzano G, Colavitti L, Pacor F, Spallarossa D. Empirical correlations between a 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.
[50] Forte G, Chioccarelli E, De Falco M, Cito P, Santo A, Iervolino I. Seismic soil classification of Italy based on surface geology and shear-wave velocity measurements. Soil Dynam Earthq Eng 2019;122:79–93.
[51] DPC, Dipartimento della Protezione Civile "Commissione tecnica per il supporto e monitoraggio degli studi di Microzonazione Sismica" (ex art.5, OPCM3907/10). WebMs; WebCLE. A cura di: Maria Sole Benigni, Fabrizio Bramerini, Gianluca Carbone, Sergio Castenetto, Gian Paolo Cavinato, Monia Coltella, Margherita Giuffrè, Massimiliano Moscatelli, Giuseppe Naso, Andrea Pietrosante, Francesco Stigliano. www.webms.it; 2018. 2018.
[52] Drouet S, Bouin M-P, Cotton F. New moment magnitude scale, evidence of stress drop magnitude scaling and stochastic ground motion model for the French West Indies. Geophys J Int 2011;187(3):1625–44. https://doi.org/10.1111/j.1365-246X.2011.05219.x.
[53] Matheron G. Traité de géostatistique appliquée, vol. 1. Editions Technip; 1962. 1962.
[54] Cressie Noel AC. Statistics for spatial data. New York: John Willy and Sons." Inc.; 1993. p. 800.
[55] 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;203:103139.
[56] Oliver M, Webster R. A tutorial guide to geostatistics: computing and modelling variograms and kriging. Catena 2014;113:56–69. https://doi.org/10.1016/j.catena.2013.09.006.
[57] Tomczak M. Spatial interpolation and its uncertainty using automated anisotropic inverse distance weighting (IDW)-cross- validation/jackknife approach. J Geogr Inf Decis Anal 1998;2(2):18–30.
[58] Hofierka J, Cebecauer T, Šúri M. Optimisation of interpolation parameters using cross-validation. In: Digital terrain modelling: development and applications in a policy support environment. Berlin, Heidelberg: Springer Berlin Heidelberg; 2007. p. 67–82.
[59] Zhu C, Cotton F, Kawase H, Haendel A, Pilz M, Nakano K. How well can we predict earthquake site response so far? Site-specific approaches. Earthq Spectra 2022;38(2):1047–75. https://doi.org/10.1177/87552930211060859.
[60] Derras B, Bard P-Y, Cotton F. Vs30, slope, H800, and f0: performance of various site-condition proxies in reducing ground-motion aleatory variability and predicting non-linear site response. Earth Planets Space 2017;69:133. https://doi.org/10.1186/s40623-017-0718-z.
[61] Kamai R, Abrahamson NA, Silva WJ. VS30 in the NGA GMPEs: regional differences and suggested practice. Earthq Spectra 2016;32(4):2083–108. https://doi.org/10.1193/072615EQS121M.
[62] Seyhan E, Stewart JP. Semi-empirical nonlinear site amplification from NGA-West2 data and simulations. Earthq Spectra 2014;30:1241–56.