Bulletin of Earthquake Engineering
https://doi.org/10.1007/s10518-026-02638-y

ORIGINAL ARTICLE

On the consistency of site amplification estimates from empirical and numerical methods in Italy

Salomon Hailemikael[1] · Giovanna Cultrera[1] · Giovanni Lanzano[2] · Carla Barnaba[3] · Giovanna Laurenzano[3] · Maria Rosaria Gallipoli[4] · Guido Martini[1] · Alessandro Peloso[5] · Sara Sgobba[2] · Francesca Pacor[2]

Received: 10 February 2026 / Accepted: 8 August 2026

Abstract

The reliable estimation of site amplification is crucial for seismic hazard and risk assessment. Although the amplification estimates are traditionally site-specific, recent efforts have focused on developing site amplification models for large areas. To this aim, we compiled existing site amplification factor estimates in Italy derived from different empirical methods. We then employed goodness-of-fit metrics and statistical tests to assess their mutual consistency and their agreement with predictions from an available 1D simulation-based model. We found a high correlation between the period-dependent site amplification factor estimates, suggesting that empirical methods consistently measure the relative amplification levels of different sites. Systematic differences among empirical estimates, observed particularly in Central Italy, are primarily due to the choice of the reference ground motion and can be corrected for by applying a simple scaling factor to a common reference ground motion level, thereby confirming the robustness of empirical estimates. Conversely, the comparison between empirical data and 1D simulations revealed a poor correlation, with the latter failing to reproduce the full range of observed amplification. For rock and stiff sites, 1D simulations tend to overestimate empirical amplification. In contrast, for a smaller fraction of soft-soil sites (e.g. within large basins), 1D simulations underestimate empirical amplification, likely due to their inability to model complex 2D/3D effects. These results suggest that the development of regional or national site-amplification models should prioritize empirical site amplification estimates.

Keywords Site amplification factor · Empirical method · 1D simulation · Statistical test · Earthquake ground motion

Extended author information available on the last page of the article

Published online: 26 August 2026

1 Introduction

Site amplification may induce large variations of earthquake ground-motion over distances of several hundred meters (Hartzell et al. 1997; Cultrera et al. 2011; Mayoral et al. 2019; Kramer and Stewart 2025; among many others) and therefore is recognized as a critical factor for seismic hazard and risk assessment.

Recently, a number of studies have developed large-scale site amplification maps to improve seismic risk assessment (Crowley et al. 2019; Weatherill et al. 2022; van Ginkel et al. 2022; Sgobba et al. 2024). These studies have typically relied on empirical determination of frequency-dependent site amplification at recording sites of permanent seismic networks and extrapolation to large areas using site condition proxies, such as time-averaged shear-wave velocity in the shallower 30 m depths *Vs,30* (Thompson et al. 2014), topography (Wald and Allen 2007) and surface geology (Vilanova et al. 2018). A crucial step in such studies is the validation of the large-scale site amplification models, which can be primarily performed by comparison with empirical site amplification data not included in the calibration dataset (Bergamo et al. 2023). However, the limited number of independent empirical observations with respect to the model spatial scale often hampers a fully independent validation; in such cases, alternative resampling-based strategies, such as leave-one-out cross-validation of empirical data, can be adopted to assess the robustness and predictive performance of the model (Sgobba et al. 2024).

In order to develop a national site amplification map in Italy, a simulation-based approach has been proposed by Falcone et al. (2021). They have collected subsurface data from a large database of seismic microzonation studies (Moscatelli et al. 2020); in particular, borehole logs, shear-wave velocity (*Vs*) profiles and associated physical and mechanical properties. Such data were grouped based on a geomorphological-geological clustering of the national territory (Mori et al. 2020); then, an extensive study of about 30 million 1D equivalent-linear ground response analyses (*GRAs*; Kottke et al. 2013; Kramer and Stewart 2025; Acunzo et al. 2023), with randomization of the 1D model parameters to account for spatial variability, has been performed to produce maps of period-dependent site amplification factors (*AFs*). Recent studies focused on combining such simulation-based *AFs* with the national probabilistic seismic hazard model of Italy (*PSHM*; Stucchi et al. 2011) for seismic risk mitigation purposes (Sabetta et al. 2023; Pieruccini et al. 2024; Albarello et al. 2025).

The median predictions of *AFs* distributions from this large-scale simulation-based model have been verified against (i) the median estimates obtained by 1D *GRAs* calibrated on site-specific *Vs* profiles retrieved from local detailed seismic microzonation studies (Falcone et al. 2021; Paolucci et al. 2020) and (ii) the median values of *AFs* calculated according to the functional form of the *Vs,30*-based site amplification term of the Bindi et al. (2014) ground motion model (*GMM*). However, a systematic consistency check of the outcomes of such studies against independent empirical *AFs* estimates derived from ground-motion observations has only been partially performed (Lanzano et al. 2025).

Leveraging the availability of large databases of empirical site amplification for the national territory (Priolo et al. 2020; Lanzano et al. 2025) and prompted by the involvement in the national research project "SERENA" (Mapping Seismic Site Effects at REgional and NAtional Scale, MUR-PRIN Italian project 2023–2025, Prot.2020MMCPER), which was aimed to develop a new simulation-based site amplification model for Italy, this work addresses two primary objectives: (1) to compare the empirical estimates of *AF*s from different methods in order to evaluate their agreement and (2) to perform a consistency check of the simulation-based *AF*s against independent empirical data. In particular, comparing empirical and simulation estimates allows for a better understanding of how well simplified 1D simulations describe site amplification, assuming that the experimental estimates are unbiased with respect to the broad range of site amplification effects (2D and 3D resonances, focusing and defocusing of seismic waves due to topographical irregularities and basin-edge generated surface waves; Kramer and Stewart 2025; Anderson 2007).

While the comparison of empirical site amplification data from different techniques has generally shown a satisfactory agreement (Grendas et al. 2025; Bindi and Kotha 2020; Pilz et al. 2009; Parolai et al. 2004), numerous studies have compared the empirical site amplification as measured by recordings from downhole array sites in US, Japan and Greece with predictions from 1D GRAs, questioning the wide applicability of 1D simulations (Afshari and Stewart 2019; Kaklamanos et al. 2013; Kaklamanos and Bradley 2018; Pilz and Cotton 2019; Tao and Rathje 2020; Hallal et al. 2022). Among such studies, those investigating a large number of KiK-net sites in Japan (more than 100, Thompson et al. 2012; Laurendau et al. 2018; Pilz and Cotton 2019), showed that only a variable fraction of these sites (from 16% to 70%) is well modeled using 1D GRA, suggesting that the 1D approximation is often inadequate to fully describe site amplification. Pilz and Cotton (2019), which investigated 354 sites, found that approximately 45% of the investigated sites exhibited significant 2D and 3D effects, which extended the site resonance frequency towards shorter periods and produced additional amplification beyond that predicted by impedance contrasts alone.

In the present study, the available empirical estimates for the Italian territory include site amplification data (Priolo et al. 2020) obtained from the Standard Spectral Ratio (*SSR*, Borcherdt 1970) and the Generalized Inversion Technique (*GIT*, Andrews 1986; Castro et al. 1990; Boatwright et al. 1991; Ameri et al. 2011; Oth et al. 2011; Klin et al. 2018; Shible et al. 2022; among others). In addition, the significant expansion of ground-motion databases allowed to conveniently estimate the frequency dependent amplification by analyzing site-to-site ground motion residuals (*δS2S*) for the national territory (Lanzano et al. 2025); this term measures the repeatable systematic deviation of seismic ground-motion from a reference partially or fully non-ergodic *GMM* (Rodriguez-Marek et al. 2011; Luzi et al. 2014; Kotha et al. 2018; Weatherill et al. 2022; among others). A major difference between the previously mentioned empirical methods lies in the reference ground-motion used to measure site amplification: (i) the recording at a reference site considered unaffected by site amplification, in the case of *SSR* and *GIT*, or (ii) the median *GMM* prediction for a reference condition, generally corresponding to outcropping rock, for *δS2S*.

In the following sections, we first describe the available empirical site amplification estimates. Then, we present a framework for the comparison of estimates from different methodological approaches using goodness-of-fit metrics and statistical tests. Finally, we use this framework to evaluate the robustness of empirical estimates and analyze residual distributions between empirical and simulation-based site amplification data.

2 Empirical site amplification factor data

In this study, we considered *AFs* based on three empirical methods for site amplification estimation: *SSR*, *GIT* and *δS2S*. These methods are detailed below.

SSR method (Borcherdt 1970) isolates the site amplification term at a specific location by identification of a nearby reference site unaffected by local site amplification (in general on outcropping bedrock conditions). This is a non-trivial task as many observations proved that outcropping rock sites often show high frequency ground-motion amplification which leads to biased estimates of the site amplification term for the specific site under investigation (Steidl et al. 1996). Assuming that the reference site contains the same source and propagation effects of the site of interest, the ratio of the recordings obtained for the same earthquake at the site of interest and that at the reference site completely describes the site amplification as a function of period (or frequency):

$$lnA_k\left(f\right) = \sum\nolimits_i \left[ln(FAS_{ik}(f)) - ln\left(FAS_{ir}(f)\right)\right]/(N-1) \tag{1}$$

where $A_k\left(f\right)$ is the site amplification function at *k-th* target site, FAS_{ik} and FAS_{ir} are the Fourier Amplitude Spectra of the recording from the *i-th* earthquake at *k-th* target and *r* reference sites, respectively, and *N* is the total number of earthquakes (Borcherdt 1970).

In general, this method is applied to data from dense, local arrays (Parolai et al. 2004) as it is difficult to find a nearby reference site for which the ratio between site-to-reference distance and epicentral distance is minimised.

In the *GIT* approach, the site amplification function is calculated by solving the following linear system of equation (Andrews 1986):

$$lnU_{ik}(f,R) = lnS_i\left(f\right) + lnA_k\left(f\right) + lnP_{ik}(f,R) \tag{2}$$

where $U_{ik}(f,R)$ represents the spectral amplitude of the recording obtained at *k-th* site at hypocentral distance *R* from the *i-th* earthquake with source function $S_i\left(f\right)$, $A_k\left(f\right)$ is the site amplification function and $P_{ik}(f,R)$ is a function accounting for amplitude attenuation. As this system of equations has two undetermined degrees of freedom, it needs to be constrained by selecting one (or more) reference site and fixing the attenuation at a given reference distance to be solved. Suitable solving schemes include parametric and non-parametric methods and single or multi-step inversion; an in-depth discussion of all solving strategies is beyond the scope of this paper but can be found in Shible et al. (2022). In general, when dealing with large ground-motion datasets, *GIT* approach is considered to provide robust estimates of the site amplification term but dependent on the identification of a suitable reference site, whereas regional variations of attenuation and source terms are affected by a higher variability and, therefore, may be less constrained.

In the *δS2S* approach, residuals are calculated as the difference between the logarithm of the observed ground motions at the site of interest and the logarithm of the predictions from a chosen *GMM*. Then, the site amplification term, *δS2S* is calculated according to the following decomposition scheme (Al Atik et al. 2010; Lanzano et al. 2025):

$$R_{es} = \delta_0 + \delta B_e\left|event + \delta S2S\right|site + \delta W_{es} \tag{3}$$

Where, R_{es} represents the residual, δ_0 is the residual offset[1], δB_e and $\delta S2S$ are the repeatable between-event and site-to-site components and δW_{es} is the event- and site-corrected residual. In this study, $\delta S2S$ represents the systematic deviations between the observed spectral acceleration (*Sa)* and the median prediction of the Italian *GMM* (ITA18, Lanzano et al. 2019) calibrated for shallow crustal earthquakes in Italy. The reference condition (hereinafter also referred to as generic rock condition) for the estimation of the site amplification is that of the outcropping engineering-bedrock ($Vs,30$ = 800 m/s) according to the current building regulations (EC8, CEN 2005; NTC18, NTC 2018), hence the site amplification term estimated through this method is denoted as $\delta S2S_{800}$ hereinafter.

In this study, the site amplification factor *(AF)* is used as the key parameter for the comparison between empirical and simulation-derived estimates. The *AF* is an integral parameter describing average amplification level in given period bands, extensively used for seismic microzonation studies at local and national level (Moscatelli et al. 2020). It is calculated according to the Italian Guidelines for Seismic Microzonation (SM Working Group, 2008) as the ratio between the averaged values of 5%-damped acceleration response spectra (*Sa*) over the *j-th* period band, according to the following formula:

$$AF_j = \frac{\Delta T_j \int_{\Delta T_j} Sa^{out} dT}{\Delta T_j \int_{\Delta T_j} Sa^{in} dT} \; for\; j = 1,2,3 \tag{4}$$

where Sa^{out} is the site-specific response spectrum at the surface, Sa^{in} is calculated for outcropping bedrock conditions and *j* represents three period intervals (ΔT_1= [0.1–0.5] s; ΔT_2= [0.4–0.8] s; ΔT_3= [0.7–1.1] s). The 3 amplification factors AF_1, AF_2, and AF_3 are named accordingly and they are intended to reflect the characteristic oscillator periods of the Italian ordinary building stock. For structures with a fundamental natural period greater than 1.1 s (e.g., high-rise buildings, large infrastructure, or special isolated structures) generalized regional site amplification models are not suitable and site-specific seismic hazard assessment should be performed according to current engineering practice.

The *AFs* data used in this study were compiled from previously published research. In the following we detail the procedures adopted for their computation and provide the relevant literature references.

In the case of *SSR*, as a slight modification of the concept expressed by Eq. 1, *AF* is calculated directly as the ratio of the recordings obtained for the same earthquake at the site of interest Sa^{out} and that at the reference site Sa^{in} as a function of period (or frequency); the final *AF* is obtained as the geometric average on all the events. Two *AF* datasets have been considered in this study (Fig. 1; Table 1), based on different seismic networks: (i) 3 A network in the Amatrice area installed during the Central Italy seismic sequence 2016-17 (Cara et al. 2019; Felicetta et al. 2021) and (ii) the 3 H network installed in the Norcia municipality during the L'Aquila Mw 6.1, 2009 seismic sequence (Luzi et al. 2019).

In the case of *GIT* (Eq. 2), *AFs* were calculated according to Eq. 4 as the ratio between the average values of the output and input response spectra over the *j-th* period band. The input spectrum Sa^{in} was defined as the average of seven 5%-damped acceleration response spectra compatible with the Uniform Hazard Spectrum (*UHS*) corresponding to a 10% prob-

[1] The offset is non null when the dataset for residual analysis is different from those used for model calibration.

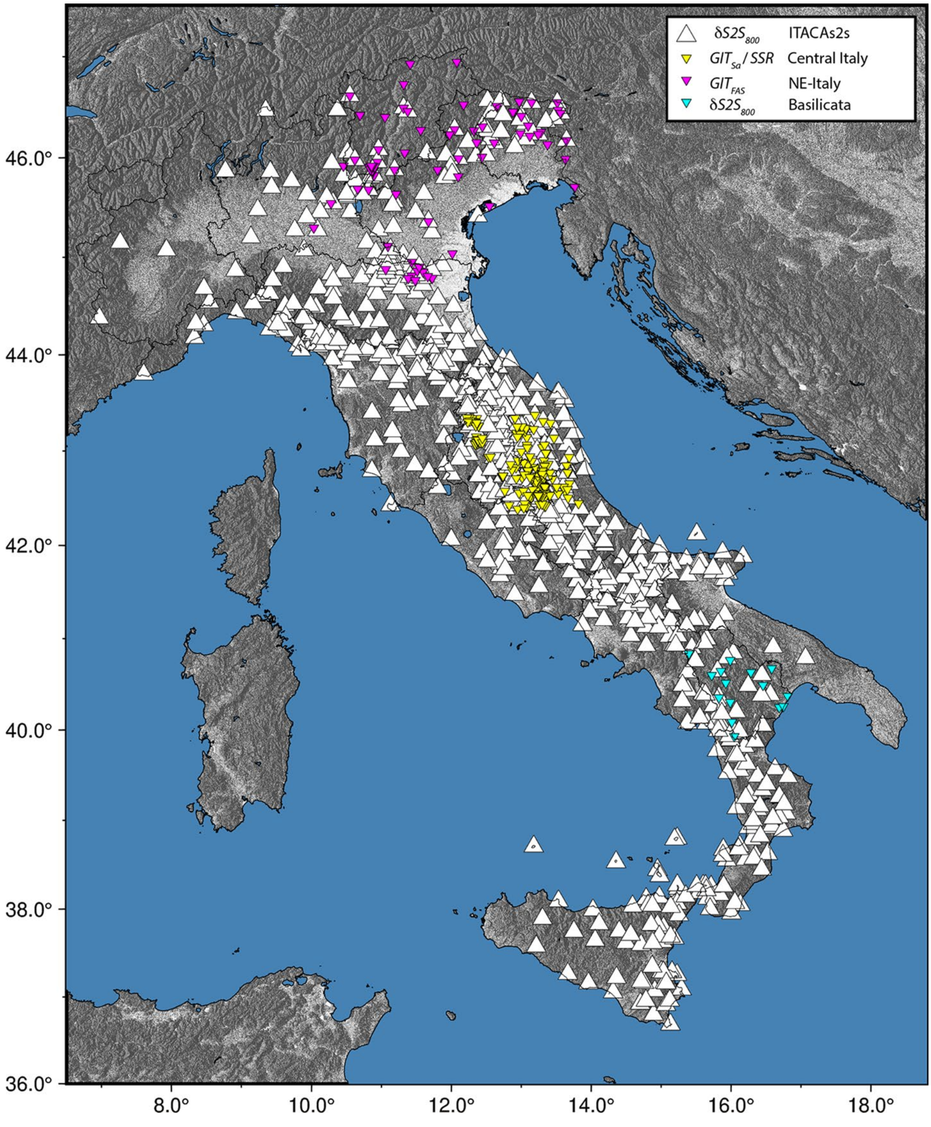

Fig. 1 Distribution of sites across the Italian territory for which empirical ground-motion amplification estimates are available, distinguished by estimation method. Regarding $\delta S2S_{800}$ stations, "ITACAs2s" refers to the sites of Lanzano et al. (2025) and "Basilicata" to the ones computed in this study. Refer to Figs. 2, 4, and 3, and for dedicated regional focus maps that explicitly display the individual stations with different estimation methods

ability of exceedance (*PoE*) in 50 years, derived from the *PSHM* for Central Italy (Luzi et al. 2020). When the *GIT* was derived from Fourier Amplitude Spectra (*FAS*), the output spectrum Sa^{out} was computed from site-specific accelerations obtained by multiplying the *GIT* spectral amplification curve (GIT_{FAS}) by the Fourier amplitude spectrum of the seismic input while preserving the phase information of the input signal, as described in Priolo et al. (2020). This is the case of the stations in North-East Italy (Klin et al. 2021; Laurenzano

Table 1 Main features of the AF empirical evaluation used in this study. References indicate the source papers for the *AFs*

Method	SSR	SSR	GIT_{FAS}	GIT_{Sa}	δS2S
Area	Central Italy	Central Italy	North-East Italy	Central Italy	Italy
Region/Municipality	Amatrice municipality	Norcia municipality	Friuli Venezia Giulia, Veneto Trentino regions, parts of Lombardia and Emilia Romagna regions	Lazio, Umbria, Marche, Abruzzo regions	all
stations	39	12	85	141	916+14
networks	3A	3H	FV, NI, OL, OX, ZX	IT, IV, XJ-2009, XO, 3 A, 3 H	Several permanent and temporary networks
reference site(s)	IV.T1299	3H.NO11	TN02 for ZX network, FAU for the others	average of response sites of 4 stations (IT.LSS, IT.SLO, IV.T1217, IV.T1221)	ITA18 predictions at $Vs{,}30$ = 800 m/s
References	Cara et al. 2019; Felicetta et al. 2021	Luzi et al. 2019	Klin et al. 2021; Laurenzano et al. 2023	Priolo et al. 2020	Lanzano et al. 2025 for 916 sites; this study for the 14 sites in the Basilicata region

et al. 2023) (Fig. 1; Table 1). As stated in Table 1, different reference sites were used since they belong to different microzonation areas.

Differently, when the *GIT* was calculated in terms of response spectra (GIT_{Sa}), the output spectrum Sa^{out} was estimated by multiplying the reference spectrum on rock with the GIT_{Sa} to obtain the corresponding site-specific response spectrum (Pacor et al. 2016). This is the case of the *AFs* for the Amatrice and Accumoli area and for most of the other Central Italy seismic stations (Priolo et al. 2020; Fig. 1; Table 1).

It is worth noting that *GIT* and *SSR* results were computed on a similar ground motion dataset collected for the 2016 Central Italy seismic sequence, adopting analogous processing schemes.

Finally, in the case of $\delta S2S_{800}$ (Eq. 3), *AF* values were calculated assuming that the integrals of Eq. 4 can be approximated by averaging the $\delta S2S_{800}$ values across the selected period intervals (further details in Lanzano et al. 2025). The considered dataset refers to the $\delta S2S_{800}$ and *AF* estimations for 925 accelerometric and velocimetric stations across Italian territory: 916 of them are available in Lanzano et al. (2025), whereas estimates for 14 sites from the temporary 1 V network in Basilicata region (Southern Italy; Gallipoli 2003) were computed in this study with the same procedure (Fig. 1; Table 1).

The empirical *AF* datasets, whose key features are summarized in Table 1, were collected, verified, and structured within a dedicated spatial data infrastructure that publishes Open

Geospatial Consortium web services. In particular, each *AF* estimate was associated with the latitude and longitude of the corresponding recording station, such that multiple estimates from different methods at the same location could be directly compared. Analysis was then performed using open-source Python libraries within a Jupyter-notebook environment.

3 Quantitative assessment of the agreement between *AF* estimates

To quantitatively evaluate the consistency of *AF* estimates obtained by different empirical methods as well as their agreement with the predictions from the 1D simulation-based model by Falcone et al. (2021) (in terms of geometric average and associated standard deviation), we employed several goodness-of-fit (*GoF*) metrics, which are described in the following. As a preliminary step, since ground motions are assumed to be log-normally distributed, natural logarithm transformation is applied to *AF* values, either empirically or numerically derived, before calculation of statistical parameters.

3.1 Agreement at individual sites

When two distinct methods are used to estimate *AF* for the same group of sites, consistent results among methods would imply a similar general trend in these estimates. Hence, the assessment of the consistency between *AF* estimates obtained from different methods was performed by the Spearman's ρ (Spearman 1907) and Kendall's τ (Kendall 1970) non-parametric correlation coefficients. Similarly to Pearson's r linear correlation coefficient, ρ and τ inform about the strength of the relation between two variates, taking values ranging between −1 (perfect anti-correlation) and 1 (perfect correlation). Differently from r, they do not require the data being normally distributed and are sensitive to non-linear (monotonic) relationships between data. These *GoFs* are calculated as:

$$\rho = \frac{cov(r_x, r_y)}{\sigma_{rx}\sigma_{ry}} \tag{5}$$

$$\tau = \frac{2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left[sgn(x_i - x_j)sgn(y_i - y_j)\right]}{n(n-1)} \tag{6}$$

where x_i and y_i are the (nat. log) values of the amplification factor estimated by two distinct methods for the *i*-th site in a sample with numerosity *n*, r_x and r_y represent the ranks of the amplification values, with standard deviations σ_{rx} and σ_{ry} (nat. log units), cov denotes the covariance and sgn is the sign function. Such *GoF* metrics find extensive usage in the literature (Thompson et al. 2012; Zhu et al. 2022).

Among the *GoF*s, ρ and τ are simple parameters to adopt but are insensitive to systematic deviations of groups of estimates, even in case of significant correlation. Therefore, we also use the chi-squared χ^2 (Scherbaum et al. 2004). This is calculated as the sum of normalized residuals (defined later in this section, Eqs. 6 and 7) between any two estimates of site amplification, divided by the number of considered sites, to account for the sample size: a value near zero indicates perfect agreement, while systematic deviations between groups of individual estimates result in higher values.

3.2 Similarity through statistical testing of distributions

While previous *GoF*s analyze the agreement of groups of *AF* estimates in a general sense, a more complete understanding of the consistency of these estimates can be achieved by examining the shape of their distributions. For this reason, the Kolmogorov-Smirnov (*KS*) goodness-of-fit test was employed to assess the agreement between two distributions of *AF* estimates. This statistical test evaluates the probability of observing the maximum value of the absolute difference between two distinct Cumulative Distribution Functions (*CDF*) against the null hypothesis of the amplification values estimated by different methods being drawn from the same population. Being a non-parametric statistical test, it does not require the data following a specific distribution and can be reliably used also in cases of small samples. In addition, this metric checks for the similarity of the distributions not only in general terms, e.g. the similarity of *CDF* shapes, but also for more specific features, such as the most deviant values of the criterion variable (Press et al. 2001; Scherbaum et al. 2004). *KS* is an exact test which does not require binning of the data. However, it is worth noting that its statistical power is inherently higher for detecting differences near the central values (median) of the distributions than in the tails (the extreme minimum and maximum values). Given its robustness, *KS* test has been extensively used in the field of ground motion selection and testing against seismic hazard prediction (Bradley 2012; Chandramohan et al. 2016).

3.3 Discrepancies through residual analysis

The examination of residual distributions can offer further understanding of *AF* estimates derived from various methods. Residuals can reveal systematic patterns based on their spatial arrangement or the classification of site conditions. This information can then provide clues about the source of potential systematic errors between estimates and suggest possible corrections.

To account for the individual uncertainty in *AF* estimates, we normalized the residuals by their dispersion according to the error propagation rule, such that the residual between *AF* values obtained by method *x* and method *z* for site *i* were calculated as:

$$Res = \frac{x_i - z_i}{\sigma} \tag{7}$$

with

$$\sigma = \sqrt[2]{\sigma_x^2 + \sigma_z^2 + 2cov(\sigma_x, \sigma_z)} \tag{8}$$

In the above equations, x_i is the natural log of *AF* from method *x* at site *i*, z_i is the estimate from method *z* at the same site *i* and σ is the standard deviation of the natural log amplification values.

Ideally, in case of equal estimates from two independent methods x and z, the normalized residuals (hereinafter only residuals) should be normally distributed with a mean value of zero and unit variance. After calculating residuals, their spatial distribution can be examined for clusters which may reveal regional dependencies. Residuals were grouped according to different site classification schemes, considering geological, morphological or geophysical information available for the sites, and the resulting distributions were further analyzed. Specifically, Kruskal-Wallis (KW) approach (Kruskal and Wallis 1952) was used to identify significant differences among group means; this is a non-parametric statistical test similar to analysis of variance (ANOVA) which operates on the rank of data and does not require to fulfill the strict assumptions of ANOVA regarding homogeneity of the distributions' variances and equal sample size (Vidakovic 2011). Since this test is unable to identify dissimilar distributions among those examined, if significant differences are detected across all considered distributions, subsequent statistical procedures, such as the Conover test (Conover and Iman 1979; Terpilowski 2019), were utilized to pinpoint the dissimilar groups via multiple pairwise comparisons.

4 Results

4.1 Comparison of empirical *AF* estimates

To assess the reliability of empirical estimates, we employed the *GoF*s metrics introduced in the previous section. These metrics enabled the comparison of empirical *AFs* estimates derived from different datasets at various spatial scales. For this study, the $\delta S2S_{800}$ estimates will serve as reference values due to their homogeneous widespread calculation across the entire national territory. We then compared this reference dataset against other empirical estimates obtained at regional and municipal levels. In order for the comparison to be meaningful, we contrasted only those estimates obtained for the exact same geographical location, namely two empirical estimates from different methods calculated for the same recording station.

4.1.1 Regional comparison: North-East Italy

In the case of NE Italy (Table 1), the regression between (nat. log) *AFs* estimated from $\delta S2S_{800}$ and GIT_{FAS} at 21 sites indicates that GIT_{FAS} overpredict $\delta S2S_{800}$ by 10 to 20% on average, depending on the considered period band (Fig. 2). However, the regression parameters alone provide an incomplete picture of the methods agreement, especially in the case of a small sample size and considering only the *AF* median values. Indeed, when evaluating the broader set of GoFs introduced in the Sect. 3, the comparison shows a remarkably close agreement, with values of ρ above 0.9 and χ^2 below 0.15 for all the considered period bands (Table 2). For these two datasets, the two-tail *KS*-test result supports the conclusion that the *CDF* distribution of median estimates from the two methods are equivalent (Table 2; Fig. 3).

4.1.2 Regional comparison: Central Italy

In the case of Central Italy (Table 1), which represent the area with the largest number of sites available for comparison (roughly 100 sites), we still found a significant correlation between $\delta S2S_{800}$ and GIT_{Sa} *AF* estimates (ρ above 0.9 and τ above 0.75 for all the considered period intervals, Table 3). However, a systematic deviation was observed between the two methods, particularly for the AF_1 (χ^2 of 0.64), for which the regression indicates

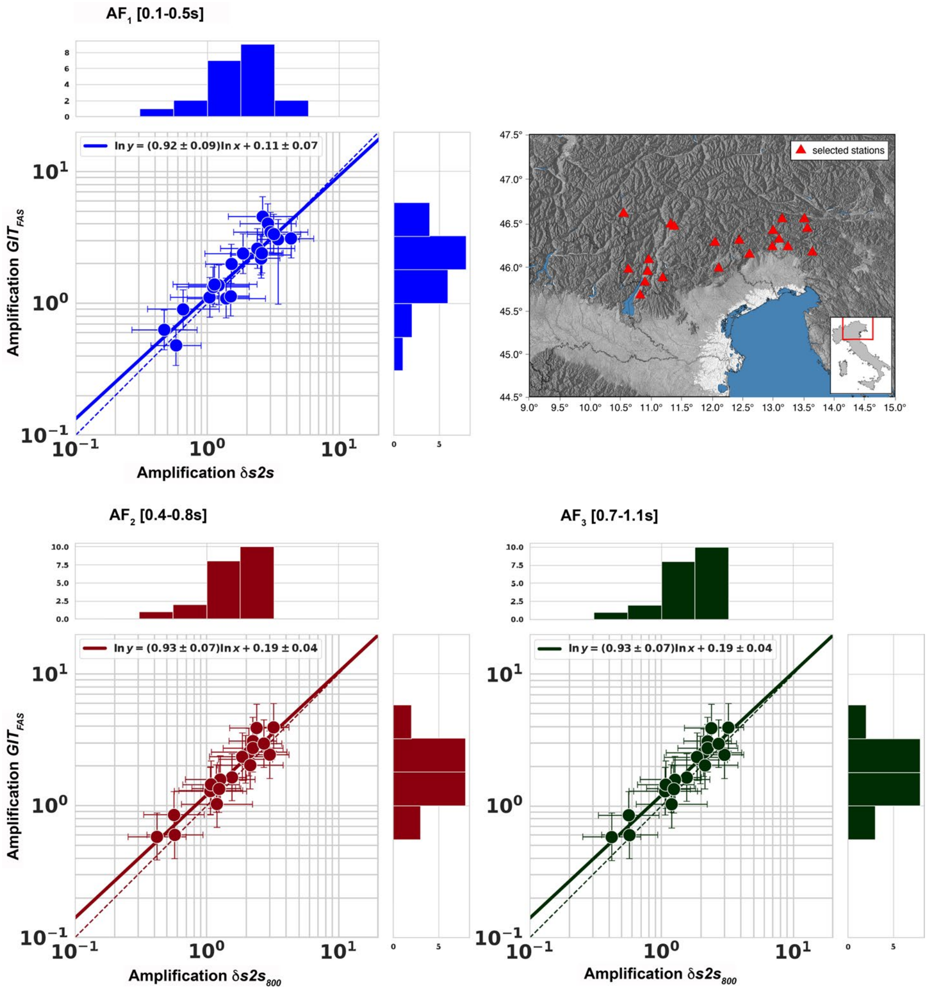

Fig. 2 Map of stations in North-East Italy used for *AFs* comparison (top right panel). Scatter plot of *AF* estimates obtained from $\delta S2S_{800}$ (in abscissa) versus those obtained from GIT_{FAS} (in ordinates), in the short- (AF_1 in blue, top left panel), intermediate- (AF_2 in red, bottom left panel) and long-period intervals (AF_3 in green, bottom right panel). In the scatter plots, the linear fit of log *AF* values (solid line) and 1:1 line are also shown

Table 2 *GoFs* values calculated from the comparison of *AF* estimates obtained through $\delta S2S_{800}$ and GIT_{FAS} at 21 sites in North-East Italy

	ρ	τ	χ^2	KS	*p*-value
AF_1	0.91	0.72	0.15	0.14	99%
AF_2	0.94	0.82	0.13	0.29	36%
AF_3	0.94	0.82	0.13	0.29	36%

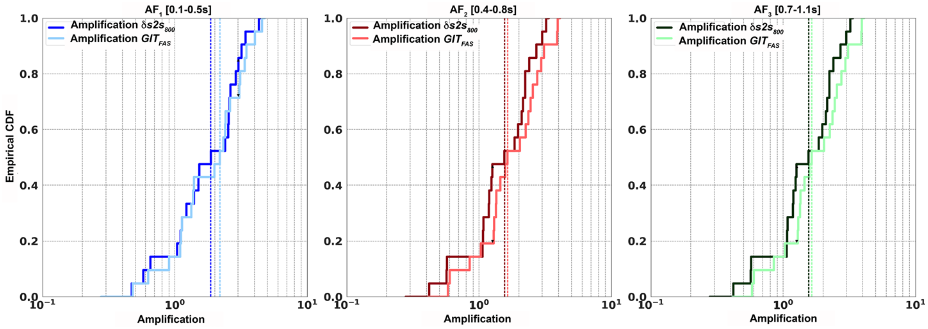

Fig. 3 Comparison of AF_1 (left panel), AF_2 (middle panel) and AF_3 (right panel) empirical cumulative distribution functions (CDF) for the sites in NE Italy obtained by $\delta S2S_{800}$ method (blue, red and green) and GIT_{FAS} (light blue, light red and light green). The vertical dashed lines indicate the median values of the two distributions. The results of two-samples *KS*-test are reported in Table 2

Table 3 *GoF* values calculated from the comparison of *AF* estimates obtained through $\delta S2S_{800}$ and GIT_{Sa} at 99 sites in Central Italy

	ρ	τ	χ^2	KS	*p*-value
AF_1	0.90	0.75	0.64	0.39	>>0.1%
AF_2	0.91	0.76	0.45	0.25	0.4%
AF_3	0.91	0.76	0.38	0.26	0.2%

that GIT_{Sa} provides *AF* estimates about 65% larger than $\delta S2S_{800}$ on average (Fig. 4). These deviations are larger for the short-period interval and decrease for the intermediate and long-period intervals (χ^2 values in Table 3). This mismatch stems from a fundamental difference in the reference ground motion definitions adopted by the two methods. As also shown for the local comparison in the next Sect. (4.1.3) and thoroughly quantified and discussed later in Sect. 4.2, the local reference stations used in the GIT_{Sa} exhibit significant lower ground-motion amplitudes (up to 55% at short periods) relative to the reference used in $\delta S2S_{800}$, which inherently drives the *AF* estimates from GIT_{Sa} toward higher values.

4.1.3 Local-scale comparison: Amatrice and Norcia municipalities, Central Italy

In this section, we show the comparison of $\delta S2S_{800}$ and *SSR* estimates of *AFs* obtained at a local scale, namely for those stations belonging to the temporary networks within the Norcia and Amatrice municipalities in Central Italy (Fig. 5).

The comparison of the estimates from the two methods, considering all sites from both the Norcia and Amatrice networks, shows a rather poor correlation, particularly in the short-period range (ρ =0.42, Table 4). However, distinguishing the *SSR* estimates in two subsets based on the network, a distinct pattern of significantly higher correlations was observed (Table 4; Fig. 5), as in the previous cases. The scatter plots show two distinct clusters of values: the Norcia cluster, with higher positive deviation between *SSR* and $\delta S2S_{800}$ estimates (χ^2 around 6, Table 4) and the Amatrice cluster, with lower deviations (χ^2 around 0.15 for AF_2 and AF_3, Table 4). As further discussed in the next Sect. 4.2, the cause of this difference is due to the choice of the local reference site, which was different between the two local

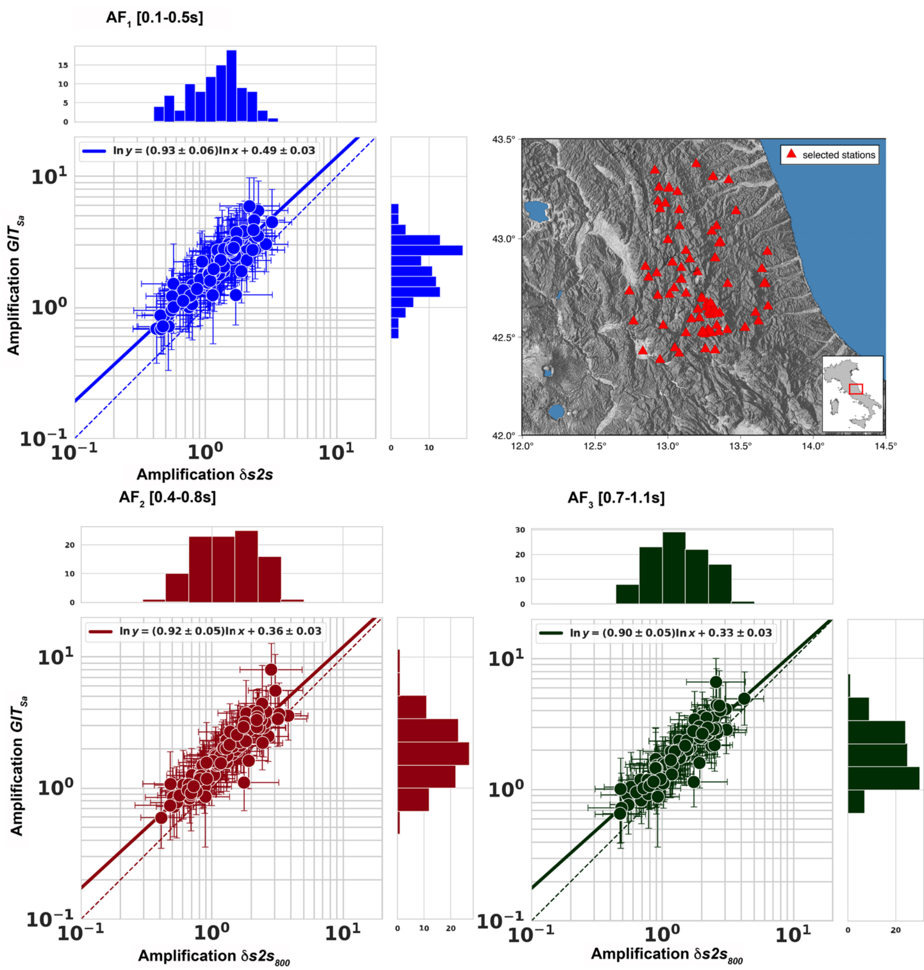

Fig. 4 Same as Fig. 2 but for the comparison between *AF* estimates obtained from $\delta S2S_{800}$ and GIT_{Sa} for the stations in Central Italy

networks: the systematic offset observed between the two groups of sites in terms of SSR values (Fig. 5) suggests that the reference site selection has a major role on the mismatch between estimates of amplification level.

4.2 On the role of the reference ground motion in the mismatch between experimental *AF* estimates

The discrepancies in experimental *AF* values may stem from various factors, such as the characteristics of the ground-motion database or the underlying assumptions of the estimation method.

However, the comparison in Sect. 4.1.3 highlighted the significant influence of the reference motion used to evaluate site amplification. To further explore this, it is crucial to quan-

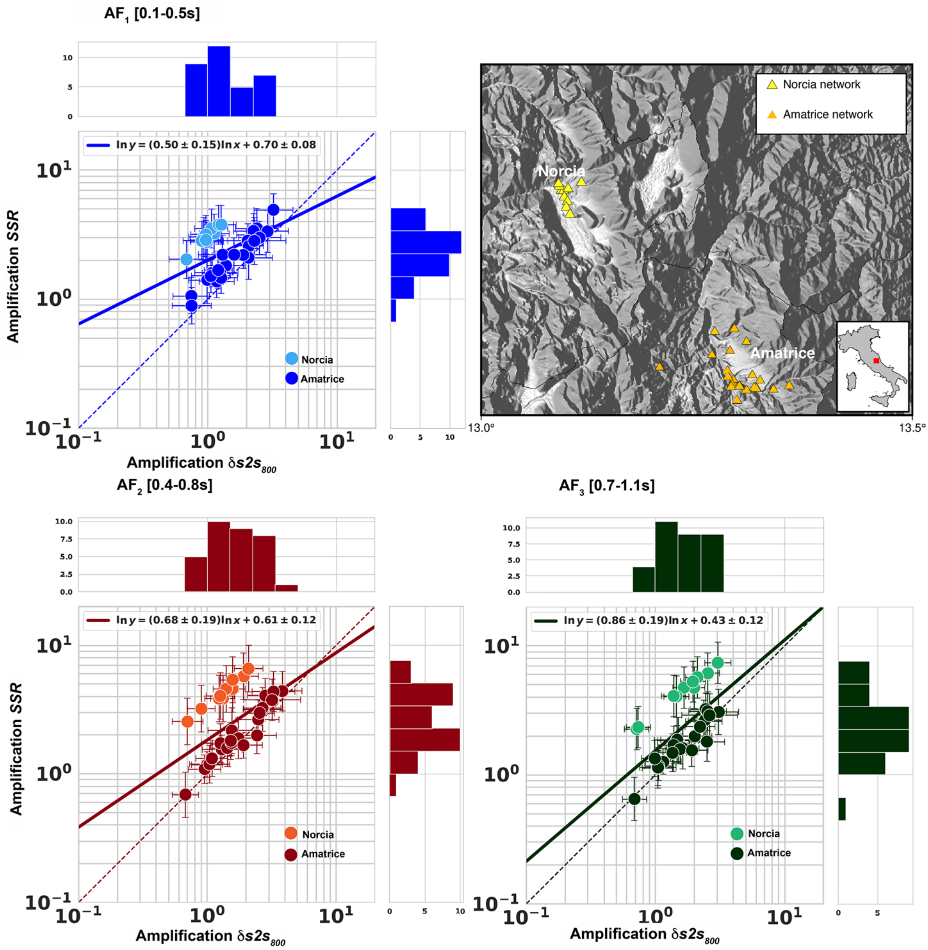

Fig. 5 Same as Fig. 2 but for the comparison between *AF* estimates obtained from $\delta S2S_{800}$ and *SSR* technique for the stations in Amatrice and Norcia municipalities, Central Italy

tify the differences among the reference motions employed by each estimation method and understand if the choice of a particular reference ground motion could introduce a potential mismatch. To this aim, the most suitable dataset is the Central Italy one (Sect. 4.1.2), which includes a greater number of sites and where we observed a systematic deviation between empirical estimates obtained by $\delta S2S_{800}$ and GIT_{Sa} (Fig. 4).

The $\delta S2S_{800}$ values of the reference sites used for GIT_{Sa} analysis in Central Italy (Table 1) are shown in Fig. 6a. These sites, on average, show site terms remarkably lower than the generic rock for periods above 0.1 s, thus implying that the generic rock conditions include several stations with non-negligible ground motion amplifications. In particular, the ground motion mean reduction with respect to the generic rock is greater for the short periods (up to 55% for IT.LSS and IV.T1221 at about 0.15 s), but it also characterized by a large variability (amplification at 0.05s for IV.T1217). At longer periods the difference is lower (around 25% around 1 s) with less variability. Similarly, the two reference stations used in Amatrice (IV.

Table 4 *GoF* values calculated from the comparison of *AF* estimates obtained through $\delta S2S_{800}$ and *SSR* at 33 sites in Amatrice and Norcia municipalities, Central Italy

	ρ	τ	χ^2	KS	*p*-value
AF_1	0.42	0.32	2.21	0.55	>>0.1%
AF_2	0.53	0.45	1.92	0.42	0.5%
AF_3	0.65	0.49	1.82	0.33	5.1%
Norcia network					
AF_1	0.83	0.69	6.32	1.00	>>0.1%
AF_2	0.99	0.96	5.88	1.00	>>0.1%
AF_3	0.95	0.87	5.68	0.80	0.2%
Amatrice network					
AF_1	0.94	0.78	0.43	0.35	12.4%
AF_2	0.96	0.89	0.17	0.30	24.1%
AF_3	0.89	0.74	0.14	0.22	66.0%

T1299) and Norcia (N3H.NO11) municipalities show site terms generally lower than the generic rock for periods in the range 0.1–1.1 s (Fig. 6b). Consequently, the SSR technique overestimates the *AFs* compared to the $\delta S2S_{800}$ in both municipalities. However, because the reference site in Norcia has a site term remarkably lower than that in Amatrice, the resulting *SSR* overestimation is systematically larger in Norcia (Fig. 5).

Lanzano et al. (2022) developed an empirical correction function for generic-to-reference ground-motion scaling, δ_{ref}, by computing the residuals between the ITA-18 *GMM* for generic rock conditions and the ground motion recorded at a group of reference sites in the national territory, chosen based on strict seismological criteria (Lanzano et al. 2020). This term describes the log factor that must be added (with sign) to the predictions of existing *GMM* for generic rocks to provide median ground motions of the selected reference sites in Italy. Thus, the predicted ground-motion amplitude $Sa_{ITA18-ref}$ for reference rock sites is:

$$\log_{10} Sa_{ITA18-ref}(T, M, R, SF) = \log_{10} Sa_{ITA18}(T, M, R, SF) + \delta\ (T) \tag{9}$$

where the generic rock prediction (Sa_{ITA18}) is dependent on period T and various explanatory variables, such as the moment magnitude M, distance R and style of faulting SF. In turn, the correction factor δ is scenario independent and depends only on T, is larger at short periods and decreases at lower periods. Note that Lanzano et al. (2022) also proposed a parameterized version of the correction term dependent on $Vs{,}30$ and κ_0 (near-site high frequency attenuation parameter; Anderson and Hough 1984), but in this study their simplified version was used.

The physical basis of such period-dependent corrective term relies on the distinction between generic rock conditions and "true" reference sites. The former is assumed by standard *GMM* and is generally calibrated on recordings obtained at sites with $Vs{,}30 \cong 800$ m/s (Lanzano et al. 2022). Being $Vs{,}30$ an incomplete predictor of site amplification (Castellaro et al. 2008), it has been proved that such condition does not ensure a negligible site amplification, instead such sites are often affected by high-frequency amplification resulting from weathering and fracturing of surface rock outcrops (Steidl et al. 1996; Cadet et al. 2010). Conversely, site-specific methods, such as *SSR* and *GIT*, utilize local reference sites lying on outcropping rock with measured negligible site amplification over a wide frequency range (Steidl et al. 1996). However, the selection of reference sites in local datasets can be challenging and is difficult to generalize: the differences between reference ground-motion

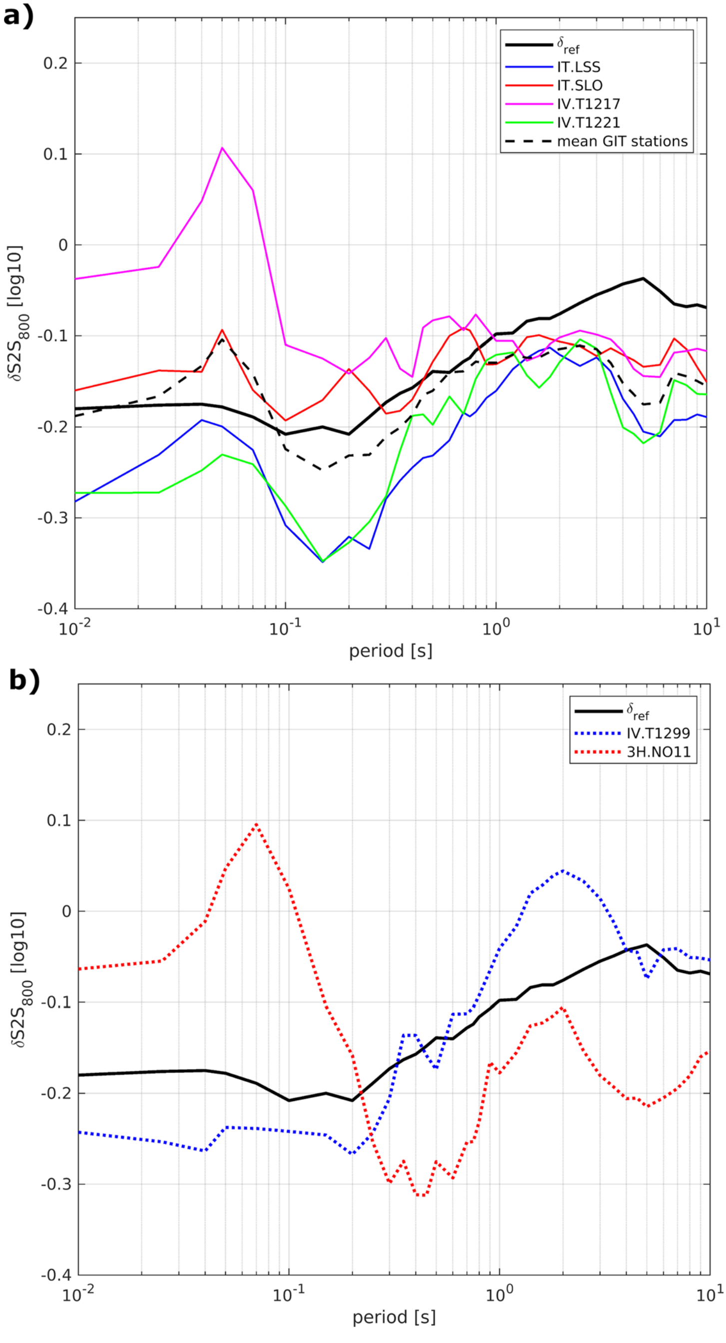

Fig. 6 $\delta S2S_{800}$ of the reference sites used to estimate site amplification as a function of period through: (**a**) GIT_{Sa} method in Central Italy and (**b**) *SSR* method in Amatrice and Norcia (colored curves). For both panels, the generic-to-reference rock scaling curve is also shown (δ_{ref}, solid black line)

levels could lead to systematic bias in the *AF* estimates, like in Central Italy, whereas this effect is not relevant in other areas, like North-East Italy.

It is then important to define the reference level using at least one nationwide dataset in order to allow for comparison across a sufficient number of samples. This is the case of the empirical correction function δ_{ref}, which is calibrated on a carefully selected group of Italian "true" reference hard-rock sites based on careful examination and weighting of multiple site proxies. It allows to systematically remove the amplification included in the generic rock condition of the *GMM*, harmonizing the generic rock and "true" reference ground motion levels and ensuring that the derived *AFs* estimates are comparable.

Application of δ_{ref} at the *AF* estimates in Central Italy (hereinafter $\delta S2S_{ref}$) compensates for the systematic deviations previously observed between $\delta S2S_{800}$ and the GIT_{Sa} derived values (Figs. 4 and 7). In fact, the $\delta S2S_{ref}$ distribution are shifted towards higher values with respect to those of $\delta S2S_{800}$ after application of the generic-to-reference rock correction, so that the χ^2 values are largely reduced and the two-tail *KS*-test results support the conclusion that the distribution of median estimates from $\delta S2S_{ref}$ and GIT_{Sa} are equivalent (Table 5).

These results demonstrate that the eventual mismatch between *AF* estimates calculated by different empirical methods is mainly controlled by the assumption on the reference ground-motion level, such that any systematic deviation obtained by empirical methods can be corrected by accounting for the scaling between the utilized reference motions.

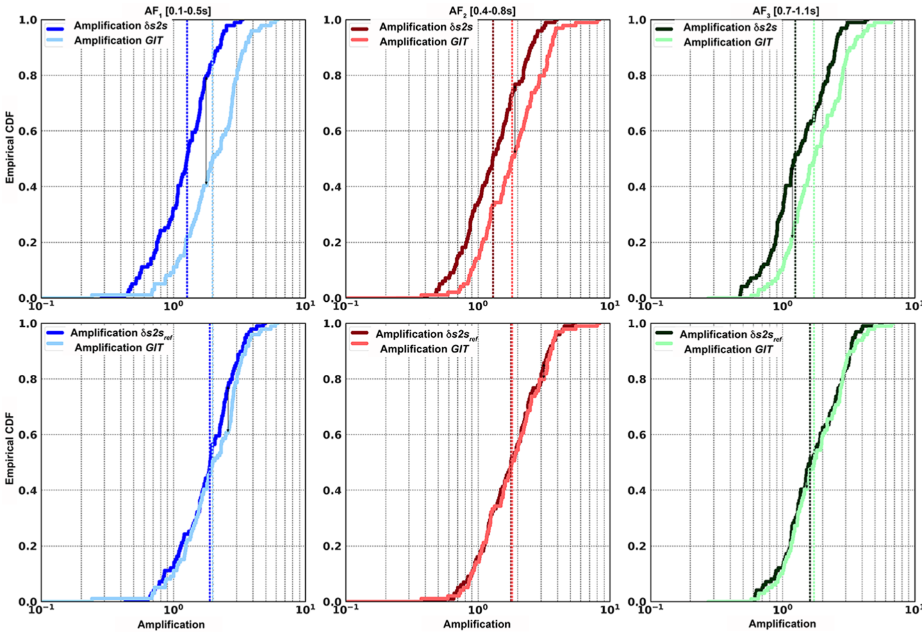

Fig. 7 Comparison of AF_1 (left panels), AF_2 (middle panels) and AF_3 (right panels) empirical cumulative distribution functions (*CDFs*) for the sites in Central Italy obtained by $\delta S2S_{800}$ method (blue, red and green) and GIT_{Sa} (light blue, light red and light green). Bottom row: similar to upper row for the comparison of *CDFs* from $\delta S2S_{ref}$ and GIT_{Sa} methods. The median values of the two distributions (vertical dashed lines) are also shown; the results of two-samples *KS* test are reported in Tables 2 and 4

Table 5 *GoF* values calculated from the comparison of *AF* estimates obtained through $\delta S2S_{ref}$ and GIT_{Sa} techniques at 99 sites in Central Italy

	ρ	τ	χ^2	KS	*p*-value
AF_1	0.90	0.75	0.14	0.17	10.8%
AF_2	0.91	0.76	0.14	0.06	99.4%
AF_3	0.91	0.76	0.13	0.07	96.7%

In the case of NE Italy dataset, application of the correction factor worsens the previously observed fit between $\delta S2S_{800}$ and GIT_{FAS} only in the short period interval (χ^2 from 0.15 to 0.58 for AF_1; Figure S1 and Table S1 in the Supplementary material), whereas a similar fit is observed in terms of AF_2 and AF_3. This evidence suggests that the ground motion of the reference site for the GIT_{FAS} in NE Italy is consistent with the ITA-18 GMM prediction for generic rock conditions.

4.3 Comparison between empirical and 1D simulation *AF* estimates

Having shown that empirical *AF* estimates are mutually consistent once scaled to a common reference condition, we now compare them with the 1D numerical simulation estimates recently proposed for the Italian territory (Falcone et al. 2021). This comparison is useful for a better understanding of how well simplified 1D simulations describe site amplification assessed by the experimental estimates.

In particular, Falcone et al. (2021) developed *AFs* maps for Italy based on a rather complex procedure including: (1) zonation of Italy into 42 clusters based on geological-geomorphological criteria; (2) subclustering of each zone to identify "typical" 1D subsurface models, namely stacks of horizontal uniform layers, based on the available geological, geotechnical and geophysical data from seismic microzonation studies (Moscatelli et al. 2020), using standard modulus reduction and material damping curves from literature studies to describe nonlinear soil behavior and accounting for the uncertainty in the model parameters by stochastically varying the distribution of soil units and their parameter values above the bedrock; (3) implementation of a large number of 1D equivalent-linear visco-elastic *GRAs* based on the Random Vibration Theory (Kottke et al. 2013), using *UHSs* input corresponding to an exceedance probability of 10% in 50 years of the available *PSHM* calculated by the Istituto Nazionale di Geofisica e Vulcanologia (Stucchi et al. 2011); (4) development of *AFs-Vs*,*30* correlations (16°, 50° and 84° percentiles) as a function of the input motion level for each of the 42 initial zones and (5) development of national *AFs* maps on a regular 50 × 50 m grid. It is worth noting that these *AF* estimates, which were calculated according to Eq. 4, do not reflect any site-specific Vs profile or soil layering but rather an average estimate from a set of 1D models representative of a wide area, with randomization of *Vs* and other model layer parameters to account for subsurface spatial variability in the considered subcluster.

We followed the same approach of Sect. 4 for the comparison between the 1D numerical simulation-based estimates (hereinafter 1D simulation) obtained by Falcone et al. (2021) and the empirical estimates, by archiving and analyzing them through the same spatial infrastructure used for the empirical *AF* data. In particular, we compared our empirically derived estimates at seismic recording sites with those estimated in the *AFs* national maps for the same geographical locations. We used the estimates from the *AFs* maps instead of the *AFs*-

Vs,30 correlations as the latter require the knowledge of *Vs,30* at the site of interest, which has been measured only for a minor fraction of seismic recording sites (about 30%, Lanzano et al. 2019).

To maximize the number of sites, thus exploring the variability of the comparison as a function of diverse local conditions, we considered the empirical *AFs* obtained at the national level by $\delta S2S$ method at 925 sites. As the *AFs* from 1D simulations were calculated with respect to the engineering-bedrock condition of the building regulations (*Vs,30*=800 m/s; CEN 2005; NTC 2018), we considered here the $\delta S2S_{800}$. As already shown by Lanzano et al. (2025), the empirical and numerical estimates are uncorrelated in terms of AF_1 and weakly correlated for AF_2 and AF_3 (Fig. 8 and Table S2 in the Supplementary material). The 1D simulations cannot reproduce the range of observed *AFs*, which spans a wider range of values (0.2-7). It is also worth noting that the uncertainty in empirical estimates is considerably larger than that of 1D simulations. Therefore, we neglected the uncertainty in the latter when calculating the residuals to a first approximation.

The distribution of residuals calculated for short-period amplification factor (AF_1) showed a negative tail (skewness of about −0.5) and mean value around −1 (Fig. 9, bottom-left panel). This indicates that most sites, mainly located along the Apennine belt and in the Tyrrhenian side of the Italian peninsula (bluish triangles in Fig. 9, top-left panel), display negative residuals, meaning that 1D simulations overestimate the empirical AF_1.

For about 30% of sites the residuals have values around zero, meaning that AF_1 estimated by the 1D simulation approach is consistent with the experimental measures (white triangles in Fig. 9). Finally, only a small fraction of sites (about 5%) showed residuals larger than 1, meaning that 1D simulation values underestimate the empirical AF_1; all these sites are located in the central and eastern side of the peninsula and in eastern Sicily, often within deep basins, like the Po plain, the Fucino basin and the Adriatic foredeep (reddish triangles in Fig. 9, top-left panel).

Focusing on the long-period site amplification factor (AF_3), the distribution of residuals showed a shape close to standard-normal, with a slightly more pronounced positive tail (skewness of 0.25, Fig. 9). With respect to AF_1, the fraction of sites for which 1D simulations overestimate empirical AF_3 decreases from 65% to 34%, while the percentage of sites with comparable AF_3 slightly increases from 30 to 35%. Notably, a more consistent

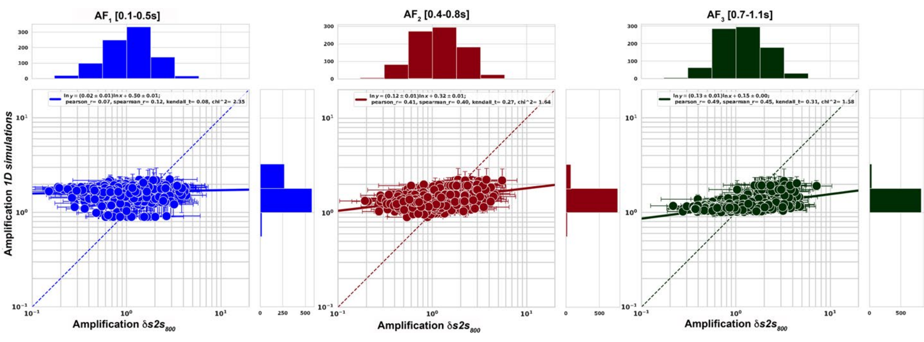

Fig. 8 Scatter plot of *AF* estimates from $\delta S2S_{800}$ method (in abscissa) versus those obtained from 1D simulations, calculated in the short-period interval (blue, left panel), intermediate- (red, middle panel) and long-period intervals (green, right panel). The linear fit of log *AF* values (solid line) and 1:1 line are also shown

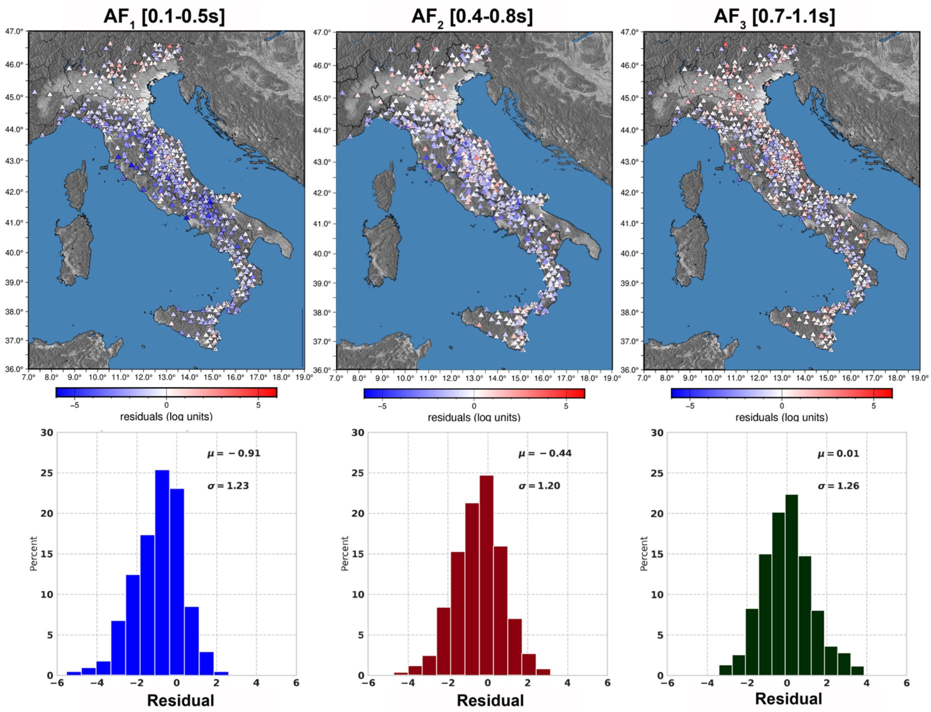

Fig. 9 Residual between site AF estimates from $\delta S2S_{800}$ method and 1D simulations for AF_1 (left), AF_2 (middle) and AF_3 (right): spatial distribution maps (top panels) and statistical distributions (bottom panels)

percentage of sites (around 16%) showed a significant underestimation of empirical AF_3. In general, the spatial distribution of AF_3 residuals follows the geographic patterns described for AF_1, with severe underestimation of empirical AF_3 in large deep basins, like the Po plain (Fig. 9, top-right panel and Figure S2 in Supplementary material). For AF_2, the statistical and spatial distribution of residuals exhibited characteristics that were intermediate between those observed for the short and long-period amplification factors.

To investigate any systematic pattern, residual distributions were analyzed as a function topographical (Mascandola et al. 2021a, b) and $Vs,30$-based site classification adopted in the current building regulations (CEN 2005; NTC 2018). Specifically, NTC18 defines 5 primary site classes (A to E) based on decreasing $Vs,30$ and 4 topographical classes (T1 to T4) corresponding to increasing slope inclination (T4 refers to slope inclination > 30°). These classifications are associated with an aggravation factor of the ground motion for $Vs,30$ < 800 m/s and inclination > 15°. In addition, distributions were also grouped according to a simplified geological classification, adapted from the national map by Bucci et al. (2022). Even if not very detailed, given the scale of the map and the crudeness of the proposed classification, it is worthwhile assessing whether geological data can account for any significant difference observed in the distributions.

For all the considered period intervals, the KW statistical test indicated significant differences among the residual distributions for all classification schemes (Table 6).

Table 6 Results of Kuskal-Wallis rank test for the null hypothesis that the residual distribution median values classified according to different schemes are equal (see main text for more details on classification)

	Topographic classification KW	Topographic classification *p*-value	Vs,30-based classification KW	Vs,30-based classification *p*-value	geological classification KW	geological classification *p*-value
AF_1	23.42	<<0.1%	83.14	<<0.1%	76.70	<<0.1%
AF_2	34.59	<<0.1%	87.02	<<0.1%	49.30	<<0.1%
AF_3	58.61	<<0.1%	105.54	<<0.1%	61.14	<<0.1%

Consequently, additional statistical testing is necessary to distinguish which are the residual distributions that significantly differ within each classification scheme.

For AF_1, when distinguishing residual values by topographical classification, the distributions showed similar median values and variance, with 84% of sites having negative residuals (Fig. 10, top-left panel). The Conover posterior statistical test revealed significant differences only between the distribution pairs T1-T2 and T1-T3 (Fig. 10, top-right panel). These results suggest that the mismatch between 1D simulation and experimental AF_1 estimates is lower for sites lying on flat topographies (T1) with respect to sites on slopes (T2, T3) but similar to that for sites lying of very inclined slopes; therefore, it is difficult interpreting the mismatch as a function of surface morphological features.

In terms of simplified geological classification, clear distinctions can be observed between sites on clays, for which the distribution showed positive residuals in more than 50% of cases, and all other classes, which showed predominantly negative residuals (Fig. 10, bottom panels). This suggests that 1D simulations underestimate empirical AF_1 for sites on predominant clayey soils and overestimate empirical data for sites in other classes. The posterior statistical test revealed significantly lower median residuals for sites on rock (or altered rock) compared to those on loose deposits (sands and gravels, Fig. 10, bottom-right panel). A similar consideration follows from observing residual distributions as a function of EC8 soil classification, which includes information on both geophysical and geological properties of the subsurface. EC8 classes A and B, corresponding to stiffer sites with higher *Vs,30*, showed negative medians with values significantly lower than classes C and E, with lower *Vs,30* (Fig. 10, intermediate panels). This suggests that 1D simulations systematically overpredict AF_1 for classes A and B, which represent the largest fraction of sites (about 70%). At longer periods, distributions tend to shift toward higher residual values but conclusions similar to those for the short period intervals can be drawn from the residual analyses for AF_2 and AF_3 (Figures S4 and S5 in the supplementary material).

These results suggest that EC8 classification and basic geological information have some explanatory capability of the observed residuals: on average, for soft soil sites, 1D simulations underestimate empirical *AFs*, particularly at long periods, whereas they overestimate observed *AFs* for rock sites, even if a large variability is observed for class B sites.

Although a wide dispersion in the dataset is observed, particularly for AF_1, analysis of residuals as a function of *Vs,30* further supports this interpretation, revealing a weak negative correlation, namely a trend of *AFs* overprediction by 1D simulations for sites with high *Vs,30* and underprediction for sites with lower *Vs,30*. This correlation is slightly more pronounced for AF_3 (Fig. 11).

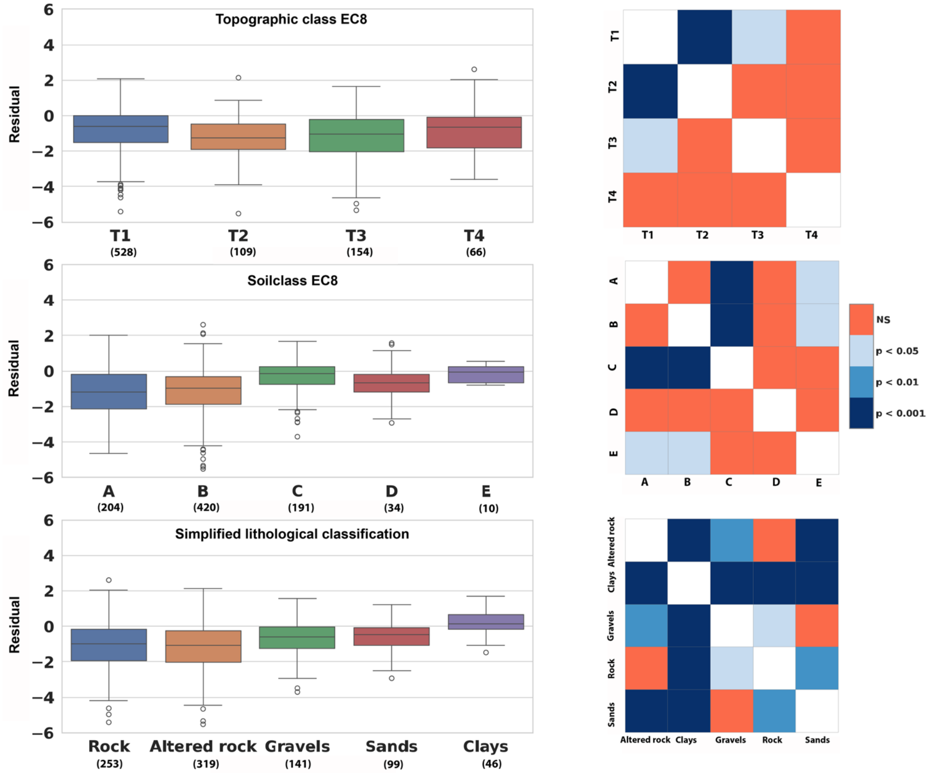

Fig. 10 Residual distributions calculated for AF_1 based on three different site classification schemes (left column): NTC18 topographic classification (top), a simplified geological classification adapted from Bucci et al. (2022; middle), and EC8 and NTC18 *Vs,30*-based classification (bottom). The number of data points for each distribution is shown in parentheses along the x-axis. Heatmap matrices showing the statistical significance of posterior pairwise comparisons among the residual distributions for each classification scheme (right column). The color code indicates conventional statistical significance levels based on *p-values*, ranging from 0.1% to 5% (blue colors, indicating significance) or *p-values* > 5% (red color, indicating non-significance)

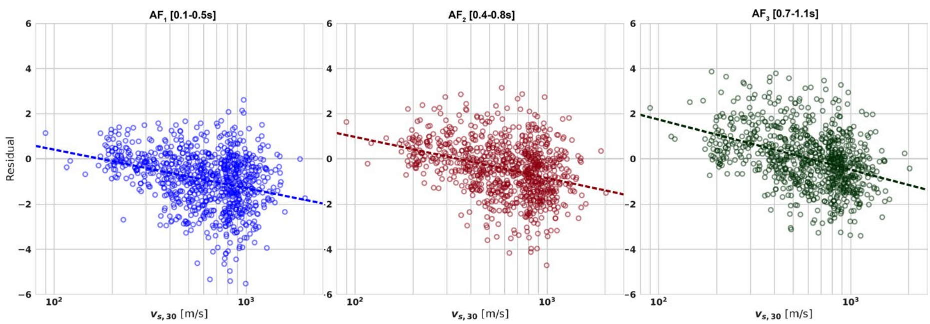

Fig. 11 Scatter plot of AF_1 (left panel), AF_2 (middle panel) and AF_3 (right panel) residuals as a function of *Vs,30*. In each panel, the dashed line of the linear fit between residual and log *Vs,30* is also shown

5 Discussion and conclusions

In this study, we established a quantitative framework of *GoF* metrics and statistical tests for evaluating the consistency between period-dependent site amplification factor estimates (*AFs*) obtained from different empirical methods and with those from a simplified 1D simulation approach, which is commonly employed for considering site effects in seismic hazard assessment. The statistics is based on a large number of Italian sites up to 925 sites, depending on the methods.

Regarding the empirical estimates, our analyses revealed a significant correlation (Spearman's coefficient ρ generally larger than 0.9) between *AFs* from different methodologies, namely *δS2S* (Lanzano et al. 2025), *GIT* (Priolo et al. 2020; Klin et al. 2021) and *SSR* (Priolo et al. 2020; Luzi et al. 2019), applied to various ground-motion datasets from temporary and permanent seismic networks. This implies that each empirical method allows for obtaining a measure of the difference in the *AFs* level between pairs of sites which is in close agreement with the other methods, as also shown by previous studies (Parolai et al. 2004; Pilz et al. 2009; Bindi and Kotha 2020).

However, in Central Italy, empirical estimates from $\delta S2S_{800}$ and GIT_{Sa} methods exhibited systematic deviations. On average, values differed by up to 50% in terms of AF_1 and up to 30% in terms of AF_3. The primary cause of the mismatch is the selection of the reference ground-motion level, while other factors may potentially contribute to these systematic deviations:

i) Magnitude and source-to-site distance distributions of the analyzed earthquake catalogue, which control the calibration of source, attenuation and site terms in *GMM* (Baltay et al. 2017) and *GIT* schemes (Shible et al. 2022);
ii) Path-dependent regional attenuation, which may be difficult to constrain with an insufficient sample of source-to-site paths, so that this unmodelled effect may bias the source and site term estimates (Baltay et al. 2020);
iii) Nonlinear site effects, which induce a decrease of ground-motion amplitude and a shift of the site resonance frequency towards lower frequencies, so that the linear site term may be biased (Bradley 2015).

Excluding the contribution of soil nonlinear behavior, which is negligible in the Italian dataset used in this study (Lanzano et al. 2025), our results support the hypothesis of the main role of the reference ground-motion in determining systematic deviations between empirical *AF* estimates. In fact, the application of the generic rock-to-reference scaling factor to the *δS2S* estimates compensates for the observed systematic underestimation of GIT_{Sa} in Central Italy. Such corrective term accounts for the period-dependent site amplification included in the ITA18-*GMM* calibrated for the engineering bedrock condition, namely for EC-8 site class A site ($V_{S,30} \geq 800$ m/s), with respect to which the residual decomposition is calculated (Lanzano et al. 2025).

Empirical estimates are therefore particularly suitable for determining site amplification, owing to their high correlation and the potential to correct for any systematic difference by scaling them to a common reference ground-motion level. However, in the site-specific seismic hazard practice, the site amplification is typically accounted for by performing 1D simulations. This is the case for Italy, where 1D equivalent-linear simulations have been extensively used to develop *AF* models at regional and national levels (Falcone et al. 2021; Pieruccini et al. 2024; Albarello et al. 2025). Hence, to improve seismic hazard assessment, comparing simulation-based estimates with those from experimental methods is useful for evaluating how effectively 1D simulations capture the observed site amplification.

Our results showed that the 1D simulations cannot reproduce the range of observed median *AFs* estimates. It is worth noting that the latter have significantly larger uncertainties (standard deviations) than the former, due to several sources of variability which are not taken fully into account in 1D simulations, such as *(i)* the variability in the azimuth and incidence angle of the wavefield embedded in the recordings used to calculate *AFs* (Maufroy et al. 2017), *(ii)* the spatial variability of topography and subsurface structure at the site of interest (Matsushima et al. 2017) and *(iii)* the temporal variability of the site properties induced by environmental factors, such as temperature and soil saturation (Roumelioti et al. 2019; Bonilla and Ben-Zion 2021). Recent studies have also shown that the standard deviation of empirical site response estimated from local low-magnitude events is particularly higher for stiffer sites and for those sites characterized by rough topography, which represent the majority of sites considered within this study (Zhu et al. 2021).

Overall, numerical and empirical estimates are uncorrelated for AF_1, and only weakly correlated in the intermediate- and long-period bands (AF_2 and AF_3). Bearing in mind that the 1D simulation-based *AF* estimates are averaged from an ensemble of 1D models where parameters were randomized to reflect the subsurface variability of a wider area, the mismatch and the lower variability of these simulations may be partly driven by this randomization and averaging effect. In fact, this generally reduces the computed site amplification (Rathje et al. 2010), particularly at short periods and for higher-amplitude input motions (Pehlivan et al. 2015), due to the more smoothed and attenuated shape of the median theoretical transfer function calculated from the multiple 1D randomized models with respect to the site-specific 1D subsurface model (Kaklamanos et al. 2020).

Another potential contribution for the bias between empirical and 1D simulation estimates is related to soil nonlinear behavior, which causes a reduction of site amplification. While the analyses of site-specific proxies for soil nonlinearity on worldwide earthquake databases highlighted its limited impact on recorded motions (Guéguen et al. 2019), the comparison of observed and predicted ground motion at Japanese KiK-net sites has shown that 1D equivalent-linear models become inaccurate in predicting short-period site amplification for induced maximum shear strain in the soil between 0.1% and 0.4%; at longer periods (T>0.5s) the accuracy of linear-equivalent and nonlinear 1D simulations is similar (Kaklamanos et al. 2018). Assumed peak ground acceleration (PGA) thresholds for induced soil nonlinear behavior are in the range 0.1 to 0.3 g (Thompson et al. 2012; Kaklamanos et al. 2013; 2018). Thus, considering the reference hazard level for the *AFs* model by Falcone et al. (2021), we expect a limited impact of soil nonlinear behaviour, which should cause underestimation of AF_1 values from 1D simulations in those areas characterized by PGA with 10% PoE in 50 years > 0.2–0.3 g, namely the central Apennines, Eastern Alps and Eastern Sicily. This is not the case for the comparison with the empirical estimates, for which we observed a general overestimation of AF_1 from 1D simulations (Fig. 9, left panel).

From the analysis of the residuals between the empirical and 1D simulation *AF* estimates, emerges a satisfactory agreement (residuals around zero) for 30–35% of sites, varying as a function of the considered period band, suggesting that the 1D simulations provide a good approximation of observed *AFs* for one third of the dataset (Figure S6 in the supplementary material). Such percentages are consistent with findings from similar studies (Thompson et al. 2012; Pilz and Cotton 2019).

The estimates disagree for the remaining 65 to 70% of sites. In particular, the larger group of rock and stiff sites (30 to 60% depending on the considered period band) are characterized by negative residuals: they belong to EC8 classes A and B (with higher *Vs,30*) and are distributed along the Apennine belt and in the Tyrrhenian side of the Italian peninsula. These results suggest that for rock and stiff site conditions, 1D simulations may provide conservative estimates of *AFs* for seismic hazard assessment. For these classes of sites, for which a limited impact of soil nonlinearity is expected, a reason different from the effect of randomization on 1D simulations should be taken into account for explaining the mismatch between estimates, particularly at short periods, such as inaccurate and/or oversimplified 1D models (Kaklamanos and Bradley 2018; Li et al. 2018). In particular, it has been demonstrated that the near surface attenuation (i.e. κ_0), which is often poorly known for stiff/rock sites, has a strong influence on short periods site amplification (Douglas et al. 2009).

On the contrary, for a smaller group of soft soil sites (5 to 15% depending on the considered period band) we observed positive residuals, implying that 1D simulations underestimate empirical *AFs*. Even if this observation may be related to the reduction of simulation-based *AFs* due to 1D model parameter randomization and soil nonlinearity, it is worth noting that these sites are located in the central and eastern side of the peninsula and in eastern Sicily, often within deep basins, like the Po plain, the Fucino basin and the Adriatic foredeep. Therefore, we suggest that the mismatch is likely due to the limitations of 1D models in accounting for more complex site effects, such as basin edge generated surface waves, 2D and 3D resonances, and it is consistent with the studies on the Po Plain, which showed a significant role of 3D basin amplification in the observed ground motion and in the regional seismic hazard assessment (Mascandola et al. 2021a, b, 2023). In similar cases, an empirical determination of *AFs* is more suitable for seismic hazard assessment. However, it is also worth underscoring that the relatively small percentage of sites exhibiting such underestimation is primarily driven by the dataset spatial distribution, which is inherently underrepresented in terms of stations lying on soft sites within sedimentary basins. The lack of a sufficiently statistical sample for these specific site conditions also hinders the possibility of calibrating an average statistically robust correction factor for the *AFs* from 1D simulations based on empirical data, which in principle could lead to the development of a regional site amplification model bridging the gap between the two approaches. Consequently, more thorough future analyses leveraging a large, targeted dataset specifically focused on such site conditions are needed to properly address the potential limitations of 1D simulations in accounting for complex 3D site effects.

In conclusion, our analyses showed that empirical estimates are particularly suitable for determining *AFs* in a seismic hazard perspective, whereas 1D simulations may underestimate *AFs* in areas characterized by complex site effects, such as large deep basins. Therefore, national or regional site-amplification models that combine empirical estimates with proxy-based spatial interpolation should be preferred over 1D simulations. Nevertheless, since the empirical dataset, albeit large, is non-uniformly sampling all the Italian territory, this approach will require a large number of recordings and sites, covering diverse soil conditions and geographical areas, and a set of easily obtainable site-amplification proxies for the spatial extrapolation of single-point empirical estimates.

In particular, the selection of reliable site proxies for spatial extrapolation remains an active area of research (Cultrera et al. 2025; Bergamo et al. 2023). Traditional indirect proxies, such as geological classifications, soil classes, or topographic parameters (e.g., slope and curvature) used for first-order approximations of $Vs,30$ often exhibit low reliability when applied to geographical areas different from those they were originally developed upon (Bergamo et al. 2021 and references therein). To overcome such limitations, recent literature suggests shifting toward more robust direct or advanced indicator frameworks (Cultrera et al. 2021; Böse et al. 2024). These include vector valued indicators, such as frequency-dependent quarter-wavelength parameters (time-averaged shear-wave velocity and seismic impedance contrast down to depth of one-fourth of the corresponding wavelengths; Bergamo et al. 2020) or set of scalar proxies (e.g. $Vs,30$-bedrock depth proxy), particularly those incorporating the site fundamental frequency of resonance (f_0), which alone can capture a major component of site effects (Zhu et al. 2020). Such novel framework has been proposed for the Second Generation of Eurocode 8 (FprEN 1998-1-1:2024), which considers $Vs,30$, H_{800} (depth beyond which the shear-wave velocity exceeds 800 m/s) and f_0 as reliable site amplification proxies.

In Italy, attempts to derive site amplification models for vast-areas based on observed data were proposed by Sgobba et al. (2024), who combined empirical measurements and a $Vs,30$ map to develop a site amplification model for Central Italy, and by Laurenzano et al. (2023), who used dense seismic noise data as a proxy for site amplification in an Alpine valley. At a broader scale, Kawase et al. (2018) demonstrated the feasibility of a novel approach for converting seismic noise Horizontal-to-Vertical spectral ratio into site amplification functions, offering a potential pathway for the development of consistent national-scale amplification models. Nevertheless, the development of vast-area site amplification models still needs further research on reliable site amplification proxies to ensure that regional models accurately capture the full spectrum of locally observed amplification effects. Future research should therefore focus on enhancing the diversity of local conditions represented by empirical site amplification data, and on exploring a broader range of suitable site amplification proxies and spatial interpolation methods, including Machine Learning algorithms.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10518-026-02638-y.

Acknowledgements The authors acknowledge the financial support in the framework of the National project PRIN-SERENA (Mapping Seismic Site Effects at REgional and NAtional Scale, MUR-PRIN Italian project 2023–2025, Prot. 2020MMCPER). The authors would like to thank Prof. Dario Albarello for the fruitful discussions which helped improve this study. The authors also thank three anonymous reviewers for their constructive comments which helped improve the quality of this paper. The support of Dr. Luca Minarelli and Dr. Giulio Brunelli during the project is gratefully appreciated.

Author contributions Conceptualization: Salomon Hailemikael, Giovanna Cultrera, Giovanni Lanzano, Carla Barnaba, Giovanna Laurenzano, Maria Rosaria Gallipoli, Francesca Pacor; Methodology: Salomon Hailemikael, Giovanna Cultrera; Software: Salomon Hailemikael, Alessandro Peloso, Guido Martini, Giovanna Cultrera; Validation: Salomon Hailemikael, Alessandro Peloso, Guido Martini; Formal analysis: Salomon Hailemikael; Visualization: Salomon Hailemikael, Giovanna Cultrera, Giovanni Lanzano, Alessandro Peloso, Guido Martini; Resources: Giovanni Lanzano, Carla Barnaba, Giovanna Laurenzano, Maria Rosaria Gallipoli, Alessandro Peloso, Guido Martini, Sara Sgobba, Francesca Pacor; Data Curation: Giovanna Cultrera, Giovanni Lanzano, Carla Barnaba, Giovanna Laurenzano, Maria Rosaria Gallipoli, Alessandro Peloso, Guido Martini, Sara Sgobba; Writing - Original Draft: Salomon Hailemikael; Writing - Review & Editing: Salomon Hailemikael, Giovanna Cultrera, Giovanni Lanzano, Carla Barnaba, Giovanna

Laurenzano, Maria Rosaria Gallipoli, Francesca Pacor; Supervision: Giovanna Cultrera; Project administration: Giovanna Cultrera; Funding acquisition: Giovanna Cultrera.

Funding Open access funding provided by Istituto Nazionale di Geofisica e Vulcanologia within the CRUI-CARE Agreement. This work has been financially supported by the National project PRIN-SERENA (Mapping Seismic Site Effects at REgional and NAtional Scale, MUR-PRIN Italian project 2023–2025, Prot.2020MMCPER).

Data availability The datasets generated during and/or analyzed in the current study are available from the corresponding author upon request.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

Acunzo G, Falcone G, di Lernia A et al (2023) NC92Soil: a computer code for deterministic and stochastic 1D equivalent linear seismic site response analyses. Comput Geotech 165:105857. https://doi.org/10.1016/j.compgeo.2023.105857

Afshari K, Stewart JP (2019) Insights from California vertical arrays on the effectiveness of ground response analysis with alternative damping models. Bull Seismol Soc Am 109(4):1250–1264. https://doi.org/10.1785/0120180292

Al Atik L, Abrahamson N, Bommer JJ et al (2010) The variability of ground-motion prediction models and its components. Seismol Res Lett 81(5):794–801. https://doi.org/10.1785/gssrl.81.5.794

Albarello D, Carfagna N, Fantozzi P (2025) The impact of 1D seismostratigraphical amplification effects on probabilistic seismic hazard maps at regional scale: the case of Tuscany (Central Italy). Bull Earthq Eng 23:3043–3055. https://doi.org/10.1007/s10518-025-02167-0

Ameri G, Oth A, Pilz M et al (2011) Separation of source and site effects by generalized inversion technique using the aftershock recordings of the 2009 L'Aquila earthquake. Bull Earthq Eng 9:717–739. https://doi.org/10.1007/s10518-011-9248-4

Anderson JG (2007) Physical processes that control strong ground motion. Treatise on Geophysics. Earthquake Seismology, Elsevier, pp 513–565

Anderson JG, Hough SE (1984) A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. Bull Seismol Soc Am 74(5):1969–1993

Andrews DJ (1986) Objective determination of source parameters and similarity of earthquakes of different size. In: Das S, Boatwright J, Scholz CH (eds) Earthquake Source Mechanics, Geophysical Monograph Series, vol. 37. American Geophysical Union, Washington, DC, pp. 259–267. https://doi.org/10.1029/GM037p0259

Baltay AS, Hanks TC, Abrahamson NA (2017) Uncertainty, variability, and earthquake physics in ground-motion prediction equations. Bull Seismol Soc Am 107:1754–1772. https://doi.org/10.1785/0120160164

Baltay AS, Abrahams LS, Hanks TC (2020) When source and path components tradeoff in Ground-Motion prediction equations. Seismol Res Lett 91:2259–2267. https://doi.org/10.1785/0220190379

Bergamo P, Hammer C, Fäh D (2020) On the Relation between Empirical Amplification and Proxies Measured at Swiss and Japanese Stations: Systematic Regression Analysis and Neural Network Prediction of Amplification. Bull Seismol Soc Am 111(1):101–120. https://doi.org/10.1785/0120200228

Bergamo P, Hammer C, Fäh D (2021) 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 112(2):1008–1030. https://doi.org/10.1785/0120210225

Bergamo P, Fäh D, Panzera F et al (2023) A site amplification model for Switzerland based on site-condition indicators and incorporating local response as measured at seismic stations. Bull Earthq Eng 21(13):5831–5865. https://doi.org/10.1007/s10518-023-01766-z

Bindi D, Kotha SR (2020) Spectral decomposition of the engineering strong motion (ESM) flat file: regional attenuation, source scaling and Arias stress drop. Bull Earthq Eng 18:2581–2606. https://doi.org/10.1007/s10518-020-00796-1

Bindi D, Massa M, Luzi L et al (2014) Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset. Bull Earthq Eng 12:391–430. https://doi.org/10.1007/s10518-013-9525-5

Böse M, Danciu L, Papadopoulos A, Clinton J, Cauzzi C, Dallo I, Mizrahi L, Diehl T, Bergamo P, Reuland Y, Fichtner A, Roth P, Haslinger F, Massin F, Valenzuela N, Blagojević N, Bodenmann L, Chatzi E, Fäh D, Glueer F, Han M, Heiniger L, Janusz P, Jozinović D, Kästli P, Lanza F, Lee T, Martakis P, Marti M, Meier M-A, Mena Cabrera B, Mesimeri M, Obermann A, Sanchez-Pastor P, Scarabello L, Schmid N, Shynkarenko A, Stojadinović B, Giardini D, Wiemer S (2024) Towards a dynamic earthquake risk framework for Switzerland. Nat Hazards Earth Syst Sci 24:583–607. https://doi.org/10.5194/nhess-24-583-2024

Boatwright J, Fletcher JB, Fumal TE (1991) A general inversion scheme for source, site, and propagation characteristics using multiple recorded sets of moderate-sized earthquakes. Bull Seismol Soc Am 81(5):1754–1782. https://doi.org/10.1785/BSSA0810051754

Bonilla LF, Ben-Zion Y (2021) Detailed space–time variations of the seismic response of the shallow crust to small earthquakes from analysis of dense array data. Geophys J Int 225:298–310

Borcherdt RD (1970) Effects of local geology on ground motion near San Francisco Bay. Bull Seismol Soc Am 60:29–61. https://doi.org/10.1785/BSSA0600010029

Bradley BA (2012) A ground motion selection algorithm based on the generalized conditional intensity measure approach. Soil Dyn Earthq Eng 40:48–61. https://doi.org/10.1016/j.soildyn.2012.04.007

Bradley BA (2015) Systematic ground motion observations in the Canterbury earthquakes and region-specific non-ergodic empirical ground motion modeling. Earthq Spectra 31(3):1735–1761. https://doi.org/10.1193/053013EQS137M

Bucci F, Santangelo M, Fongo L et al (2022) A new digital lithological map of Italy at the 1:100 000 scale for geomechanical modelling. Earth Syst Sci Data 14:4129–4151. https://doi.org/10.5194/essd-14-4129-2022

Cadet H, Bard P-Y, Rodriguez-Marek A (2010) Defining a standard rock site: Propositions based on the KiK-net database. Bull Seismol Soc Am 100(1):172–195

Cara F, Cultrera G, Riccio G et al (2019) Temporary dense seismic network during the 2016 Central Italy seismic emergency for microzonation studies. Sci Data 6:182. https://doi.org/10.1038/s41597-019-0188-1

Castellaro S, Mulargia F, Rossi PL (2008) VS30: Proxy for seismic amplification? Seismol. Res Lett 79(4):540–543. https://doi.org/10.1785/gssrl.79.4.540

Castro RR, Anderson JG, Singh SK (1990) Site response, attenuation and source spectra of S waves along the Guerrero, Mexico, subduction zone. Bull Seismol Soc Am 80:1481–1503

CEN (2005) EN 1998-1:2005 Eurocode 8: Design of structures for earthquake resistance – Part 1: General rules, seismic actions and rules for buildings. European Committee for Standardization (CEN), Brussels

Chandramohan R, Baker JW, Deierlein GG (2016) Impact of hazard-consistent ground motion duration in structural collapse risk assessment. Earthq Eng Struct Dyn 00:1–22. https://doi.org/10.1002/eqe.2711

Conover WJ, Iman RL (1979) On multiple-comparisons procedures. Tech Rep LA-7677-MS, Los Alamos Scientific Laboratory

Crowley H, Weatherill G, Riga E et al (2019) SERA deliverable D26.4: Methods for estimating site effects in risk assessments. Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe. http://static.seismo.ethz.ch/SERA/JRA/SERA_D26.4_%20Site_Amplification.pdf. Accessed Dec 2025

Cultrera G, Mercuri A (2025) Statistical analysis of site effect indicators at the Italian seismic network: inside the site characterization database CRISP. Bull Earthq Eng 23:4365–4390. https://doi.org/10.1007/s10518-025-02212-y

Cultrera G, Mucciarelli M, Parolai S (2011) The L'Aquila earthquake–a view of site effects and building behavior from temporary networks. Bull Earthq Eng 9(3):691–695. https://doi.org/10.1007/s10518-011-9270-6

Cultrera G, Cornou C, Di Giulio G (2021) Indicators for site characterization at seismic station: recommendation from a dedicated survey. Bull Earthq Eng 19:4171–4195. https://doi.org/10.1007/s10518-021-01136-7

Douglas J, Gehl P, Bonilla LF, Scotti O, Régnier J, Duval AM, Bertrand E (2009) Making the most of available site information for empirical ground-motion prediction. Bull Seismol Soc Am 99(3):1502–1520
Falcone G, Acunzo G, Mendicelli A et al (2021) Seismic amplification maps of Italy based on site-specific microzonation dataset and one-dimensional numerical approach. Eng Geol 289:106170. https://doi.org/10.1016/j.enggeo.2021.106170
Felicetta C, Mascandola C, Spallarossa D et al (2021) Quantification of site effects in the Amatrice area (Central Italy): Insights from ground-motion recordings of the 2016–2017 seismic sequence. Soil Dyn Earthq Eng 142:106565. https://doi.org/10.1016/j.soildyn.2020.106565
Gallipoli M (2003) Temporary Accelerometric Network in Basilicata region for Site Effect studies [Data set]. Int Federation Digit Seismogr Networks. https://doi.org/10.7914/nhc9-gg46
Grendas I, Hollender F, Perron V et al (2025) The SSRc method as a practical solution for evaluating site response using far apart stations: applications in Western Greece and South-Eastern France. Bull Earthq Eng 23:5341–5377. https://doi.org/10.1007/s10518-025-02221-x
Guéguen P, Bonilla LF, Douglas J (2019) Comparison of soil nonlinearity (in situ stress-strain relation and G/Gmax reduction) observed in strong-motion databases and modeled in ground-motion prediction equations. Bull Seismol Soc Am 109(1):178–186. https://doi.org/10.1785/0120180169
Hallal MM, Cox BR, Vantasel JP (2022) Comparison of state-of-the-art approaches used to account for spatial variability in 1D ground response analyses. J Geotech GeoEnviron Eng 148(5):04022019
Hartzell S, Cranswick E, Frankel A et al (1997) Variability of site response in the Los Angeles urban area. Bull Seismol Soc Am 87:1377–1400. https://doi.org/10.1785/BSSA0870061377
Kaklamanos J, Bradley BA (2018) Challenges in predicting seismic site response with 1D analyses: Conclusions from 114 KiK-net vertical seismometer arrays. Bull Seismol Soc Am 108:2816–2838
Kaklamanos J, Bradley BA, Thompson EM, Baise LG (2013) Critical parameters affecting bias and variability in site-response analyses using KiK-net downhole array data. Bull Seismol Soc Am 103(3):1733–1749. https://doi.org/10.1785/0120120166
Kaklamanos J, Bradley BA, Moolacattu AN, Picard BM (2020) Physical hypotheses for adjusting coarse profiles and improving 1D site-response estimation assessed at 10 KiK-net sites. Bull Seismol Soc Am 110:1338–1358. https://doi.org/10.1785/0120190263
Kawase H, Nagashima F, Nakano K, Mori Y (2018) Direct evaluation of S-wave amplification factors from microtremor H/V ratios: Double empirical correction to "Nakamura" method. Soil Dyn Earthq Eng 126:105067
Kendall MG (1970) Rank Correlation Methods, 4th edn. Charles Griffin & Co, London
Klin P, Laurenzano G, Priolo E (2018) GITANES: A MATLAB package for estimation of site spectral amplification with the generalized inversion technique. Seismol Res Lett 89(1):182–190. https://doi.org/10.1785/0220170080
Klin P, Laurenzano G, Barnaba C et al (2021) Site Amplification at Permanent Stations in Northeastern Italy. Bull Seismol Soc Am 111(4):1885–1904. https://doi.org/10.1785/0120200361
Kotha SR, Cotton F, Bindi D (2018) A new approach to site classification: Mixed-effects ground motion prediction equation with spectral clustering of site amplification functions. Soil Dyn Earthq Eng 110:318–329. https://doi.org/10.1016/j.soildyn.2018.01.051
Kottke A, Wang X, Rathje EM (2013) Technical manual for Strata. https://peer.berkeley.edu/sites/default/files/web_peer810_albert_r._kottke_ellen_m._rathje.pdf. Accessed Nov 2025
Kramer SL, Stewart JP (2025) Geotechnical earthquake engineering, 2nd edn. CRC, Abingdon
Kruskal WH, Wallis WA (1952) Use of ranks in one-criterion variance analysis. J Am Stat Assoc 47(260):583–621. https://doi.org/10.1080/01621459.1952.10483441
Lanzano G, Luzi L, Pacor F et al (2019) A revised ground motion prediction model for shallow crustal earthquakes in Italy. Bull Seismol Soc Am 109(2):525–540. https://doi.org/10.1785/0120180210
Lanzano G, Felicetta G, Pacor F et al (2020) Methodology to identify the reference rock sites in regions of medium-to-high seismicity: an application in Central Italy. Geophys J Int 222(3):2053–2067. https://doi.org/10.1093/gji/ggaa261
Lanzano G, Felicetta C, Pacor F et al (2022) Generic-to-reference rock scaling factors for seismic ground motion in Italy. Bull Seismol Soc Am 112:1583–1606. https://doi.org/10.1785/0120210063
Lanzano G, Brunelli G, Sgobba S, Pacor F (2025) Estimation of the seismic microzonation amplification factors from the GMMs site-to-site residuals: an application to the Italian strong-motion stations in ITACA 4.0. Bull Earthq Eng 23:5229–5257. https://doi.org/10.1007/s10518-025-02249-z
Laurendeau A, Bard PY, Hollender F, Perron V, Foundotos L, Ktenidou OJ, Hernandez B (2018) Derivation of consistent hard rock (1000 < VS < 3000 m/s) GMPEs from surface and down-hole recordings: Analysis of KiK-net data. Bull Earthq Eng 16:2253–2284
Laurenzano G, Garbin M, Parolai S et al (2023) High-resolution local seismic zonation by cluster and correlation analysis. Soil Dyn Earthq Eng 173:108122. https://doi.org/10.1016/j.soildyn.2023.108122

Li G, Motamed R, Dickenson S (2018) Evaluation of one-dimensional multi-directional site response analyses using geotechnical downhole array data in California and Japan. Earthq Spectra 34:349–376

Luzi L, Bindi D, Puglia R et al (2014) Single-station Sigma for Italian strong-motion stations. Bull Seismol Soc Am 104(1):467–483. https://doi.org/10.1785/0120130089

Luzi L, D'Amico M, Massa M, Puglia R (2019) Site effects observed in the Norcia intermountain basin (Central Italy) exploiting a 20-year monitoring. Bull Earthq Eng 17:97–118. https://doi.org/10.1007/s10518-018-0444-3

Luzi L, Pacor F, Lanzano G et al (2020) 2016–2017 Central Italy seismic sequence: strong-motion data analysis and design earthquake selection for seismic microzonation purposes. Bull Earthq Eng 18:5533–5551. https://doi.org/10.1007/s10518-019-00676-3

Mascandola C, Luzi L, Felicetta C, Pacor F (2021a) A GIS procedure for the topographic classification of Italy, according to the seismic code provisions. Soil Dyn Earthq Eng 148:106848. https://doi.org/10.1016/j.soildyn.2021.106848

Mascandola C, Barani S, Massa M, Albarello D (2021b) New Insights into Long-Period (>1 s) Seismic Amplification Effects in Deep Sedimentary Basins: A Case of the Po Plain Basin of Northern Italy. Bull Seismol Soc Am 111:2071–2086. https://doi.org/10.1785/0120200315

Mascandola C, Barani S, Albarello D (2023) Impact of site-response characterization on probabilistic seismic hazard in the Po plain (Italy). Bull Seismol Soc Am 113(3):1269–1285. https://doi.org/10.1785/0120220177

Matsushima S, Kosaka H, Kawase H (2017) Directionally dependent horizontal-to-vertical spectral ratios of microtremors at Onahama, Fukushima, Japan. Earth Planet Space 69:96

Maufroy E et al (2017) Source-related variability of site response in the Mygdonian basin (Greece) from accelerometric recordings and 3D numerical simulations. Bull Seism Soc Am 107:787–808

Mayoral JM, Assimaki D, Tepalcapa S et al (2019) Site effects in Mexico City basin: past and present. Soil Dyn Earthq Eng 121:369–382. https://doi.org/10.1016/j.soildyn.2019.02.028

MIT (2018) Norme tecniche per le costruzioni. Decreto Ministeriale 17 gennaio 2018, Gazzetta Ufficiale della Repubblica Italiana n. 42 del 20 febbraio 2018, S.O

Mori F, Mendicelli A, Moscatelli M et al (2020) 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

Moscatelli M, Albarello D, Scarascia Mugnozza G, Dolce M (2020) The Italian approach to seismic microzonation. Bull Earthq Eng 18:5425–5440. https://doi.org/10.1007/s10518-020-00856-6

NTC (2018) Ministero delle Infrastrutture e dei Trasporti - Norme tecniche per Le costruzioni. Decreto Ministeriale Del 17 Gennaio 2018, supplemento ordinario Alla Gazzetta ufficiale N. 42 del 20 febbraio 2018 (in italian).

Oth A, Bindi D, Parolai S, Giacomo DD (2011) Spectral analysis of K-NET and KiK-net data in Japan, part II: On attenuation characteristics, source spectra, and site response of borehole and surface stations. Bull Seismol Soc Am 101:667–687. https://doi.org/10.1785/0120100135

Pacor F, Spallarossa D, Oth A et al (2016) Spectral models for ground motion prediction in the L'Aquila region (Central Italy): evidence for stress-drop dependence on magnitude and depth. Geophys J Int 204:697–718. https://doi.org/10.1093/gji/ggv448

Paolucci E, Tanzini A, Peruzzi G et al (2020) Empirical testing of a simplified approach for the estimation of 1D litho-stratigraphical amplification factor. Bull Earthq Eng 18:1285–1301. https://doi.org/10.1007/s10518-019-00772-4

Parolai S, Bindi D, Baumbach M et al (2004) Comparison of different site response estimation techniques using aftershocks of the 1999 Izmit earthquake. Bull Seism Soc Am 94(3):1096–1108. https://doi.org/10.1785/0120030086

Pehlivan M, Hashash Y, Harmon J, Rathje EM, Stewart JP, Silva WJ, Campbell KW, Nikolaou S (2015) Influence of shear wave velocity reversals on one-dimensional site response of spatially varied profiles. Paper presented at the 6th International Conference on Earthquake Geotechnical Engineering, Christchurch, New Zealand, 1–4 November 2015, pp. 1–8

Pieruccini P, Fantozzi PL, Carfagna N et al (2024) Mapping 1D seismic amplification effects in the range of periods of engineering interest based on geological data. Eng Geol 341:107701. https://doi.org/10.1016/j.eggeo.2024.107701

Pilz M, Cotton F (2019) 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 35(2):883–905. https://doi.org/10.1193/050718eqs113m

Pilz M, Parolai S, Leyton F et al (2009) A comparison of site response techniques using earthquake data and ambient seismic noise analysis in the large urban areas of Santiago de Chile. Geophys J Int 178(2):713–728. https://doi.org/10.1111/j.1365-246X.2009.04195.x

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2001) Numerical Recipes in Fortran 77: The Art of Scientific Computing, 2nd edn. Cambridge University Press, Cambridge

Priolo E, Pacor F, Spallarossa D et al (2020) Seismological analyses of the seismic microzonation of 138 municipalities damaged by the 2016–2017 seismic sequence in central Italy. Bull Earthq Eng 18:5553–5593. https://doi.org/10.1007/s10518-019-00652-x
Rathje EM, Kottke AR, Trent WL (2010) Influence of input motion and site property variabilities on seismic site response analysis. J Geotech Geoenviron Eng 136:607–619
Rodriguez-Marek A, Montalva G, Cotton F, Bonilla F (2011) Analysis of Single-Station standard deviation using the KiK-net data. Bull Seismol Soc Am 101:31242–31258. https://doi.org/10.1785/0120100252
Roumelioti Z, Hollender F, Gueguen P (2019) Rainfall-induced variation of seismic waves velocity in soil and implications for soil response: what the ARGONET (Cephalonia, Greece) vertical array data reveal. Bull Seism Soc Am 110:441–451
Sabetta F, Fiorentino G, Bocchi F et al (2023) Influence of local site effects on seismic risk maps and ranking of Italian municipalities. Bull Earthq Eng 21:2441–2468. https://doi.org/10.1007/s10518-023-01619-9
Scherbaum F, Cotton F, Smit P (2004) On the use of response spectral-reference data for the selection and ranking of ground-motion models for seismic-hazard analysis in regions of moderate seismicity: The case of rock motion. Bull Seismol Soc Am 94(6):2164–2185. https://doi.org/10.1785/0120030147
Sgobba S, Felicetta C, Bortolotti T et al (2024) A Geostatistical modelling of empirical amplification functions and related site proxies for shaking scenarios in central Italy. Soil Dyn Earthq Eng 179:108496. https://doi.org/10.1016/j.soildyn.2024.108496
Shible H, Hollender F, Bindi D et al (2022) GITEC: A generalized inversion technique benchmark. Bull Seismol Soc Am 112:850–877. https://doi.org/10.1785/0120210242
SM Working Group (2008) Guidelines for seismic microzonation. Conference of Regions and Autonomous Provinces of Italy, Civil Protection Department, Rome (Original Italian Edition: Gruppo di lavoro MS, Indirizzi e criteri per la microzonazione sismica, Conferenza delle Regioni e delle Province autonome - Dipartimento della protezione civile, Roma, 2008, 3 vol. e Dvd). https://www.centromicrozonazionesismica.it/documents/18/GuidelinesForSeismicMicrozonation.pdf. Accessed 6 February 2026
Spearman C (1907) Demonstration of formulæ for true measurement of correlation. Am J Psychol 18:161–169. https://doi.org/10.2307/1412408
Steidl JH, Tumarkin AG, Archuleta RJ (1996) What is a reference site? Bull Seismol Soc Am 86(6):1733–1748
Stucchi M, Meletti C, Montaldo V et al (2011) Seismic hazard assessment (2003–2009) for the Italian Building code. Bull Seismol Soc Am 101(4):1885–1911. https://doi.org/10.1785/0120100130
Tao Y, Rathje E (2020) Taxonomy for evaluating the site-specific applicability of one-dimensional ground response analysis. Soil Dyn Earthq Eng 128:105865. https://doi.org/10.1016/j.soildyn.2019.105865
Terpilowski M (2019) scikit-posthocs: Pairwise multiple comparison tests in Python. J Open Source Softw 4(36):1169. https://doi.org/10.21105/joss.01169
Thompson EM, Baise LG, Tanaka Y, Kayen RE (2012) A taxonomy of site response complexity. Soil Dyn Earthq Eng 41:32–43. https://doi.org/10.1016/j.soildyn.2012.04.005
Thompson EM, Wald DJ, Worden CB (2014) A VS30 Map for California with Geological and Topographic Constraints. Bull Seism Soc Am 104:2313–2321. https://doi.org/10.1785/0120130312
van Ginkel J, Ruigrok E, Stafleu J, Herber R (2022) Development of a seismic site-response zonation map for the Netherlands. Nat Hazards Earth Syst Sci 22:41–63. https://doi.org/10.5194/nhess-22-41-2022
Vidakovic B (2011) Statistics for Bioengineering Sciences: with MATLAB and WinBUGS Support. Springer Science & Business Media, New York
Vilanova SP, Narciso J, Carvalho JP et al (2018) Developing a geologically based VS30 site-condition model for Portugal: methodology and assessment of the performance of Proxies. Bull Seismol Soc Am 108(1):322–337
Wald DJ, Allen TI (2007) Topographic slope as a proxy for seismic site conditions and amplification. Bull Seismol Soc Am 97(5):1379–1395. https://doi.org/10.1785/0120060267
Weatherill G, Crowley H, Roullé A et al (2022) Modelling site response at regional scale for the 2020 European Seismic Risk Model (ESRM20). Bull Earthq Eng 21:665–714. https://doi.org/10.1007/s10518-022-01526-5
Zhu C, Pilz M, Cotton F (2020) 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 18:797–820. https://doi.org/10.1007/s10518-019-00738-6
Zhu C, Cotton F, Kwak DY, Ji K, Kawase H, Pilz M (2021) Within-site variability in earthquake site response. Geophys J Int 229:1268–1281
Zhu C, Cotton F, Kawase H et al (2022) How well can we predict earthquake site response so far? Site-specific approaches. Earthq Spectra 38:1047–1075. https://doi.org/10.1177/87552930211060

Authors and Affiliations

Salomon Hailemikael[1] · **Giovanna Cultrera**[1] · **Giovanni Lanzano**[2] · **Carla Barnaba**[3] · **Giovanna Laurenzano**[3] · **Maria Rosaria Gallipoli**[4] · **Guido Martini**[1] · **Alessandro Peloso**[5] · **Sara Sgobba**[2] · **Francesca Pacor**[2]

✉ Salomon Hailemikael
salomon.hailemikael@ingv.it

[1] Istituto Nazionale di Geofisica e Vulcanologia (INGV), Sezione Roma1, Italy

[2] Istituto Nazionale di Geofisica e Vulcanologia (INGV), Sezione Milano, Italy

[3] Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS, Trieste, Italy

[4] Consiglio Nazionale delle Ricerche (CNR-IMIOT), Tito Scalo, Italy

[5] Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (ENEA), Frascati, Italy