Geophysical Journal International

Geophys. J. Int. (2023) **233,** 51–68
Advance Access publication 2022 November 11
GJI Seismology

https://doi.org/10.1093/gji/ggac445

Empirical correlations between an FAS non-ergodic ground motion model and a GIT derived model for Central Italy

P. Morasca [1], M. D'Amico [1], S. Sgobba [1], G. Lanzano [1], L. Colavitti [1], F. Pacor [1] and D. Spallarossa [1,2]

[1] *Istituto Nazionale di Geofisica e Vulcanologia, Milan, Italy. E-mail: maria.damico@ingv.it*
[2] *Istituto Nazionale di Geofisica e Vulcanologia, INGV, Via A. Corti 12, 20133, Milan, Italy. Dipartimento di Scienze della Terra dell'Ambiente e della Vita, Università di Genova, UNIGE, Viale Benedetto XV, 5, 16132, Genoa, Italy*

Accepted 2022 November 9. Received 2022 September 22; in original form 2022 May 17

SUMMARY

In this study, we investigate the correlation between the residuals of a neGMM (non-ergodic Ground Motion Model) and the physics-based parameters obtained using a non-parametric GIT approach (Generalized Inversion Technique) to lay the groundwork for the implementation of an *ad-hoc* FAS (Fourier Amplitude Spectra) neGMM for the Central Italy region. This region is particularly suitable for data-driven methodologies as those applied in this work because of the large amount of available data due to the recent multiple main shock–aftershock sequences occurred in this area. Both neGMM and GIT models are developed for Fourier spectra in the frequency range between 0.5 and 25 Hz and using the same reference sites. The comparison of the non-ergodic terms with the source, path and site spectral parameters provides interesting results. First, we find a strong correlation between the source parameters, stress drop $\Delta\sigma$ and decay k_{source} and the source neGMM corrective terms (the combination of the between-event δBe and the location-to-location terms δL2L). This correlation is frequency dependent and, at high frequency, is remarkably positive for $\Delta\sigma$ and negative for k_{source}. Concerning the attenuation terms, the path-to-path residuals (δP2P) are clearly associated with the deviations from the regional Q estimates obtained from GIT analysis. This indicates that the neGMM properly captures the properties of the anelastic attenuation and that the corrective terms δP2P can be used to account for differences in travel paths across different crustal domains. Finally, adopting the same reference sites for neGMM and non-parametric GIT, we observe that the systematic site terms (δS2S$_s$) and the GIT-derived amplification functions are in good agreement. The next step for an appropriate modelling is to identify the physical parameters (e.g. $V_{S,30}$ and k_0) describing the empirical amplification curves to be introduced as explanatory variables in the ground motion model.

Key words: Central Italy; Computational seismology; Earthquake ground motions; Generalized Inversion Technique; Seismic attenuation.

1 INTRODUCTION

One of the most addressed topics in engineering seismology concerns the adaptability of empirical ground motion models (GMMs) to describe different seismological contexts for a variety of applications, such as Probabilistic Seismic Hazard Analyses (PSHA) and engineering purposes. In common practice, the expected shaking level at a specific site is given by GMMs featured, at least, by earthquake magnitude, source-to-site distance and site condition *proxies*. These models are often calibrated over ground motion data sets of earthquake waveforms that are typically recorded in small regions characterized by high-seismicity rate for moderate to large earthquakes and then applied also at global scale under the ergodic assumption (Anderson & Brune 1999). As a consequence, GMMs predictions are affected by large variability that also reflects in the estimates of PSHA. Furthermore, despite major efforts made to increase the prediction capability of GMMs, through more sophisticated functional forms, just in a few cases it was attempted to correlate intensity measures of the ground shaking with physical parameters related to fault rupture, propagation path and site effects (Bora *et al.* 2015; Baltay *et al.* 2017; Bindi *et al.* 2018c; Sahakian *et al.* 2019).

In light of these critical aspects of the ground motion modelling, in this study we aim to investigate the physical connection between Fourier Amplitude Spectra (FAS) and physics-based parameters inferred by Generalized Inversion Technique (GIT, e.g. Oth *et al.* 2011) performed in the highly sampled region of Central Italy. Although engineering applications of predictive models often deal with acceleration

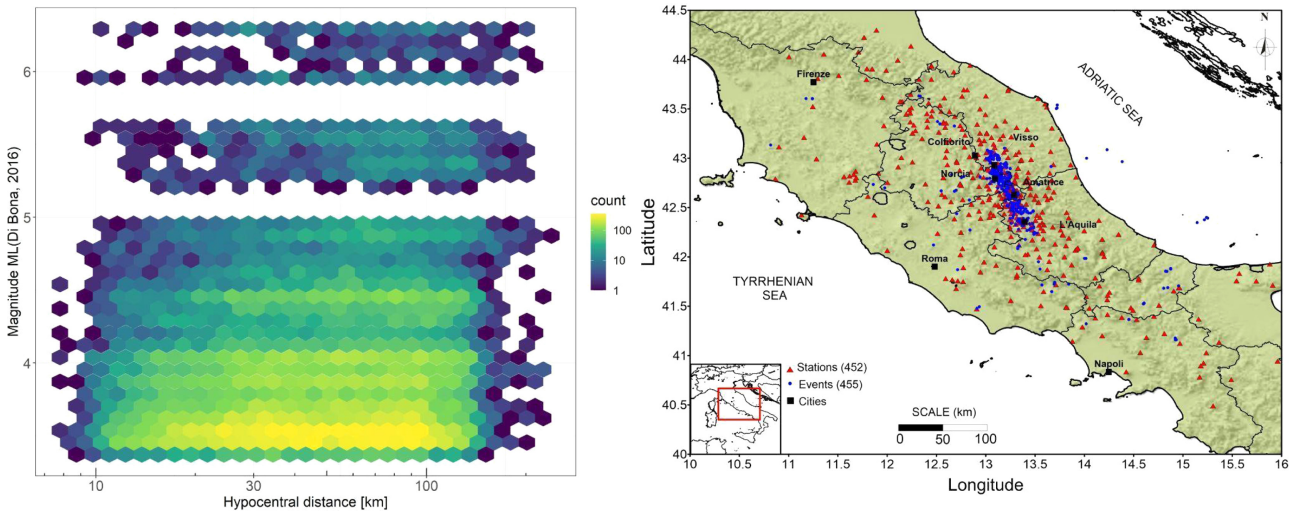

Figure 1. Left: magnitude–distance distribution as function of the number of recordings and right: maps showing the distribution of events and stations.

response spectra (or spectral amplitude, SA), we decided to use the FAS because the link between the seismological parameters and ground motion can be more easily explored in the Fourier domain (Bora *et al.* 2015; Bindi *et al.* 2018a, b). Moreover, the large amount of site-specific data in Central Italy after the 2016–2017 seismic sequence, gave us the opportunity to isolate the systematic contributions due to source, path and site effects calibrating an *ad-hoc* FAS model in a non-ergodic framework (FAS-neGMM). The adopted methodology to estimate these contributions is based on the residuals decomposition technique (Al-Atik *et al.* 2010) aiming to identify the systematic terms of variability into event-, source-, site- and path effects. Examples of fully non-ergodic models can be found in Lin *et al.* (2011), Baltay *et al.* (2017), Lanzano *et al.* (2017), Kuehn & Abrahamson (2019), Sgobba *et al.* (2019, 2021) and Parker *et al.* (2020). All these random effects act as adjustment terms of the median prediction, while moving part of the aleatory variability into epistemic uncertainty (Anderson & Brune 1999; Al-Atik *et al.* 2010; Anderson & Uchiyama 2011). In this framework, we develop an FAS-neGMM on 69 ordinates of the Fourier spectrum in the frequency range from 0.5 to 25 Hz. The model grounds on the same data-driven methodology and data set proposed by Sgobba *et al.* (2021) for predicting the peak ground acceleration (PGA) and the ordinates of the 5% damped-acceleration response spectra SA in the same range of vibration periods.

These systematic terms of uncertainty are then related to the physics-based parameters obtained from GIT analysis based on a non-parametric technique (Oth *et al.* 2011). The inversion analysis is performed starting from the same data set used for the FAS-neGMM and further selected for inversion robustness reasons as explained in the following sections. Furthermore, the same reference stations are used to constraint the site terms.

The correlation between the systematic contributions of the source, site and path variability of the FAS-neGMM with physically based GIT parameters allows to improve the accuracy of the model capturing more epistemic uncertainties. This is possible because the large amount of available data in this region implies robust calibrations for both approaches without the need of any *a priori* assumption.

2 DATA SET

The collection of records for this study is composed by accelerometric and velocimetric three-components waveforms of events occurred between 2008 and 2018 in Central Italy (Fig. 1), and includes the latest major seismic sequences that occurred in the study area, that is, the 2009 L'Aquila seismic sequence (Chiarabba *et al.* 2009) and the 2016–2017 Amatrice–Visso–Norcia sequence (Improta *et al.* 2019).

The data set was originally developed in the framework of seismic microzonation studies carried out in Central Italy after the 2016 M_w 6.0 Amatrice earthquake (Priolo *et al.* 2020), but it was extensively updated and exploited by other authors to study both the temporal and spatial variability of the ground motion in the area (e.g. Bindi *et al.* 2018c; Sgobba *et al.* 2021, and references therein). It is composed of more than 30 000 waveforms relative to 456 earthquakes in the local magnitude range 3.2–6.3 (for larger events also moment magnitude is available), recorded by about 460 stations within 250 km from the hypocentres (Fig. 1 to the left).

The region is very well sampled especially in the distance R [10–120 km] and magnitude ranges M [3.2–4.5] also thanks to the contributions due to the stations of the temporary networks (Cara *et al.* 2019). In fact, more than 100 temporary stations were installed in the region during the last 2016–2017 seismic sequence by the Italian Department of Civil Protection (DPC) and several academic and research institutions including the Istituto Nazionale di Geofisica e Vulcanologia (INGV).

From the whole waveform data set, we calculated the FAS of S-wave windows selected on the base of a frequency-dependent threshold on the signal-to-noise ratio (SNR, e.g. SNR = 10 for 0.2–0.4 Hz, SNR = 5 for 0.4–15 Hz and SNR = 10 for $f > 15$ Hz). The time windows used to calculate the spectral amplitudes were selected considering the fraction of cumulated energy (i.e. 90% for data recorded within an hypocentral distance of 25 km; 80% for distances between 25 and 50 km and 70% for larger distances), imposing a minimum duration of 4 s.

The FAS are smoothed using the Konno & Ohmachi (1998) algorithm where the smoothing parameter b was set to 40. Details about the data selection and processing are provided in several papers (Pacor *et al.* 2016; Bindi *et al.* 2018a, b; Castro *et al.* 2021; Spallarossa *et al.* 2021).

3 REFERENCE ROCK SITES

Response site analyses based on empirical techniques, such as standard spectral ratio, GMMs, and GIT, need the definition of the reference motion that is the ground motion not affected by site effects (i.e. the influence of the soil response at the ground due to the occurrence of a seismic event). The identification of reference recording sites is a critical issue, since there are several factors which can contribute to alter the ground motion recorded at a station, due to specific local conditions (e.g. variability of near-surface geology, topographic effects and seismic waves polarization).

In this study, the reference site selection is mainly based on the study of Lanzano *et al.* (2020), where the authors identified 36 recording stations as reference rock sites by means of the application of the multiproxies technique RRIM (Reference Rock Identification Method). The latter consists in the application of a decision matrix (Pugh 1981), allowing to perform the selection including several *proxies* that influence the site response, that is, (i) the outcropping geology; (ii) the installation features; (iii) the shear wave velocity $V_{S,30}$; (iv) the site topography, (v) the horizontal-to-vertical spectral ratios obtained from noise measurements or recordings and (vi) the repeatable site term obtained from residual analysis (δS2S), which defines the systematic bias of ground motions recorded at each station with respect to the reference predictions. The reference rock sites identified by Lanzano *et al.* (2020) have also been used to constrain the reference ground motion in the non-ergodic model calibrated by Sgobba *et al.* (2021) for the 5% damped ordinates of the acceleration response spectra (SA).

Fig. 2(a) represents the δS2S of the 36 reference rock sites which will be used in this study for a first stage calibration of an FAS-neGMM model (described in Section 5 and in particular in eq. 9). From a physical standpoint, these estimates of δS2S can be considered as the site amplification of each station, relative to the mean site response of the 36 reference rock sites. As shown in Fig. 2(a), the median of these δS2S is zero mean, but the variability dramatically increases as frequency increases.

In order to control such large variability, we decide to perform a further selection based on additional parameters for site characterization. First of all, we divide the set of stations into subgroups by means of a clusterization k-means algorithm (David & Vassilvitskii 2007). In order to select the number of clusters, we compute the percentage of the explained variance (PVE), to measure the discrepancy between the k-means model and the actual data: although the optimal number of clusters is 6 (PVE$\sim$100%), we fixed the number of cluster to 4, which corresponds to PVE $>$ 70%, with the aim of have populous classes on which to perform a subsequent skimming. Indeed, the population of each cluster is analysed on the basis of the value of the high-frequency attenuation parameter (Anderson & Hough 1984; Ktenidou *et al.* 2014; Castro *et al.* 2022), computed on the FAS, $\kappa_{0,\mathrm{fas}}$. Fig. 2(a) also reports the median curves obtained for each group: groups no. 3 (9 sites) and no. 4 (7 sites) have positive δS2S values at higher frequencies, while for groups no.1 (11 sites) and no. 2 (9 sites) are negative or null, respectively.

For most of the reference rock stations, the $\kappa_{0,\mathrm{fas}}$ was estimated according to the procedure described in Lanzano *et al.* (2022a) and varies over a wide range (between 0.007 and about 0.05 s Fig. 2b) reflecting the large variability at high frequencies. However, the lowest values that are expected to be representative of the reference rock sites (e.g. Hashash *et al.* 2014 proposed a threshold value of 0.006 s for very hard rock sites in central-eastern America) are included in the groups 3 and 4.

The number of stations included in the latter groups are 16 and the associated δS2Ss are reported in Fig. 3(a). The selected stations are furtherly separated in three subgroups, characterized by different intervals of $\kappa_{0,\mathrm{fas}}$: $\kappa_{0,\mathrm{fas}} < 0.015$ s (Fig. 3b); 0.015s$<\kappa_{0,\mathrm{fas}} < 0.025$ s (Fig. 3c) and $\kappa_{0,\mathrm{fas}} > 0.025$ s (Fig. 3d). The final selection includes the five stations with $\kappa_{0,\mathrm{fas}} < 0.015$ s, namely LSS, MNF, SLO, SNO and SDM, plus a sixth station (NRN), which does not have a robust estimate of $\kappa_{0,\mathrm{fas}}$ but has a δS2S trend very similar to the average of the five selected sites.

In principle, an ideal reference site, characterized by flat seismic response over the entire frequency range and sampling a wide magnitude and distance interval, can meet the demand to figure out the trade-off between the site and source terms in the GIT analysis. As a matter of fact, we made a trial analysis using only a single site as reference. We selected LSS station since it has flat response and the lowest $\kappa_{0,\mathrm{fas}}$ (0.007 s). However, the inversion results were unstable, providing source spectra that decay at high frequency and stress-drop values inconsistent with other studies in the area (Pacor *et al.* 2016; Bindi *et al.* 2020). For this reason, all the six stations here identified as reference rock sites will be used also for the generalized inversion of the FAS as described in the further section.

4 GENERALIZED INVERSION TECHNIQUE

In order to derive a suite of region-specific spectral parameters (i.e. source, attenuation and site parameters), we applied the non-parametric GIT (Andrews 1986; Castro *et al.* 1990; Oth et al., 2011) to the FAS of data recorded in Central Italy.

Being a data-driven approach (Bindi *et al.* 2020), it is important to select a robust data set in terms of number of recordings per station, distance, and frequency range coverage. For this reason, we carefully analysed and further selected the original FAS data set described in Section 2. We selected only spectral amplitudes in the frequency range 0.5–25 Hz (discretized in 69 bins equally spaced in logarithmic scale), and only stations with at least 10 recordings for each frequency bin. This reduced the number of stations to 283 homogeneously distributed in the region but not the number of events that remained 456.

The inversion analysis was performed considering a hypocentral distance range 10–120 km because, as shown in Fig. 1 (left), this represents a good sampling distance on a wide magnitude range for this data set. The distance range has been discretized into 62 bins, each

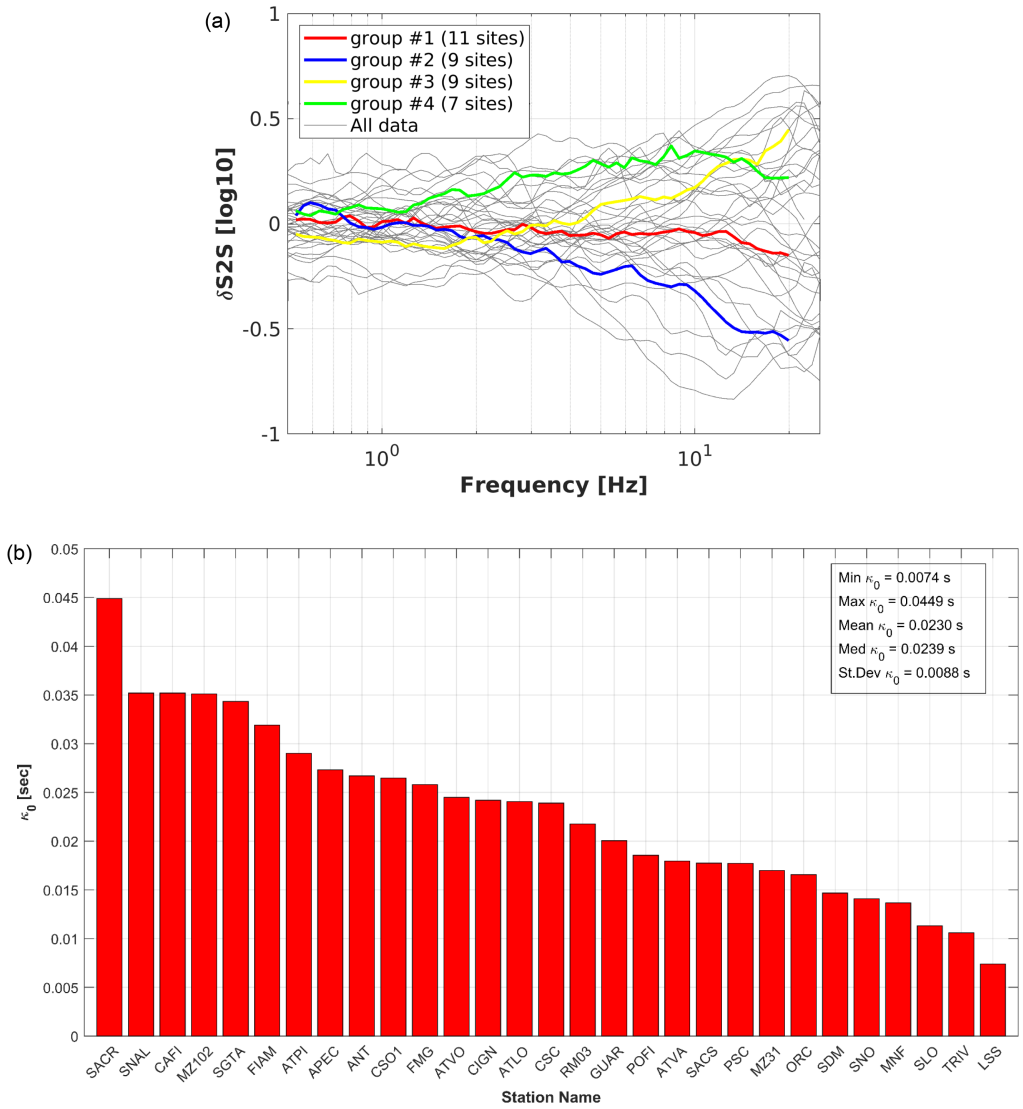

Figure 2. (a) Repeatable station residuals (δS2S) of the reference rock sites obtained by Lanzano *et al.* (2020). The coloured curves are the result of a clusterization analysis of the δS2S curves. (b) Values of k_0 for some of the 36 reference sites identified by Lanzano *et al.* (2020) were provided by applying the technique of Lanzano *et al.* (2022a) and shown here.

one is 2 km wide. Fig. 4 shows the cumulative distribution of events (left) and stations (right) of the final data set. About 50% of events used for the GIT inversion were recorded by at least 60 stations and about 50% of stations recorded at least 70 events.

The spectral amplitude decomposition procedure applied in this study consists of a one-step non-parametric GIT as described by Oth *et al.* (2011) and briefly summarized here.

The observed FAS derive from the combination of source, path and site contributions and can be described by a linear system of equations:

$$\log_{10}\mathrm{FAS}_{ij}\left(\mathrm{f}, \mathrm{R}_{ij}, \mathrm{M}_i\right) = \log_{10}\mathrm{Source}_i\left(\mathrm{f}, \mathrm{M}_i\right) + \log_{10}\mathrm{Path}\left(\mathrm{f}, \mathrm{R}_{ij}\right) + \log_{10}\mathrm{Site}_j\left(\mathrm{f}\right) \quad (1)$$

where R_{ij} is the hypocentral distance associated to the event i recorded by station j and f is the frequency.

The solution of a linear combination of three terms requires two *a priori* constraints because the system is inherently affected by trade-offs due to two unresolved degrees of freedom. The first constraint is applied to the attenuation function, setting it to unity at a reference distance of 10 km for all frequencies. Consequently, the source spectra are shifted to the same reference distance. Typically, the reference distance is assumed to be as small as possible compatible with the data distribution to avoid possible near-field effects on source spectra (Oth *et al.* 2011; Oth 2013).

In order to eliminate the linear dependence between the source and site terms, the average site amplification of the six reference stations (LSS, MNF, SLO, SNO, SDM e NRN) indicated in Section 2 was set to one.

The overdetermined linear system of equations is then solved in a constrained least-squares sense (Paige & Saunders 1982) to isolate the different contributions. For the stability of inversion solutions, we performed 200 bootstraps for each frequency point. No specific functional

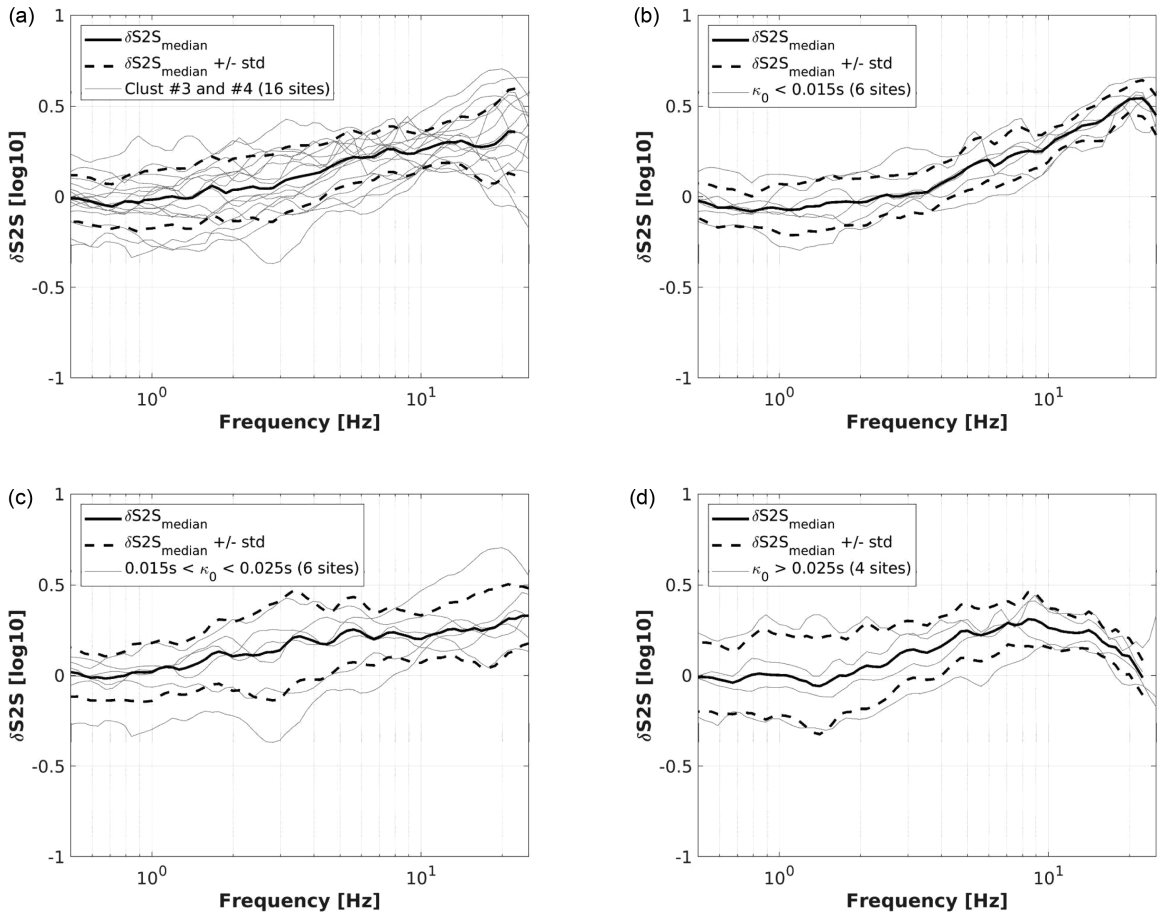

Figure 3. δS2S of group 3 and 4 versus frequency: (a) all (16 sites), (b) $\kappa_{0,\mathrm{fas}} < 0.015$ s (6 sites), (c) $0.015\,\mathrm{s} < \kappa_{0,\mathrm{fas}} < 0.025$ s (6 sites) and (d) $\kappa_{0,\mathrm{fas}} > 0.025$ s (4 sites).

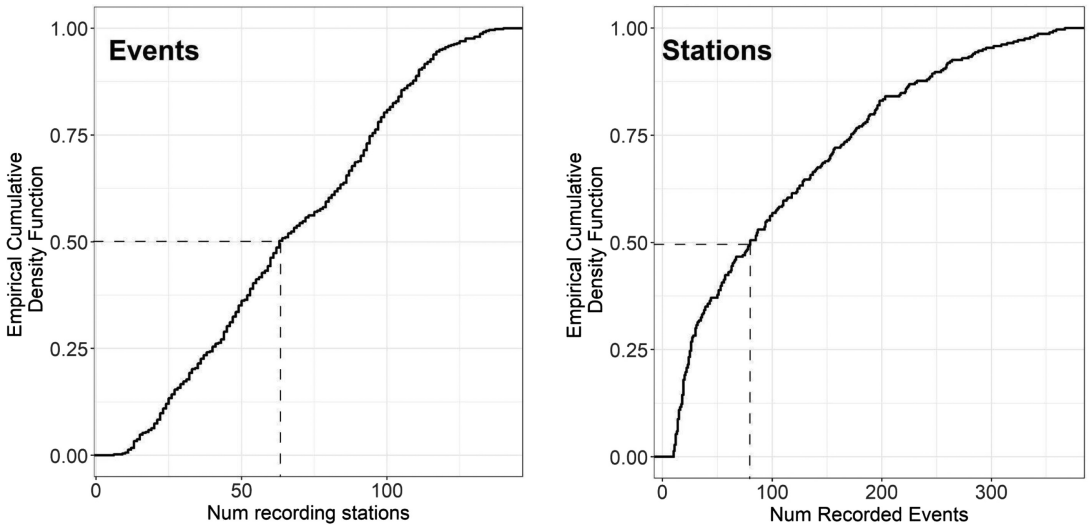

Figure 4. Cumulative distribution of events and stations used for the GIT analysis.

form is used for the unknown terms, but they are completely data-driven, resulting in non-parametric source and attenuation functions that will be l in the following sections to obtain those parameters useful for ground motion predictions (i.e. stress drop, kappa values, quality factor and geometrical spreading).

Finally, to demonstrate the reliability of the proposed GIT models, we simulated the Fourier spectra and compared to the observed FAS. Some examples are shown in Fig. A1 of the Appendix A.

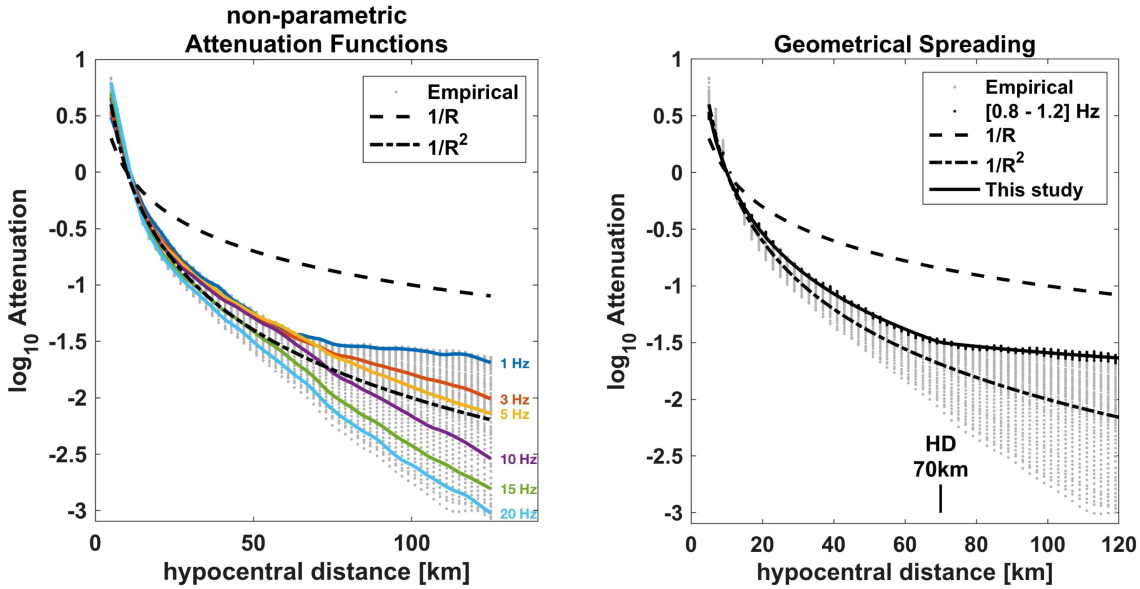

Figure 5. (a) Non-parametric spectral attenuation as a function of the hypocentral distance; (b) geometrical spreading fitting the non-parametric attenuation curves in the frequency range 0.8–1.2 Hz by means of a bilinear model with a hinge distance at 70 km. The grey dots represent the attenuation of the spectral ordinates of the Fourier spectrum for frequencies between 0.5 and 25 Hz. The coloured lines are the attenuation curves at selected frequencies, dashed and black lines indicate the R^{-1} and the R^{-2} decay, respectively.

4.1 Source model

We considered the following ω^2 model (Brune 1970, 1971) to fit the GIT derived non-parametric source spectra to provide the source parameters for each event and then to describe the regional scaling relationships.

$$S(f) = \frac{\Re_{\theta\varphi} A}{4\pi\rho\beta^3 R_0} M_0 (2\pi f)^2 \frac{1}{1 + \left(\frac{f}{f_C}\right)^2} \tag{2}$$

In this equation, $\Re_{\theta\varphi}$ represents the average S-wave radiation pattern set to 0.55 (Boore & Boatwright 1984); the parameter A accounts for the free-surface amplification factor set to 2. $\rho = 2.8$ g cm^{-3} and $\beta = 3.2$ km s^{-1} are the assumed density and the S-wave velocity for the Central Italy region. Finally, R_0 indicates the reference distance set to 10 km according to the GIT analysis. The seismic moment M_0 and the corner frequency f_c are then derived through a nonlinear regression by applying an iterative least-squares algorithm.

Once obtained these two parameters, as the relation between f_c and $\Delta\sigma$ is model dependent, the Brune stress drop is calculated as follows:

$$\Delta\sigma = \frac{7M_0}{16r^3} \tag{3}$$

where r is the source radius of a circular fault model (Brune 1970, 1971):

$$r = \frac{2.34\beta}{2\pi f_C} \tag{4}$$

The seismic moments were also used to derive moment magnitudes M_W following Hanks & Kanamori (1979). For the large events ($M_W > 5.5$), M_W values are constrained to the values retrieved from literature (Bindi *et al.* 2020) because the bandwidth limitations toward low frequencies do not allow observation of the corresponding corner frequencies.

The derived source parameters of the analysed earthquakes (i.e. seismic moment, M_0; corner frequency, f_c; stress drop, $\Delta\sigma$ and kappa source, k_{source}) are available for further analysis (see Data Resources). In addition, in Appendix B we describe the scaling relationships obtained analysing their behaviours.

4.2 Attenuation model

The non-parametric attenuation functions obtained from the GIT inversions at different frequencies as function of the hypocentral distance (Fig. 5a), have been fitted assuming a standard model including a distance-dependent geometrical spreading and a frequency-dependent attenuation term:

$$P\left(f,\ R_{ij}\right) = G\left(R\right) \exp\left[\frac{-\pi f\left(R - R_0\right)}{Qr\left(f\right)\beta}\right] \exp\left[-\pi k f\right] \tag{5}$$

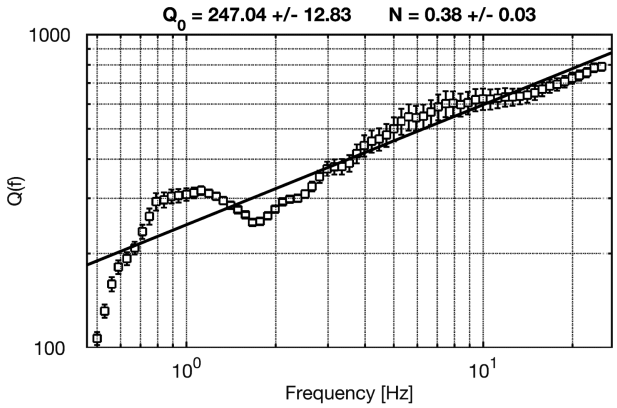

Figure 6. Fitting of the frequency-dependent quality factor.

where $G(R)$ is the geometrical spreading, β is the average shear wave velocity in the crustal medium (assumed equal to 3.2 km s^{-1}), R_0 is the reference distance used in the GIT inversion (10 km in this study), Q_r is the frequency-dependent quality factor and κ is the parameter introduced by Anderson & Hough (1984) to describe the high-frequency attenuation. First, we isolated the geometrical spreading contribution assuming that for the non-parametric attenuation around 1 Hz the anelastic attenuation contribution is small although not negligible. Hence, we performed a regression of the non-parametric attenuation curves from 0.8 to 1.2 Hz, considering the following bilinear model with hinge distance at 70 km (Fig. 5b):

$$G(R) = \begin{cases} \left(\frac{R_0}{R}\right)^{n_1} & \text{for } R \leq 70\,\text{km} \\ \left(\frac{R_0}{70}\right)^{n_1}\left(\frac{70}{R}\right)^{n_2} & \text{for } R > 70\,\text{km} \end{cases} \tag{6}$$

Finally, we derived the frequency-dependent attenuation term correcting the non-parametric attenuation functions from 20 to 120 km for the geometrical spreading $G(R)$ [$n_1 = 1.77$ and $n_2 = 0.56$] to estimate the frequency dependent quality factor $Q(f) = Q_0 f^{\alpha}$, fitting the following model:

$$AK(f) = \exp\left[\frac{-\pi f(R - R_0)}{\beta Q_0 f^{\alpha}}\right]\exp[-\pi k f] \tag{7}$$

Applying the logarithm of both sides of eq. (7) and considering $\kappa = 0$ s, we evaluate the Q_0 and α values through a least-squares regression in the frequency range 0.5–25 Hz (Fig. 6). Although $Q_r(f)$ is fitted by a monotonically increasing function, its frequency dependence shows a more complex trend characterized by a relative maximum around 1 Hz and a linear trend above 5 Hz, respectively (Fig. 6). The obtained best-fitting model is given by:

$$Q_r(f) = (247 +/- 12.8)\, f^{(0.38+/-0.03)} \tag{8}$$

5 NON-ERGODIC GMM IN FAS

In order to properly account for the relationship between seismological parameters inferred by GIT and intensity measures of the ground shaking, we calibrated a fully non-ergodic GMM for the Central Italy region based on 69 ordinates of the FAS in the frequency range from 0.5 to 25 Hz (FAS-neGMM). The model grounds on the same data-driven methodology and data set proposed by Sgobba *et al.* (2021) for predicting the PGA and the ordinates of the 5% damped-acceleration response spectra SA in the same range of vibration periods. The ergodic assumption is relaxed via a mixed-effect regression (Bates et al. 2015), providing the estimation of different repeatable effects of the seismic motion (i.e. source, site and path) for the Central Italy region.

The assumed functional form for FAS amplitudes at each frequency Y is:

$$\log_{10} Y = a + F_{\rm M}(M_{\rm w}) + F_{\rm R}(M_{\rm w}, R) + \delta{\rm Be} + \delta{\rm S2S_s} + \delta{\rm L2L_r} + \delta{\rm P2P_{esp}} + \delta W_0 \tag{9}$$

where a is the offset, while the magnitude $F_M(M)$ and distance $F_R(M,R)$ scalings are treated adopting standard dependencies, as:

$$F_{\rm M}(M_{\rm W}) = b_1(M_{\rm W} - M_{\rm h})(M_{\rm W} <= M_{\rm h}) + b_2(M_{\rm W} - M_{\rm h})(M_{\rm W} > M_{\rm h}) \tag{10}$$

$$F_R(M_W, R) = [c_1(M_W - M_{\rm ref}) + c_2]\ \log_{10}\sqrt{R^2 + h^2}/R_{\rm ref} + c_3\left(\sqrt{R^2 + h^2} - R_{\rm ref}\right) \tag{11}$$

The explanatory variables are the moment magnitude $M_{\rm w}$ and the source-to-site distance R. The latter is equal to the Joyner and Boore distance ($R_{\rm JB}$), available for the strongest events $M_{\rm w} > 5.5$, that is computed from the fault geometries published in the ITACA (ITalian ACcelerometric Archive, https://itaca.mi.ingv.it) database. For events $M_{\rm w} < 5.5$, the epicentral distance $R_{\rm epi}$ is used, assuming point-like sources.

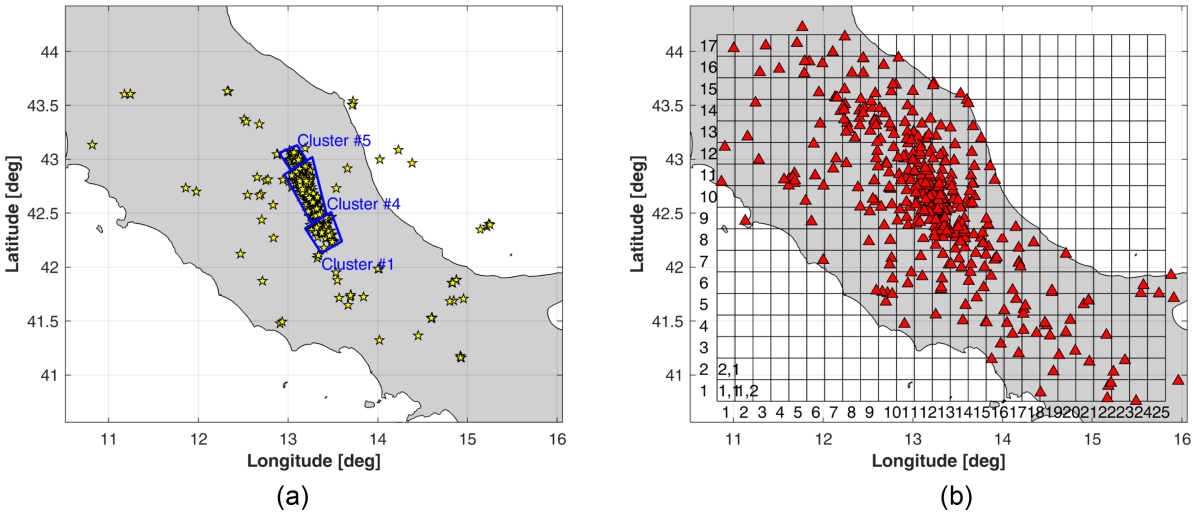

Figure 7. (a) Source clusters used for δL2Lr and δP2Pp calibration: no. 1 (main event: L'Aquila 06/04/2009—01:36 UTC), no. 4 (main event: Amatrice 24/08/2016–01:32 UTC) and no. 5 (main event: Muccia 10/04/2018—03:11 UTC) and (b) map of stations (red triangles) overlapped to the reference grid for δP2P calibration.

Some parameters are fixed in a first-stage non linear regression, that is, the hinge magnitude $M_h = 5.0$, the reference distance $R_{ref} = 1$ km and the pseudo-depth $h = 6$ km. The reference magnitude M_{ref} is instead to be frequency-dependent from nonlinear regression and varies between 5.45 ($f \sim 1$ Hz) and 3.3 ($f \sim 7.5$ Hz).

Eq. (9) includes the corrective factors introduced to remove the ergodic assumption (Stafford 2014), and that are estimated using the random effects on the systematic terms of events e (δB_e), stations s ($\delta S2S_s$), source regions r ($\delta L2L_r$) and region-to-site paths p ($\delta P2P_p$). The leftover residual δW_0 reflected the remaining variability not completely modelled by eq. (9): this term could include non-systematic effects related to the source-site configurations, for example potential rupture-directivity effects (Colavitti *et al.* 2022).

The coefficients a, b_1, b_2, c_1, c_2 and c_3 and the random effects (δB_e, $\delta S2S_s$, $\delta L2L_r$ and $\delta P2P_{esp}$) are obtained by a linear mixed effects regression (Stafford 2014; Bates *et al.* 2015) to robustly estimate the repeatable terms and their variability (see Data Resources).

According to Sgobba *et al.* (2021), the site terms $\delta S2S_s$ represent the empirical site effects computed with respect to the average ground motion level at the six reference sites. The source terms (i.e. the location-to-location residuals) $\delta L2L_s$ have been identified with a clustering approach, where the events are aggregated within polygonal source areas identified on the basis of space–time criteria. In the following, we refer to the most sampled clusters, that is, no. 1 (area of L'Aquila and Campotosto), no. 4 (Amatrice-Norcia) and no. 5 (Muccia). A map of the clusters is reported in Fig. 7(a).

Taking into account the source-to-site paths, we divide the whole region into squared cells (0.2° spaced), following the approach by Dawood & Rodriguez-Marek (2013), which allows to capture the spatial distribution of the attenuation behaviour (cell-specific attenuation), that depends on the differences in travel paths across heterogeneous geological layers or main structural discontinuities, such as the presence of faults (namely the Sibillini thrust system for Central Italy, see Buttinelli *et al.* 2021; Di Bucci *et al.* 2021 in the frame of RETRACE-3-D project). As a result, we obtain the mesh shown in Fig. 7(b) used to sample the cluster-cell paths; major details can be found in Sgobba *et al.* (2021).

6 RESULTS COMPARISON

The obtained non-ergodic random terms defined in the previous section (i.e. δB_e, $\delta S2S_s$, $\delta L2L_r$ and $\delta P2P_{esp}$) are physically based components of ground motion variability and thus can be linked with the spectral functions obtained from the non-parametric GIT approach. In doing this comparison, we consider that the two approaches are intrinsically related, so we aim to look for possible seismological connections with the underlying source and geophysical properties. A schematic view of the possible relations between the spectral parameters retrieved from GIT and FAS-neGMM systematic residuals is given in Table 1.

Hence, we compare here the non-ergodic terms with the parameters from GIT, with the aim to establish some empirical correlations useful to improve the median prediction of the FAS-neGMM model. Namely, we explore the correlations between δB_e and $\delta L2L_r$ with the source-related parameters from GIT (the stress drop $\Delta\sigma$ and the high-frequency attenuation parameter of the source spectrum k_{source}). Later, we compare the site-to-site systematic terms $\delta S2S_s$ with the non-parametric amplification functions from GIT. Finally, the correlation between the path-terms $\delta P2P_{esp}$ and the residuals of the quality factor Q (δQ2Q) with respect to the median of the empirical model $Q(f)$ reported in eq. (8) is also investigated. This latter correlation is the least studied in the literature due to difficulties both in estimating path-to-path residuals, owing to the need to dispose of dense data set for non-ergodic implementation, and also in capturing the variations of crustal properties in terms of residuals of Q. In the following section, we analyse the trend and corresponding statistical correlation for each individual pair of investigated parameters.

Table 1. Relationship between seismological parameters inferred by GIT (see Data Resources) and residuals from FAS-neGMM.

GIT parameters	FAS-neGMM residuals	Significance
$\Delta\sigma$, k_{source}	δB_e, $\delta L2L_r$	Source physics due to energy release, tectonic setting (Ameri *et al.* 2009; Bindi *et al.* 2018a)
Non-parametric site amplification function	$\delta S2S_s$	Local site amplification due to deep and shallow geology, 2D-3D effects (Bindi *et al.* 2017)
$\delta Q2Q$	$\delta P2P_{esp}$	Anelastic attenuation due to crustal properties, path-specific heterogeneities (Sahakian *et al.* 2019)

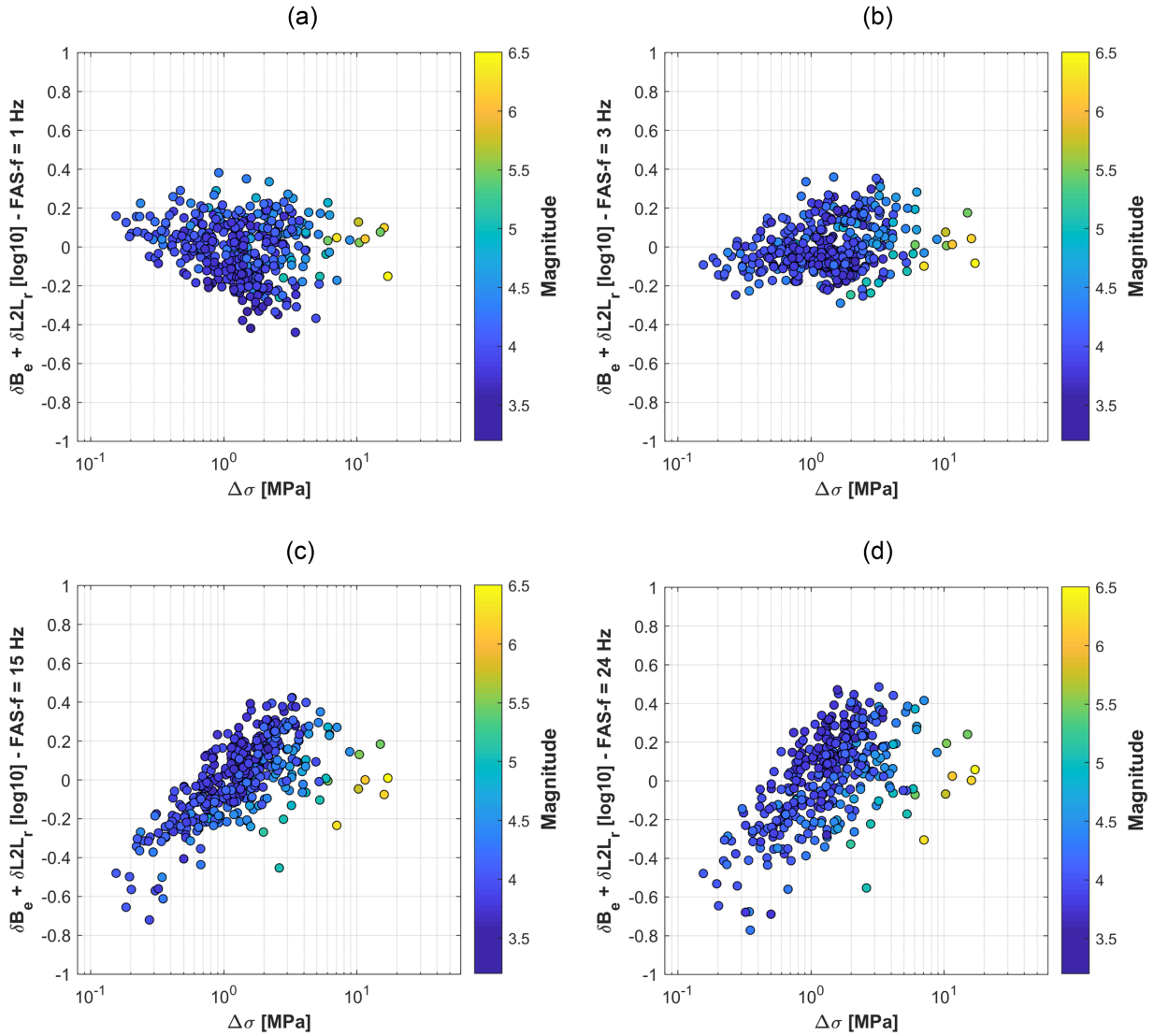

Figure 8. δB_e and $\delta L2L_r$ versus stress drop $\Delta\sigma$ for (a) $f = 1$ Hz; (b) $f = 3$ Hz; (c) $f = 15$ Hz and (d) $f = 24$ Hz. Colour intensity is proportional to the magnitude.

6.1. Source parameters versus between event

Fig. 8 shows the comparison between the sum of the source components $\delta B_e + \delta L2L_r$ and $\Delta\sigma$ released by the earthquakes included in the data set for 1, 3, 15 and 24 Hz.

As already observed in past investigations performed in Central Italy (e.g. Bindi *et al.* 2018c; Wang *et al.* 2019), we can observe that the largest stress drop values correspond to the events with highest magnitude (in our case $M = 6$–6.5, with yellowish colours in the colour bar). Regarding the dependence between the plotted parameters, for the lowest frequency (1 Hz, Fig. 8a), we can observe a weak negative correlation between stress drop values and the sum between $\delta B_e + \delta L2L_r$. At 3 Hz, the correlation becomes negligible (Fig. 8b), with values of $\delta B_e + \delta L2L_r$ ranging between -0.3 and $+0.4$ log10 units for the entire stress drop range (from 0.1 up to 100 MPa). At higher frequencies,

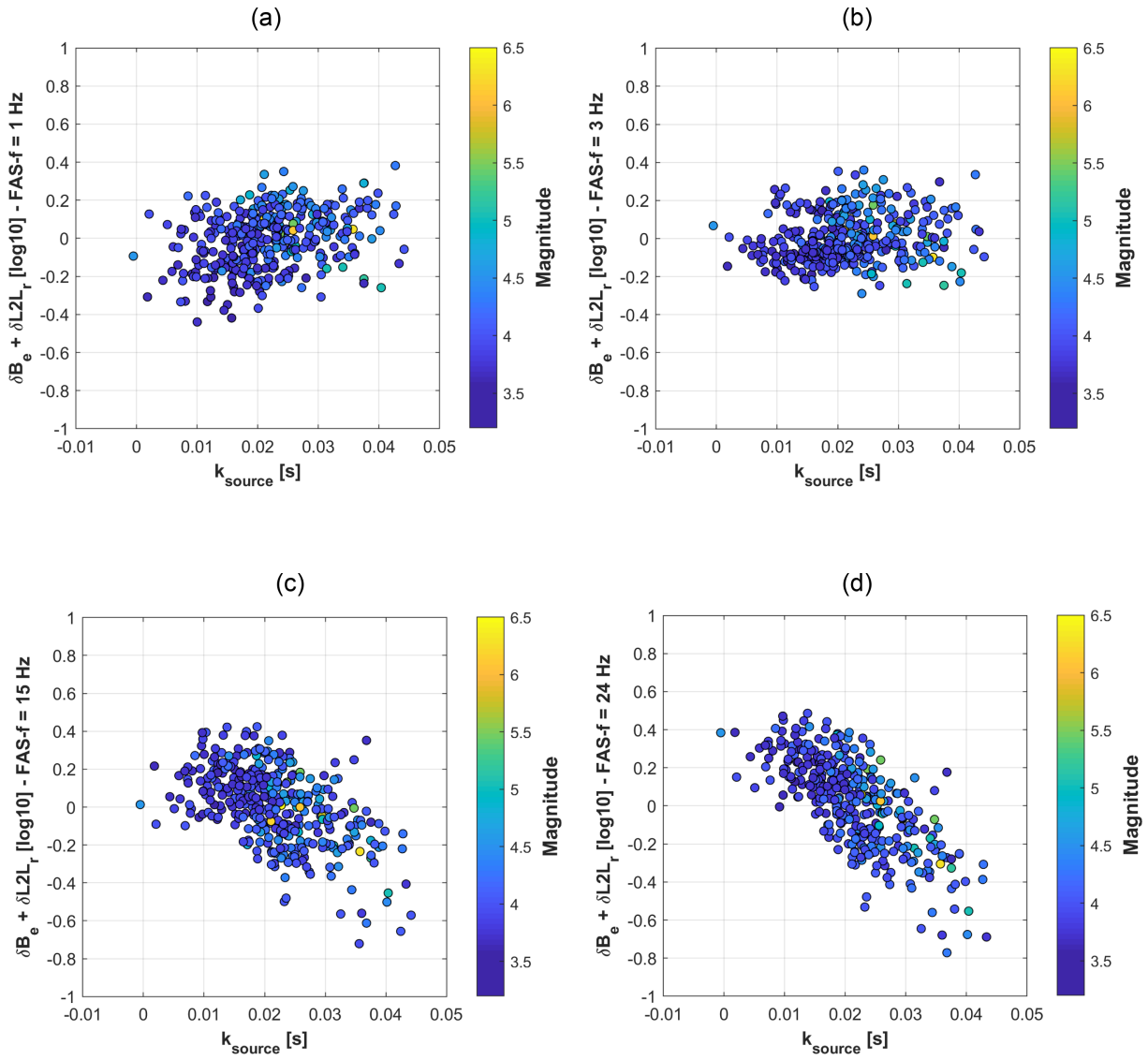

Figure 9. δB_e and $\delta L2L_r$ versus k_{source} for (a) $f = 1$ Hz; (b) $f = 3$ Hz; (c) $f = 15$ Hz and (d) $f = 24$ Hz. Colour intensity is proportional to the magnitude.

the correlation becomes remarkably positive as frequency increases, that is, as $\Delta\sigma$ increases, the value of the sum of the corrective factors increases (Figs 8c and d at 15 and 24 Hz, respectively).

Conversely, If we compare the sum of $\delta B_e + \delta L2L_r$ with the k_{source} parameter (Fig. 9), we observe an opposite trend. In fact, for lowest frequency (1 Hz, Fig. 9a), we observe a weak positive correlation between k_{source} and the sum of δB_e and $\delta L2L_r$. This slight dependence disappears at 3 Hz (Fig. 9b), indicating no correlation between the two parameters. Finally, at higher frequency values (15 and 24 Hz, see Figs 9c and d, respectively), we observe a negative correlation, which becomes stronger as frequency increases.

The findings are in agreement with the observations of Bindi *et al.* (2018c) and show that the correlations between the sum of the corrective terms with $\Delta\sigma$ and k_{source} are exactly specular. As a matter of fact, this is an indication of the negative trade-off between the two physical parameters, suggesting that only one of them should be introduced into the GMM modelling. For example, Bora *et al.* (2015) explicitly introduced only M_{W} and $\Delta\sigma$ as explanatory variables of source effects in the calibration of an FAS GMM for Europe.

6.2 Attenuation parameters versus path terms

In this section, we investigate the correlation between the path-terms $\delta\text{P2P}_{\text{esp}}$ and the residuals of the quality factor Q (δQ2Q) with respect to the median of the empirical model $Qr(f)$ calibrated in this study. In eq. (11), the terms accounting for the ground motion attenuation are defined through the coefficients c_1 and c_2 for the magnitude-dependent geometrical spreading and c_3 for the distance-dependent amount of anelastic attenuation. These corresponding terms of the GIT analysis are denoted by the non-parametric attenuation functions (eq. 5).

Assuming that the median attenuation characteristics are captured by these terms, we investigate the correlation between the random terms of the ground motion model, tha is, the path-to-path residuals (hereinafter referred to δP2P) in eq. (9) and the deviations of the quality

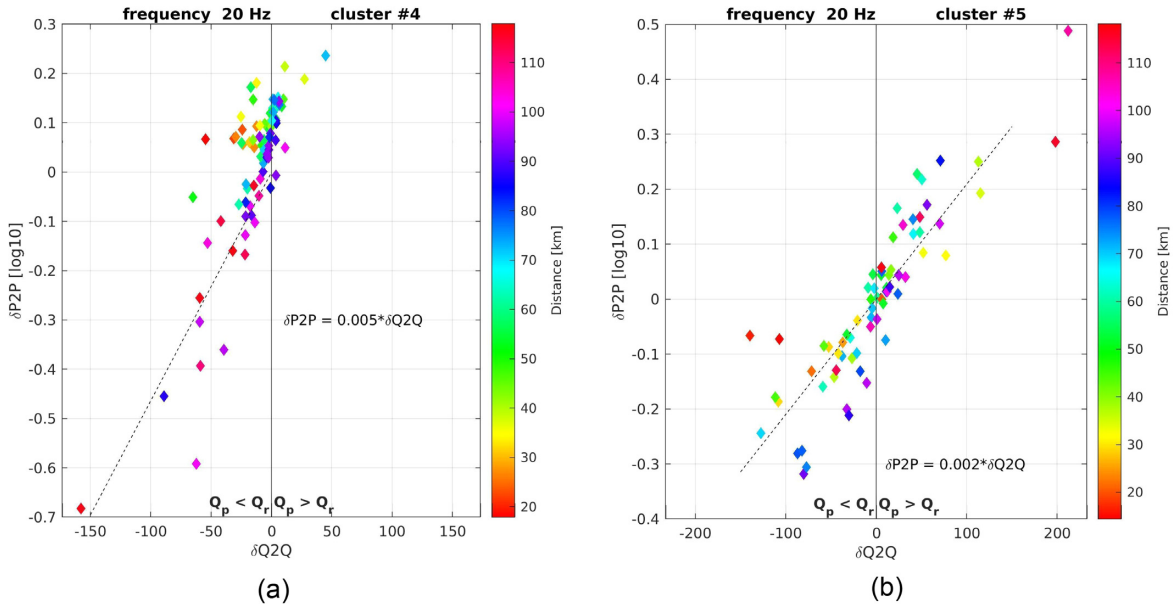

Figure 10. Scatter plot of the path terms δP2P from the non-ergodic FAS model against the quality factor residuals δQ2Q from GIT at 20 Hz for (a) cluster no. 4 and (b) for cluster no. 5, respectively. Colour bar indicates the different source-to-site distances.

factor specific of each source-to-site path $Q_p(f)$ from the median model $Q_r(f)$ reported in eq. (8), (i.e. the residuals Q-to-Q, δQ2Q), which can be regarded as crustal *proxies* of the path-specific heterogeneities that cause anisotropic attenuation over the study area.

The procedure to calculate the δQ2Q residuals is described in the following: the attenuation specific of each source-to-site path is computed by removing source and site contribution from FAS of recorded motion. In order to evaluate to what extent the specific path attenuation Path (f,R_{ij}) deviates from the regional behaviour we calculate the residual RES_{ij} (f,R) = $\log_{10}$ [Path (f,R_{ij})/P(f,R_{ij})], where P(f,R_{ij}) is the non-parametric attenuation derived by GIT (Fig. 5a). Assuming that the geometrical spreading and the S-wave velocity are the same for the regional model and the specific path, the residual RES_{ij} (f,R) is linearly dependent on $Q_r(f)^{-1}$–$Q_p(f)^{-1}$ (eq. 12) by the constant c (eq. 13):

$$\log_{10}\mathrm{RES}_{ij} = c\left[\frac{1}{Q_r(f)} - \frac{1}{Q_p(f)}\right] \tag{12}$$

$$c = \frac{-\pi f R \log_{10} e}{\beta} \tag{13}$$

The residuals δQ2Q = $Q_p(f)$—$Q_r(f)$ are then computed with reference to each source cluster identified in Section 5 and to each recording cell in the reference grid (0.2° spaced), following the same cell-specific approach applied by Sgobba *et al.* (2021) for δP2P sampling in non-ergodic GMM. In this way, the residuals δQ2Q represent the average deviations of the quality factor for each cluster-cell pair from the median isotropic model referred to the entire study area. Both δP2P and δQ2Q include effects of scattering and intrinsic anelastic attenuation. Here we report results of clusters 4 and 5, where we sampled a sufficient number of cells for comparison (94 and 76 cells, respectively). The path terms δP2P versus δQ2Q residuals are plotted in Fig. 10 at the highest frequency (20 Hz) where the correlation is stronger, while the spatial distributions of δQ2Q and δP2P are shown in Fig. 11.

One can observe a linear trend of δP2P for both cluster nos 4 and 5 (Fig. 10). In the case of cluster no. 4 (Fig. 10a) the path-specific quality factor is generally lower than the regional model; this means that higher levels of attenuation, marked by larger variations of δQ2Q, correspond to de-amplifications of ground motions (decrease in δP2P). At high frequency, for cluster no. 5 the zero mean distributions of the terms δP2P and δQ2Q (Fig.10b) denote similar spatial patterns, clearly correlated with the tectonic structure of the region, which is mainly influenced by the presence of the Sibillini thrust system, a west-dipping structure dividing the investigated area into hangingwall and footwall blocks (Fig. 11). Indeed, when ray paths move from the source cluster through the hangingwall block, we observe an increasing ground motion with relative positive adjustments of δP2P. Conversely, the paths travelling from the source towards the heterogeneous geological layers of the footwall in the Eastern region, crossing the Sibillini thrust, produce a decreasing ground motion (negative δP2P). These findings are in agreement with the empirical observations of Sgobba et al (2021) and tectonic investigations (Buttinelli *et al.* 2021) that have highlighted the presence of two main crustal domains and their propagation effects in the region with respect to the major Sibillini thrust system.

6.3. Amplification function versus site terms

The GIT analysis provides the non-parametric attenuation and source models described in the previous sections calibrated on 456 events and 283 stations. This allows to derive the site amplifications for each event and station by simply correcting the FAS using the GIT non-parametric terms. Finally, for each station, an average over all events is computed to obtain a more robust station amplification curve (Fig. 12 shows

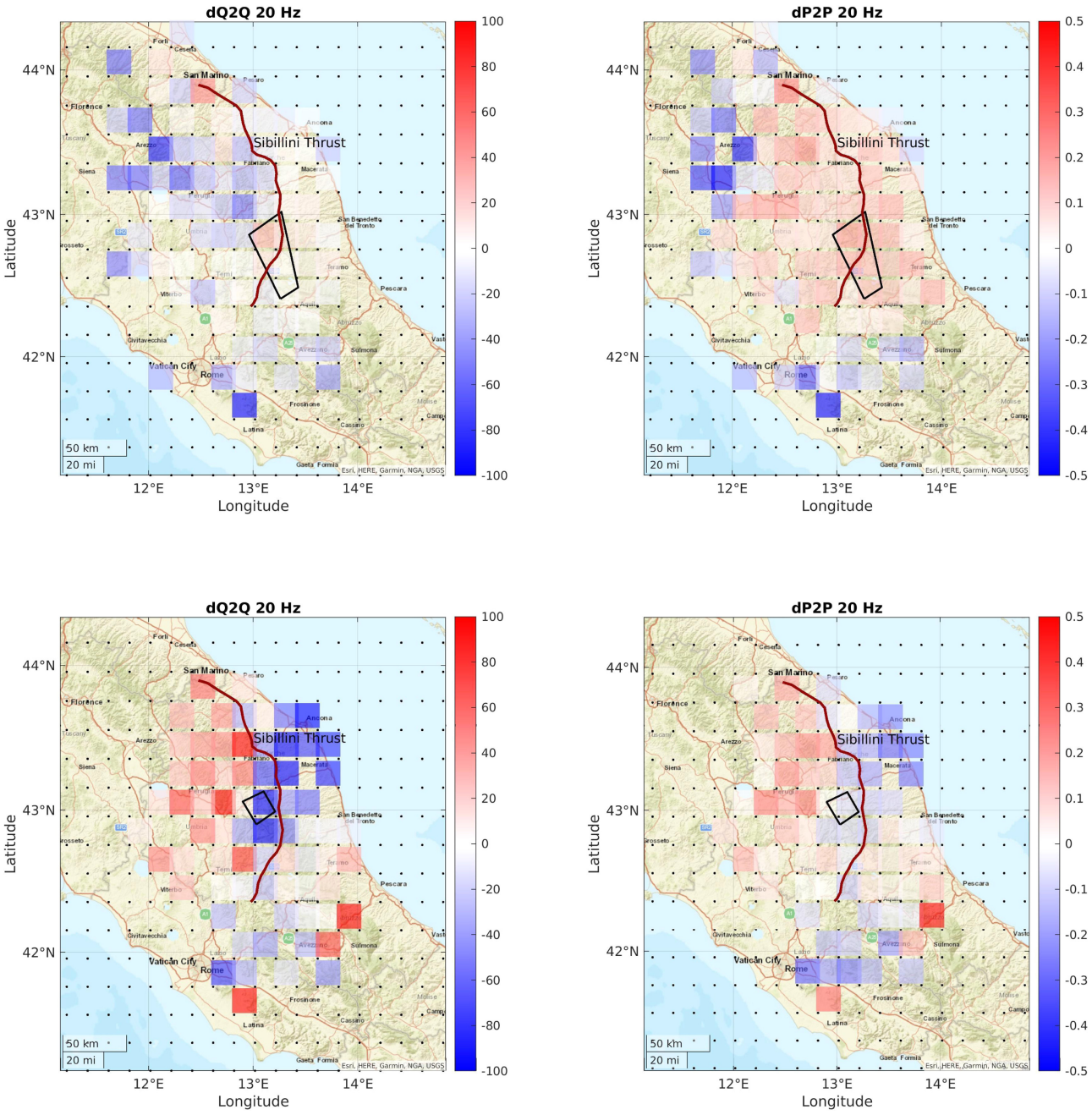

Figure 11. Left: residuals between the source-to-site quality factor (Q_p) and the regional model (Q_r) and right: path terms δP2P. The results are for cluster no. 4 (top) and cluster no. 5 (bottom) at 20 Hz.

some examples). Although the GIT is calibrated using a limited number of stations (283) with respect to the ∼460 available, the attenuation and source models are also used to correct the FAS for the remaining stations to obtain the corresponding site amplifications.

For more than 200 stations the site amplification curve is averaged over at least 50 events, and for 48 stations over more than 200 events. The number of events for the six reference stations (LSS, MNF, SLO, SNO, SDM e NRN) is 259, 163, 73, 94, 64 and 23, respectively. Although NRN has only 23 events, the site amplification curve has an acceptable average standard deviation of 0.17 on the log amplification.

Fig. 12 shows some examples of GIT-derived amplification curves together with $\delta S2S_s$ results from FAS-neGMM models. The comparison shows a general agreement between GIT and FAS-neGMM, especially at low frequencies where curves are overlapped. At high frequencies there are some differences between δS2S and GIT amplification functions, but the trends remain very similar. The standard deviation of the difference between GIT site response and $\delta S2S_s$, considering the whole range of frequency, varies between 0.02 and 0.05 for the shown stations. Analysing all the available stations the standard deviation remains below 0.1 for almost all stations. Only 11 stations have slightly larger values but in any case within 0.22.

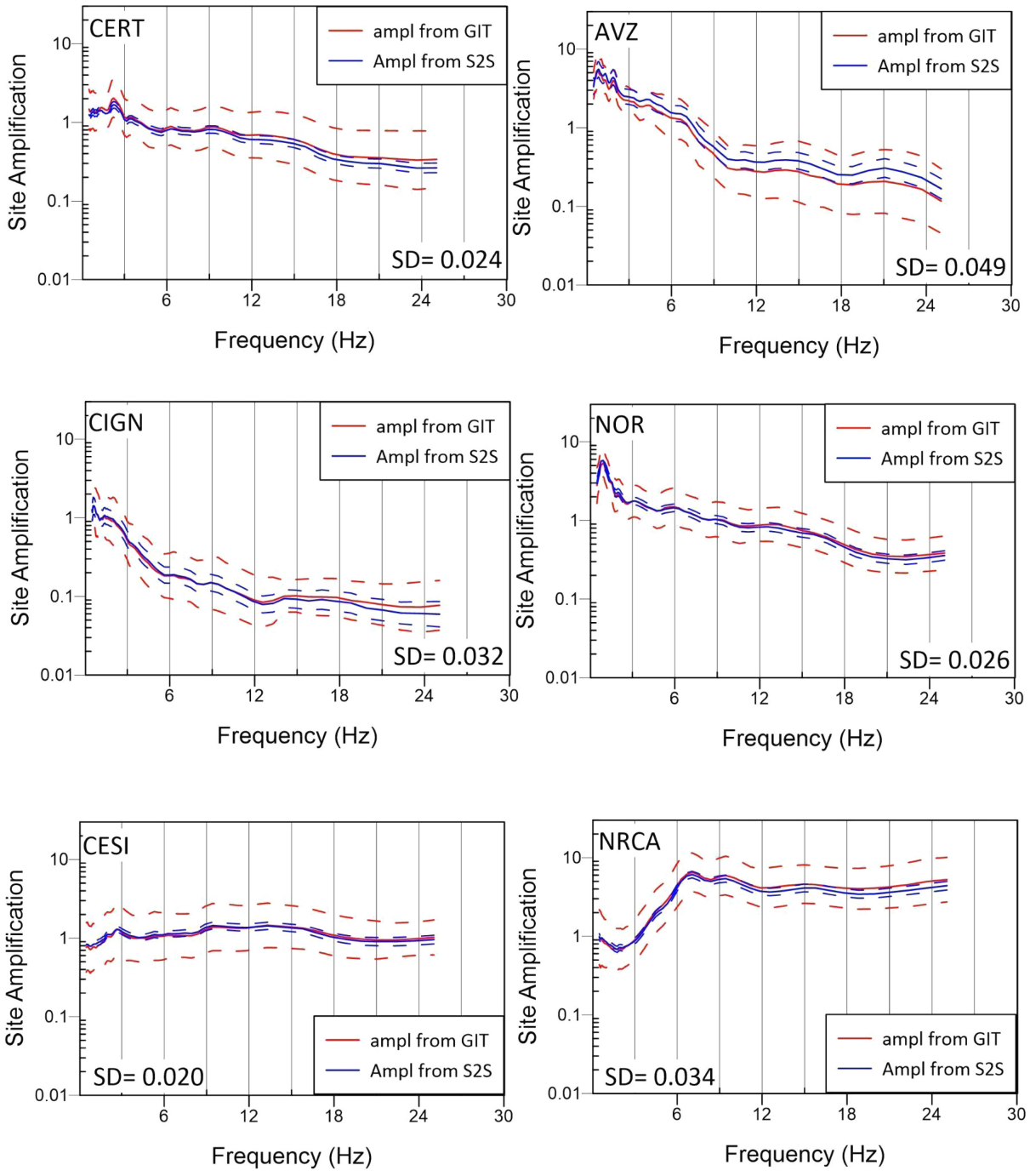

Figure 12 Examples of site amplifications obtained from GIT analysis (red) compared to $\delta S2S_s$ (blue) and the error associated with them (dashed curves). The standard deviation (SD) of the difference between the amplification curves obtained using the two models is also reported.

7 CONCLUSIONS

The large availability of data due to the multiple mainshock-aftershock sequences that have occurred in the last years in Central Italy (L'Aquila 2009; Amatrice-Visso-Norcia, 2016–2017), makes this region ideal for data-driven approaches. Both the non-parametric spectral amplitude decomposition approach (GIT) and the non-ergodic FAS-GMM exploit the possibility to analyse data sets with redundant information to effectively isolate the different contributions of the GIT and compute a robust estimate of the residual terms from the FAS-neGMM.

The analysis of the empirical correlations between the residuals of a non-ergodic GMM and the physical parameters obtained using a non-parametric GIT approach provided the basis for implementing an *ad-hoc* FAS neGMM for the Central Italy region. The main relevant observations are:

(i) the between-event δBe and the location-to-location terms δL2L can be described in terms of GIT-derived source parameters (stress drop and k_{source}); the correlation between the source parameters and the GMM corrective terms is frequency-dependent, and, at high frequency is remarkably positive for $\Delta\sigma$ and negative for k_{source};

(ii) the random terms of the neGMM model, represented by the path-to-path residuals, are associated with the deviations from the regional Q model. The resulting patterns show that the deviations of the quality factor specific of each source-to-site path $Q_p(f)$ from the regional isotropic model $Q_r(f)$ in, are properly captured by our model. The two crustal *proxies* of the path-specific heterogeneities (δP2P and δQ2Q) correlate well at higher frequency, demonstrating that differences in travel paths across different crustal domains can be explained by the corrective term δP2P.

The GIT amplification function and the $\delta S2S_s$ term describe exactly the same effect, when referred to the same reference seismic motion. This is consistent with the findings of Bindi & Kotha (2020) who observed that 'the site amplifications obtained through the non-parametric GIT inversion can be used to validate (and interpret) the site specific amplification terms obtained when developing GMMs'. Our findings could be useful to improve the median predictions and assess the standard deviation of empirical GMMs by explicitly introducing additional explanatory variables related to the source properties (e.g. the stress drop), regional crustal properties as well as the high-frequency attenuation at the site.

DATA RESOURCES

At the link https://shake.mi.ingv.it/central-italy/# we provide the following data sets:

CI-GIT contains GIT derived results such as source parameters; kappa values, calculated on different frequency bands for the stations used for the inversion, and the amplification curves for each station used to calibrate the GIT model (Morasca *et al.* 2022).

CI-FAS_GMM provides the coefficients, the residual decomposition terms (δBe, δL2Lr, δS2Ss, δP2Pesp and δW0) and the associated standard deviations of the non-ergodic GMM available for PGA and 69 ordinates of the FAS (Lanzano *et al.* 2022b).

ACKNOWLEDGMENTS

This research is supported by INGV in the frame of the project Pianeta Dinamico (Working Earth) - Geosciences for the Understanding of the Dynamics of the Earth and the Consequent Natural Risks (CUP code D53J19000170001), founded by the Italian Ministry of University and Research (MIUR) in the Task S3–2021–Seismic attenuation and variability of seismic motion.

This study has also benefited from funding provided by the agreement B1, DPC-INGV 2019–2021 between INGV and the Italian Dipartimento della Protezione Civile (DPC).

The authors are very grateful to the editor Dr. Eiichi Fukuyama and the reviewers Dr. Chuanbin Zhu and Prof. Yefei Ren for their valuable comments that improved this manuscript.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

REFERENCES

Al-Atik, L., Abrahamson, N. A., Bommer, J. J., Scherbaum, F., Cotton, F. & Kuehn, N., 2010. The variability of ground-motion prediction models and its components, *Seismol. Res. Lett.,* **81**(5), 794–801.

Ameri, G. *et al.*, 2009. The 6 April 2009 Mw 6.3 L'Aquila (Central Italy) Earthquake: strong-motion observations, *Seismol. Res. Lett.,* **80**(6), 951–966.

Anderson, J.G. & Brune, J. N., 1999. Probabilistic seismic hazard analysis without the ergodic assumption, *Seismol. Res. Lett.,* **70,** 19–28.

Anderson, J.G. & Hough, S., 1984. A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *Bull. seism. Soc. Am.,* **74,** 1969–1984.

Anderson, J.G. & Uchiyama, Y., 2011. A Methodology to Improve Ground-Motion Prediction Equations by Including Path Corrections, *Bull. seism. Soc. Am.,* **101**(4), 1822–1846.

Andrews, D. J., 1986. Objective determination of source parameters and similarity of earthquakes of different size, in *Earthquake Source Mechanics,* Das, S., Boatwright, J. & Scholz, C. H.(Editors), American Geophysical Union, Washington, pp. 259–267.

Baltay, A. S., Hanks, T. C. & Abrahamson, N. A., 2017. Uncertainty, variability, and earthquake physics in ground -motion prediction equations, *Bull. seism. Soc. Am.,* **107**(4), 1754–1772.

Baltay, A., Abercrombie, R. E. & Taira, T., 2022 The SCEC/USGS community stress drop validation study using the 2019 Ridgecrest earthquake sequence data, in *Presented at the SSA 2022 Meeting, Technical Sessions,* Seismological Society of America, 19–23 April, Bellevue, Washington, doi:10.1785/0220210025.

Bates, D., Mächler, B., Bolker, B. & Walker, S., 2015. Fitting linear mixed-effects models using lme4, *J. Stat. Software,* **67**(1), 1–48.

Bindi, D., Cotton, F., Spallarossa, D., Picozzi, M. & Rivalta, E., 2018c. Temporal variability of ground shaking and stress drop in Central Italy: a hint for fault healing?, *Bull. seism. Soc. Am.,* **108**(4), 1853–1863.

Bindi, D. & Kotha, S.R., 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.

Bindi, D., Picozzi, M., Spallarossa, D., Cotton, F. & Kotha, R., 2018b. Impact of magnitude selection on aleatory variability associated with ground-motion prediction equations: part II—analysis of the between-event distribution in Central Italy, *Bull. seism. Soc.Am.,* **109**(1), 251–262.

Bindi, D., Spallarossa, D. & Pacor, F., 2017. Between-event and between-station variability observed in the Fourier and response spectra domains: comparison with seismological models, *Geophys. J. Int.,* **210**(2), 1092–1104.

Bindi, D., Spallarossa, D., Picozzi, M. & Morasca, P., 2020. Reliability of source parameters for small events in Central Italy: insights from spectral decomposition analysis applied to both synthetic and real data, *Bull. seism. Soc. Am.,* **110** 1–19. https://doi.org/10.1785/0120200126.

Bindi, D., Spallarossa, D., Picozzi, M., Scafidi, D. & Cotton, F., 2018a. Impact of magnitude selection on aleatory variability associated with ground-motion prediction equations: part I—local, energy, and moment magnitude calibration and stress-drop variability in Central Italy, *Bull. seism. Soc. Am.,* **108**(3A), 1427–1442.

Boore, D. M. & Boatwright, J. 1984, Average body-wave radiation coefficients, *Bull. seism. Soc. Am.,* **74**(5), 1615–1621.

Bora, S. S., Scherbaum, F., Kuehn, N., Stafford, P. & Edwards, B., 2015. Development of a response spectral ground-motion prediction equation (GMPE) for seismic-hazard analysis from empirical Fourier spectral and duration models, *Bull. seism. Soc. Am.,* **105**(4), 2192–2218.

Brune, J. N., 1970. Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. geophys. Res.,* **75**(26), 4997–5009.

Brune, J. N., 1971. Seismic sources, fault plane studies and tectonics, *Eos,* **52**(5), 178–187.

Buttinelli, M. *et al.*, 2021. The impact of structural complexity, fault segmentation, and reactivation on seismotectonics: constraints from the upper crust of the 2016-2017 Central Italy seismic sequence area, *Tectonophysics,* **810,** 228861.

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

Castro, R. R., Anderson, J. G. & Singh, S. K., 1990. Site response, attenuation and source spectra of S waves along the Guerrero, Mexico, subduction zone, *Bull. seism. Soc. Am.,* **80,** 1481–1503.

Castro, R. R., Colavitti, L., Vidales-Basurto, C. A., Pacor, F., Sgobba, S. & Lanzano, G., 2022. Near-source attenuation and spatial variability of the spectral decay parameter kappa in Central Italy, *Seismol. Res. Lett.,* **93,** 2299–2310.

Castro, R. R., Pacor, F. & Spallarossa, D., 2021. Depth-dependent shear-wave attenuation in Central Apennines, Italy, *Pure appl. Geophys.,* **178,** 2059–2075.

Chiarabba, C. *et al.*, 2009. The 2009 L'Aquila (Central Italy) Mw6.3 earthquake: main shock and aftershocks, *Geophys. Res. Lett.,* **36,** L18308.

Colavitti, L., Lanzano, G., Sgobba, S., Pacor, F. & Gallovic, F., 2022. Empirical evidence of directivity patterns for small-to-moderate earthquakes on normal faults, *J. Geophys. Res.,* **127**(6), https://doi.org/10.1029/2021JB023498.

David, A. & Vassilvitskii, S., 2007. K-means++: the advantages of careful seeding, in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics Philadelphia, PA, USA. pp. 1027–1035.

Dawood, H. M. & Rodriguez-Marek, A., 2013. A method for including path effects in ground-motion prediction equations: an example using the Mw 9.0 earthquake aftershocks, *Bull. seism. Soc. Am.,* **103**(2B), 1360–1372.

Di Bona, M., 2016. A local magnitude scale for crustal earthquakes in Italy, *Bull. seism. Soc. Am.,* **106,** 242–258.

Di Bucci, D., Buttinelli, M., D'Ambrogi, C., Scrocca, D. &the RETRACE-3D Working Group, 2021. RETRACE-3D project: a multidisciplinary collaboration to build a crustal model for the 2016-2018 central Italy seismic sequence, *Boll. di Geofis. Teor. ed Appl.,* **62**(1), 1–18.

Drouet, S., Bouin, M.-P. & Cotton, F., 2011. New moment magnitude scale, evidence of stress drop magnitude scaling and stochastic ground motion model for the French West Indies, *Geophys. J. Int.,* **187**(3), 1625–1644.

Hanks & Kanamori, 1979. A moment-magnitude scale, *J. geophys. Res.,* **84,** 2348–2350.

Hashash, Y. M. *et al.*, 2014. Reference rock site condition for central and eastern North America, *Bull. seism. Soc. Am.,* **104**(2), 684–701.

Improta, L. The Bollettino Sismico Italiano Working Group *et al.*, 2019 Multi-segment rupture of the 2016 Amatrice-Visso-Norcia seismic sequence (central Italy) constrained by the first high-quality catalog of Early Aftershocks, *Sci. Rep.,* **9,** 6921.

Kanamori, H. & Rivera, L., 2004 Static and dynamic scaling relations for earthquakes and their implications for rupture speed and stress drop, *Bull. seism. Soc. Am.,* **94**(1), 314–319.

Konno, K. & Ohmachi, T., 1998 Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor, *Bull. seism. Soc. Am.,* **88,** 228–241.

Ktenidou, O. -J., Cotton, F., Abrahamson, N. A. & Anderson, J. G., 2014. Taxonomy of κ:a review of 285 definitions and estimation approaches targeted to applications, *Seismol. Res. Lett.,* **85**(1), 135–146.

Kuehn, N. M. & Abrahamson, N. A., 2019. Spatial correlations of ground motion for non-ergodic seismic hazard analysis, *Earthq. Eng. Struct. Dyn.,* **49,** 4–23.

Lanzano, G., Colavitti, L., Sgobba, S., Spallarossa, D. & Pacor, F., 2022b. CI-FAS_GMM: ground motion model of the fourier amplitude spectrum ordinates for the shallow active crustal events in Central Italy (Version 1.0) [Data set], *Istituto Nazionale di Geofisica e Vulcanologia (INGV)*. https://doi.org/10.13127/CI_dataset/CI-FAS_GMM.

Lanzano, G., Felicetta, C., Pacor, F., Spallarossa, D. & Traversa, P., 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.

Lanzano, G., Felicetta, C., Pacor, F., Spallarossa, D. & Traversa, P., 2022a. Generic-to-reference rocks scaling factors for the seismic ground motion in Italy, *Bull. seism. Soc. Am.,* **112,** 1583–1606.

Lanzano, G., Pacor, F., Luzi, L., D'Amico, M., Puglia, R. & Felicetta, C., 2017. Systematic source, path and site effects on ground motion variability: the case study of Northern Italy, *Bull. Earthq. Eng.,* **15**(11), 4563–4583.

Lin, P. S., Chiou, B. S., Abrahamson, N. A., Walling, M., Lee, C. T. & Cheng, C. T., 2011. Repeatable source, site, and path effects on the standard deviation for empirical ground-motion prediction models, *Bull. seism. Soc. Am.,* **101**(5), 2281–2295.

Malagnini, L. & Mayeda, K., 2008 High-stress strike-slip faults in the Apennines: an example from the 2002 SanGiuliano earthquakes (southern Italy), *Geophys. Res. Lett.,* **35,** L12302, doi:10.1029/2008GL034024.

Malagnini, L. & Munafò, I., 2018. On the Relationship between ML and Mw in a Broad Range: An Example from the Apennines, Italy, *Bull. seism. Soc. Am.,* **108**(2), 1018–1024.

Malagnini, L., Scognamiglio, L., Mercuri, A., Akinci, A. & Mayeda, K., 2008 Strong evidence for non-similar earthquake source scaling in central Italy, *Geophys. Res. Lett.,* **35,** L17303, doi:10.1029/2008GL034310.

Morasca, P., D'Amico, M. & Spallarossa, D., 2022. *CI-FAS_GIT: Seismological parameters and amplification functions derived by the Generalized Inversion Technique in Central Ita (Version 1.0) [Data set]*. Istituto Nazionale di Geofisica e Vulcanologia (INGV).

Morasca, P., Walter, W. R., Mayeda, K. & Massa, M., 2019. Evaluation of earthquake stress parameters and its scaling during the 2016 Amatrice sequence, *Geophys. J. Int.,* **218,** 446–455.

Muggeo, V. M. R., 2003. Estimating regression models with unknown breakpoints, *Stat. Med.,* **22,** 3055–3071.

Oth, A., 2013, On the characteristics of earthquake stress release variations in Japan, *Earth planet. Sci. Lett.,* **377–378,** 132–141.

Oth, A., Bindi, D., Parolai, S. & Di Giacomo, D., 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. seism. Soc. Am.,* **101,** 667–687.

Pacor, F. *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.

Paige, C. C. & Saunders, M. A., 1982, LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math Softw.,* **8**(1), 43–71.

Parker, G. A., Baltay, A. S. & Thompson, E. M., 2020. Repeatable source, path, and site effects from the 2019 M 7.1 Ridgecrest earthquake sequence, *Bull. seism. Soc. Am.,* **110**(4), 1530–1548.

Priolo, E. *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.

Pugh, S., 1981. Concept selection: a method that works, In: Hubka, V.(ed.), *Review of Design Methodology,* Proceedings International Conference on Engineering Design, Rome, pp. 497–506.

Sahakian, V. J., Baltay, A., Hanks, T. C., Buehler, J., Vernon, F. L., Kilb, D. & Abrahamson, N. A., 2019. Ground motion residuals, path effects, and crustal properties: a pilot study in Southern California, *J. geophys. Res: Solid Earth,* **124,** 5738–5753.

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

Sgobba, S., Lanzano, G., Pacor, F., Puglia, R., D'Amico, M., Felicetta, C. & Luzi, L., 2019. Spatial correlation model of systematic site and path effects for ground-motion fields in Northern Italy, *Bull. seism. Soc. Am.,* **109**(4), 1419–1434.

Spallarossa, D., Cattaneo, M., Scafidi, D., Michele, M., Chiaraluce, L., Segou, M. & Main, I. G., 2021. An automatically generated high-resolution earthquake catalogue for the 2016-2017 Central Italy seismic sequence, including P and S phase arrival times, *Geophys. J. Int.,* **225,** 555–571.

Stafford, P. J., 2014. Source-scaling relationships for the simulation of rupture geometry within probabilistic seismic-hazard analysis, *Bull. seism. Soc. Am.,* **104**(4), 1620–1635.

Wang, H., Ren, Y., Wen, R. & Xu, P, 2019. Breakdown of earthquake self-similar scaling and source rupture directivity in the 2016–2017 central Italy seismic sequence, *J. geophys. Res.: Solid Earth,* **124,** 3898–3917.

APPENDIX A. COMPARISON BETWEEN OBSERVED FAS AND GIT MODELLING

Fig. A1 shows some examples of comparison between observed FAS (geometric mean of the horizontal components, represented by the red curves) and corresponding theoretical results. The GIT derived models are verified by calculating the theoretical Fourier spectra for each event and for each source-station pair through Monte Carlo analysis by applying 250 simulations.

APPENDIX B. GIT-DERIVED SOURCE PARAMETERS AND THEIR RELATIONSHIPS

The source parameters obtained for each earthquake (see Data Resources) are shown in Fig. B1 and used to derive regional relationships. Moment magnitudes are compared to the available local magnitudes, Ml, based on Di Bona (2016) model (Fig. B1a). The best least-square fit is described by a bilinear model through a breakpoint regression (Muggeo 2003) with slopes of 0.7420 (± 0.0296) and 1.03441 (± 0.55), and a breakpoint at Ml = 4.546 (± 0.16655). The comparison with Malagnini and Munafò (2018) confirms the bilinear trend of the relation between *M*l and M_W with similar breakdown point, although some differences can be observed especially for small events due to the different approach to estimate both M_w and *M*l.

The scaling relationship between seismic moment and corner frequency is shown in Fig. B1(b) together with the constant stress drop lines. Through linear regressions on observed data in this region, many authors (Malagnini & Mayeda 2008; Malagnini *et al.* 2008; Pacor *et al.* 2016; Morasca *et al.* 2019) show an increase of stress drop with logM_0 (i.e. with M_w). This general behaviour is clearly evidenced also for our GIT-derived source parameters, nevertheless, our data are better fitted considering small and large events separately. A linear regression on the whole range of magnitude implies an extrapolation of unrealistically large stress drops for the major events (Drouet *et al.* 2011). To better capture the stress drop variability we prefer to apply two linear regressions as shown in Fig. B1(b). Following Kanamori & Rivera (2004), we calculated the corresponding ε values giving information on the deviation from the self-similarity. For small events ($M_0 < 10^{17}$) an $\varepsilon = 0.70852 \pm 0.028$ is in good agreement with observations from other studies in Central Italy showing an increase of stress drop with seismic moment. On the other hand, the rate of increase of stress drop for large events ($M_0 > 10^{17}$) decreases significantly as demonstrated by the small ε value (0.0293 ± 0.069) suggesting an almost constant stress drop.

Fig. B1(c) shows the scaling relationship between stress drop and logM_0 described by the following functions:

$$\log_{10}(\Delta\sigma) = 0.380413\,(\pm 0.145) * \log_{10}(M_0) - -5.7496.025\,(\pm 2.2379)\,(\text{for log } M_0 < 16.8) \quad \text{(b1)}$$

$$\log_{10}(\Delta\sigma) = 0.0808\,(\pm 0.23208) * \log_{10}(M_0) - 0.545 + 1.080\,(\pm 4.1773.728)\,(\text{for log } M_0 > 17.0) \quad \text{(b2)}$$

where is $\Delta\sigma$ in MPa and M_0 in N·m.

The comparison with other studies (Bindi *et al.* 2017; Wang *et al.* 2019 and Pacor *et al.* 2016) confirms the general observation of a breakdown in self similarity in Central Italy, but it is also important to highlight the high variability of stress drop estimations using different methodologies. Infact, this is a well-known issue and the main item of the SCEC/USGS Community Stress Drop Validation Study (Baltay *et al.* 2022) involving many research groups in the ongoing effort to understand and resolve differences in stress drop estimations.

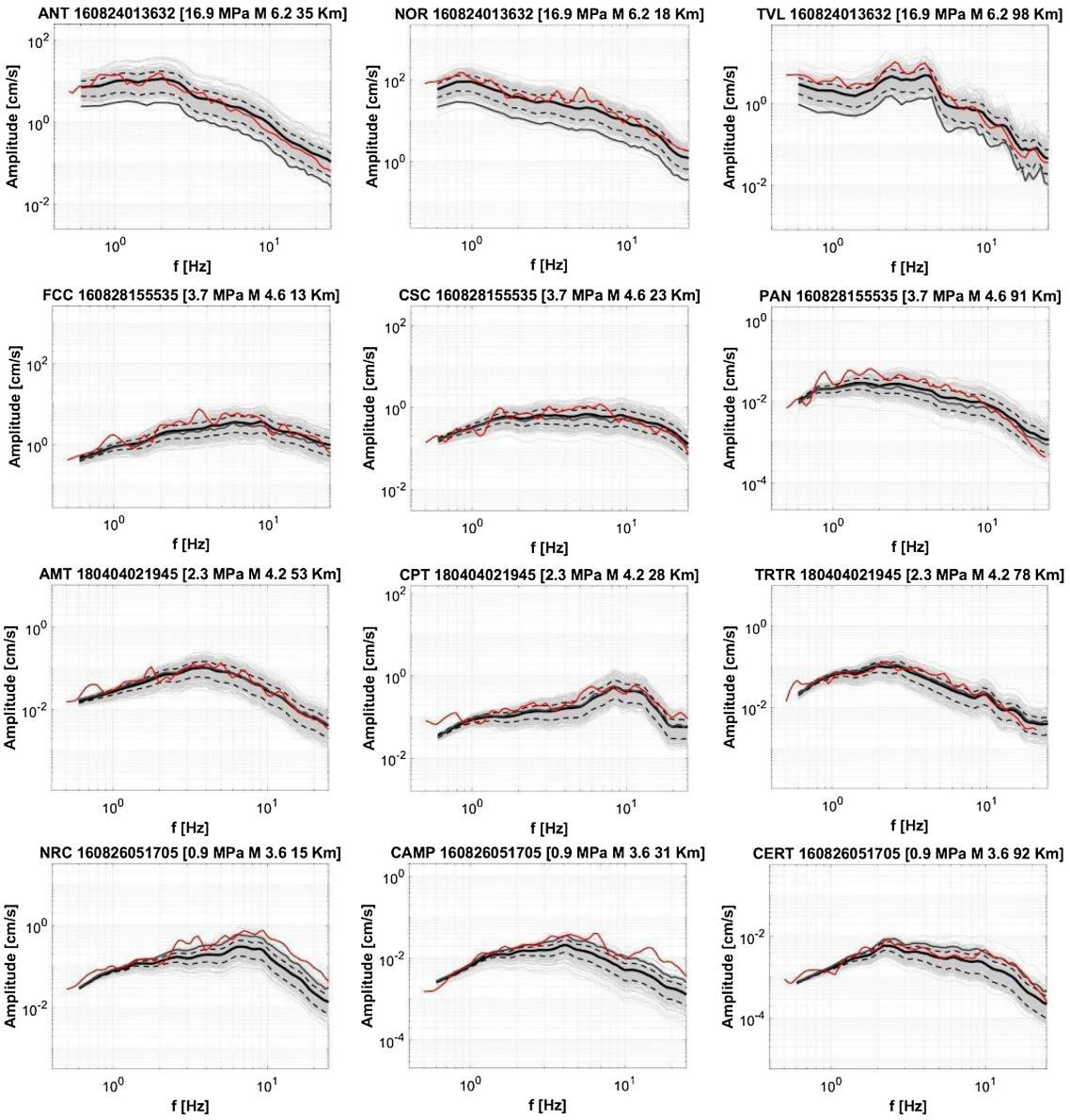

Figure A1. Comparison between observed FAS (geometric mean of the horizontal components, represented by the red curves) and corresponding theoretical results. The thin grey curves represent the theoretical FAS calculated with the Monte Carlo method using a variable stress drop (eqs b1 and b2 of Appendix B). The solid black curve represents their mean, while the dashed ones indicate the mean ± standard deviation. Finally, the thick grey curve is the theoretical mean FAS calculated with a constant stress drop of 3.1 MPa (mean value estimated on the data).

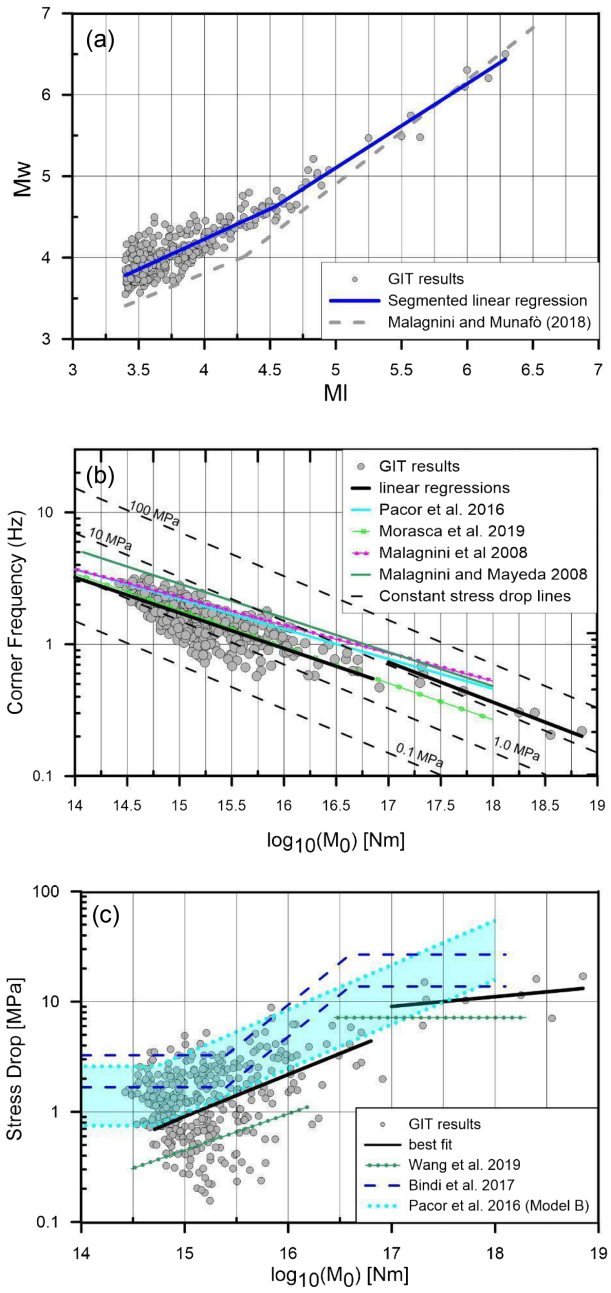

Figure B1. Scaling relationships obtained analysing the source parameters derived from fitting GIT non-parametric spectra using Brune (1970) in Central Italy. (a) Moment magnitude versus local magnitude (Di Bona 2016). The black line represents the best piecewise linear model on data (grey circles) compared to the bilinear model (grey lines) proposed by Malagnini and Munafò (2018). (b) Scaling of seismic moment versus corner frequency. The black lines represent the models for the corner frequency scaling in terms of the ε parameter for our GIT data (grey circles) for small and large events. Colour lines indicate the linear regressions obtained in other studies. (c) Stress-drop scaling with seismic moment fitted with two separate linear models for small and large events. Results are compared with other studies in the same region: Wang *et al.* (2019, green line), Bindi *et al.* (2017, top blue dashed line is for event deeper than 8 km, bottom blue dashed line is for shallower events) and Pacor *et al.* (2016, the colour band between the two dotted light blue lines includes the models for the different depths).