On the non uniqueness of the source, propagation and site effects decomposition

Dino Bindi[1], Kevin Mayeda[2], Daniele Spallarossa[3], Matteo Picozzi[4], Paola Morasca[5], Adrien Oth[6], and William R Walter[7]

[1]Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences
[2]Air Force Techical Application Center
[3]University of Genoa
[4]University of Naples Federico II
[5]Istituto Nazionale di Geofisica e Vulcanologia INGV
[6]European Center for Geodynamics and Seismology
[7]Lawrence Livermore National Laboratory (DOE)

November 22, 2022

Abstract

Although the non-uniqueness of the solution is commonly mentioned in the context of studies that perform spectral decompositions to separate source and propagation effects, its impact on the interpretation of the results is often overlooked. The purpose of this study is to raise awareness on this important subject for modelers and users of the models and to evaluate the impact of strategies commonly applied to constrain the solution. In the first part, we study the connection between the source-station geometry of an actual data set and the properties of the design matrix that defines the spectral decomposition. We exemplify the analyses by considering a geometry extracted from the data set prepared for the benchmark Community Stress Drop Validation Study (Baltay et al., 2021). In the second part, we analyze two different strategies followed to constrain the solutions. The first strategy assumes a reference site condition where the average site amplification for a set of stations is constrained to values fixed a-priori. The second strategy consists in correcting the decomposed source spectra for unresolved global propagation effects. Using numerical analysis, we evaluate the impact on source scaling relationships of constraining the corner frequency of magnitude 2 events to 30 Hz when the true scaling deviates from this assumption. We show that the assumption can not only shift the overall seismic moment versus corner frequency scaling but can also affect the source parameters of larger events and modify their spectral shape.

On the non uniqueness of the source, propagation and site effects decomposition

D. Bindi[1], K. Mayeda[2], D. Spallarossa[3], M Picozzi[4], P. Morasca[5], A. Oth[6], W. R. Walter[7]

[1]German Research Centre for Geoscience GFZ, Potsdam, Germany
[2]AFTAC, Patrick AFB, FL, USA
[3]University of Genova, DISTAV, Genova, Italy
[4]University of Naples Federico II, DIFI, Naples, Italy
[5]INGV, Milan, Italy
[6]European Center for Geodynamics and Seismology, Walferdange, Luxembourg
[7]LLNL, Livermore, CA, USA

Key Points:

- We discuss the non-uniqueness of the spectral decomposition using a data set extracted from Community Stress Drop Validation Study benchmark
- We evaluate the impact of a-priori constraints applied to restore the uniqueness of the solution
- We use numerical tests to quantify the distortions introduced in source scaling when constraining the corner frequency of small events

Corresponding author: Dino Bindi, `bindi@gfz-potsdam.de`

Abstract

Although the non-uniqueness of the solution is commonly mentioned in the context of studies that perform spectral decompositions to separate source and propagation effects, its impact on the interpretation of the results is often overlooked. The purpose of this study is to raise awareness on this important subject for modelers and users of the models and to evaluate the impact of strategies commonly applied to constrain the solution. In the first part, we study the connection between the source-station geometry of an actual data set and the properties of the design matrix that defines the spectral decomposition. We exemplify the analyses by considering a geometry extracted from the data set prepared for the benchmark Community Stress Drop Validation Study (Baltay et al., 2021). In the second part, we analyze two different strategies followed to constrain the solutions. The first strategy assumes a reference site condition where the average site amplification for a set of stations is constrained to values fixed a-priori. The second strategy consists in correcting the decomposed source spectra for unresolved global propagation effects. Using numerical analysis, we evaluate the impact on source scaling relationships of constraining the corner frequency of magnitude 2 events to 30 Hz when the true scaling deviates from this assumption. We show that the assumption can not only shift the overall seismic moment versus corner frequency scaling but can also affect the source parameters of larger events and modify their spectral shape.

Plain Language Summary

Source properties at seismogenic depths cannot be measured directly. Therefore, the characterization of the size and strength of the source has to rely on our ability to isolate the footprint left by the source into seismic recordings acquired at the Earths surface. Since propagation effects are not known in detail, questions about absolute measure of the earthquake source have not unique answers. An approach in which source, propagation and site effects are isolated within observations is called a decomposition approach. In this study, we discuss the non-uniqueness of the decomposition and the impact of the trade-offs affecting the different decomposed terms. Using a data set collected in Southern California during the 2019 Ridgecrest sequence, we connect non-uniqueness and trade-offs to properties of matrices describing the problem such as the design, covariance and resolution matrices. We also develop numerical analysis to quantify the impact on source scaling of the constraints used to resolve the non-uniqueness. Overall, this study raises awareness on the consequences that different strategies applied to deal with the non-uniqueness can have on the final outcomes of the decomposition, an impact that should not be overlooked when comparing conclusions drawn from different studies performed in the same area.

1 Introduction

The retrieval of the moment rate function or spectrum of an earthquake from a set of recordings requires the solution of an inverse problem. In the model framework where the seismic waves propagate through the Earth's interior as elastic waves, the transfer function of the system describes the effects of seismic wave propagation from the hypocenter to the stations, including near-surface site effects. Therefore, under the assumption of linear and time-invariant systems, the source function retrieval is a deconvolution problem (Helmberger & Wiggins, 1971; Hartzell, 1978; Mueller, 1985; Bertero et al., 1997). Two main limitations affect the retrieval process: first, propagation effects are also unknown, introducing a system identification problem; second, the unknown propagation effects filter the input signal and source-related information can be recorded only within a limited bandwidth. The consequence of the first issue is that solving the inverse problem involves either the determination of both the source and the propagation terms (Andrews, 1986) or requires assumptions about the propagation effects, such as within the empirical Green's function approach (e.g., Hartzell, 1978; Bertero et al., 1998, among others). Regarding the filtering issue, since source characteristics outside the data bandwidth cannot be reconstructed only using observations (Ide & Beroza, 2001), the detection of high-frequency source parameters for small events is hampered by the limited bandwidth and by the competing source and attenuation effects. The detection limits of the source parameters for small events have been investigated in a number of studies using numerical simulations and theoretical models (e.g., Kwiatek & Ben-Zion, 2016; Chen & Abercrombie, 2020; Bindi, Spallarossa, et al., 2020). Although several factors contribute to determine the magnitude threshold (e.g., source-station geometry, severity of the near surface attenuation, stress drop values for small events, sampling rate, among others), the determination of reliable source parameters below magnitude 2 is challenging with surface monitoring networks.

The analysis of data sets with a high level of redundancy in the cross-sampling between stations and events has been proposed to simultaneously determine source, propagation and site effects in the spectral domain (Andrews, 1986; Castro et al., 1990; Boatwright et al., 1991; Shearer et al., 2006; Oth et al., 2011; Bindi, Zaccarelli, & Kotha, 2020; Shible et al., 2022). The possibility to observe the same source spectrum at different distances and azimuths, with each site term combined with different sources, allows to set up an overdetermined system of equations that can be solved in a least-squares sense to isolate the different terms. A peculiarity of such an approach is that the solution is not unique since, regardless of the number of available recordings, the model assumptions, and the wavefield selected for analysis (e.g., S-waves or coda), the design matrix is rank deficient (Andrews, 1986). Specific solutions are obtained by imposing a-priori constraints to the solution, causing difficulties in comparing results from different studies even when performed in the same area. In addition to the non-uniqueness generated by the convolution model, other factors limit our ability to resolve source, propagation and site effects from recordings. We identify two major limitations: the first limitation is connected to the physics of the seismological problem, where the different terms show some degree of correlation (for example, the same high-frequency spectral content could be obtained by different combinations of source parameters and near-surface attenuation); the second limitation depends on the assumptions made over the different terms. Regarding the latter, examples are the assumption of dealing with spherically averaged source spectra (i.e., neglecting radiation pattern and directivity effects) and of assuming a 1D spectral attenuation depending only on hypocentral distance (i.e., neglecting lateral velocity and attenuation variabilities; not accounting explicitly for depth dependencies), which are typical assumptions for decomposition studies. The impact of these two hampering factors can be exacerbated or mitigated by the geometry of the analyzed problem (translated into characteristics of the design matrix) and, in general, it depends on frequency. Moreover, the impact depends on the analyzed wavefield since, for example, seismic coda is less sensitive to azimuthal variability of the source radiation than direct waves (Aki & Chouet, 1975; Aki, 1981; Mayeda & Malagnini, 2010).

The target of this study is to discuss the connections between the source-station geometry of an actual data set and the properties of the design matrix defining the source, propagation and site spectral decomposition. In particular, this study is motivated by the ongoing effort to develop a community stress drop initiative for comparing the source parameters estimated for the shared Ridgecrest data set (Baltay et al., 2021; Trugman, 2020) by applying a wide variety of approaches. Refreshing the connection between the non-uniqueness of the spectral decomposition approach and the trade-offs affecting the source, propagation and site terms can support the comparison of spectral decomposition results with other approaches, and increase awareness of the impact of non-uniqueness of the solution on the interpretation of results (see also Chen & Abercrombie, 2020; Zhang et al., 2022). For this reason, we also discuss strategies applied to select specific source spectra from the infinite set of possible solutions, evaluating their impact on source parameters estimation.

2 Spectral decomposition of the Fourier amplitude spectra

The Generalized Inversion Technique (GIT) aims at isolating the source, propagation and site terms from the Fourier Amplitude Spectra (FAS) of the recorded P- or S-waves. Large data sets with a high level of redundancy in the station and earthquake sampling are decomposed under the assumption that source, propagation and site effects combine through a convolution product expressed in the Fourier domain as

$$O_{ij}(f) = S_i(f) + P(R_{ij}, f) + Z_j(f) \tag{1}$$

where the O_{ij}(f) is the logarithm of the spectral amplitude at frequency f for earthquake i recorded at station j; $S_i(f)$ is the logarithm of the source spectrum for event i; $P(R_{ij}, f)$ is the logarithm of the propagation term for the hypocentral distance R_{ij}; $Z_j(f)$ is the logarithm of the site term for station j. In equation 1, the propagation effects are assumed to be isotropic and controlled by the hypocentral distance. Two strategies can be applied to isolate source, propagation and site effects: the parametric (Castro et al., 1990; Boatwright et al., 1991) and the non-parametric (Andrews, 1986) approaches. In the former, source and attenuation terms are described in terms of standard seismological models, such as a parametrization of the source spectra in terms of seismic moments and corner frequencies and a parametrization of the propagation in terms of geometrical spreading and t^*; in the non-parametric approach, equation 1 is solved without imposing a-priori parametric models on the different terms. In this study, we focus only on the non-parametric approach, although the non-uniqueness of the solution is intrinsic to both the parametric and non-parametric approaches and, more in general, to any decomposition approach applied in the time or frequency domains to isolate source, propagation and site contributions to recorded time histories or Fourier spectra. By analyzing each frequency separately, equation 1 generates a set of overdetermined linear systems that can be solved in a least-squares sense. In the following, we provide a description of the coefficient matrix considering an extraction of the data set based on the 2019 Ridgecrest sequence (Trugman, 2020) compiled for the Community Stress Drop Validation study (Baltay et al., 2021). We consider 554 earthquakes recorded by 94 stations (counting co-located sensors separately) generating 11064 spectral amplitudes at f=4.2 Hz; magnitudes are in the range 2.5-7.1 and hypocentral distances cover the range 5-111.4 km. For the attenuation term $P(R, f)$, the range of distances from 4 to 112 km is discretized by considering 37 nodes spaced 3 km apart. The geometry of the data set is shown in Figure 1.

3 Matrix of coefficients, covariance and resolution

The design matrix of the linear system generated by equation 1 is a sparse matrix with in general four non-zero elements for each row. The matrix has dimension MxN, where M is the number of recordings and N is the sum of the number of events N_e, of stations N_s and the number of nodes N_d used to discretize the distance range. The structure of the matrix can be expressed in terms of dummy variables used to select the event, station and distance bin relative to each recording. If the recording in the m$-th$ row is associated with the event i recorded by station j and the distance R_{ij} lies between nodes k and $k+1$ (i.e., $R_k \leq R_{ij} < R_{k+1}$), equation 1 can be written as

$$O_{m,n} = \delta_{n,i} S_i + \delta_{n,N_e+j} Z_j + (1 - w_k)\delta_{n,k+N_e+N_s} P_k + w_k \delta_{n,k+1+N_e+N_s} P_{k+1} \tag{2}$$

where S_i with $i = 1, ..., N_e$, Z_j with $j = 1, ..., N_s$, and P_k with $k = 1, ..., N_d$ define the vector of unknowns. In our example, since N_e=554, N_s=94, and N_d=37, the total number of unknowns for a given frequency is N= 685, and the number of data counted at f=4.2 Hz is M=11064. In equation 2, the attenuation term is linearly interpolated between nodes k and $k+1$, i.e., $w_k = (R_{ij} - R_k)/(R_{k+1} - R_k)$. Therefore, all entries in each row of the design matrix are zero except those for one specific column in the event-column group, one specific column in the station-column group and two adjacent columns

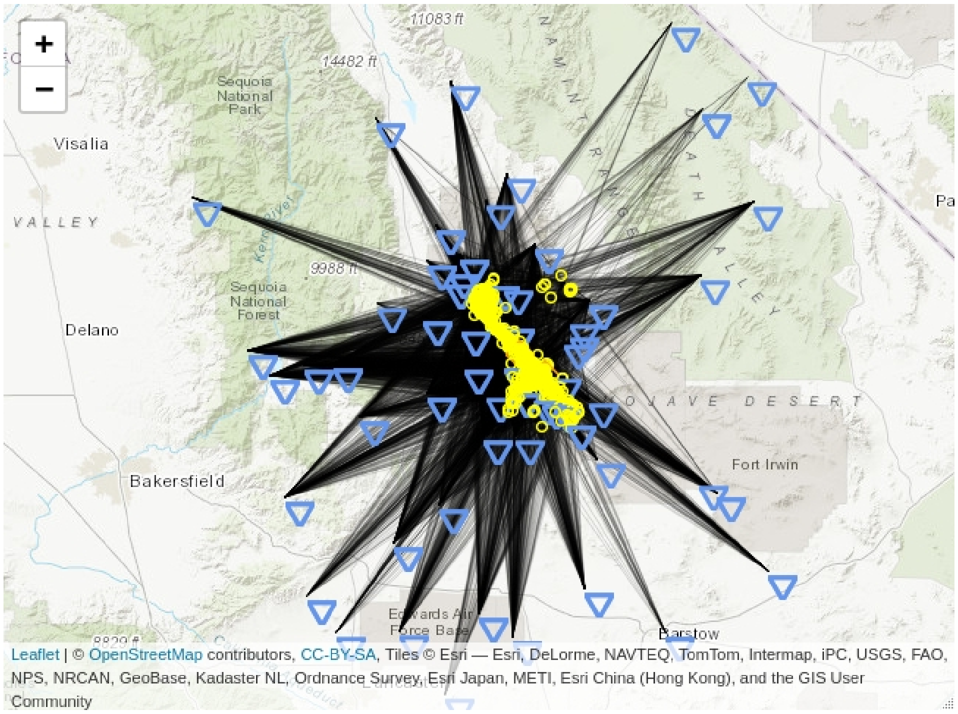

Figure 1. Event-station geometry of the data set considered in this study, extracted from the Community Stress Drop Validation study (Baltay et al., 2021): black lines connect epicenters (circles) of the considered earthquakes with the location of the recording stations (triangles)

in the attenuation-column group (the latter reduce to one if the distance R_{ij} is coincident with one of the nodes R_k). The design matrix for the geometry in Figure 1 is shown in Figure 2, where the recordings are ordered per event (all recordings of the same event are consecutive). This matrix is based on the data available at 4.2 Hz. In the lower panel of Figure 2, which shows a detail of the matrix for rows 10991 to 10998 corresponding to two different earthquakes, non-null elements are shown by vertical bars (the value is 1 for the event and station columns, two values between 0 and 1 and summing to 1 for the two adjacent attenuation columns). The least-square solution of system 1 is not unique because there are two unresolved degrees of freedom: since we are summing three terms, we can add a constant to one term and remove the same constant from another and the sum will not change (Andrews, 1986). To test the rank deficiency of the design matrix, we perform its singular value decomposition

$$G = USV^T \tag{3}$$

where G is the MxN design matrix, U is the MxM orthogonal matrix whose columns generate the data space; V is the NxN orthogonal matrix whose columns generate the model space; S is an MxN diagonal matrix with non-negative diagonal elements called singular values. The singular values for the geometry of Figure 1 are shown in Figure 3. As expected, among the 685 singular values, two are numerically close to zero (i.e. of the order of 10^{-16}), confirming that the model null space (kernel) has dimension (nullity) equal to 2. The right panel of Figure 3 shows the columns $V_{,684}$ and $V_{,685}$ of the matrix V, which are associated with the null singular values (i.e., the singular values are ordered in decrescent order). These two columns form an orthonormal basis for the kernel of the design matrix. Each element of the basis consists of a constant value on each column block related to event, station and attenuation, with the sum of the three constant values equal to 0. They represent the trade-off existing among the source, station

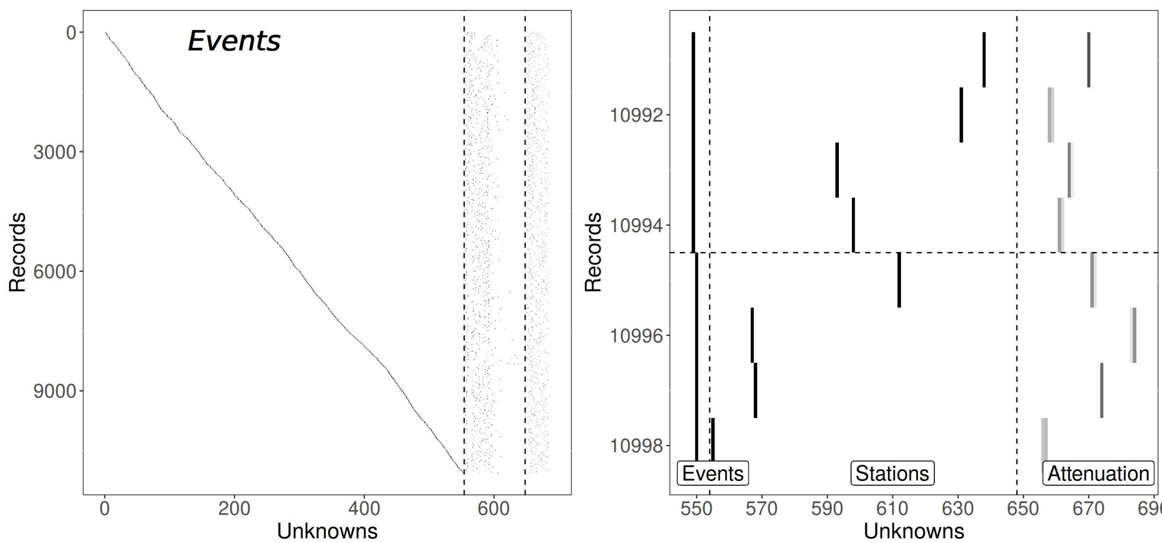

Figure 2. Design matrix (left) for the network geometry shown in Figure 1. Columns from 1 to 554 are relevant to events; columns from 555 to 648 to stations; columns from 649 to 685 to attenuation. The right panel shows a detail of the design matrix for selected rows and columns.

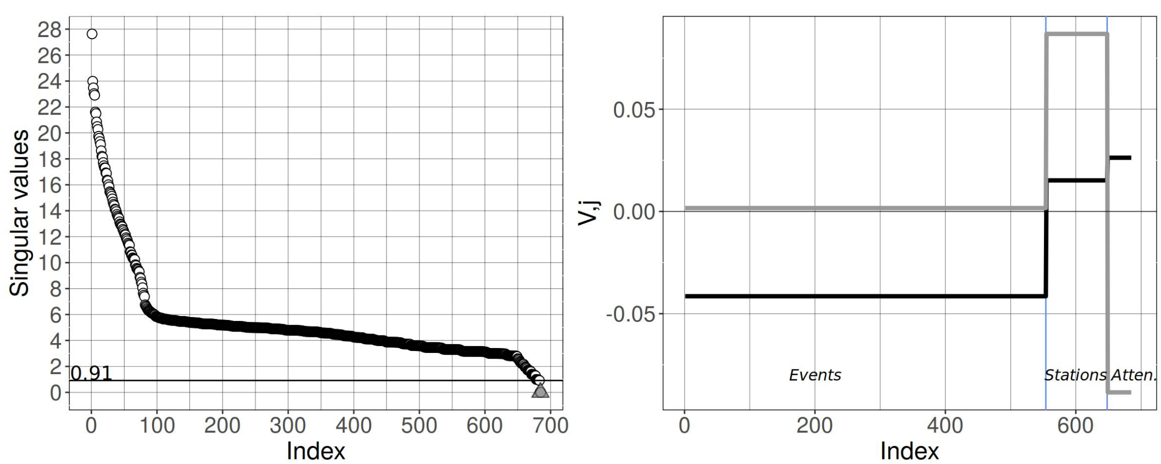

Figure 3. Left. Singular values of the design matrix compiled for the data set in Figure 1; the two null singular values are shown as gray triangles. Right. Base vectors for the null space corresponding to columns $V_{,684}$ and $V_{,685}$, where V is the right orthonormal matrix of the SVD decomposition (equation 3).

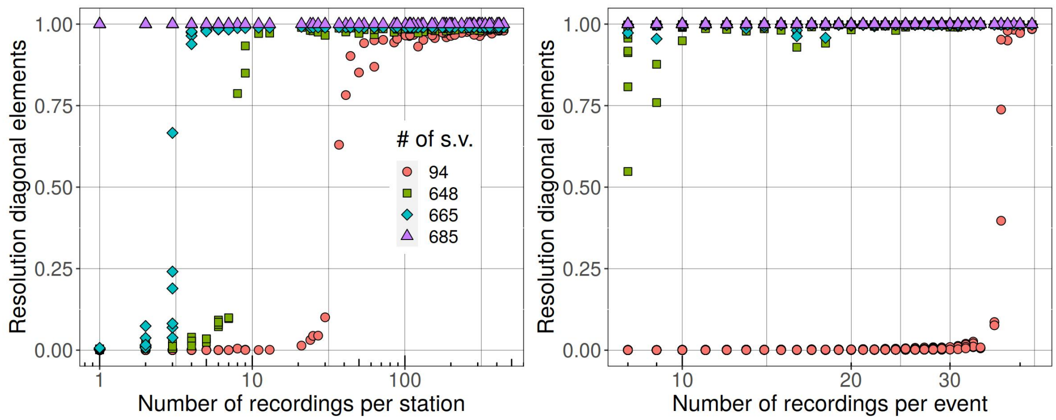

Figure 4. Diagonal elements of the resolution matrix versus number of recordings for each station (left) and event (right) considering different cutoff thresholds applied to singular values.

and propagation terms and without additional information (constraints), the solution of the system 1 is not unique. This point is further discussed later. The non-null singular values vary from 0.91 to 27.63 (Figure 3). To investigate the connection between singular values and the resolution of the model parameters, Figure 4 shows the diagonal elements of the resolution matrix R computed considering only the largest q singular values (truncated SVD):

$$R = V_q V_q^T \tag{4}$$

where V_q is the Nxq matrix composed by the first q columns of V (equation 3). For q=n=685, R is the identity matrix (overdetermined system), as well as for q=684 (since the last two singular values are zero). For q <684, the truncation removes model space basis vectors associated with small singular values (regularization). In Figure 4, we consider the cases q=94, 648, and 665 where the first two cut-offs correspond approximately to the sharp changes in the convexity of the distribution of singular values (Figure 3, left). Figure 4 shows that the regularization affects mostly those model parameters (station on the left and events on the right) with low sampling; the first 94 elements allow to resolve well model parameters with more than 40 recordings; when the number of singular values is increased to 648, parameters with at least about 10 recordings are reasonably well resolved; considering 665 singular values, model parameters with at least 5 recordings are resolved. In solving system 1 a threshold is generally applied on the minimum number of recordings per station and per event. For example, for the decomposition performed in the framework of the Stress Drop benchmark, we set the threshold to 6. Other important information associated with the GIT design matrix is provided by the covariance matrix (Boatwright et al., 1991; Bindi et al., 2006). Figure 5(a) shows the diagonal elements of the covariance matrix, i.e., the sample variances of the model parameters. Since the elements of the design matrix in the event and station columns are either 0 or 1, and since the number of rows M is much larger than the total number of ones in each column, the variances increase almost linearly with the sampling of each station or event. The off-diagonal elements of the covariance matrix provide information about the trade-offs (correlation) between different model parameters that jointly contribute to generating the observations. These trade-offs are determined by the geometry of the problem. Figure 5(b) shows that there is a trade-off between station and attenuation model parameters. For example, Figure 5(c) focuses on model parameter 560, corresponding to station CI.WBM.HH. The recordings available for this station are in the distance range of 25 to 50 km, with median and mode hypocentral distance equal to about 35 km. Figure 5(d) confirms that there is a trade-off between the station term and the attenuation coefficients in the distance range sampled by the station, in particular for the mostly sam-

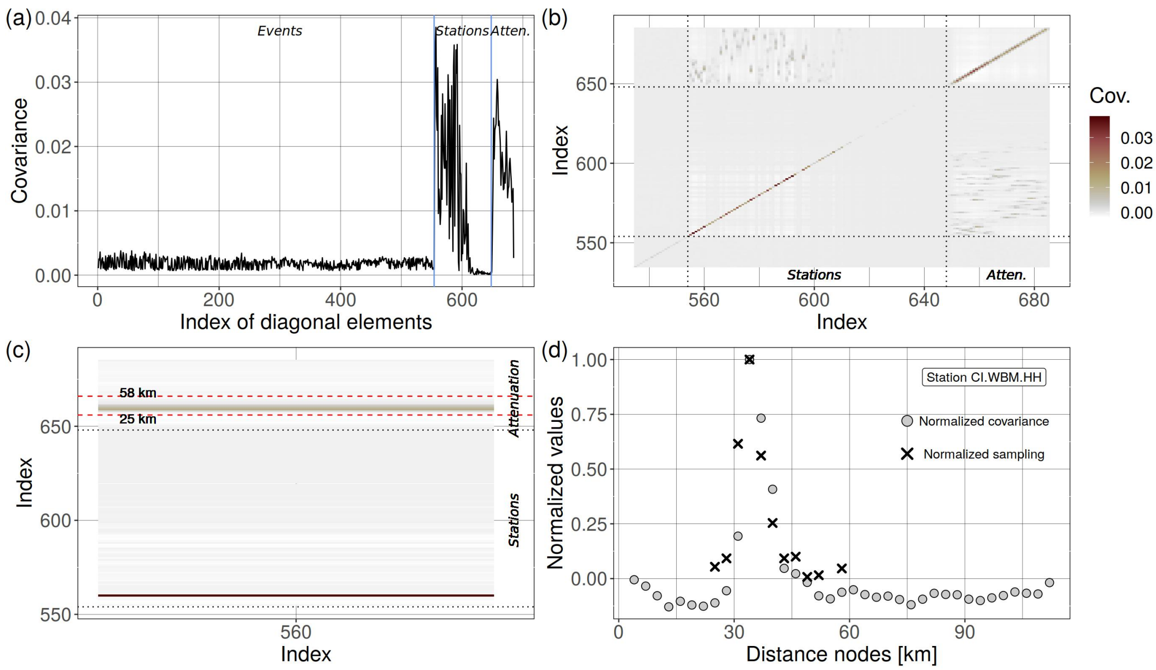

Figure 5. (a) Diagonal elements of the covariance matrix. (b) detail of the covariance matrix considering the station and attenuation columns. (c) detail of the covariance matrix for station CI.WBM.HH (column 560) considering only the station and attenuation portion of the matrix. Red dashed lines indicate the distance range sampled by recordings at CI.WBM.HH. (d) circles indicate the normalized covariance off-diagonal entries between station CI.WBM.HH and the attenuation columns; crosses indicate the normalized number of recordings for station CI.WBM.HH available within each distance bin.

pled distances. In order to limit the trade-offs, it is important that different stations sample the same distance bins, avoiding that a few stations dominate specific distance intervals, and that each station provides recordings over wide distance and magnitude intervals (Shible et al., 2022).

4 Approaches to remove the non-uniqueness of the solution

Different strategies can be followed to select a specific solution among the infinite possible ones. In the following, we discuss two widely used approaches: the first one consists in constraining the site term within the design matrix; the second one makes assumptions on the source parameters for some reference events that in turn are used to correct the source spectra provided by the unconstrained decomposition.

4.1 Constraining site and attenuation terms

The first strategy consists in adding some rows to the design matrix, forcing the solution to satisfy specific conditions that make the matrix full rank. A typical choice is to constrain to zero the average of the logarithm of all site amplifications. In this way, the site effects do not bias the observations on average but contribute to the overall variability. A similar constraint is often applied when calibrating a local magnitude scale using data a seismological network (with stations mostly installed on rock) by setting to zero the average of the station magnitude corrections (Bakun & Joyner, 1984; Savage

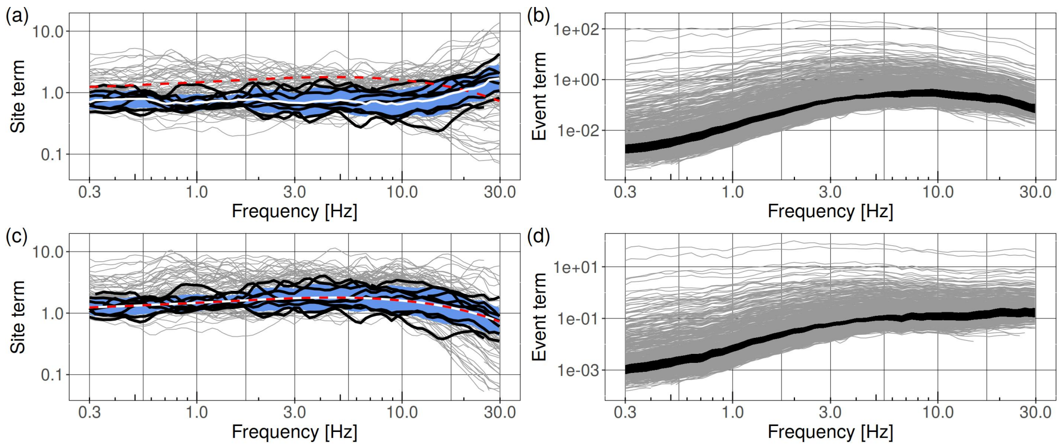

Figure 6. Impact on the source spectra of constraints applied to the site term. a) Site amplifications obtained by setting to 1 the average of all site terms. b) source spectra relevant to the site terms as in panel a. c) site amplifications obtained by constraining the average (white curve) of selected rock sites (black curves) to be identical to the spectral function shown as a dashed line; the same reference sites (black) and their average amplification (white) ± one standard deviation are also shown in panel a. d) source spectra relevant to the site terms in panel c. In panels b and d, the black curves are source spectra of magnitude 3 events.

& Anderson, 1995; Langston et al., 1998; Baumbach et al., 2003). Since there are two null singular values, a second constraint is applied to the attenuation by requiring that the attenuation assumes a given value at a reference distance, i.e, logP_k=0 for $R_k = R_{ref}$. As a consequence, the source spectra is scaled at R_{ref}. For the case study analyzed in this study, we compare the impact of applying two different constraints on the site term, and we set to 0 the logarithm of the attenuation at 10 km. The first site constraint coincides with the standard requirement that the average of the logarithm site amplifications is 0 for all frequencies. As expected, this requirement together with the constraint applied to the attenuation term, removes the two null singular values and the condition number of the design matrix is now $\sim$ 2543. Since the considered stations are installed in different geological settings, the site amplifications (Figure 6a) show a large variability, in particular above 10 Hz. The applied constraint implies that the site terms represent the site amplifications with respect to the network average. If the average amplification deviates from the imposed flat spectral behavior, the deviation from the true average is moved to the source spectra (Figure 6b). The acceleration source spectra show an average high frequency decay different from the flat asymptote predicted by the omega-square source model, and the absorbed average site amplification contributes to the observed high-frequency spectral fall-off. In Figure 6a, the amplification of 6 stations installed on rock (vs30 values measured or inferred from geology above 710 m/s) are shown in black. They have a flat amplification lower than the network average at frequencies smaller than about 10 Hz, and then they show a positive amplification trend with frequency. The positive trend indicates that the near-surface attenuation for these stations (k_0 parameter, (Anderson & Hough, 1984)) is probably weaker than the average near surface attenuation of the overall network. Therefore, following Bindi, Zaccarelli, and Kotha (2020), in the second reference site strategy we constrain the average amplification of the six stations to the crustal amplification proposed by Campbell and Boore (2016) for the National Earthquake Hazards Reduction Program (NEHRP) B/C boundary multiplied by an exponential term with k_0=0.034 s (Figure 6c, dashed line). The source spectra obtained using this constraint (Figure 6d) show an average flat high frequency spectral level, although single spectra have still small positive and negative slopes (k_{source}) that have been shown to correlate well with the between-event ground motion variability at high frequencies (Bindi, Zaccarelli, & Kotha, 2020). If we compare the logarithm of the ratio between the average spectra for magnitude 3 events with the logarithm of the inverse average site amplifications for the reference stations, they are identical (Figure S1 of the Supplements). This confirms that the constraint applied to the site amplifications breaks the trade-off between source and site terms but the constrained solutions provide the same predictions as the unconstrained solutions and the analysis of the residuals alone is not sufficient to discriminate among the quality of solutions obtained by applying different constraints.

4.2 Reference source conditions

A different approach followed in the context of source studies is to correct the obtained source spectra $S_j(f)$ after the decomposition. The correction is performed by using additional information about source parameters and making assumptions about the source spectral shape (e.g., Baltay et al., 2011; Walter et al., 2017; Trugman & Shearer, 2017; Morasca et al., 2022). For example, if the source spectra are assumed to be omega-square, the source spectra can be scaled by fixing the seismic moment and the corner frequency of one or more calibration events (reference source condition). In the following, we discuss some aspects connected to recently applied reference source conditions (e.g., Shearer et al., 2022). It is worth noting that if the source spectra are described in terms of a spectral model, then the model can be introduced in equation (1) making the GIT approach parametric. In this case, the system can be solved to directly determine the source parameters (e.g. seismic moment and corner frequency) and the constraint on seismic moment and corner frequency can be included directly in the design matrix (e.g., Moya & Irikura, 2003). In this case, the complexity of the source models generally makes the system non linear.

5 Impact of constraining the corner frequency of calibration events

In the reference source strategy, reference moment magnitudes are used to remove the bias on the estimated seismic moments (shift of the source spectra), whereas an Empirical Correction Spectrum (ECS) is determined to correct the shape of the decomposed source spectra (e.g., Shearer et al., 2006; Trugman & Shearer, 2017; Shearer et al., 2022). The spectral shape of the source spectra is described in terms of a standard model (e.g., ω-square) and the corner frequency of selected events is assumed to be known. Typically, the corner frequency is fixed for small events, hereinafter referred to as global empirical Green's functions (gEGF). For example, in recent applications for southern California, the corner frequency of the gEGF with local magnitude 1.5 was fixed to 30 Hz (Shearer et al., 2022) whereas Chen and Abercrombie (2020); Zhang et al. (2022) let the data to select the constant value.

We consider a simplified approach where the errors arising from the decomposition are neglected, as well as the propagation of experimental errors from data to solutions. We assume that the source spectra isolated by the spectral decomposition are given by the product between the true source spectra S and the unknown ECS. The latter is estimated by considering small earthquakes (also called calibration events) representing the gEGF, whose corner frequency is a-priori fixed to a certain value, that is:

$$ECS_{est} = \frac{gEGF \cdot ECS}{g\widehat{EGF}(f_c = \hat{f}_c; shape = \omega^2)} \tag{5}$$

In equation 5, the term $gEGF{\cdot}ECS$ represents the $gEGF$ source spectrum as provided by the decomposition; therefore, an estimate of the ECS (indicated as ECS_{est}) is obtained by removing $g\widehat{EGF}$ from this term. The quantity $g\widehat{EGF}$ represents the a-priori assumed spectrum for the $gEGF$ (in our case, an omega-square spectrum with corner frequency fixed to $\hat{f}_c$ and a known seismic moment). The ECS_{est} spectrum is then used to correct the spectra of the target events provided by the decomposition and get their source spectra, that is:

$$S_{est} = \frac{S \cdot ECS}{ECS_{est}} \tag{6}$$

where S is the true source spectrum of the target event; $S{\cdot}ECS$ is the source spectrum of the target event as provided by the decomposition; S_{est} is the estimated source spectrum of the target event. Substituting equation 5 into equation 6, we obtain

$$S_{est} = \frac{g\widehat{EGF}_{\hat{f}}}{gEGF} S \tag{7}$$

where $g\widehat{EGF}_{\hat{f}}$ is the spectrum assumed for the calibration event and $gEGF$ its true spectrum. Equation 7 shows that in our numerical tests the shape of the ECS does not play any role (since it affects the target and the calibration spectra in the same way) and the quality of the retrieved target source spectrum is controlled by the ratio between the assumed and the true spectra of the calibration event. Even if the shape of the ECS is not entering in equation 7, in order to show its impact on the ECS and source target estimations (as in equations 5 and 6) in the following we assume an exponential shape for describing the ECS:

$$ECS(f) = e^{\pi k f} \tag{8}$$

The spectral shape of equation 8 resembles typical shapes of ECS shown in literature (e.g., Shearer et al., 2006, 2022).

5.1 Numerical Test

To evaluate the impact of assuming $f_c = \hat{f}_c$ for the gEGF, we perform a numerical test considering: 1) a synthetic distribution of events; 2) uncertainty on seismic moment; 3) variability of the stress drop for the same seismic moment. The components are the following:

- Event distribution. We generate a distribution of events following a Gutenberg-Richter magnitude distribution with minimum magnitude M_{min}=1.8 and maximum magnitude M_{max}=6.5; we generate a catalog of 10000 earthquakes with b-value equal 0.6 (we set the b to a low value to increase the number of larger events). The distribution is shown in Figure S2 of the Supplements.
- Uncertainty on magnitude. We perturb the magnitude values by adding random numbers extracted from a normal distribution centered on the true magnitude value and with standard deviation equal to 0.05 magnitude units.
- Variability on $\Delta\sigma$. We allow the stress drop to vary for the same seismic moment. The applied perturbation is as follows

  1. for each Mw, we compute $Mw_{err} = \mathcal{N}(Mw, 0.05)$ (Figure 7a); from Mw and Mw_{err}, we compute the seismic moments Mo and Mo_{err}, respectively.
  2. we assume three different scaling models, referred to A, B and C, between seismic moment and corner frequency as shown in Figure 7b; each scaling corresponds to a different dependency of the stress drop $\Delta\sigma$ on seismic moment, being scaling A deviating with the largest error and scaling C being compatible on average with the gEGF assumption.
  3. for each scaling, we compute the true corner frequency from Mw and $\Delta\sigma$ using the formula for circular rupture model with uniform stress drop (Brune, 1970, 1971);
  4. we add a variability over $\Delta\sigma$ as $\Delta\sigma_{var} = \mathcal{N}(\Delta\sigma, 0.15{*}\Delta\sigma)$ (Figure 7c) . The normal distribution is truncated below $0.5{*}min(\Delta\sigma)$ and above $2{*}max(\Delta\sigma)$.

Figure 7d shows the scaling considering the errors on M_o and the variability on $\Delta\sigma$.

- average EGF. Following previous studies (e.g. Shearer et al., 2022), a set of candidate calibration events is selected by considering events located close to the target event and with magnitudes similar to the selected reference magnitude. We assume that the reference magnitude is 2 and we select from the catalog events having magnitude in the range 1.99-2.01 (we do not introduce errors in the numerical analysis errors due to differences in location). The generated catalog contains 202 candidates. Since we use Mw_{err} to perform the selection, the error on magnitude contributes to the variance of the final residuals. Among the 202 candidates, we randomly select 10 candidates and we average their spectra.
- ECS. For producing figures of intermediate results, we consider $k = 0.025$s (equation 8) and we assume $\hat{f}_c$=30 Hz for the magnitude 2 gEGF. The 202 available gEGFs, the 10 selected, and the average gEGF are shown in Figure S3 considering the scaling case A of Figure 7.

5.2 Results of the numerical test

Figure 8 shows the results obtained considering the three different source scaling of Figure 7. As expected, when the corner frequency assumed for the $\widehat{gEGF}_{\hat{f}}$ deviates from the true unknown value (as in the case of scaling A and B), the overall source scaling is biased producing, in our case, corner frequencies higher than those used to generate the synthetic data. The overestimation is propagated also to magnitude larger than the magnitude of the $gEGF$, as shown in Figure 8b: for magnitudes close to the $gEGF$ magnitude, the overestimation is determined by the suitability of the $\widehat{gEGF}_{\hat{f}}$ assumption; for magnitudes larger than the $gEGF$ magnitude, the overestimation decreases and converge to an asymptotic average relative error of the order of 70 and 25 % for scaling A and B, respectively. This apparent rotation of the overall source scaling caused by a larger bias for small than for large events, impacts also the seismic moment. Whereas the variability for a given magnitude is controlled by the error added to Mw, Figures 8d and 8f show that the impact of the wrong assumption on the $\widehat{gEGF}_{\hat{f}}$ is propagated to larger magnitudes, with scaling A and B generating an average Mo underestimation above magnitude 5 of about 40 and 10 %, respectively. Figure 9 exemplifies the spectral fitting for two synthetic events with magnitude 2 (left) and 5.9 (right). Results for source scaling A are shown in the top panels, those for scaling C in the bottom ones. The top left panel shows the worst case where the wrong assumption on f_c affects the $\widehat{gEGF}_{\hat{f}}$ spectrum at frequencies close to the corner frequency of the source S. Moreover, the errors added to Mw introduce an overall bias to the retrieved source spectrum S_{est}. Indeed, the final fit is very good (small residuals) but with significant deviations on both f_c and Mo with respect to the true values. In other words, the precision on S_{est} is high but its accuracy with respect to S is low. When scaling C is considered (bottom left panel), $\widehat{gEGF}_{\hat{f}}$ is similar to the true $gEGF$ and the high frequencies of the source spectra S are well retrieved. For magnitude 5.9 (right panels), even if the corner frequency of the source is much smaller (0.7 Hz for the case shown in Figure 9) than f_c of $\widehat{gEGF}_{\hat{f}}$, Figure 8b and 8d already showed that scaling A and B produce biased values for both Mo and f_c. Whereas the added random error on Mw controls the variability of the Mo estimates, the error on the $\widehat{gEGF}_{\hat{f}}$ corner frequency has an impact on the overall shape of the source spectrum, leading to biased estimates for both fc and Mo (Figure 9, top right panel). Figure 10 exemplifies the impact of the f_c assumption on the spectral shape considering two events with magnitude 3.14 and 4.08. The left panels show the results obtained considering $n = 2$ (omega square), i.e., the same value used for generating S; in the right panels, the fit performed by lowering n to 1.6 improves the quality of the fit being the standard deviation of the residuals reduced by about a factor one-third (e.g., from 0.09147 to 0.03032 in the case of the magnitude 4.68 event). Therefore, the error on $\widehat{gEGF}_{\hat{f}}$ led

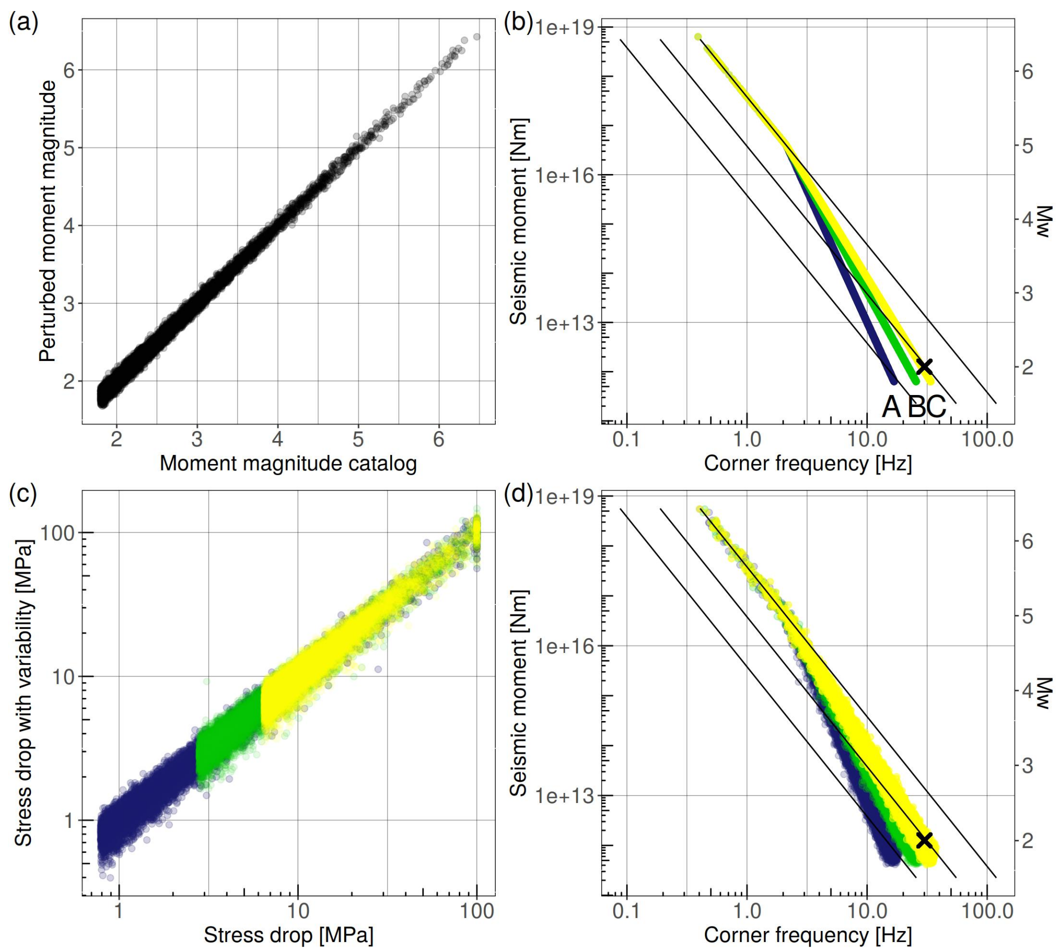

Figure 7. Source parameters used for the numerical test on the impact of constraining the corner frequency of small events. (a) uncertainty added to the magnitude values generated from the Gutenberg-Richter distribution in Figure **??**; (b) assumed source parameters for the gEGF (cross) and three different stress drop scaling used for the numerical test (A is the most deviating scaling from the cross assumption; C is the most compatible scaling; B is the intermediate case); (c) variability added to the stress drop values for the same seismic moment; colors represent the three scaling in panel (b). (d) scaling accounting for the stress drop variability (panel c) and magnitude uncertainty (panel a).

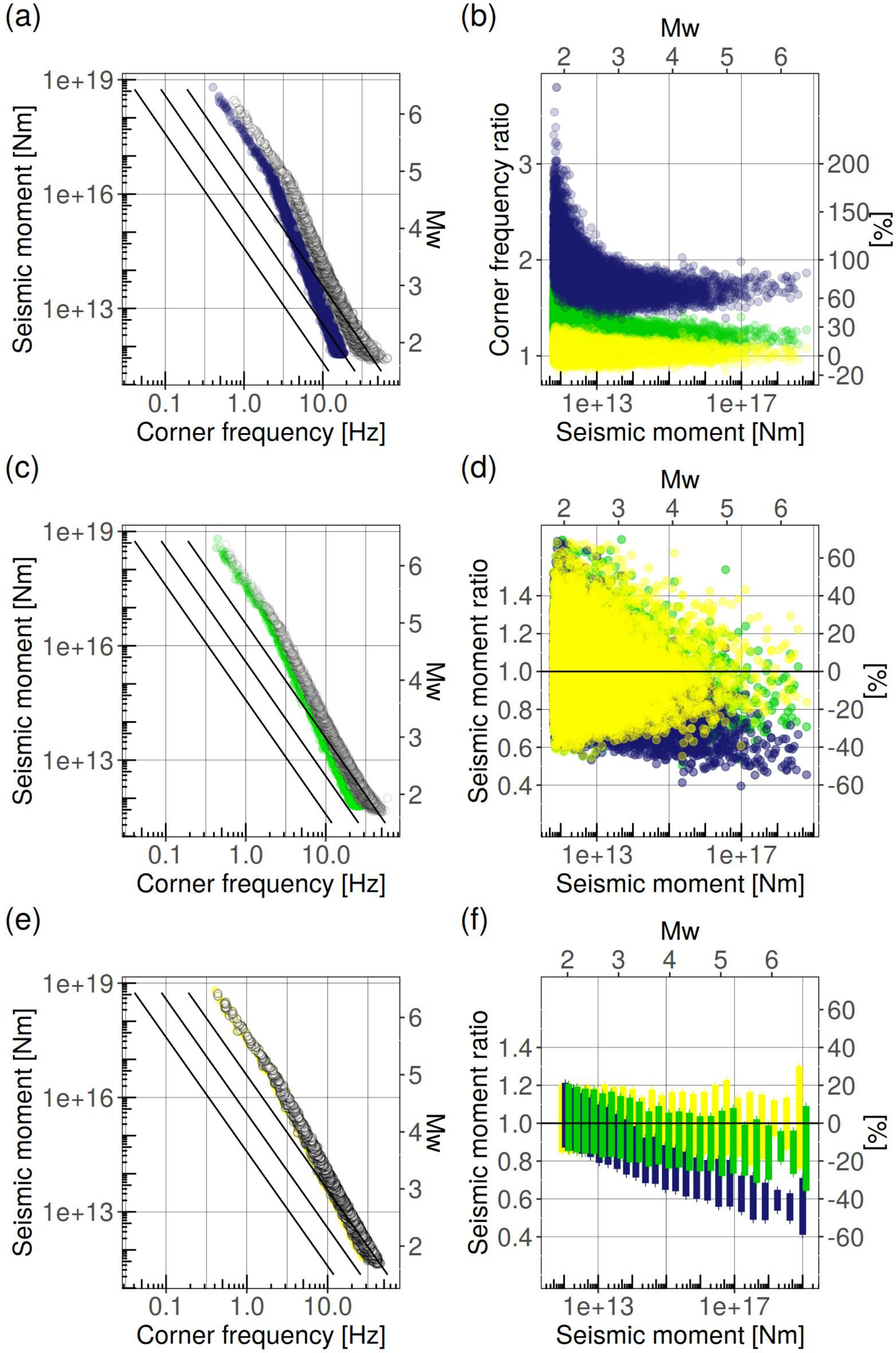

Figure 8. Panels (a), (c) and (e) show the comparison between the retrieved scaling (white) and the simulated ones (colors represent scaling A, B and C as in Figure 7); results in terms of relative error for the corner frequency and seismic moment are shown in panels (b) and (d), respectively. Panel (f) shows the mean ± 1 standard deviation of the relative errors on seismic moment computed over different seismic moment bins.

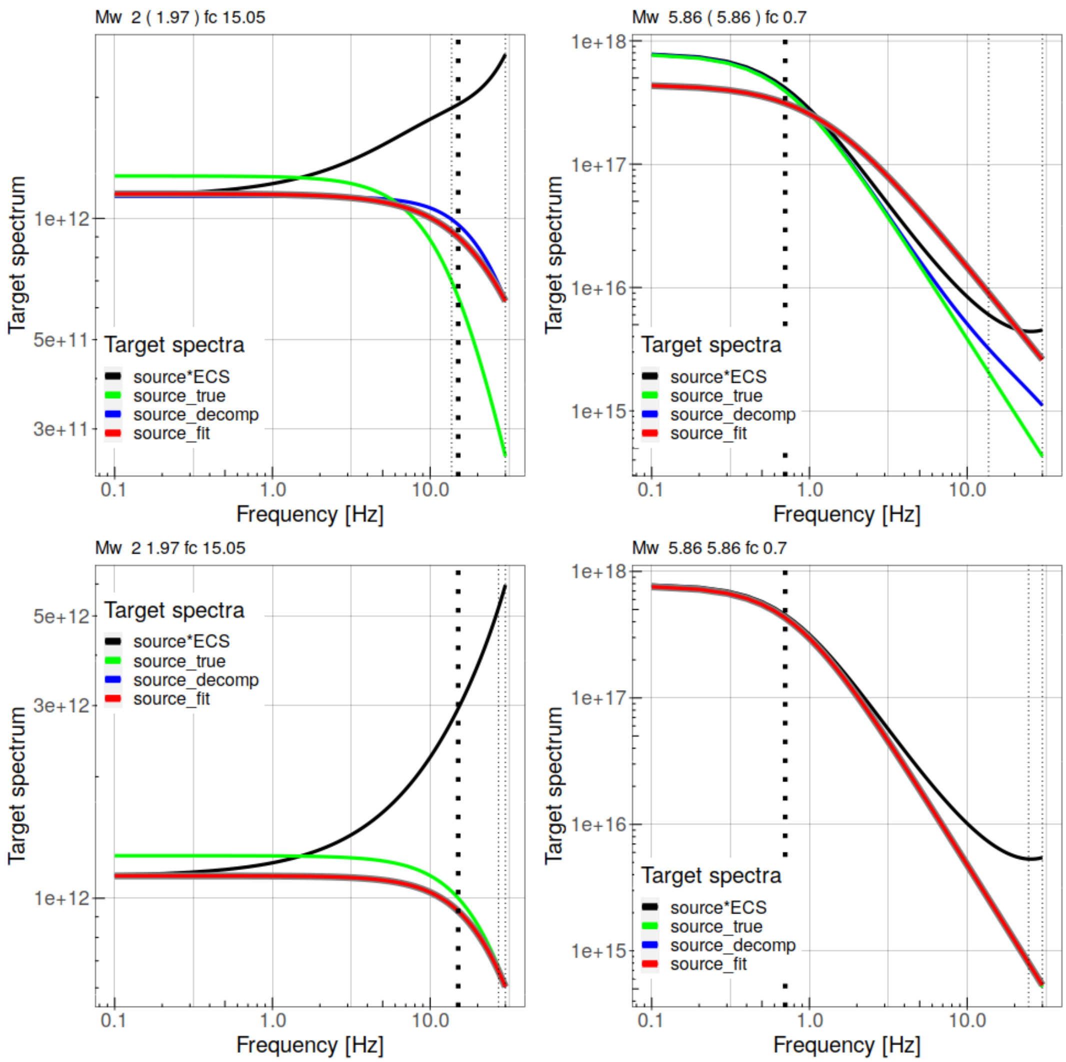

Figure 9. Examples of the numerical test performed to evaluate the impact of constraining the corner frequency of small events. Results are exemplified in terms of spectral fitting considering a synthetic magnitude 2 (left) and a magnitude 5.9 (right) event. Top: results for the source scaling A of Figure 7 (worst case); bottom: results for the source scaling C of Figure 7 (best case). In the legend, source*ECS indicates the numerator of equation 6, right hand side; source_true is the true source spectrum S; source_decomposed is S_{est} (left hand side of equation 6); source_fit is the final results, i.e., the best fit omega-square model of S_{est}

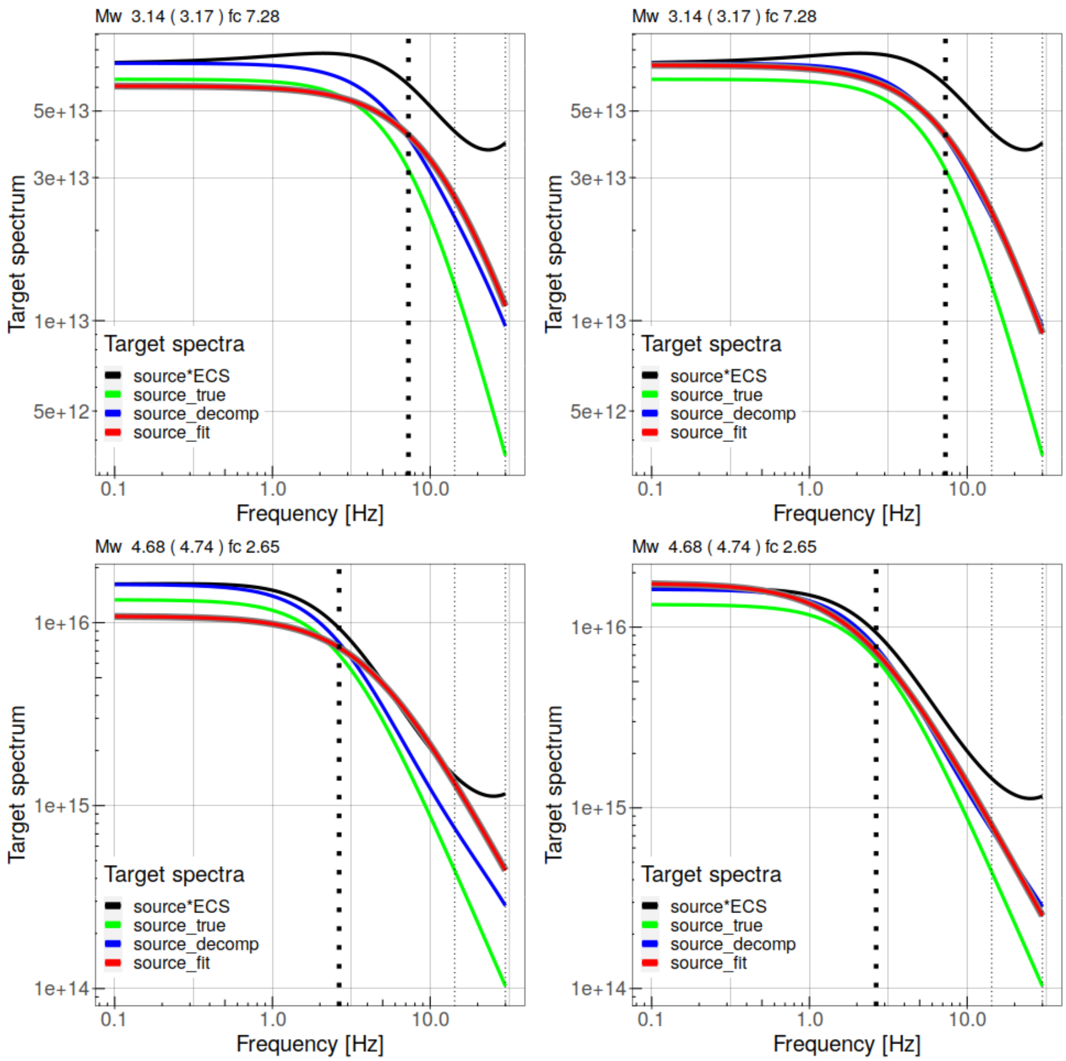

Figure 10. Impact of considering $n = 2$ (left) and $n = 1.6$ (right). As in Figure 9, the true source spectrum S is indicated in legend as source_true, the estimated source spectrum S_{est} as source_decomp, the best fit source spectrum as source_fit, and the source*ECS in black. Using $n = 1.6$, the quality of the fits improves although the correct source model used to generate the synthetics is $n = 2$.

to source spectra S_{est} with high-frequency slopes that deviate from the (true) omega-square model, allowing wrong models to fit the empirical source spectra better than models having the correct shape. This result confirms that we cannot discriminate among alternative source models by looking only at the quality of the fit.

6 Discussion

The main goal of this study was to recall some peculiarities of the non-parametric spectral decomposition approach and to discuss the consequences of strategies applied to select specific solutions. We developed our discussions by analyzing the characteristics of the design matrix compiled using a data set extracted from the Ridgecrest Stress Drop benchmark (Baltay et al., 2021). The most striking characteristic of the considered decomposition approach is the block-wise trade-offs among the source, propagation, and site terms, which results in a rank-deficient design-matrix with two null singular values. This characteristic is intrinsic to the decomposition approach, regardless of the number of available recordings or the model assumptions imposed on the unknown terms. By performing the singular value decomposition of the design matrix, we depicted the trade-off for the considered network geometry and earthquake catalog by looking at the singular vectors associated to null singular values. Since the rank deficiency implies the non-uniqueness of the solution (i.e., the matrix has a non-trivial kernel), solutions can only be selected relative to a-priori assumptions which generally vary from study to study. This is especially critical when comparing the results of different studies performed on the same data set, as in the case of the Ridgecrest stress drop benchmark. To avoid systematic differences due to the assumption of different reference conditions, it is necessary to consider procedures for comparing results in terms of relative quantities calculated within each study (Pennington et al., 2021), in addition to comparing absolute results between studies. Constraints to restore the uniqueness of the solution can be applied either to the source or to the site terms and, as a consequence of the block-wise trade-offs, they affect the unconstrained blocks as well. We have shown that the site constraint applied to the average amplification of selected reference stations corresponds to an average bias applied to the source spectra. Whereas information about site classification and empirical analysis can be used to design reliable site constraints, and to anchor the spectral decomposition models to site conditions of interest for engineering seismology applications, additional information on source characteristics are needed to assess, and possibly correct, biases transferred to the source. For example, seismic catalogs of moment magnitudes can be used as reference to anchor, on average, the seismic moments measured from the source spectra, by adding an overall vertical offset to the spectra (e.g., Bindi, Zaccarelli, & Kotha, 2020); the assumption of reference source spectral shapes (e.g., omega-square) can be used to evaluate the impact of the reference site condition at high frequencies by measuring the deviation from the expected acceleration spectral trend (e.g., Oth et al., 2011). Previous studies measured the deviation from the reference source model by coupling the omega-square model with a frequency dependent exponential term in which the exponent k_{source} parameter is controlling the high-frequency spectral decay. Investigations showed a good correlation between event-specific ground motion residuals and the k_{source} distribution (Bindi et al., 2019) suggesting that near source attenuation effects or differences in high-frequency radiation efficiencies contribute to the aleatory variability associated with ground motion prediction equations. Another common strategy for constraining the spectral decomposition solutions is to add a-priori information about the sources. To make solutions compatible with a-priori knowledge, source parameters of selected events (sometimes called reference events, ground truth events, or empirical Greens function) are imposed. Since parameters such as seismic moment, corner frequency, stress drop, and apparent stress are generally considered, spectral shapes and rupture models are also assumed, such as the omega-square and Brune models. We investigated the impact on the overall source scaling of constraining the corner frequency of small reference events by performing numerical simulations. Besides the expected propagation of biases to corner frequencies and stress drops of small events when the assumption deviates from the correct values, numerical tests showed that even the source parameters of larger magnitudes are affected by the assumptions applied to small events. Even when the corner frequencies of large events are small enough to fall far below the spectral range affected by the approximations on the corner frequency of small events, wrong assumptions for small events introduce a distortion of the source spectra of large events. As a result, a fit performed using the correct spectral shape results in source parameters different from those simulated. In addition, the quality of the fit from the minimization of the residuals point of view can be improved by modifying the source model. It follows that incorrect source models might be preferred to true models, suggesting false deviations from the omega-square shape assumed for the simulations. Finally, it is worth mentioning the impact of model assumptions related to different terms in the decomposition, particularly those concerning source and propagation. As shown in equation 1, the spectral decomposition provides a single source spectrum corresponding to the least-squares solution with respect to recordings available at several stations. Since azimuthal dependencies are not taken into account, the retrieved spectra are assumed to represent spherically averaged quantities. Therefore, a good azimuthal and distance coverage is important to avoid systematic biases at low frequencies due to unbalanced sampling of radiation pattern effects and to avoid corner frequency biases due to variability in source geometry, rupture directivity and rupture speeds (Kaneko & Shearer, 2015). Comparison of GIT source spectra with the distribution of the apparent source spectra obtained by correcting the observations for GIT propagation and site effects can provide useful indications for identifying possible azimuthal and distance-dependent effects (Bindi, Spallarossa, et al., 2020; Morasca et al., 2022). Because the back-scattering effects that dominate the coda wavefield make coda analysis less sensitive to radiation pattern effects (Aki, 1981; Mayeda & Malagnini, 2010), comparing GIT results with those from coda is another effective strategy to quantify the impact of the spherical averaging hypothesis. Regarding the propagation term, although few studies have implemented more complex parametrizations for describing the attenuation term (e.g., Scherbaum, 1990; Edwards et al., 2008; Koulakov et al., 2010), it is generally assumed to be controlled only by frequency and hypocentral distance. In the presence of significant lateral crustal variability, unbalanced station distributions and clustered seismicity could favor the transfer of attenuation effects to either the source or site terms. The analysis of the GIT residuals can reveal the presence of spatial anomalies that could, for example, hint towards the presence of areas with attenuation different from the average attenuation in the studied region. Considering the region in Figure 1, stronger than average attenuation can be expected for the Coso volcanic area and weaker for the Mojave desert along the Garlock fault (Hauksson & Shearer, 2006; Bindi et al., 2021, their Figure 10). Besides anisotropy in the propagation medium characteristics, other phenomena not captured by the models used in equation 1 can generate coherent azimuthal patterns in the residuals, such as directivity of the rupture. Recent investigations have also detected directivity effects for small and moderate earthquakes (e.g., Galloviĉ, 2015; Pacor et al., 2016; Colavitti et al., 2022) with significant modification of the source spectrum around the corner frequency. Therefore, good station arrangement and high redundancy of the station and event sampling are needed to mitigate the impact of the model assumption and misinterpretation. The characteristics of the data set contribute also to determine the properties of the covariance matrix. We have shown that the off-diagonal elements of the covariance matrix inform about trade-offs between existing between different couples of unknowns. For example, we showed that the trade-offs between attenuation and site amplification concentrate over the distance range sampled by the considered station. Again, the quality of the data set in terms of recordings well distributed with distance and azimuth and high level of cross-information between stations and events are fundamental to mitigate the impact of trade-offs generated by competing physical effects.

7 Conclusions

We investigated the non-uniqueness of the spectral decomposition approach applied to isolate source, propagation and site contribution to measured S-waves spectra. We related the non-uniqueness to the characteristics of the design matrix exemplifying the analysis for a data set extracted from the Community Stress Drop Validation Study (Baltay et al., 2021). We also discussed the impact on the final retrieved source scaling of constraints on site and the source terms applied to select specific solutions. We can summarize the main conclusions as follows:

- The problem of isolating source, propagation and site effects starting from a large compilation of S-wave amplitude Fourier spectra has a non-unique solution due to two unresolved degrees of freedom, regardless of the number of available recording, characteristics of the source catalog and network geometry; the non-uniqueness of the decomposition is intrinsic to any model framework where seismic measurements in either time or frequency domains are expressed as combination of source, propagation and site contributions;
- solutions can be only retrieved relative to a-priori constraints applied to break the numerical trade-offs affecting the three blocks (i.e., to fix the scale of the different blocks by assuming some reference anchoring points). Since the retrieved solutions are relative to these assumptions, comparison among results provided by different studies should be performed evaluating relative quantities (ratios or log differences) in order to remove the impact of the assumptions on the absolute values;
- Non-parametric spectral decompositions require data sets with high level of redundancy, i.e., each event recorded by several stations and each station recording several events; the covariance matrix allows to highlight the physical trade-offs between different unknowns, such as the site amplifications at any station and the attenuation coefficients for the distance range sampled by the station;
- Constraints applied to the site term have an impact on the source term; the selection of the reference stations and the imposed average amplification determines the average amplitude level of the source block: independent seismic moment values from a reference catalog should be used to correct for a possible bias. Moreover, the spectral shape of the reference amplification function has an impact on the average source spectral shape (e.g., on the average high frequency fall-off) that should be taken into account when the source spectra are fitted to a specific source spectral model.
- Numerical tests showed that assuming an a-priori value for the corner frequency (or stress drop) of small events (empirical Greens functions or reference events) can have a strong impact on the overall seismic moment versus corner frequency source scaling, introducing a bias for small events; the bias is also propagated to larger magnitudes but with reduced amplitude, generating an apparent rotation of the overall source scaling that could lead to wrong interpretations about the self-similarity of the rupture scaling. Moreover, the corner frequency assumption can modify the spectral shape of larger events, leading to best-fit source model shapes different from the correct one (e.g. with high frequency fall-off different from the values used for generating the synthetic data); consequently, alternative source models cannot be discriminated against solely on the basis of the quality of their fit.

In conclusion, we have shown that the well-known property of the decomposition approach of not having a unique solution can have a strong impact on the final interpretation of results, thus motivating our effort to raise awareness on this important topic for modelers and model users.

Open Research

This work has been inspired by the Community Stress Drop Validation Study (Baltay et al., 2021) supported by the Southern California Earthquake Centre SCEC (https://www.scec.org/research/stress-drop-validation). The data set used in this study is disseminated through the benchmarks web-page. Routines for the simulations used for testing the source constraint have been prepared with R software (R Core Team, 2020) and are available on request.

Acknowledgments

We thank R. Abercrombie, A. Baltay and Taka'aki Taira for organizing and managing the Community Stress Drop Validation Study (https://www.scec.org/research/stress-drop-validation).

References

Aki, K. (1981, 12). Source and scatering effects on the spectra of small local earthquakes. *Bulletin of the Seismological Society of America*, *71*(6), 1687-1700. Retrieved from https://doi.org/10.1785/BSSA0710061687 doi: 10.1785/BSSA0710061687

Aki, K., & Chouet, B. (1975). Origin of coda waves: Source, attenuation, and scattering effects. *Journal of Geophysical Research (1896-1977)*, *80*(23), 3322-3342. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB080i023p03322 doi: https://doi.org/10.1029/JB080i023p03322

Anderson, J. G., & Hough, S. E. (1984, 10). A model for the shape of the fourier amplitude spectrum of acceleration at high frequencies. *Bulletin of the Seismological Society of America*, *74*(5), 1969-1993. Retrieved from https://doi.org/10.1785/BSSA0740051969 doi: 10.1785/BSSA0740051969

Andrews, D. J. (1986). Objective determination of source parameters and similarity of earthquakes of different size. In *Earthquake source mechanics* (p. 259-267). American Geophysical Union (AGU). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GM037p0259 doi: https://doi.org/10.1029/GM037p0259

Bakun, W. H., & Joyner, W. B. (1984, 10). The ML scale in central California. *Bulletin of the Seismological Society of America*, *74*(5), 1827-1843. Retrieved from https://doi.org/10.1785/BSSA0740051827 doi: 10.1785/BSSA0740051827

Baltay, A., Abercrombie, R., & Taira, T. (2021). A community stress drop validation study using the 2019 ridgecrest earthquake dataset, SSA annual meeting..

Baltay, A., Ide, S., Prieto, G., & Beroza, G. (2011). Variability in earthquake stress drop and apparent stress. *Geophysical Research Letters*, *38*(6). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011GL046698 doi: https://doi.org/10.1029/2011GL046698

Baumbach, M., Bindi, D., Grosser, H., Milkereit, C., Parolai, S., Wang, R., ... Zschau, J. (2003, 10). Calibration of an ML Scale in Northwestern Turkey from 1999 Izmit Aftershocks. *Bulletin of the Seismological Society of America*, *93*(5), 2289-2295. Retrieved from https://doi.org/10.1785/0120020157 doi: 10.1785/0120020157

Bertero, M., Bindi, D., Boccacci, P., Cattaneo, M., Eva, C., & Lanza, V. (1997, apr). Application of the projected landweber method to the estimation of the source time function in seismology. *Inverse Problems*, *13*(2), 465–486. Retrieved from https://doi.org/10.1088/0266-5611/13/2/017 doi: 10.1088/0266-5611/13/2/017

Bertero, M., Bindi, D., Boccacci, P., Cattaneo, M., Eva, C., & Lanza, V. (1998, aug). A novel blind-deconvolution method with an application to seismology. *Inverse Problems*, *14*(4), 815–833. Retrieved from https://doi.org/10.1088/0266-5611/14/4/004 doi: 10.1088/0266-5611/14/4/004

Bindi, D., Parolai, S., Grosser, H., Milkereit, C., & Karakisa, S. (2006, 02). Crustal Attenuation Characteristics in Northwestern Turkey in the Range from 1 to 10 Hz. *Bulletin of the Seismological Society of America*, *96*(1), 200-214. Retrieved from https://doi.org/10.1785/0120050038 doi: 10.1785/0120050038

Bindi, D., Picozzi, M., Spallarossa, D., Cotton, F., & Kotha, S. R. (2019, 01). Impact of Magnitude Selection on Aleatory Variability Associated with Ground-Motion Prediction Equations: Part IIAnalysis of the BetweenEvent Distribution in Central Italy. *Bulletin of the Seismological Society of America*, *109*(1), 251-262. Retrieved from https://doi.org/10.1785/0120180239 doi: 10.1785/0120180239

Bindi, D., Razafindrakoto, H. N. T., Picozzi, M., & Oth, A. (2021, 06). Stress Drop Derived from Spectral Analysis Considering the Hypocentral Depth in the Attenuation Model: Application to the Ridgecrest Region, California. *Bulletin of the Seismological Society of America*, *111*(6), 3175-3188. Retrieved from https://doi.org/10.1785/0120210039 doi: 10.1785/0120210039

Bindi, D., Spallarossa, D., Picozzi, M., & Morasca, P. (2020, 07). Reliability of Source Parameters for Small Events in Central Italy: Insights from Spectral Decomposition Analysis Applied to Both Synthetic and Real Data. *Bulletin of the Seismological Society of America*, *110*(6), 3139-3157. Retrieved from https://doi.org/10.1785/0120200126 doi: 10.1785/0120200126

Bindi, D., Zaccarelli, R., & Kotha, S. R. (2020, 12). Local and Moment Magnitude Analysis in the Ridgecrest Region, California: Impact on Interevent GroundMotion Variability. *Bulletin of the Seismological Society of America*, *111*(1), 339-355. Retrieved from https://doi.org/10.1785/0120200227 doi: 10.1785/0120200227

Boatwright, J., Fletcher, J. B., & Fumal, T. E. (1991, 10). A general inversion scheme for source, site, and propagation characteristics using multiply recorded sets of moderate-sized earthquakes. *Bulletin of the Seismological Society of America*, *81*(5), 1754-1782. Retrieved from https://doi.org/10.1785/BSSA0810051754 doi: 10.1785/BSSA0810051754

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research (1896-1977)*, *75*(26), 4997-5009. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB075i026p04997 doi: https://doi.org/10.1029/JB075i026p04997

Brune, J. N. (1971). Correction [to tectonic stress and the spectra, of seismic shear waves from earthquakes]. *Journal of Geophysical Research (1896-1977)*, *76*(20), 5002-5002. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB076i020p05002 doi: https://doi.org/10.1029/JB076i020p05002

Castro, R. R., Anderson, J. G., & Singh, S. K. (1990, 12). Site response, attenuation and source spectra of S waves along the Guerrero, Mexico, subduction zone. *Bulletin of the Seismological Society of America*, *80*(6A), 1481-1503. Retrieved from https://doi.org/10.1785/BSSA08006A1481 doi: 10.1785/BSSA08006A1481

Chen, X., & Abercrombie, R. E. (2020, 06). Improved approach for stress drop estimation and its application to an induced earthquake sequence in Oklahoma. *Geophysical Journal International*, *223*(1), 233-253. Retrieved from https://doi.org/10.1093/gji/ggaa316 doi: 10.1093/gji/ggaa316

Colavitti, L., Lanzano, G., Sgobba, S., Pacor, F., & Galloviĉ, F. (2022). Empirical evidence of frequency-dependent directivity effects from small-to-moderate normal fault earthquakes in central italy. *Journal of Geophysical Research: Solid Earth*, *127*(6), e2021JB023498. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JB023498 (e2021JB023498 2021JB023498) doi: https://doi.org/10.1029/2021JB023498

Edwards, B., Rietbrock, A., Bommer, J. J., & Baptie, B. (2008, 08). The Acquisition of Source, Path, and Site Effects from Microearthquake Recordings Using Q Tomography: Application to the United Kingdom. *Bulletin of the Seismological Society of America*, *98*(4), 1915-1935. Retrieved from https://doi.org/10.1785/0120070127 doi: 10.1785/0120070127

Gallovič, F. (2015, 11). Modeling Velocity Recordings of the Mw6.0 South Napa, California, Earthquake: Unilateral Event with Weak HighFrequency Directivity. *Seismological Research Letters*, *87*(1), 2-14. Retrieved from https://doi.org/10.1785/0220150042 doi: 10.1785/0220150042

Hartzell, S. H. (1978). Earthquake aftershocks as green's functions. *Geophysical Research Letters*, *5*(1), 1-4. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GL005i001p00001 doi: https://doi.org/10.1029/GL005i001p00001

Hauksson, E., & Shearer, P. M. (2006). Attenuation models (qp and qs) in three dimensions of the southern california crust: Inferred fluid saturation at seismogenic depths. *Journal of Geophysical Research: Solid Earth*, *111*(B5). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JB003947 doi: https://doi.org/10.1029/2005JB003947

Helmberger, D., & Wiggins, R. A. (1971). Upper mantle structure of midwestern united states. *Journal of Geophysical Research (1896-1977)*, *76*(14), 3229-3245. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB076i014p03229 doi: https://doi.org/10.1029/JB076i014p03229

Ide, S., & Beroza, G. C. (2001). Does apparent stress vary with earthquake size? *Geophysical Research Letters*, *28*(17), 3349-3352. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL013106 doi: https://doi.org/10.1029/2001GL013106

Kaneko, Y., & Shearer, P. M. (2015). Variability of seismic source spectra, estimated stress drop, and radiated energy, derived from cohesive-zone models of symmetrical and asymmetrical circular and elliptical ruptures. *Journal of Geophysical Research: Solid Earth*, *120*(2), 1053-1079. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JB011642 doi: https://doi.org/10.1002/2014JB011642

Koulakov, I., Bindi, D., Parolai, S., Grosser, H., & Milkereit, C. (2010, 02). Distribution of Seismic Velocities and Attenuation in the Crust beneath the North Anatolian Fault (Turkey) from Local Earthquake Tomography. *Bulletin of the Seismological Society of America*, *100*(1), 207-224. Retrieved from https://doi.org/10.1785/0120090105 doi: 10.1785/0120090105

Kwiatek, G., & Ben-Zion, Y. (2016). Theoretical limits on detection and analysis of small earthquakes. *Journal of Geophysical Research: Solid Earth*, *121*(8), 5898-5916. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JB012908 doi: https://doi.org/10.1002/2016JB012908

Langston, C. A., Brazier, R., Nyblade, A. A., & Owens, T. J. (1998, 06). Local magnitude scale and seismicity rate for Tanzania, East Africa. *Bulletin of the Seismological Society of America*, *88*(3), 712-721. Retrieved from https://doi.org/10.1785/BSSA0880030712 doi: 10.1785/BSSA0880030712

Mayeda, K., & Malagnini, L. (2010). Source radiation invariant property of local and near-regional shear-wave coda: Application to source scaling for the mw 5.9 wells, nevada sequence. *Geophysical Research Letters*, *37*(7). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009GL042148 doi: https://doi.org/10.1029/2009GL042148

Morasca, P., Bindi, D., Mayeda, K., Roman-Nieves, J., Barno, J., Walter, W. R., & Spallarossa, D. (2022). Stable source spectra from coda envelopes and direct s-waves: Two perspectives for the determination of source parameters in central italy. *Geophysical Journal International*, *under review*.

Moya, A., & Irikura, K. (2003, 08). Estimation of Site Effects and Q Factor Using a

Reference Event. *Bulletin of the Seismological Society of America*, *93*(4), 1730-1745. Retrieved from https://doi.org/10.1785/0120020220 doi: 10.1785/0120020220

Mueller, C. S. (1985). Source pulse enhancement by deconvolution of an empirical green's function. *Geophysical Research Letters*, *12*(1), 33-36. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GL012i001p00033 doi: https://doi.org/10.1029/GL012i001p00033

Oth, A., Bindi, D., Parolai, S., & Di Giacomo, D. (2011, 04). 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. *Bulletin of the Seismological Society of America*, *101*(2), 667-687. Retrieved from https://doi.org/10.1785/0120100135 doi: 10.1785/0120100135

Pacor, F., Gallovîĉ, F., Puglia, R., Luzi, L., & D'Amico, M. (2016). Diminishing high-frequency directivity due to a source effect: Empirical evidence from small earthquakes in the abruzzo region, italy. *Geophysical Research Letters*, *43*(10), 5000-5008. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL068546 doi: https://doi.org/10.1002/2016GL068546

Pennington, C. N., Chen, X., Abercrombie, R. E., & Wu, Q. (2021). Cross validation of stress drop estimates and interpretations for the 2011 prague, ok, earthquake sequence using multiple methods. *Journal of Geophysical Research: Solid Earth*, *126*(3), e2020JB020888. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JB020888 (e2020JB020888 2020JB020888) doi: https://doi.org/10.1029/2020JB020888

R Core Team. (2020). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Savage, M. K., & Anderson, J. G. (1995, 08). A local-magnitude scale for the western Great Basin-eastern Sierra Nevada from synthetic Wood-Anderson seismograms. *Bulletin of the Seismological Society of America*, *85*(4), 1236-1243. Retrieved from https://doi.org/10.1785/BSSA0850041236 doi: 10.1785/BSSA0850041236

Scherbaum, F. (1990). Combined inversion for the three-dimensional q structure and source parameters using microearthquake spectra. *Journal of Geophysical Research: Solid Earth*, *95*(B8), 12423-12438. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB095iB08p12423 doi: https://doi.org/10.1029/JB095iB08p12423

Shearer, P. M., Abercrombie, R. E., & Trugman, D. T. (2022). Improved stress drop estimates for m 1.5 to 4 earthquakes in southern california from 1996 to 2019. *Journal of Geophysical Research: Solid Earth*, *127*(7), e2022JB024243. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022JB024243 (e2022JB024243 2022JB024243) doi: https://doi.org/10.1029/2022JB024243

Shearer, P. M., Prieto, G. A., & Hauksson, E. (2006). Comprehensive analysis of earthquake source spectra in southern california. *Journal of Geophysical Research: Solid Earth*, *111*(B6). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JB003979 doi: https://doi.org/10.1029/2005JB003979

Shible, H., Hollender, F., Bindi, D., Traversa, P., Oth, A., Edwards, B., ... Gueguen, P. (2022, 02). GITEC: A Generalized Inversion Technique Benchmark. *Bulletin of the Seismological Society of America*, *112*(2), 850-877. Retrieved from https://doi.org/10.1785/0120210242 doi: 10.1785/0120210242

Trugman, D. T. (2020, 05). StressDrop and Source Scaling of the 2019 Ridgecrest, California, Earthquake Sequence. *Bulletin of the Seismological Society* *of America*, *110*(4), 1859-1871. Retrieved from https://doi.org/10.1785/0120200009 doi: 10.1785/0120200009

Trugman, D. T., & Shearer, P. M. (2017). Application of an improved spectral decomposition method to examine earthquake source scaling in southern california. *Journal of Geophysical Research: Solid Earth*, *122*(4), 2890-2910. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017JB013971 doi: https://doi.org/10.1002/2017JB013971

Walter, W. R., Yoo, S.-H., Mayeda, K., & Gök, R. (2017). Earthquake stress via event ratio levels: Application to the 2011 and 2016 oklahoma seismic sequences. *Geophysical Research Letters*, *44*(7), 3147-3155. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL072348 doi: https://doi.org/10.1002/2016GL072348

Zhang, J., Chen, X., & Abercrombie, R. E. (2022). Spatiotemporal variability of earthquake source parameters at parkfield, california, and their relationship with the 2004 m6 earthquake. *Journal of Geophysical Research: Solid Earth*, *127*(6), e2021JB022851. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JB022851 (e2021JB022851 2021JB022851) doi: https://doi.org/10.1029/2021JB022851

JGR: Solid Earth

Supporting Information for

On the non uniqueness of the source, propagation and site effects decomposition

D. Bindi (1), K. Mayeda (2), D. Spallarossa (3), M Picozzi (4), P. Morasca (5), A. Oth (6), W. R. Walter (7)

(1) German Research Centre for Geoscience GFZ, Potsdam, Germany

(2) AFTAC, Patrick AFB, FL, USA

(3) University of Genova, DISTAV, Genova, Italy

(4) University of Naples Federico II, DIFI, Naples, Italy

(5) INGV, Milan, Italy

(6)European Center for Geodynamics and Seismology, Walferdange, Luxembourg

(7)LLNL, Livermore, CA, USA

Contents of this file

Figures S1 to S3

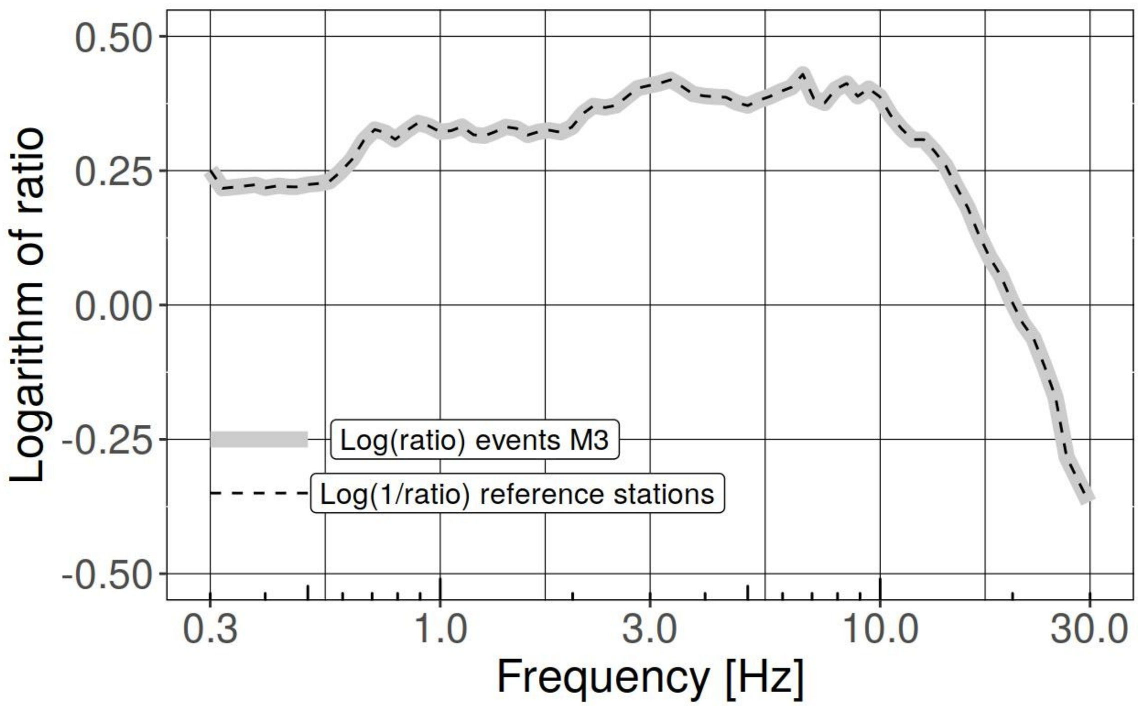

Figure S1. Impact on the source spectra of the constraint applied to the site term. The gray line corresponds to the ratio of the average source spectra of magnitude 3 events (panels b and c in Figure 6) obtained considering the two different constraints applied to the site terms; the dashed line is the inverse of the ratio computed for the average site term of the reference stations.

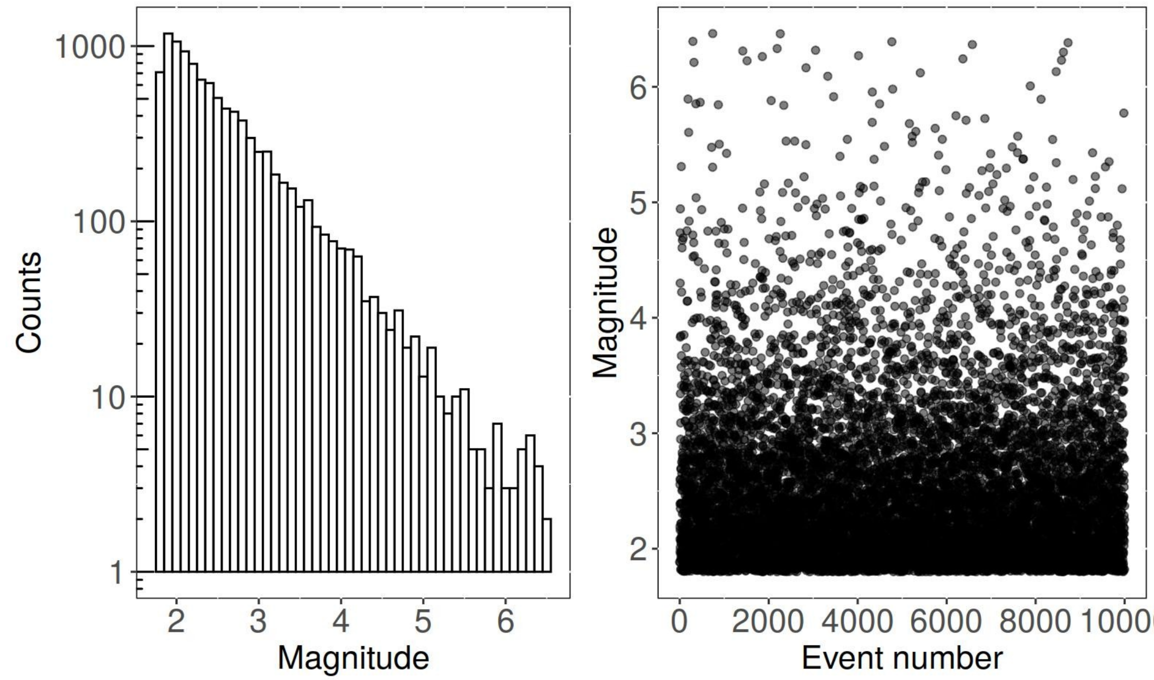

Figure S2. Synthetic magnitude catalog composed by 10000 events generated for the numerical test on the impact of constraining the corner frequency of small events. The left panel shows the distribution in the form of histogram reporting the number of events per magnitude bin.

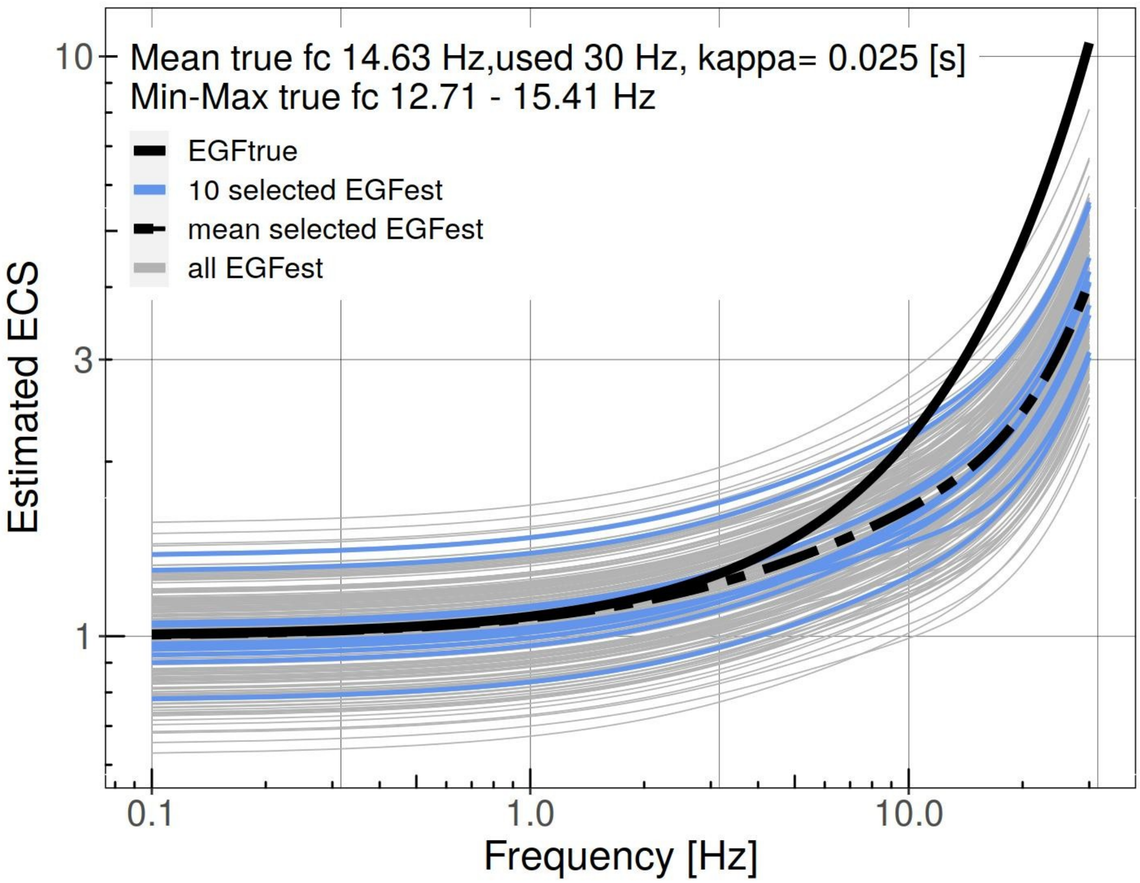

Figure S3. Estimated ECS in the numerical test performed to evaluate the impact of constraining the corner frequency of small events (scaling A of Figure 7). Thin lines are the spectra of the 202 candidate gEGFs with magnitude between 1.99 and 2.01; the average of the 10 selected gEGF is shown as dashed line; the black-solid line is the true EGF: the mean f_c of the used gEGFs is 14.6 Hz, the constrained one is 30 Hz. For the numerical ECS, *k* is fixed to 0.025 s (equation 8).