Bulletin of Earthquake Engineering

The pan-European Engineering Strong Motion (ESM) flatfile: consistency check via residual analysis

--Manuscript Draft--

Manuscript Number:	BEEE-D-18-00225R1
Full Title:	The pan-European Engineering Strong Motion (ESM) flatfile: consistency check via residual analysis
Article Type:	Original Research
Keywords:	Ground Motion Prediction Equation; residual analysis; European strong motion data
Corresponding Author:	Dino Bindi, PhD Helmholtz Centre Potsdam GFZ German Research Centre For Geosciences Potsdam, GERMANY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Helmholtz Centre Potsdam GFZ German Research Centre For Geosciences
Corresponding Author's Secondary Institution:	
First Author:	Dino Bindi, PhD
First Author Secondary Information:	
Order of Authors:	Dino Bindi, PhD
	Sreeram Reddy Kotha
	Graeme Weatherill
	Giovanni Lanzano
	Lucia Luzi
	Fabrice Cotton
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	We present the results of a consistency check performed over the flatfile extracted from the Engineering Strong Motion (ESM) database. The flatfile includes 23014 recordings from 2179 earthquakes in the magnitude range from 3.5 to 7.8 that occurred since the 1970s in Europe and Middle East, as presented in the companion article by Lanzano et al. (2018a). The consistency check is developed by analyzing different residual distributions obtained from ad-hoc ground motion prediction equations for the absolute spectral acceleration (SA), displacement (SD) and Fourier amplitude spectra (FAS). Only recordings from earthquakes shallower than 40 km are considered in the analysis. The between-event, between-station and event-and-station corrected residuals are computed by applying a mixed-effect regression. We identified those earthquakes, stations, and recordings showing the largest deviations from the GMPE median predictions, and also evaluated the statistical uncertainty on the median model to get insights on the applicable magnitude-distance ranges and the usable period (or frequency) range. We observed that robust median predictions are obtained up to 8s for SA and up to 20 Hz for FAS, although median predictions for Mw ≥7 show significantly larger uncertainties with ‘bumps’ starting above 5s for SA and below 0.3 Hz for FAS. The between-station variance dominates over the other residual variances, and the dependence of the between-station residuals on logarithm of Vs30 is well-described by a piece-wise linear function with period-dependent slopes and hinge velocity around 580 m/s. Finally, we compared the between-event residuals obtained by considering two different sources of moment magnitude. The results show that, at long periods, the between-event terms from the two regressions have a weak correlation and the overall between-event variability is dissimilar, highlighting the

	importance of magnitude source in the regression results .
Suggested Reviewers:	Carlo Cauzzi, Dr Eidgenossische Technische Hochschule Zurich carlo.cauzzi@sed.ethz.ch International expert on Ground Motion Prediction Equation and strong motion data analysis
	Laurentiu Danciu laurentiu.danciu@sed.ethz.ch International expert about the subject
	abdullah sandikkaya abdullahsandikkaya@hacettepe.edu.tr International expert about the topic
	Benedikt Halldorsson skykkur@hi.is International expert about the topic
	Susana Vilanova susana.vilanova@ist.utl.pt
Response to Reviewers:	our replies are uploaded as attachments to the manuscript.

The pan-European Engineering Strong Motion (ESM) flatfile: consistency check via residual analysis

Bindi D. (1), Kotha S-R.(1), Weatherill G.(1), Lanzano G.(2), Luzi L.(2), and F. Cotton (1,3)

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

(2) Istituto Nazionale di Geofisica e Vulcanologia INGV, Milan, Italy

(3) University of Potsdam, Potsdam, Germany

Abstract

We present the results of a consistency check performed over the flatfile extracted from the Engineering Strong Motion (ESM) database. The flatfile includes 23014 recordings from 2179 earthquakes in the magnitude range from 3.5 to 7.8 that occurred since the 1970s in Europe and Middle East, as presented in the companion article by Lanzano et al. (2018a). The consistency check is developed by analyzing different residual distributions obtained from ad-hoc ground motion prediction equations for the absolute spectral acceleration (SA), displacement (SD) and Fourier amplitude spectra (FAS). Only recordings from earthquakes shallower than 40 km are considered in the analysis. The between-event, between-station and event-and-station corrected residuals are computed by applying a mixed-effect regression. We identified those earthquakes, stations, and recordings showing the largest deviations from the GMPE median predictions, and also evaluated the statistical uncertainty on the median model to get insights on the applicable magnitude-distance ranges and the usable period (or frequency) range. We observed that robust median predictions are obtained up to 8s for SA and up to 20 Hz for FAS, although median predictions for Mw $\geq$7 show significantly larger uncertainties with 'bumps' starting above 5s for SA and below 0.3 Hz for FAS. The between-station variance dominates over the other residual variances, and the dependence of the between-station residuals on logarithm of Vs30 is well-described by a piece-wise linear function with period-dependent slopes and hinge velocity around 580 m/s. Finally, we compared the between-event residuals obtained by considering two different sources of moment magnitude. The results show that, at long periods, the between-event terms from the two regressions have a weak correlation and the overall between-event variability is dissimilar, highlighting the importance of magnitude source in the regression results .

Introduction

Most applications in Engineering Seismology and Earthquake Engineering deal with the analysis of strong-motion data. In Europe, the dissemination of uniformly processed strong-motion waveforms accompanied by event and station metadata progressed since the 1980s (Ambraseys, 1990; Ambraseys and Bommer, 1991; Ambraseys et al., 2000; Ambraseys et al 2004). The two most recent examples are the Reference Database for Seismic Ground-motion Prediction in Europe-RESORCE (Akkar et al., 2014), developed in the framework of the Seismic Hazard HARmonization in Europe project (SHARE; www.share.eu.org), and the Engineering Strong Motion (ESM) database (Luzi et al., 2016) developed in the framework of the Network of European Research Infrastructures for Earthquake Risk Assessment and Mitigation project (NERA). For many applications aimed at developing and testing Ground-Motion Prediction Equations (GMPEs), ground-motion intensity measures of engineering interest (e.g., peak ground acceleration-PGA, peak ground velocity-PGV, response spectra, Housner intensity, etc) are extracted or computed from the waveforms, and disseminated along with the relevant event and station metadata in the form of parametric tables - often referred to as flatfiles. Examples of recent flatfiles include the Pacific Earthquake Engineering Research Center (PEER) flatfile NGA-West2 (Ancheta et al., 2014) and regional flatfiles like those for Japan (Dawood et al., 2016) and Chile (Bastias and Montalva., 2016).

In this work, we focus on the flatfile generated from the ESM data-base, as described in detail in a companion paper (Lanzano et al., 2018a). In order to support the development of ground-motion models from the ESM flatfile, we perform a consistency check over the data and metadata from the perspective of the ground-motion variability, to report possible anomalies that a GMPE developer might encounter. To accomplish this task, we first calibrate a set of GMPEs for spectral acceleration (SA), spectral displacement (SD) and Fourier amplitude spectra (FAS), considering simple models describing the distance and magnitude scaling. Then, the random-effects and residual distributions obtained by performing a mixed-effects regression (Bates et al., 2015) are analyzed to detect anomalous records, events and stations. Issues that could be of relevance for ground-motion models, such as those related to the selected magnitude scale, and to ground-motion and path regionalization are discussed as well.

Data selection for developing the reference model

In this study, we use the 2018 version of the ESM flatfile (Lanzano et al., 2018b). The characteristics of the ESM flatfile are described in Lanzano et al. (2018a) whereas the waveform processing is detailed in Puglia et al. (2018). For developing the reference ground-motion model, we select only data at distances shorter than 300 km, generated by earthquakes with at least two recordings, and occurring at depths shallower than 40 km. The lower limit of hypocentral depth is set to select only the Active Shallow Crustal Earthquakes, whereas the deeper continental and subduction events are left out of this study considering their relatively complex phenomenon requiring ad-hoc studies. For SA and SD, each record used in the GMPE regression for period T should satisfy the usable period criterion $T \leq 0.8/f_{HP}$, where f_{HP} is the high-pass corner frequency of the filter applied in the waveform processing stage. For FAS computed at frequency f, only spectral amplitudes in the range $1.2\, f_{HP} \leq f \leq f_{LP}/1.2$, where f_{LP} is the low-pass corner-frequency of the filter, are considered in deriving the GMPE. The distributions of the corner-frequencies of the filter applied to all recordings are provided in Lanzano et al. (2018b). In particular, most of the f_{LP} frequencies lie in the range 30-40 Hz. Regarding the station selection, we only select stations installed at surface (sensor depth <10 m) avoiding borehole installations (identified by *housing_code* equal to "WEL") and stations installed at building floor (*installation_code* equal to "BF"). For our consistency check, we also select sensors with *location_code* equal to "00".

Regarding the magnitude, the flatfile includes multiple magnitude scales and estimates. In particular, the local magnitude (ML) and two choices for the moment magnitude (Mw), namely the moment magnitude listed in the ESM database (hereinafter referred to as Mw^{ESM}) and the moment magnitude from the EMEC catalog (Grünthal et al., 2012; hereinafter referred to as Mw^{EMEC}). In particular, while Mw^{ESM} is computed from moment tensor solutions provided by different sources selected according to an hierarchical schema (Lanzano et al., 2018b), Mw^{EMEC} is based on both moment tensor solutions and empirical regional conversions from other magnitude scales. The two Mw scales are compared in Figure 1a, where the best fit line is $Mw^{EMEC}=0.23+0.95Mw^{ESM}$, with the residual standard deviation equal to 0.16. Since neither Mw^{ESM} nor Mw^{EMEC} is available for all events, we derive an empirical relationship between ML and Mw in order to extend the residual analysis to as large a number of events as possible. These relations are not proposed for applications different from those performed in the context of this study. The best fit lines shown in Figure 1 (right) are:

$$Mw^{ESM} = 6.81 - 1.67\, ML + 0.26\, ML^2 \quad (\pm 0.3) \tag{1}$$

$$Mw^{EMEC} = 6.18 - 1.34\, ML + 0.23\, ML^2 \quad (\pm 0.3) \tag{2}$$

GMPE functional form

Mixed-effects regressions (Brillinger and Preisler, 1985; Bates et al., 2015) are performed for the natural-logarithm of SA, SD and FAS considering simple parametric functional forms describing the median scaling with distance and magnitude:

$$logY = e_1 + F(M_w) + G(R_{hypo}) + RE + \epsilon \quad (3)$$

$$F(M_w) = \begin{cases} b_1(M_w - M_{ref}) if M_w \leqslant M_H \\ b_1(M_H - M_{ref}) + b_3(M_w - M_H) otherwise \end{cases} \quad (4)$$

$$G(R_{hypo}) = \begin{cases} c_{1A} log(R_{hypo}) if R_{hypo} \leqslant R_H \\ c_{1A} log(R_H) + c_{1B} log\left(\frac{R_{hypo}}{R_H}\right) + c_3 R_{hypo} otherwise \end{cases} \quad (5)$$

where in equation (1) the random effect (RE) are

$$RE = \delta B_e + \delta S2S_s + \delta C2C_c \quad (6)$$

In equations (4) and (5), the reference and hinge magnitude and distances are set to M_{ref}=4.5, M_H=6, R_H=15km, respectively. The country-to-country random effect ($\delta C2C_c$), where *c* indicates country, is applied either on the offset term (e_1) or to c_3 (that is, two different regressions are performed). The other random-effects applied to the offset are the customary between-event (δB_e) and between-station ($\delta S2S_s$) residuals, where subscripts *e* and *s* are event and station indices. Residual ε in equation (3) represents the event- and station-corrected record-to-record variability. The standard deviations of ε, $\delta S2S_s$, $\delta C2C_c$, and δB_e random distributions are indicated with ϕ_0, ϕ_{S2S}, ϕ_{C2C} and τ, respectively. The regressions for SA and SD are performed over the orientation independent (rotD50) ground-motion measure (Boore, 2010), while the geometric-mean of the horizontal components is used for FAS GMPEs.

For the GMPEs regressed on SA considering $\delta C2C_c$ on the offset e_1, Figure 2 shows the period-dependent coefficients along with their standard-errors. Regarding the scaling with distance, coefficient c_{1B}

increases from about -2 at T<0.1s to about -1.3 for T>0.5s while coefficient c_{1A} is maximum around T=0.1s and decreases below -2 for T>4s. The term c_3 is negative for short periods and turns positive for T $\geq$ 1.5s. The positive values of c_3 suggest to investigate further segmentation of the distance scaling component to capture the effects of late–arrivals on long period ground-motion at distances beyond 80km (e.g., Lanzano et al., 2016; Kotha et al., 2018).

Regarding the scaling with magnitude, although large uncertainties affect b_3 (describing the magnitude scaling above the hinge magnitude), its values are always smaller than b_1, suggesting weaker scaling at larger magnitudes. The trend of the coefficients suggest that regressions above 8s are unstable due to a strong reduction in the number of data usable in GMPE regressions (Lanzano et al., 2018a). For example, the number of recordings used for the regressions at 0.01, 4, 8, and 10 s are 18859, 16222, 9664 and 2767, respectively. The asymptotic standard error (σ_μ) on the median predictions, quantifying the within-model statistical estimate of epistemic uncertainty (e.g., Al Atik and Youngs, 2014 ; Bindi et al., 2017a), due to the combined effect of limited data availability and complexity of implemented functional form (in terms of extent of GMPE parametrization) are shown in Figure 3. For Mw $\leq$6.5, σ_μ is smaller than 0.2 and it shows a weak dependency on magnitude for Mw<6. For Mw<6.5, σ_μ reaches the minimum values around T=2s while the sharp increase above T=8s indicates that regressions are not well-constrained for these periods. At Mw > 6.5, σ_μ increases rapidly, exceeding 0.25 at Mw=7.5. Above magnitude 7, σ_μ shows a bump around T=5s, suggesting that at periods longer than 4s GMPE median predictions for Mw > 7 are uncertain. The dependence of σ_μ on distance (here not shown) is negligible because the data appears to constrain very well the distance scaling component of the GMPE (small errors on c_{1A}, c_{1B}, c_3 in Figure 2).

Results

The results are presented in terms of event- and station-corrected residuals (ε), between-event (δB_e) and between-station ($\delta S2S_s$) distributions evaluated for SA with regional attributes $\delta C2C_c$ on e_1.

Event- and station-corrected residuals ε

The event- and station-corrected residuals for T=0.01s and T=8s are shown in Figure 4 as function of the recording Id (top), of magnitude (middle) and of distance (bottom). In the middle and bottom frames, the average residuals computed over magnitude and distance bins (red error bars) are shown as well. Please note that the number of recordings used in GMPE regression decrease towards long periods. The residuals do not show any significant trend either with distance or with magnitude. The number of residuals (in absolute value) above 3 times ϕ_0 is 60 at T=0.01 s and 20 at T=4s, as shown in Tables 1 and 2, respectively. A further check of the waveforms does not reveal any peculiarities in the recordings except for the records [IT-2016-0001, IV.INTR], which is characterized by high noise level at high frequencies, as shown in Figure 5.

Between-event random effects δB_e

The between-event random-effects for T=0.01s and T=8s are shown in Figures 6a and 6b, respectively. δB_e does not show significant trend for M> 4.5, while δB_e increases with decreasing magnitude for M<4.5. Unlike the functional form implemented in this study, an additional segmentation of F(Mw) at Mw = 4.5 (e.g, Campbell and Bozorgnia 2014, Kotha et al. 2018) would allow different scaling for small, moderate and large magnitude bins, removing the trend observed for small magnitudes. We note here that moment magnitudes below 4 are not converted from local magnitude.

Differences in observed SAs based on Style-of-Faulting (SoF) are widely debated in GMPE development. For the dataset analyzed in this study, the box and whisker plots corresponding to δB_e dependence on SoF do not show any systematic differences. Since the model included a random-effect at the country level (Italy-IT, Turkey-TK, Romania-RO, Greece-GR, other countries-Others), a dependence of δB_e on the country grouping is not expected, as shown by the box and whisker plots in Figures 6a and 6b. Events with δB_e larger than 3 times τ, in absolute value, are listed in Table 3 and 4 for T=0.01 and 8 s, respectively.

Between-station random-effects $\delta S2S$.

The between-station random-effects for T=0.01s and T=8s are shown in Figures 7a and 7b, respectively. At short periods, $\delta S2S_s$ shows a trend with Vs30, as expected (log(Vs30) is commonly used as proxy for site amplifications). The best-fit model obtained through a break-point regression (Muggeo, 2003) is

described by a piece-wise linear function with break-point at $\log_{10}$(Vs30)=(2.78 ± 0.09) (i.e., Vs30=600 m/s) and slopes (-0.6 ±0.2), (-1.7±0.40). A weaker relation exists between $\delta S2S_s$ and Vs30WA (which is an estimate of Vs30 derived from the topographic gradient, see Lanzano et al., 2018b). At long periods, the trend with measured Vs30 is strong only for Vs30<600 m/s, while considering Vs30WA a correlation over a broader velocity range is observed. For T=8s, the break-point is located at $\log_{10}$Vs30=(2.77±0.09), and the slopes are (-1.6 ±0.2) and (-0.3 ±0.2).

When split among the different countries (box and whiskers plots in Figure 7), no residual offset is detected. Often the station Vs30 distributions (mean and range) differ from region-to-region depending on the local geology over which the networks reside. As a result, systematic region-dependent offsets could be expected. In our analysis, since all the country-to-country variability is accounted by the $\delta C2C_c$ random-effects on e_1, we observe no differences in mean $\delta S2S_s$ among the countries. Nevertheless, it is worth of further investigation whether the scaling of $\delta S2S_s$ with Vs30 is identical across various regions or not. Kotha et al. (2016) for instance, observed a similar (regionally) unbiased $\delta S2S_s$ for Italy and Turkey, but with significantly different scaling with Vs30. Stations with $\delta S2S_s$ larger than 3 times ϕ_{S2S}, in absolute value, are listed in Table 5 and 6 for T=0.01 and 8 s, respectively.

Standard deviations of random-effects

The standard deviations of the random-effects are summarized in Figure 8. The between-station variability ϕ_{S2S} dominates the total aleatory variability of the dataset (σ): it shows a peak around 0.10 s and it decreases above 2s. The peak at T = 0.1s is observed in several other studies as well (e.g. Zhao et al. 2016, Cauzzi and Faccioli 2017; Kotha et al. 2018), indicating a high variability of short period site-response in response spectra domain. The between-event standard deviation τ has a peak located around 2s, otherwise almost constant with a small bump around 0.1 s; τ shows a suspicious drop at T=9s and 10s, probably due to the strong reduction in the usable recordings. The event- and station-corrected variability (ϕ_0) is close to τ for T<0.5s, while for longer periods it is smaller than the between-event variability. Finally, the between-country variability ϕ_{C2C} (region-to-region variability of offset e_1) is small (of the order of 0.25) with a trough around 2s. The anti-symmetry of τ and ϕ_{C2C} around T = 2s suggests a trade-off between event-to-event and country-to-country variability in this GMPE regression, suggesting the need of further investigating the formulation of REs in GMPEs.

Regional differences on anelastic attenuation.

An additional regression is performed by applying the country-group random-effect not on the overall offset e_1 in equation 3 but on the anelastic coefficients c_3 of equation5. Figure 9 shows the obtained random effects δc_3 , which can be interpreted as regional adjustment to the generic region-independent "anelastic attenuation" term c_3. These country-specific adjustments (to c_3) are statistically significant with small standard errors (horizontal bar in the caterpillar plot of Figure 9), suggesting considerable differences in propagation effects across these regions. For instance, since the δc_3 value for Romania is not well-constrained with the very few data from that region, its standard error is larger than that for Italy, Greece, and Turkey. For periods between 0.01 and 2 s, Greek and Italy show stronger attenuation than Turkey; for longer periods we recall that c_3 becomes positive.

Since in this case the country-based RE is not applied to the overall offset, the between-events show some variability across the countries (Figure 10). Note that since $\delta C2C_c$ is no longer applied on e_1, systematic country-to-country variability in δB_e and $\delta S2S_s$ distributions may now appear, as seen in the between-event distributions of Figure 10. For example, on average the events from Greece (GR) show a higher δB_e compared to those from Turkey (TR) and Italy (IT) (Figure 9).

δB_e obtained using either Mw^{EMEC} or Mw^{ESM} .

Several studies (e.g., Rohades, 1997; Moss, 2011; Kuehn and Abrahamson, 2018) evaluated the impact of the uncertainty in the predictor variables on the aleatory variability. The effect of magnitude uncertainty on δB_e variability has been observed in Europe for the RESORCE flatfile by comparing Mw values computed following different procedures (Kotha et al., 2016). For the ESM flatfile, we assess the impact of using Mw^{EMEC} or Mw^{ESM} by repeating the regression for each magnitude type, with $\delta C2C_c$ applied on e_1. The results are compared in Figure 11 in terms of δB_e for T=0.1 and 2.5s. The red dots in Figure 11 indicate recordings for which both Mw^{EMEC} and Mw^{ESM} are available, while the other colors indicate earthquakes for which at least one of the two Mw values is derived from ML. In particular, the green dots are for earthquakes without Mw (neither EMEC nor ESM). Figure 11 shows that at short periods δB_e for the two regressions show a good correlation although with a large scatter. On the other hand, at

longer periods (T=2.5s in Figure 11 but results for longer periods are similar) the correlation is very weak (being trivial the correlation shown by the green points). Figure 12 shows the standard deviation τ of the δB_e residuals for the two regressions, considering both actual Mw values and values derived from ML. τ for Mw^{EMS} is higher than the one obtained using Mw^{EMEC} for periods T<2.0s whereas, from T=2.0 to 8s, similar values are obtained for the two regressions.

Summary for SD and for FAS

The ESM flatfile includes several intensity measure of engineering interest, such as spectral acceleration (SA) and displacement (SD); Fourier amplitude spectra (FAS); peak ground velocity (PGV) acceleration (PGA) and displacement (PGD); Arias intensity, Housner intensity, effective duration, cumulative absolute velocity (CAV). While the check of all those parameters is beyond the aim of this work, the residual analysis have been extended to SD and FAS as well (Bindi et al., 2017b). Figure 13 summarizes the results by comparing the standard deviations of the residual distributions. The main conclusions drawn from SD and FAS analysis agree with those from SA. Regarding the uncertainty on the median model for FAS, the dependency of σ_μ on frequency is shown in Figure 14. The models are robust up to 20 Hz, for higher frequencies the number of available recording decreases due the condition applied to the usable frequency range (see Lanzano et al., 2018b, for the distribution of the low-pass corner frequencies of the filter applied during the processing phase) and the uncertainty becomes very large. Towards low frequencies, σ_μ for Mw $\geq$ 6.5 shows a bump around 0.2 Hz. In general, in agreement with the results obtained for SA, σ_μ significantly increases above magnitude 7.

Conclusion

We performed residuals analysis over the ESM flatfile (Lanzano et al., 2018b) as complementary information to the flatfile description (Lanzano et al., 2018a), for supporting those studies aiming at developing ground-motion models. The main outcomes of the consistency check performed on recordings from shallow earthquakes (hypocentral depth less than 40 km) can be summarized as follow:

1) considering the data population and the selection criteria applied to define the usable period range, robust results for response spectra have been obtained for periods up to 8s, although the median models for Mw$\geq$7 are affected by large within-model epistemic uncertainty (σ_μ) above 5s. Regarding the

model for Fourier, σ_μ increases significantly above 20Hz and shows a bump around 0.2 Hz for magnitudes above 7. The epistemic uncertainty for both response and Fourier spectra increases significantly for Mw$\geq$ 6.5;

2) the sanity check has been performed considering a simple functional form for describing the magnitude scaling and the scaling with hypocentral distance, although for larger finite-source events the Joyner-Boore or the rupture distances should be considered. Regarding the distance scaling, a piece-wise linear function of logarithm of distance was needed to avoid trends in the residuals at distances shorter than 15km (see also Kotha et al., 2018). Dependence of the hinge distance on magnitude and period is worth future investigation in order to model the magnitude dependence of geometric spreading as well;

3) regional dependencies of the attenuation with distance (so-called apparent anelastic attenuation parameter) already observed in previous studies (e.g., Kotha et al., 2016) are confirmed. In addition, it is now possible to refine the regionalization scheme given the significantly larger amount of data from a few more regions;

4) regarding the scaling of SA with magnitude, it is worth considering a second hinge magnitude at small magnitudes, around 4 (see also Kotha et al 2018);

5) the scaling of the between-station random-effects $\delta S2S_s$ with the logarithm of Vs30 is well described by a piece-wise linear function with period-dependent slopes and hinge velocity around 580 m/s;

6) magnitude is one of the two fundamental explanatory variables used to develop ground motion models, and the flatfiles includes information about different magnitude scales extracted from the ESM database. In addition, two different moment magnitude values are provided (when available): Mw^{ESM} is the moment magnitude as included in ESM without any conversion (a reference for the source is provided); Mw^{EMEC} is the moment magnitude according to the EMEC catalog (Grünthal et al., 2012). Derived through different procedures, the two magnitudes show an overall agreement (linear correlation) but with large scatter around the mean trend. The between-event residuals δB_e at long periods absorb the uncertainties in the seismic moment estimates: therefore, the weak correlation shown by the δB_e distribution obtained considering the two different moment magnitudes warns about the critical role played by the magnitude selection;

7) Records, events, and stations showing standardized residuals larger, in absolute value, than 3 are listed in tables to allow for a comparison with other models.

8) among different intensity measures, Fourier amplitude spectra allow an easier and direct comparison between components of the observed ground-motion variability and seismological interpretations.

Data and Resources

In this study, we analyzed the ESM flatfile (Lanzano et al, 2018b) with persistent identifier PID: 11099/ESM_flatfile_2018. The flatfile can be downloaded from the ESM webpage: http://esm.mi.ingv.it.

Acknowledgments

This work has been developed within the European Plate Observing System (EPOS; https://epos-ip.org/) and the Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe (SERA; http://www.sera-eu.org) projects. Comments from B. Edwards and one anonymous reviewer helped us to improve the manuscript.

References

Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BS-J, Wooddell KE, Graves RW, Kottke AR, Boore DM, Kishida T, Donahue JL (2014) NGA-West2 database. Earthquake Spectra 30(3):989–1005

Akkar S, Sandıkkaya M A, Şenyurt M, Azari Sisi A, Ay B Ö, Traversa P, Douglas J, Cotton F, Luzi L, Hernandez B and Godey S (2014) Reference database for seismic ground-motion in Europe (RESORCE), Bull. Earthquake Eng 12:311–339, DOI 10.1007/s10518-013-9506-8

Al Atik L, Youngs B (2014) Epistemic uncertainty for NGA-West2 models. Earthquake Spectra 30(3):1301–1318

Ambraseys NN (1990) Uniform magnitude re-evaluation of European earthquake associated with strong-motion records. Earthq Eng Struct Dyn 19:1–20

Ambraseys NN, Bommer JJ (1991) Database of European strong-motion records. Eur Earthq Eng 5:18–37

Ambraseys N, Smit P, Berardi R, Rinaldis D, Cotton F, Berge-Thierry C (2000) Dissemination of European strong-motion data. CD-ROM collection. European Commission, Directorate General XII, Science, Research and Development, Environment and Climate Programme, Bruxelles

Ambraseys NN, Smit P, Douglas J, Margaris B, Sigbjörnsson R, Olafsson S, Suhadolc P, Costa G (2004) Internet site for European strong-motion data. Bollettino di Geofisica Teorica ed Applicata 45:113–129

Bastías, N, Montalva GA (2016) Chile strong ground motion flatfile, Earthq. Spectra 32, 2549–2566, doi: 10.1193/102715EQS158DP.

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

Bindi D, Cotton F, Kotha S, Bosse C, Stromeyer D, Grünthal G (2017a) Application-driven ground motion prediction equation for seismic hazard assessments in non-cratonic moderate-seismicity areas, Journal of Seismology 21, 5, 1201-1218.

Bindi D, Spallarossa D, Pacor F (2017b) Between-event and between-station variability observed in the Fourier and response spectra domains: comparison with seismological models, Geophysical Journal International 210, 1092–1104, https://doi.org/10.1093/gji/ggx217

Boore DM (2010) Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. Bull Seim Soc Am 100(4):1830-1835

Brillinger DR, Preisler HK (1985) Further analysis of the Joyner-Boore attenuation data. Bull. Seism. Soc. Am. 75, no. 2, 611–614.

Campbell KW, Bozorgnia Y (2014) NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. Earthquake Spectra 30,1087–115.

Cauzzi C, Faccioli E (2017) Anatomy of sigma of a global predictive model for ground motions and response spectra. Bull Earthq Eng, 1–19. doi: 10.1007/s10518-017-0278-4

Dawood HM, Rodriguez-Marek A, Bayless J, Goulet C, Thompson E, (2016) Flatfile for the KiK-net database processed using an automated protocol. Earthq Spectra 32,1281–1302.

Grünthal G, Wahlström R (2012) The European-Mediterranean Earthquake Catalogue (EMEC) for the last millennium. J Seismol 16(3):535–570

Kotha SR, Bindi D, Cotton F (2016) Partially non-ergodic region specific GMPE for Europe and Middle-East. Bull Earthq Eng,14:1245–1263.

Kotha SR, Cotton F, Bindi D (2018) A new approach to site classification: Mixed-effects Ground Motion Prediction Equation with spectral clustering of site amplification functions, Soil Dynamics Earth Eng, doi: 10.1016/j.soildyn.2018.01.051

Kuehn NM, Abrahamson NA (2018) Uncertainty in predictor variables on the estimation of Ground-Motion Prediction Equations, Bull. Seism. Soc. Am 108, doi: 10.1785/0120170166

Lanzano G, D'Amico M, Felicetta C, Puglia R, Luzi L, Pacor F, Bindi D (2016) Ground motion prediction equations for region specific probabilistic seismic hazard analysis. Bull Seim Soc Am 106(1):73–92

Lanzano G, Sgobba S, Luzi L, Puglia R, Pacor F, Felicetta C, D'Amico M, Cotton F, Bindi D (2018a) The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and data statistics. Bull Earthquake Eng (submitted).

Lanzano G, Puglia R, Russo E, Luzi L, Bindi D, Cotton F, D'Amico M, Felicetta C, Pacor F & ORFEUS WG5 (2018b) ESM strong-motion flat-file 2018. Istituto Nazionale di Geofisica e Vulcanologia (INGV), Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum (GFZ), Observatories & Research Facilities for European Seismology (ORFEUS). PID: 11099/ESM_flatfile_2018

Luzi L, Puglia R, Russo E, D'Amico M, Felicetta C, Pacor F, Lanzano G, Çeken U, Clinton J, Costa G, Duni L, et al (2016) The Engineering Strong-Motion Database: A Platform to Access Pan-European Accelerometric Data, Seismological Research Letters 87, 987-997. doi: 10.1785/0220150278

Moss RES (2011) Reduced sigma of ground-motion prediction equations through uncertainty propagation, Bull. Seismol. Soc. Am. 101, 250–257, doi: 10.1785/0120090325

Muggeo VMR (2003) Estimating regression models with unknown break-points. Statistics in Medicine, 22, 3055-3071

Puglia R, Russo E, Luzi L, D'Amico M, Felicetta C, Pacor F, Lanzano G (2018) Strong motion Processing Service: a tool to access and analyse earthquakes strong motion waveforms. Bull Earthquake Eng., doi: 10.1007/s10518-017-0299-z.

Rhoades DA (1997) Estimation of attenuation relations for strong-motion data allowing for individual earthquake magnitude uncertainties, Bull. Seismol. Soc. Am. 87,1674 – 1678.

Zhao JX, Zhou SL, Gao PJ, Zhang YB, Zhou J, Lu M, and Rhoades DA (2016) Ground-motion prediction equations for shallow crustal and upper mantle earthquakes in Japan using site class and simple geometric attenuation functions, Bull. Seism. Soc. Am. 106, 1552-1569, doi**:** 10.1785/0120150063

Table 1. Event- and station-corrected residuals greater than 3 times ϕ_0 in absolute value ($|\varepsilon|>3\phi_0$) at T=0.01s

Event ID	Station	ε(0.01s)	Event ID	Station	ε(0.01s)
EMSC-20101109_0000039	FR.REVF	-1.943414	EMSC-20161129_0000079	IT.AQG	-1.870889
EMSC-20110707_0000048	RA.IRPV	2.168072	EMSC-20161129_0000079	IT.AQV	-1.864027
EMSC-20130125_0000039	IV.BDI	2.24683	EMSC-20161129_0000079	IT.GNU	-1.864027
EMSC-20130829_0000020	KO.GONE	-1.996238	EMSC-20161129_0000079	IT.PSC	-1.912796
EMSC-20140907_0000025	IT.BRM	1.882042	EMSC-20161129_0000079	IT.SPD	-2.056146
EMSC-20140907_0000025	IT.PRT	3.342786	EMSC-20161129_0000079	IT.SULA	-1.839875
EMSC-20141216_0000014	KO.GONE	-1.881109	EMSC-20161129_0000079	IT.SULC	-1.954815
EMSC-20141219_0000039	IT.FOC	-1.853693	EMSC-20161129_0000079	IV.BSSO	-1.938496
EMSC-20141224_0000053	IT.SULC	1.87239	EMSC-20161129_0000079	IV.CERA	-2.019819
EMSC-20150130_0000003	IV.OPPE	1.919609	EMSC-20161129_0000079	IV.MOMA	2.029745
EMSC-20150130_0000003	NI.POLC	-2.251522	EMSC-20161129_0000079	IV.RNI2	-2.023737
EMSC-20150130_0000003	RF.GEPF	-1.924863	EMSC-20161129_0000079	IV.VAGA	-1.965676
EMSC-20150205_0000061	HL.ZKR	2.124032	EMSC-20161201_0000064	IT.CNE	1.943368
EMSC-20150726_0000076	HL.PLG	2.325277	GR-1999-0001	HL.PREA	3.271508
EMSC-20150913_0000012	KO.CAME	2.257234	IT-1987-0003	IT.PNN	-2.084439
EMSC-20151101_0000023	MN.PDG	1.961717	IT-2007-0208	IT.FMG	2.013387
EMSC-20151118_0000047	KO.CAME	-2.911697	IT-2009-0186	IV.SACR	-2.090624
EMSC-20151118_0000094	KO.CAME	-2.740185	IT-2009-0189	IV.SACR	-2.126487
EMSC-20160530_0000085	IV.SACS	2.118076	IT-2010-0003	IV.NRCA	2.243845
EMSC-20160824_0000006	IT.GRD	-1.950702	IT-2011-0029	IT.GAI	-2.075673
EMSC-20160824_0000295	IV.FIAM	1.86564	IT-2012-0002	FR.REVF	2.025614
EMSC-20160824_0000295	IV.RM33	2.219246	IT-2013-0003	IT.PSC	1.974364
EMSC-20160903_0000009	IT.ACC	-1.866856	IT-2013-0003	IV.POFI	2.157508
EMSC-20160903_0000009	IT.CIT	-1.983131	IT-2013-0005	IT.MDN	2.272414
EMSC-20160903_0000009	IT.FOC	2.200381	IT-2013-0010	IV.SNTG	-2.148043
EMSC-20160903_0000009	IV.NRCA	1.919352	IT-2016-0001	IV.INTR	-3.470851
EMSC-20160903_0000009	IV.T1201	-1.887857	ME-1979-0003	EU.BAR	-2.051513

EMSC-20160903_0000009	IV.T1218	-2.022831	TK-1999-0077	TK.8101	1.865905
EMSC-20161027_0000072	IV.SNTG	2.005795	TK-2003-0180	TK.1701	-1.970150
EMSC-20161129_0000079	IT.AQF	-1.928834	TK-2006-0098	TK.1001	-2.058961

Table 2. Event- and station-corrected residuals greater than 3 times ϕ_0 in absolute value ($|\varepsilon|>3\phi_0$) at T=8s

Event ID	Station	ε(8.0s)	Event ID	Station	ε(8.0s)
EMSC-20110707_0000048	RA.IRPV	1.996037	IT-2010-0003	IV.NRCA	1.48384
EMSC-20150205_0000061	HL.ZKR	1.78805	IT-2010-0006	IV.NRCA	1.627814
EMSC-20151106_0000016	FR.REVF	-2.236079	IT-2012-0010	TV.MIR01	-1.484999
EMSC-20151106_0000016	FR.SURF	1.709347	IT-2012-0025	IT.FIN0	1.660915
EMSC-20151118_0000094	KO.CAME	-1.846812	IT-2012-0025	IV.T0800	1.737367
EMSC-20161030_0000033	IT.ACC	1.99265	IT-2012-0025	IV.T0811	2.127796
EMSC-20161129_0000079	IT.AQV	-1.492742	IT-2012-0035	IV.CPGN	-1.493063
IT-2009-0009	IT.ANT	-1.557013	IT-2012-0063	IT.SBT	1.630845
IT-2010-0001	IV.NRCA	1.59095	IT-2013-0013	IT.NAS	1.569184
IT-2010-0001	IV.SENI	-1.600929	TK-1999-0077	TK.8101	1.501051

385

Table 3. Between-event greater than 3 times τ, in absolute value ($|\delta B_e|>3\ \tau$), at T=0.01 s

Event ID	δB_e(0.01s)	N records
EMSC-20121203_0000066	1.73	18
EMSC-20150411_0000019	-1.69	10
EMSC-20161102_0000149	-1.66	72
IT-2009-0316	1.97	6

Table 4. Between-event greater than 3 times τ, in absolute value ($|\delta B_e|>3\ \tau$), at T=8.0 s

Event ID	δB_e(8.0s)	N records
EMSC-20121203_0000066	1.65	11
EMSC-20141124_0000030	1.84	5
IT-2010-0003	1.95	3

Table 5. Between-station greater than 3 times ϕ_{s2s} , in absolute value ($|\delta S2S| > 3\phi_{s2s}$), at T=0.01 s

Station	**δS2S(0.01s)**	N records
IV.T0203	-2.35	1
HL.KSL	-2.04	42
CR.VELS	2.01	1
RO.JURR	2.04	2
FR.REVF	2.05	5

Table 6. Between-station greater than 3 times ϕ_{s2s} , in absolute value ($|\delta S2S| > 3\phi_{s2s}$), at T=8.0 s

Station	**δS2S(8.0s)**	N records
IV.T0501	-1.86	3
FR.REVF	2.29	4
IV.SNTG	-1.74	15
KO.ENZZ	1.70	7

Figures

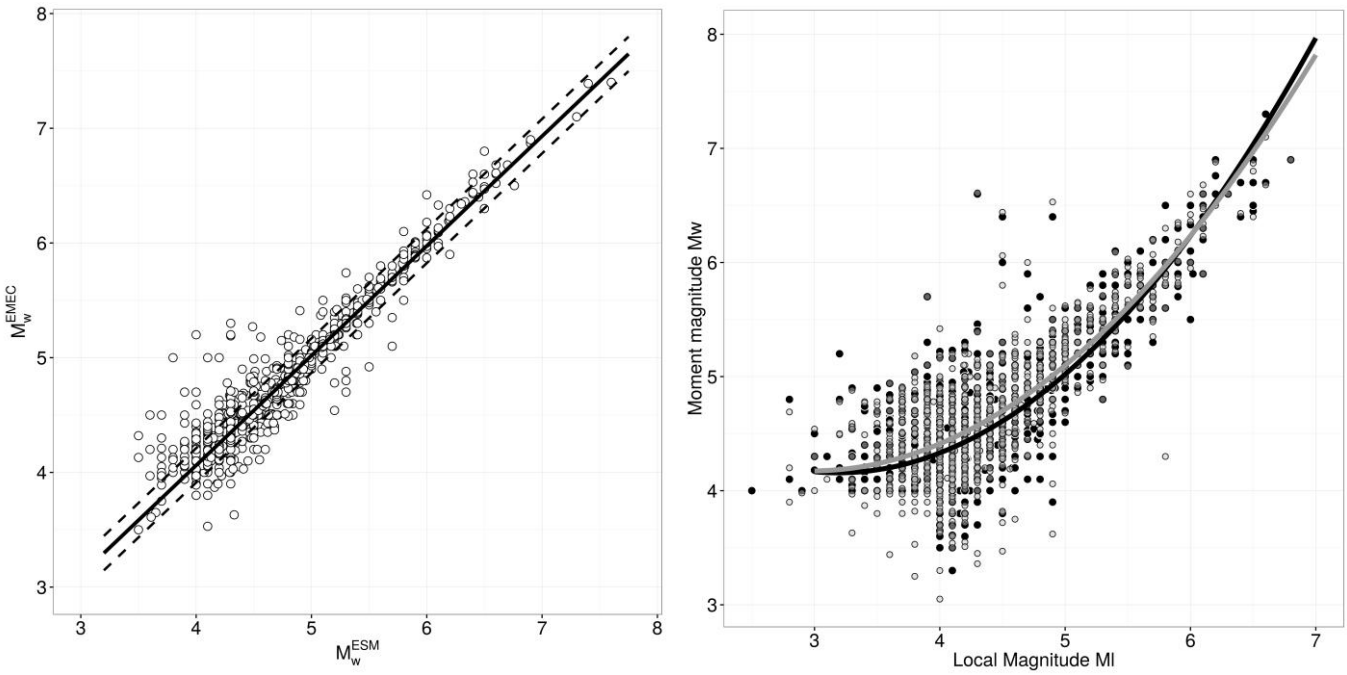
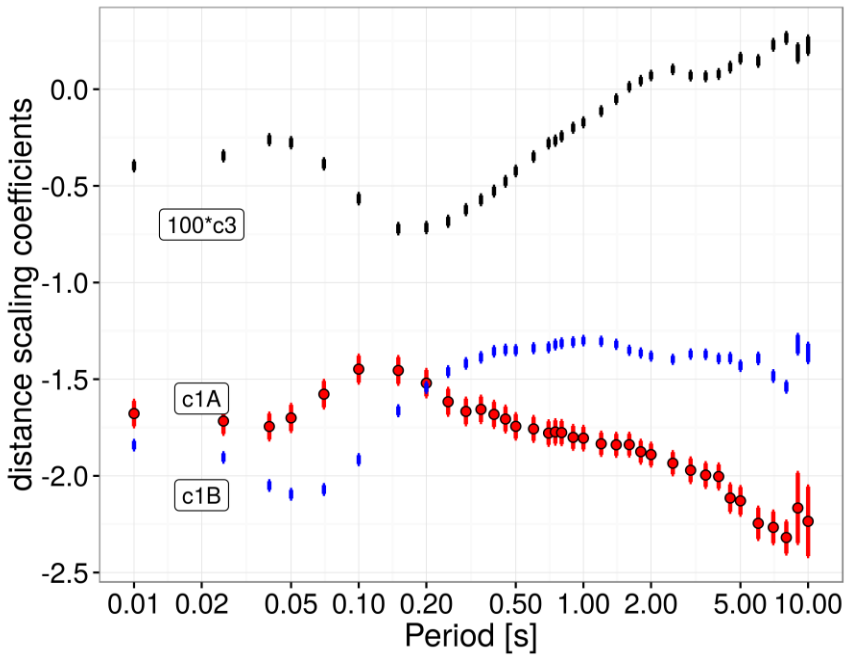


Figure 1. Left: regression between Mw^{ESM} and Mw^{EMEC}; right: quadratic regressions between Mw^{ESM} and ML (black line and points) and between Mw^{EMEC} and ML (gray line and points) (see equations 1 and 2).



405

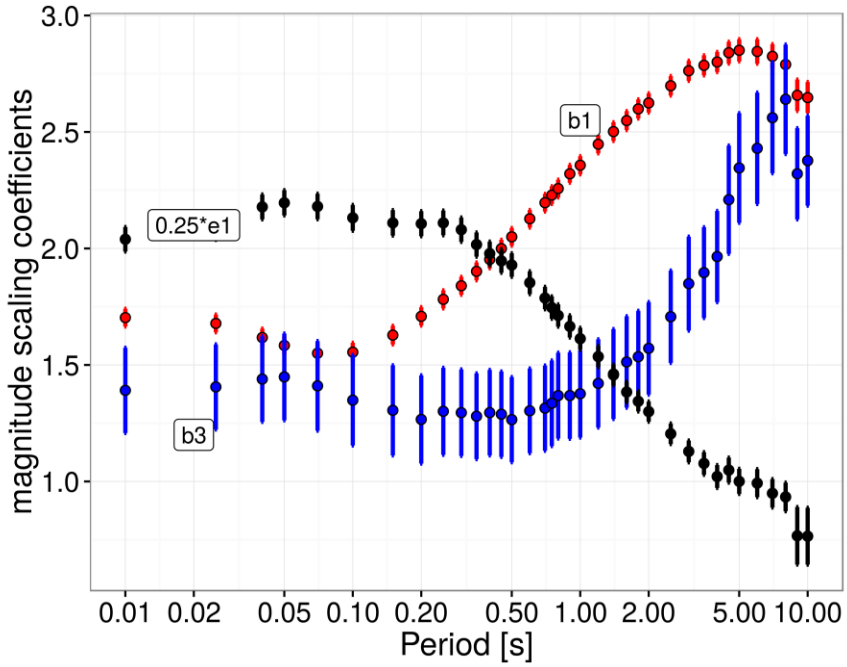
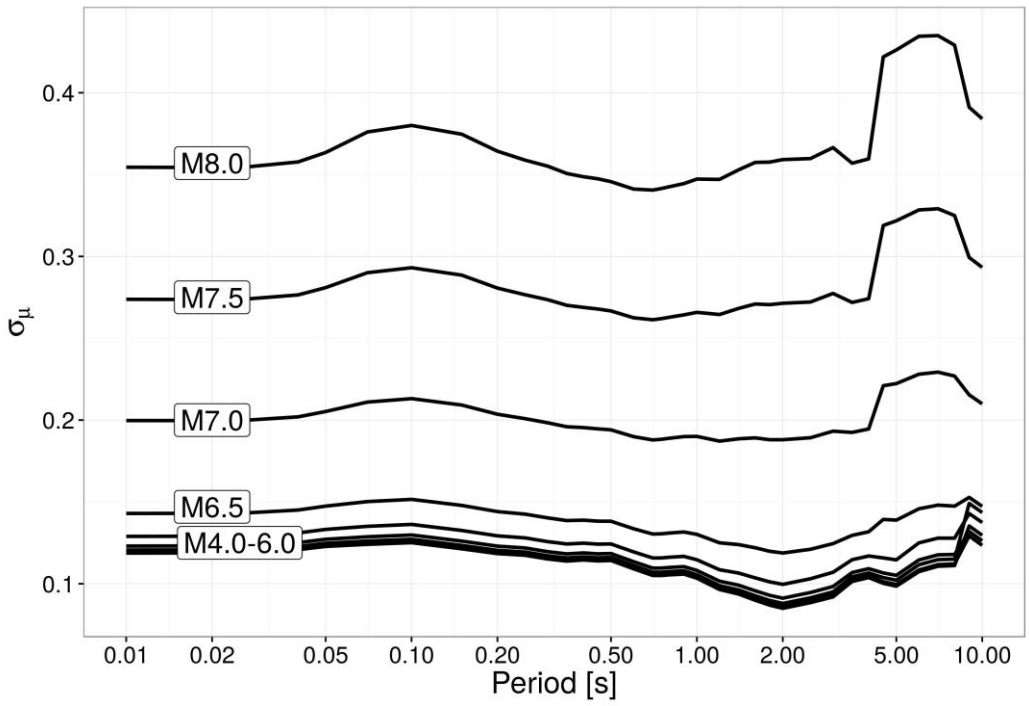
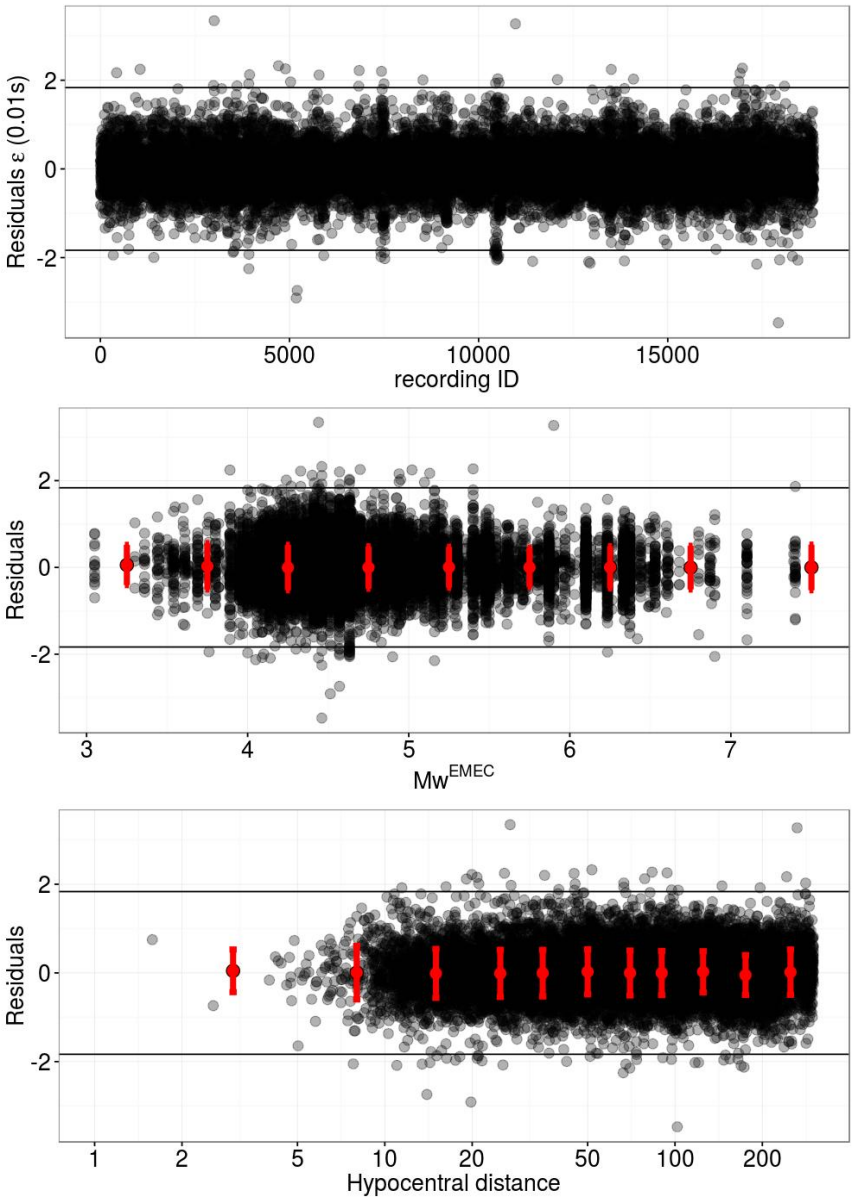


Figure 2. Coefficients (mean and standard error) of the model in equation (4) and (5) for Spectral Acceleration (SA).

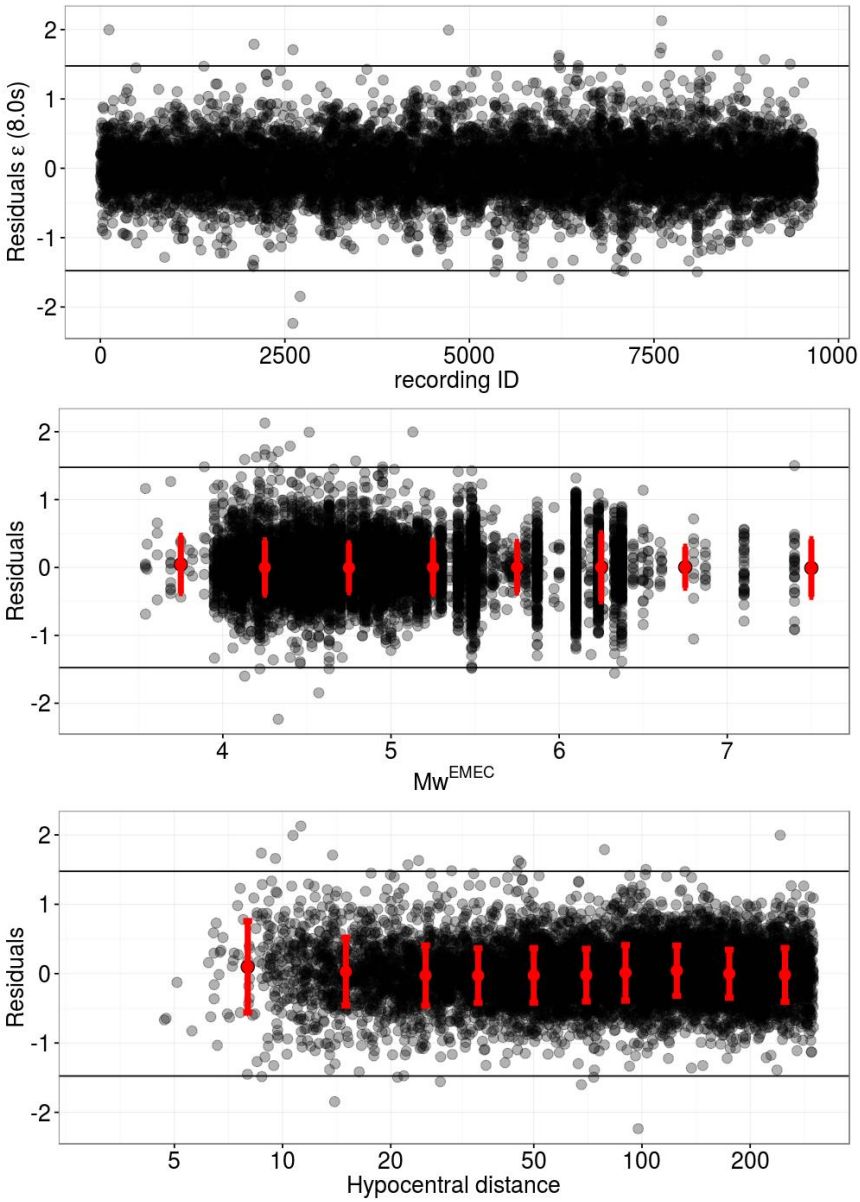


410 **Figure 3**. Period dependence of the asymptotic standard error on the median (σ_μ) for SA considering different magnitudes and a fixed distance of 30 km.



415

Figure 4a. Residuals ε computed for SA at T=0.01s. Horizontal lines correspond to $\pm 3\phi_0$. The average residuals ±1 standard deviation computed over magnitude and distance bins are shown as red bars.



420

Figure 4b. The same as Figure 4a but for T=8s.

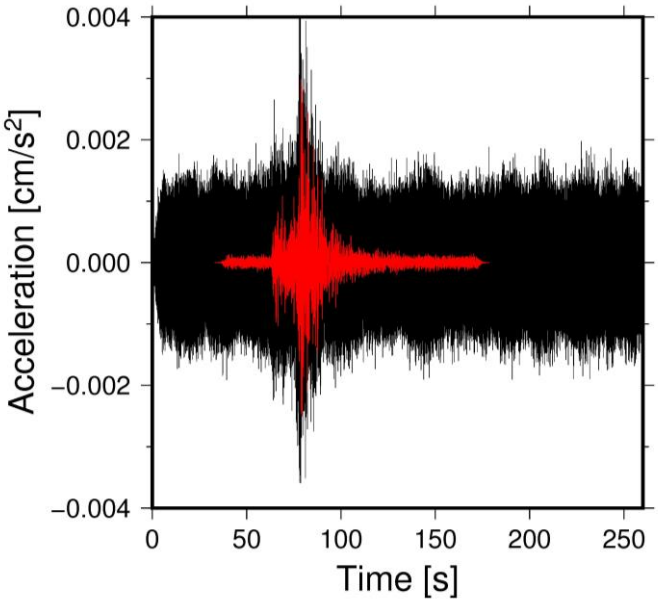


Figure 5 Record (EW component) of event IT-2016-0001 at station IV.INTR: comparison between raw data (black) and processed acceleration (red). The processed acceleration has been filtered over the bandwidth 0.15-30Hz using a second order a-causal Butterworth filter.

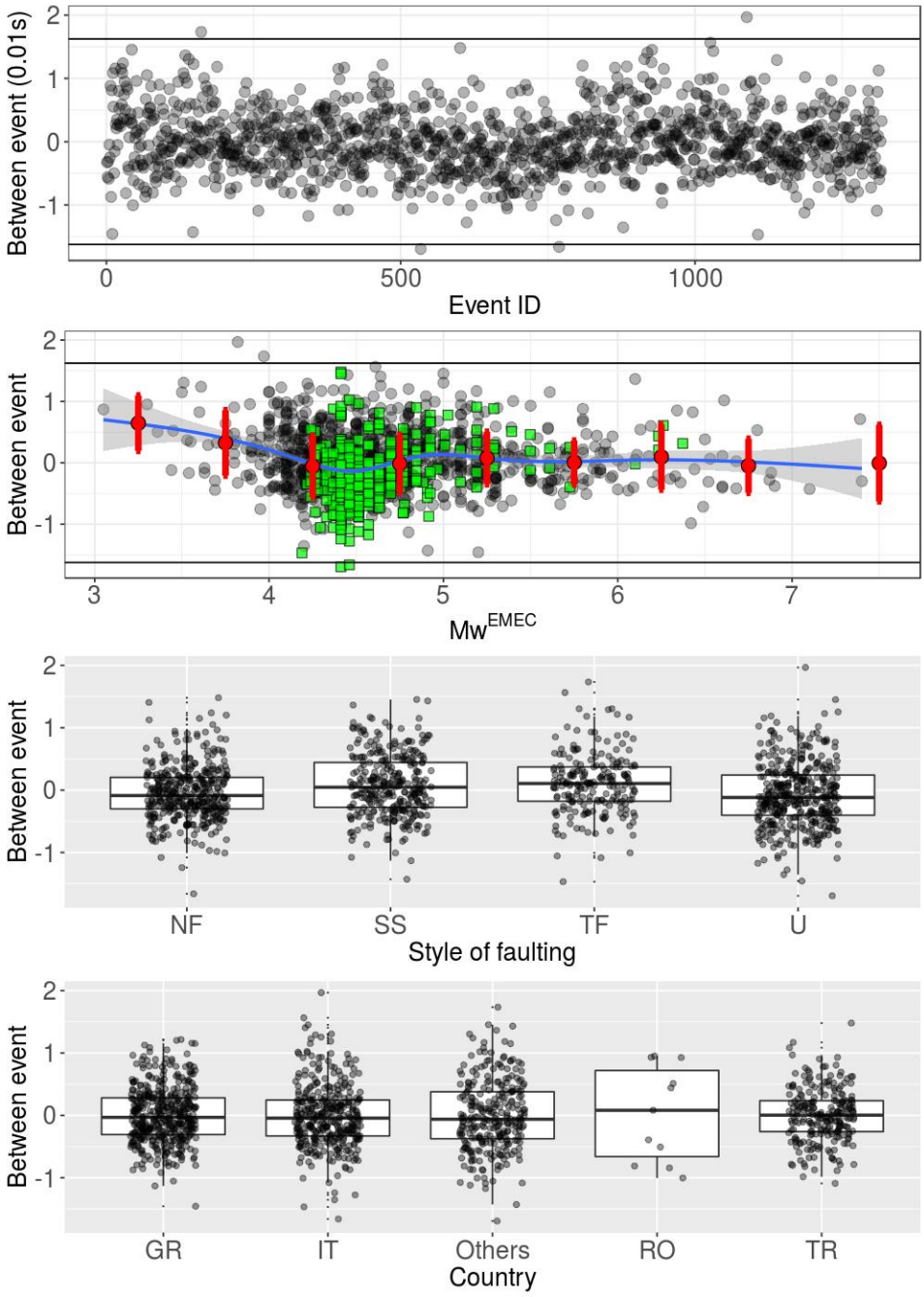


Figure 6a. Between-event residuals δB_e at T=0.01 s for SA. Green squares indicate earthquakes with Mw estimated from ML using equation (2). Horizontal lines correspond to $\pm 3\tau$. The following abbreviations are used for the style of faulting: NF- normal faulting; SS – strike slip; TF – thrust faulting; U – unknown style of faulting. The country codes are: GR- Greece; IT – Italy; RO – Romania; TR- Turkey; Others – any other country different from the previous ones.

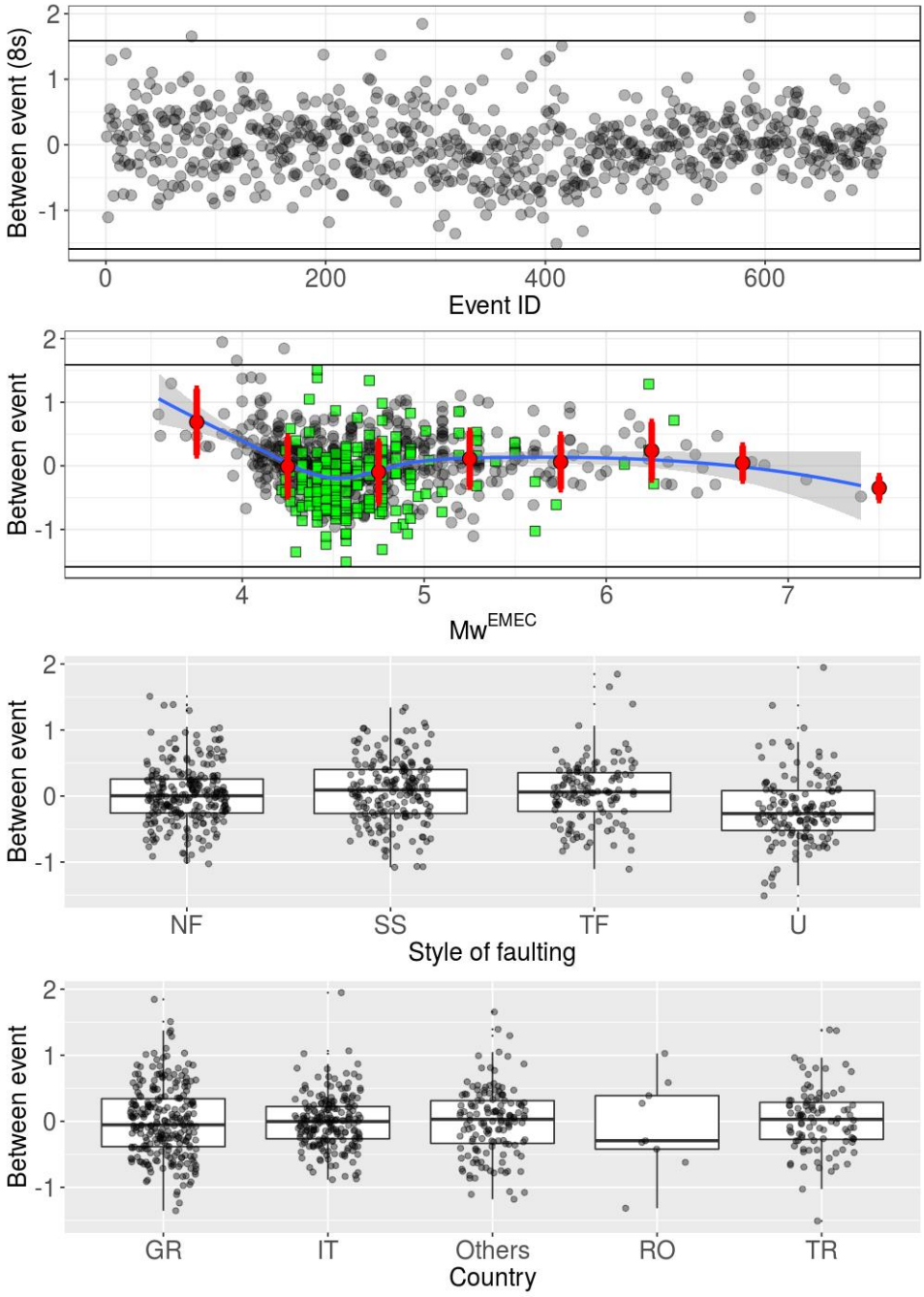


Figure 6b. The same as Figure 6a but for T=8 s

445

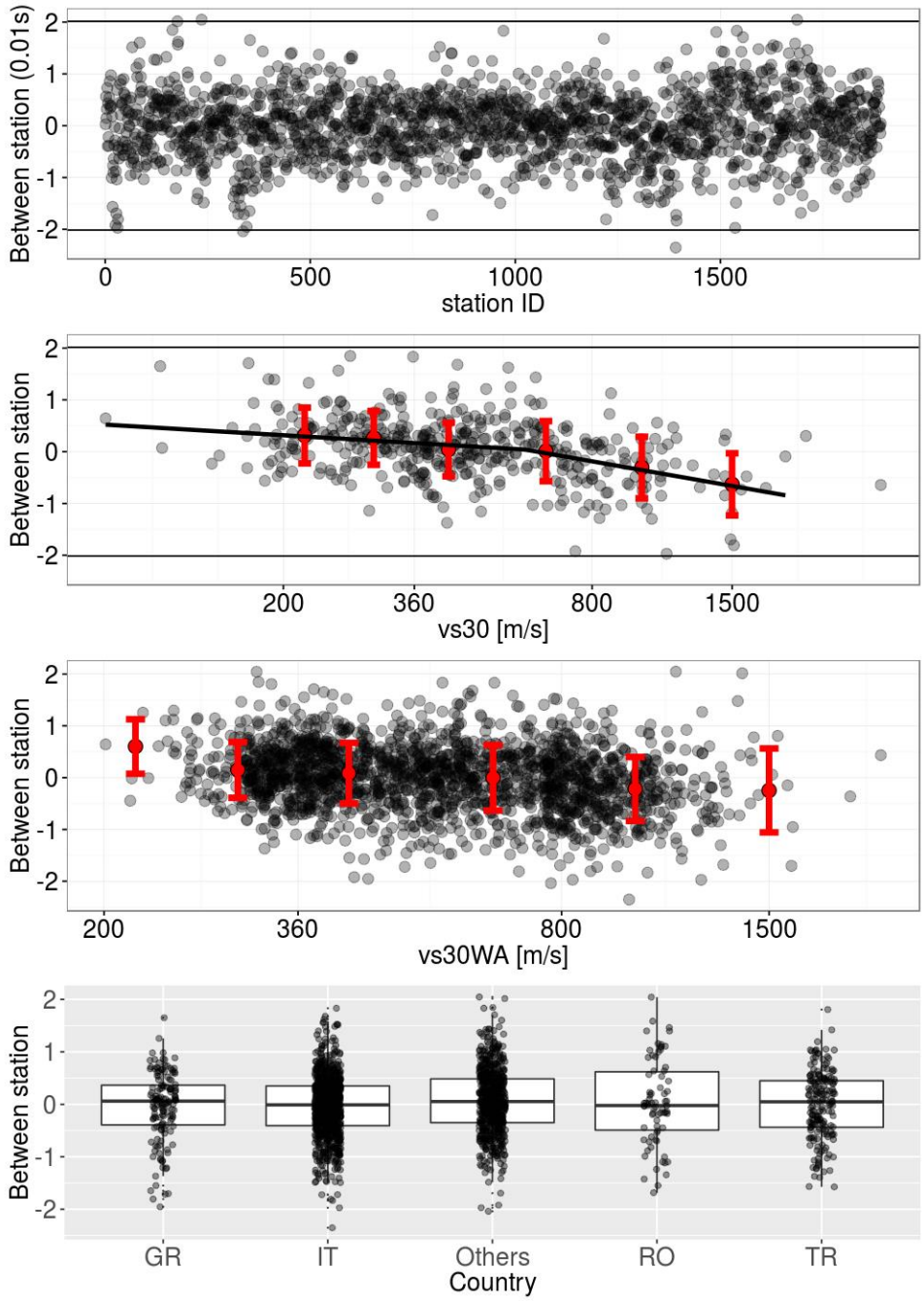
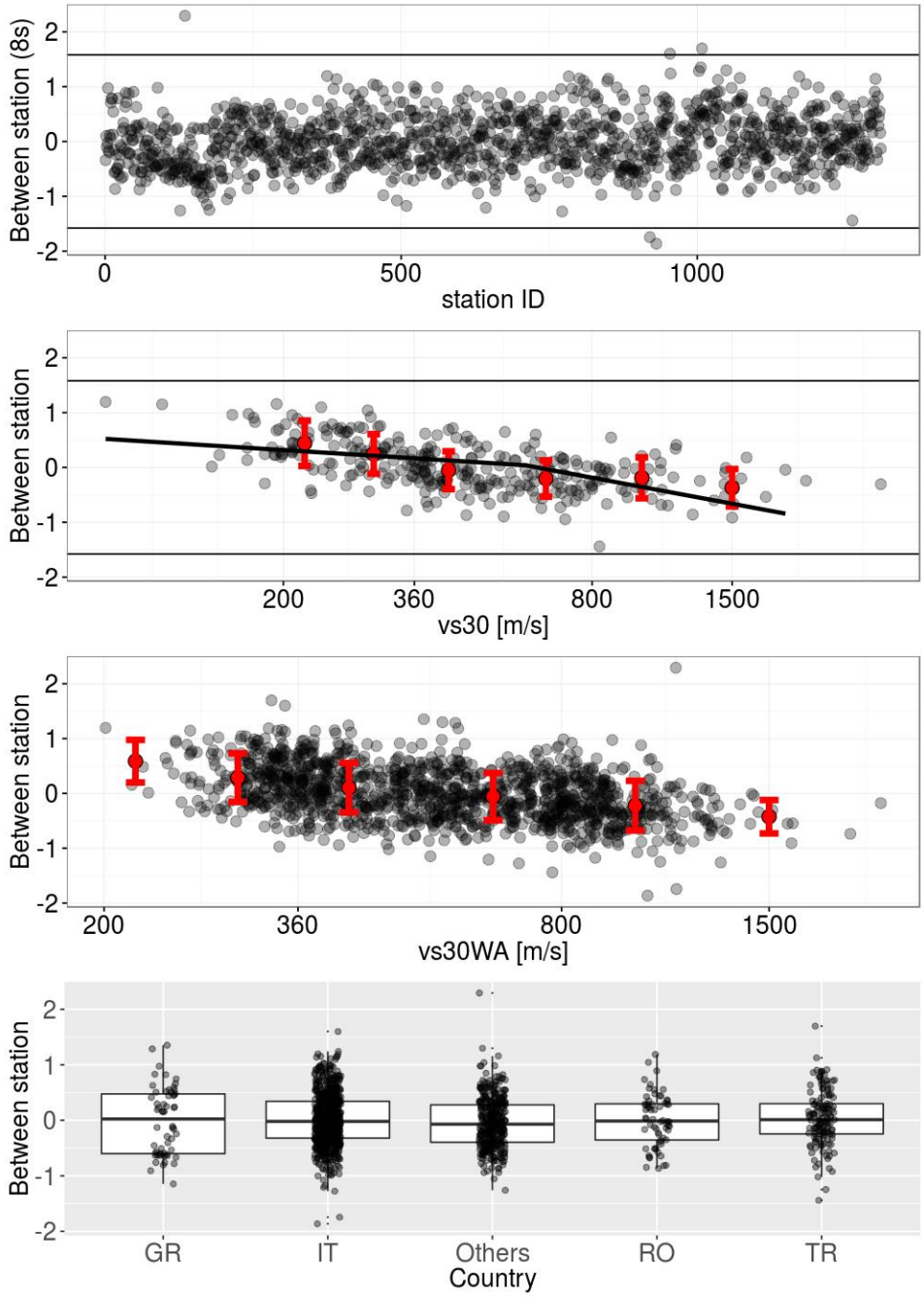


Figure 7a. Between-station residuals $\delta S2S_s$ at T=0.01 s for SA. Horizontal lines correspond to $\pm 3\phi_{S2S}$.



450

Figure 7b. The same as Figure 7a but for T= 8 s.

455

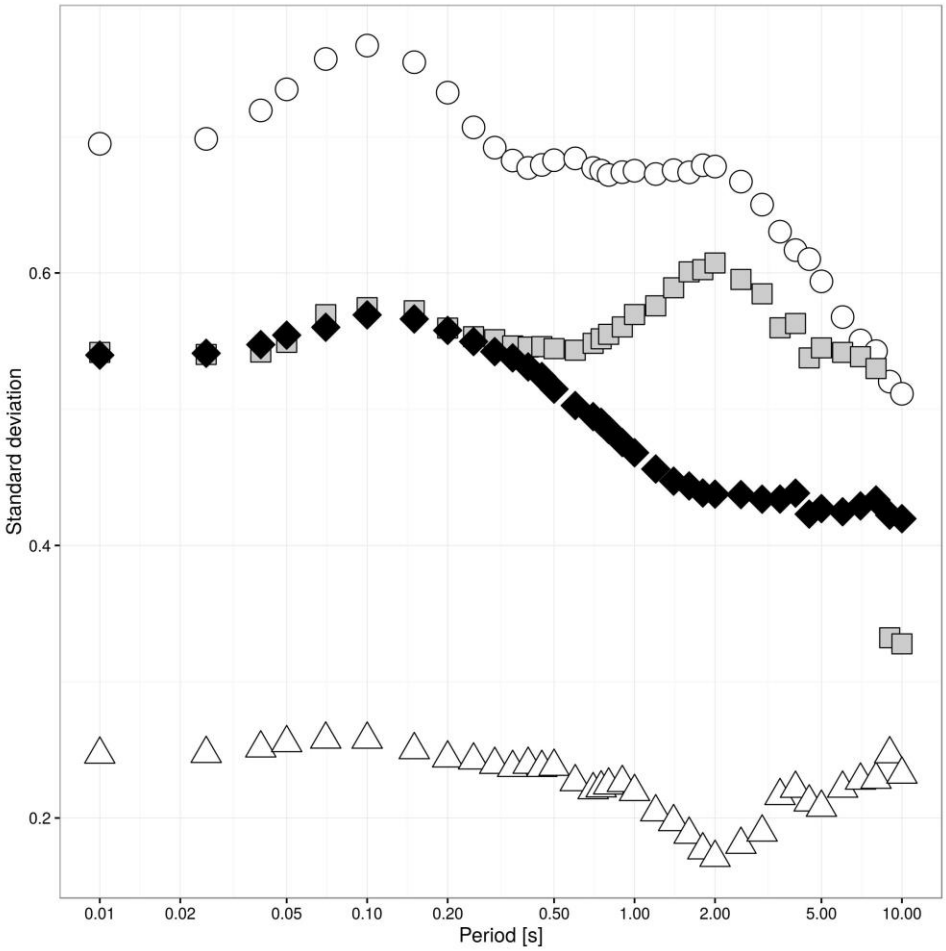


Figure 8. Standard deviation of different residual distributions for SA ($\delta C2C_c$ applied to the offset coefficient e_1). White circle: standard deviation of the between-station distribution (i.e., ϕ_{S2S}); gray square: standard deviation of the between-event distribution (i.e., τ); white triangle: standard deviation of the between-country distribution (ϕ_{C2C}); black diamond: standard deviation of the residual distribution (ϕ_0).

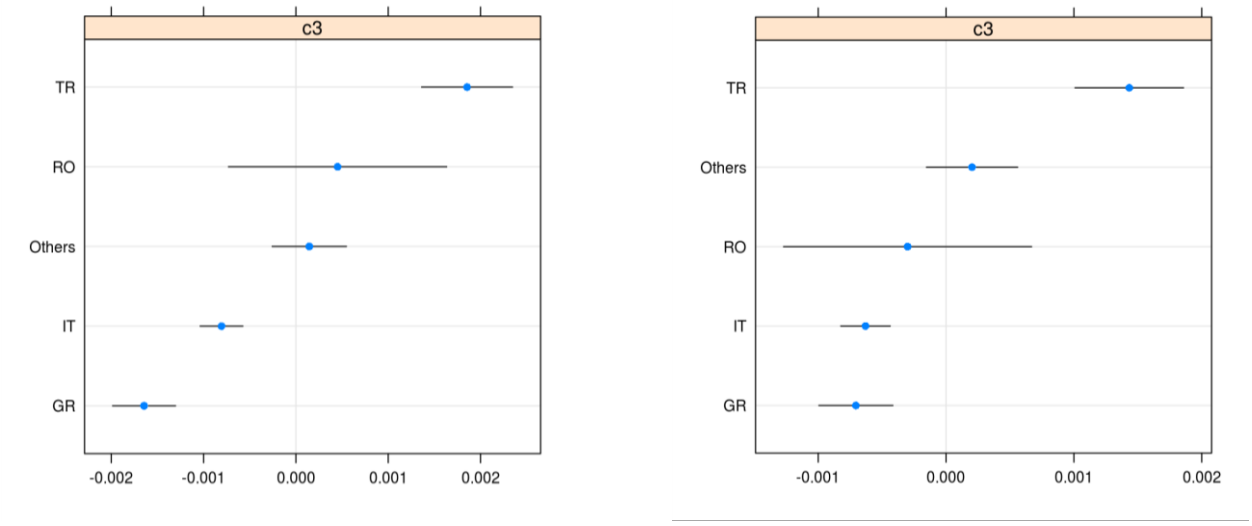


Figure 9. Random-effect on c_3 (anelestic parameter) depending on the country grouping level for T=0.01s (left) and T=2s (right), considering SA. Please, note the different order for the country classes and the different horizontal scale.

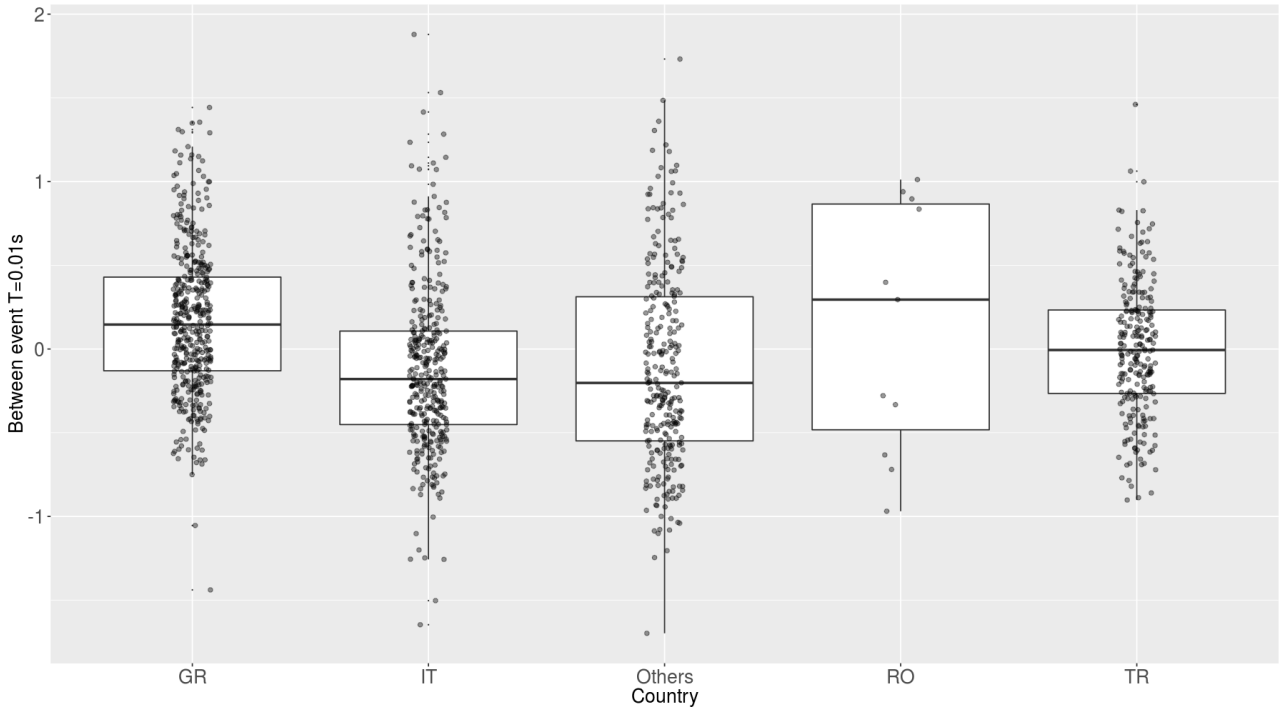


Figure 10. Between-events δB_e for SA at T=0.01s applying the country-based random-effect $\delta C2C_c$ to the anelastic attenuation parameter c_3.

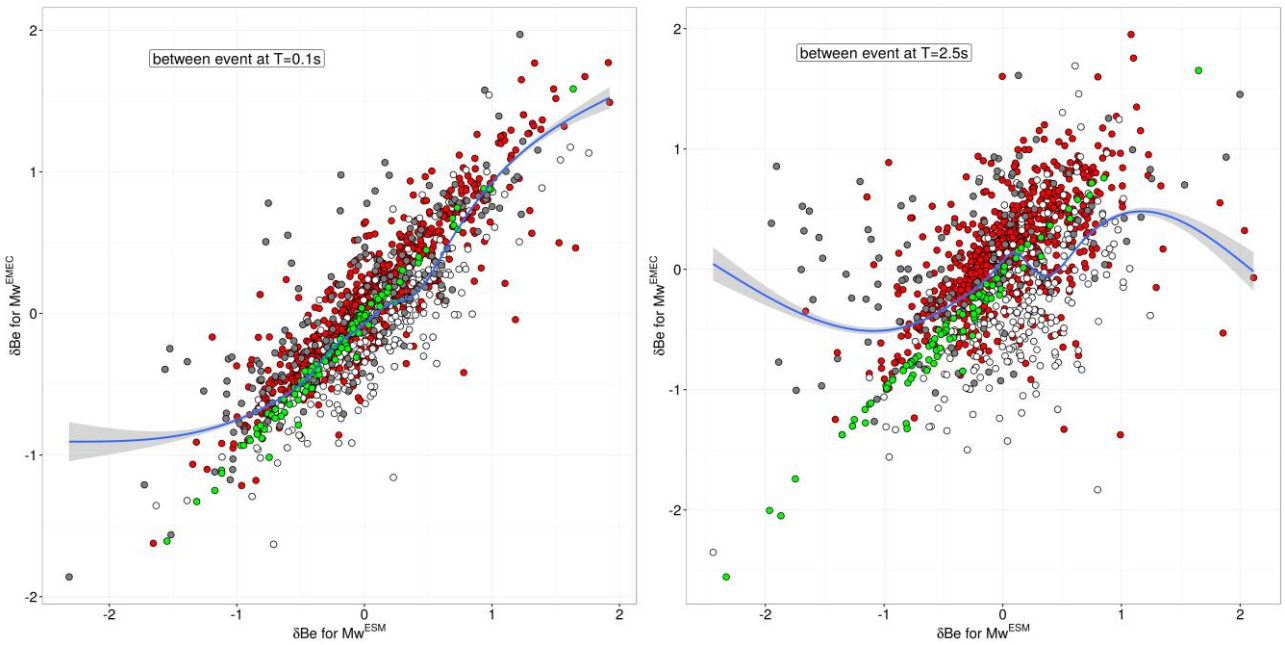


Figure 11. Comparison between δBe obtained using Mw from EMEC and ESM, considering T=0.1s (left) and 2.5s (right). The colors represent different selections on magnitude. In particular, only red represents recordings with both Mw^{EMEC} and Mw^{ESM} available, whereas the other colors indicate conversions from ML to Mw. In particular, the green dots represent earthquakes for which both Mw^{ESM} and Mw^{EMEC} are derived from the ML (and therefore showing high correlation). The blue curves are the results of a local polynomial regression (LOESS), giving an idea about the degree of correlation between the two δBe estimates.

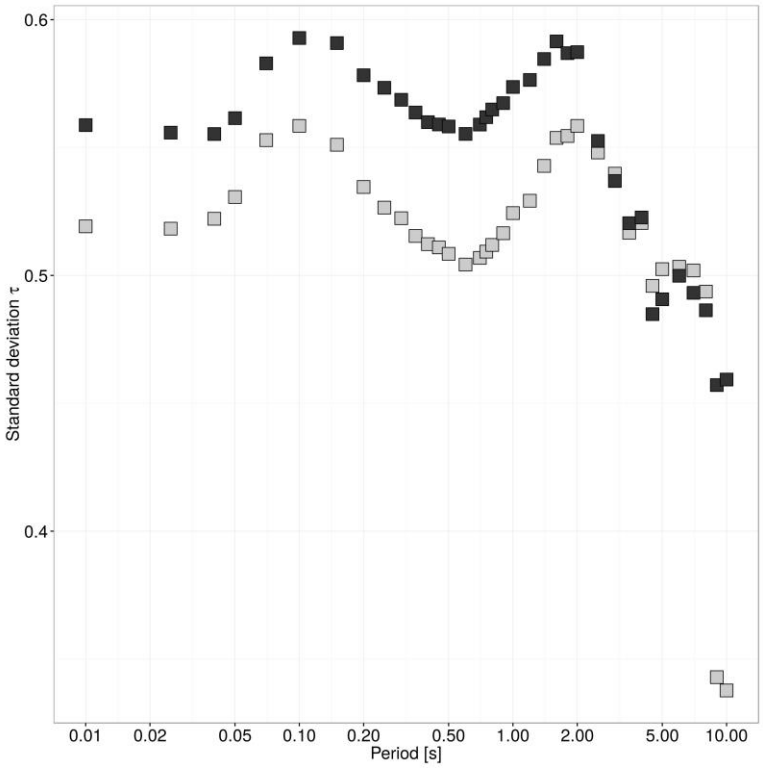


Figure 12. Between-event standard deviation τ for the regression using Mw^{EMEC} (gray) and Mw^{ESM} (black)

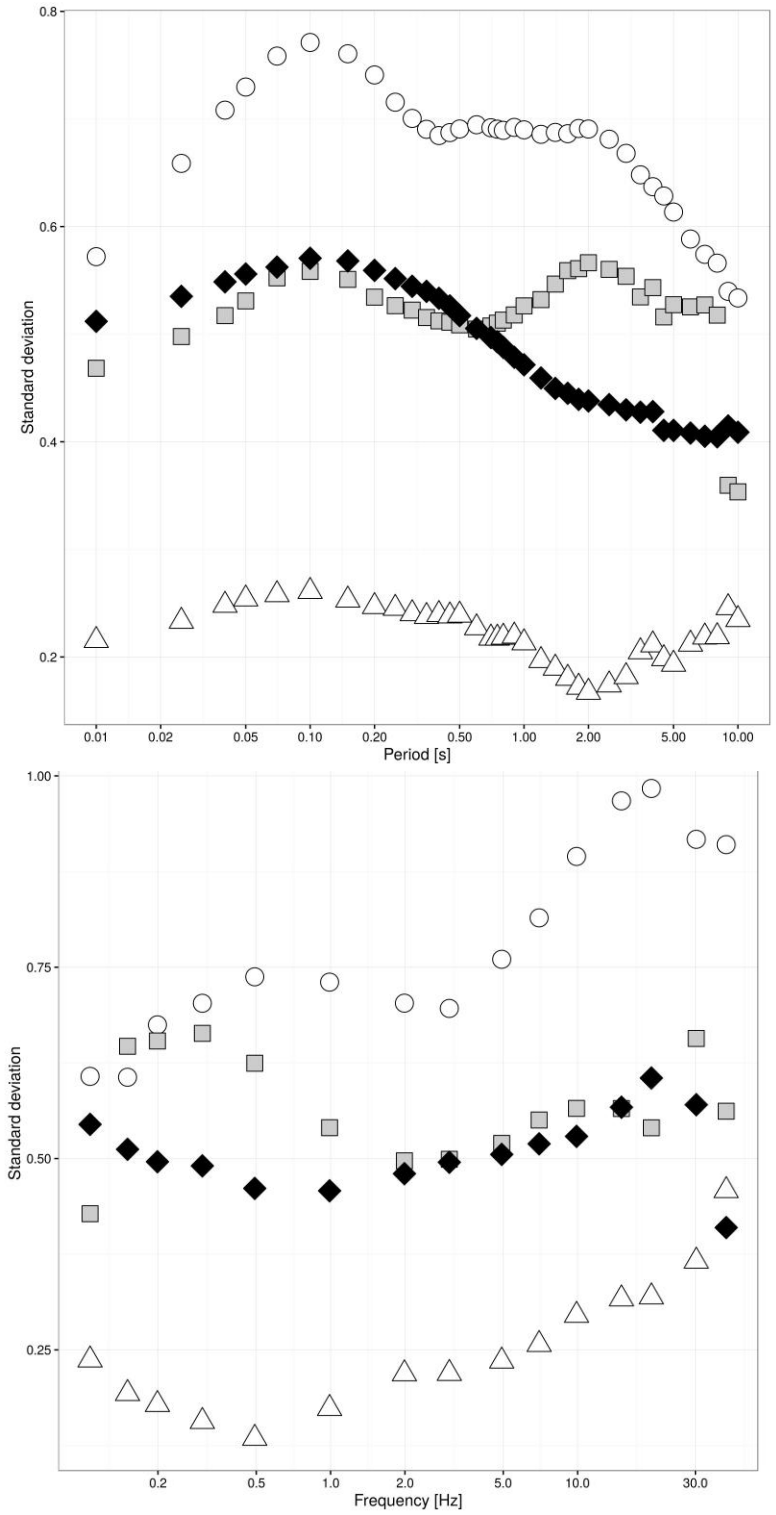


Figure 13. Standard deviation of residual distributions for spectral displacement-SD (top), according to period, and for Fourier Amplitude Spectra-FAS (bottom), according to frequency. White circle: standard deviation of the between-station distribution (ϕ_{S2S}); gray square: standard deviation of the between-event distribution (τ); white triangle: standard deviation of the between country distribution (ϕ_{C2C}); black diamond: standard deviation of the residual distribution (ϕ_0).

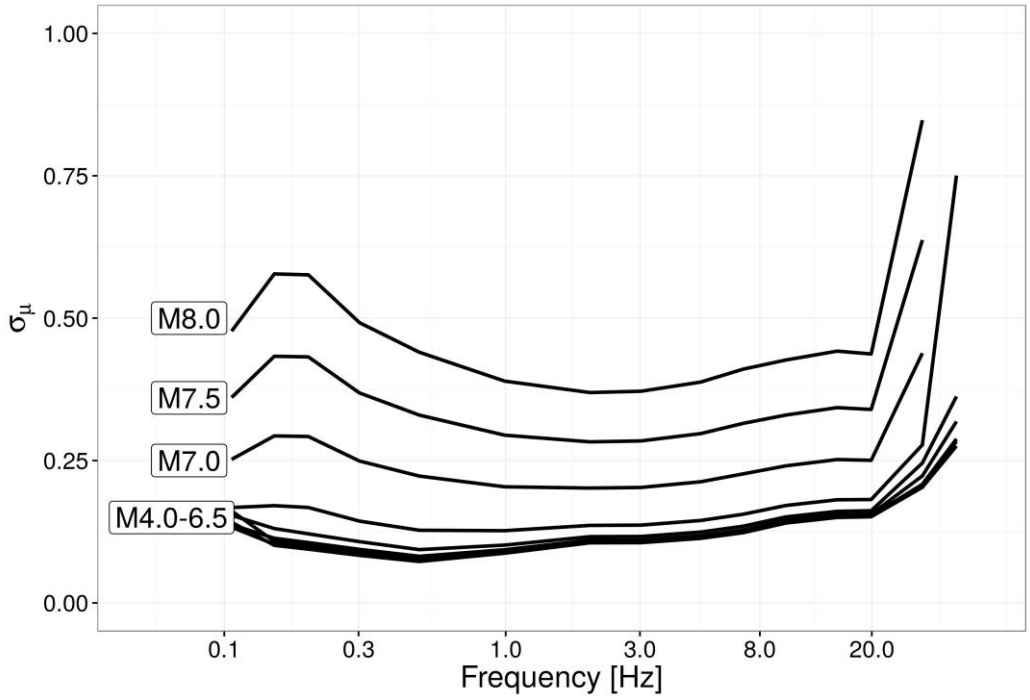


Figure 14. Frequency dependence of the asymptotic standard error on the median (σ_μ) for FAS considering different magnitudes and a distance of 30km .

The pan-European Engineering Strong Motion (ESM) flatfile: ~~sanity~~consistency check via residual analysis

Bindi D. (1), Kotha S-R.(1), Weatherill G.(1), Lanzano G.(2), Luzi L.(2), and F. Cotton (1,3)

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

(2) Istituto Nazionale di Geofisica e Vulcanologia INGV, Milan, Italy

(3) University of Potsdam, Potsdam, Germany

Abstract

We present the results of a ~~sanity~~consistency check performed over the flatfile extracted from the Engineering Strong Motion (ESM) data~~-~~base. The flatfile includes 23014 recordings from 2179 earthquakes in the magnitude range from 3.5 to 7.8 that occurred since the 1970s in Europe and Middle East, as presented in the companion article by Lanzano et al. (2018a). The ~~sanity~~consistency check is developed by analyzing different residual distributions obtained ~~by developing~~from ad-hoc ground motion prediction equations for the absolute s~~S~~pectral a~~A~~cceleration (SA), ~~D~~displacement (SD) and Fourier a~~A~~mplitude s~~S~~pectra (FAS). Only recordings from earthquakes shallower than 40 km are considered ~~for~~in the analysis. The between-event, between-station and event-and-station corrected residuals are computed by applying a mixed-effect regression. We identified those earthquakes, stations, and recordings showing the largest deviations from the ~~predicted~~GMPE median ~~models~~predictions, and ~~we~~also evaluated the statistical uncertainty on the median model to get insights ~~about~~on the applicable magnitude-distance ranges and the usable period (or frequency) range. We observed that robust median predictions~~mean~~ are obtained up to 8s for SA and up to 20 Hz for FAS, although median predictions for Mw ≥7 show significantly larger uncertainties with 'bumps' starting above 5s for SA and below 0.3 Hz for FAS. The between-station ~~variability~~variance dominates over the other residual ~~components~~variances, and the dependence of the between-station residuals on logarithm of V~~v~~s30 is well~~-~~described by a piece-wise linear function with period-dependent slopes and hinge velocity around 580 m/s. Finally, we compared the between-event residuals obtained by considering two different ~~choices~~sources of ~~for the~~ moment magnitude. The results show that, at long periods, the between-event terms from the two regressions have a weak correlation and the overall between-event variability is

~~different~~dissimilar, ~~confirming~~highlighting the importance of magnitude source ~~played~~ in the regression results ~~by the implemented moment magnitude values~~.

Introduction

Most ~~of the~~ applications in Engineering Seismology and Earthquake Engineering deal with the analysis of strong-motion data. In Europe, the dissemination of uniformly processed strong-motion waveforms accompanied by event and station meta-data progressed since the 1980s (Ambraseys, 1990; Ambraseys and Bommer, 1991; Ambraseys et al., 2000; Ambraseys et al 2004). The two most recent examples are the Reference ~~d~~Database for S~~s~~eismic G~~g~~round-motion P~~p~~rediction in Europe-RESORCE (Akkar et al., 2014), developed in the framework of the Seismic Hazard HARmonization in Europe project (SHARE; www.share.eu.org), and the Engineering Strong Motion (ESM) data-base (Luzi et al., 2016) developed in the framework of the Network of European Research Infrastructures for Earthquake Risk Assessment and Mitigation project (NERA). For many applications aim~~ed~~ing at developing and testing Ground-Motion Prediction Equations (GMPEs), ground-motion intensity measures of engineering interest (e.g., peak ground acceleration-PGA, peak ground velocity-PGV, response spectra, Housner intensity, etc) are extracted or computed from the waveforms, and disseminated~~,~~ along with the relevant event and station metadata~~,~~ in the form ~~a~~of parametric tables~~,~~ - often referred to as flatfiles. Examples of recent flat~~e~~files include the Pacific Earthquake Engineering Research Center (PEER) flatfile NGA-West-2 (Ancheta et al., 2014) and regional flatfiles like those for Japan (Dawood et al., 2016) and Chile (Bastias and Montalva., 2016).

In this work, we focus on the flatfile generated ~~by~~from the ESM data-base, as described in detail~~s~~ in a companion paper (Lanzano et al., 2018a). In order to support the development of ground-motion models from the ESM flatfile, we perform a ~~sanity~~consistency check over the data and metadata from the perspective of the ground-motion variability, to report possible anomalies that a GMPE developer might encounter. To accomplish this task, we first calibrate a set of GMPEs for spectral acceleration (SA), spectral displacement (SD) and Fourier amplitude spectra (FAS), considering simple models -describing the distance and magnitude scaling. Then, the random-effects and residual distributions obtained by performing a mixed-effects regression (Bates et al., 2015) are analyzed to detect -anomalous records, events and stations. Issues that could be of relevance for ground-motion models, such as those related to the selected magnitude scale, ~~or~~and to ground-motion and path regionalization~~,~~ are discussed as well.

Data selection for developing the reference model

In this study, we use the 2018 version of the ESM flat-file (Lanzano et al., 2018b). The characteristics of the ESM flatfile are described in Lanzano et al. (2018a) whereas the waveforms processing is detailed in Puglia et al. (2018). For developing the reference ground-motion model, we select only data at distances shorter than 300 km, generated by earthquakes with at least two recordings, and occurring at depths shallower than 40 km. The lower limit of hypocentral depth is set to select only the Active Shallow Crustal Earthquakes, whereas the deeper continental and subduction events are left out of this study considering their relatively complex phenomenon which requiringes ad-hoc studies. For SA and SD, each record used in the GMPE regression for period T should satisfy the usable period criterion $T \leq 0.8/f_{HP}$, where f_{HP} is the high-pass corner frequency of the filter applied forin the waveform processing the waveformstage. For FAS computed at frequency f, only spectral amplitudes in the range $1.2\ f_{HP} \leq f \leq f_{LP}/1.2$, where f_{LP} is the low-pass corner-frequency of the filter, are considered in deriving the GMPE. The distributions of the corner-frequencies of the filter applied to all recordings are provided in Lanzano et al. (2018b). In particular, most of the f_{LP} frequencies lie in the range 30-40 Hz. Regarding the station selection, we only select stations installed at surface (sensor depth <10 m) avoiding borehole installations (identified by *housing_code* equal to "WEL") and stations installed at building floor (*installation_code* equal to "BF"). For our sanityconsistency check, we also select sensors with *location_code* equal to "00".

Regarding the magnitude, the flatfile includes multiple magnitude scales and estimates. In particular, the local magnitude (ML) and two choices for the moment magnitude (Mw), namely the moment magnitude listed in the ESM data-base (hereinafter referred to as Mw^{ESM}) and the moment magnitude from the EMEC catalog (Grünthal et al., 2012; hereinafter referred to as Mw^{EMEC}). In particular, while Mw^{ESM} is computed from moment tensor solutions provided by different sources selected according to an hierarchical schema (Lanzano et al., 2018b), Mw^{EMEC} is based on both moment tensor solutions and empirical regional conversions from other magnitude scales. The two Mw scales are compared in Figure 1a, where the best fit line is $Mw^{EMEC}=0.23+0.95Mw^{ESM}$, beingwith the residual standard deviation equal to 0.16. Since neither Mw^{ESM} nor Mw^{EMEC} is available for all events, we derive an empirical relationship between ML and Mw in order to extend the residual analysis to as larger as possible a number of events as possible. These relations are not proposed for applications different from those performed in the context of this study. The best fit lines shown in Figure 1 (right) are:

Mw^{ESM} $=6.81-1.67\ ML+0.26\ ML^2$ (± 0.3) (1)

$Mw^{EMEC}=6.18-1.34\ ML+0.23\ ML^2$ (± 0.3) (2)

GMPE functional form

Mixed-effects regressions (Brillinger and Preisler, 1985; Bates et al., 2015) are performed for the natural logarithm of SA, SD and FAS considering simple parametric functional forms ~~for~~ describing the median scaling with distance and magnitude:

$$logY = e_1 + F(M_w) + G(R_{hypo}) + RE + \epsilon \quad (3)$$

$$F(M_w) = \begin{cases} b_1(M_w - M_{ref}) if M_w \leqslant M_H \\ b_1(M_H - M_{ref}) + b_3(M_w - M_H) otherwise \end{cases} \quad (4)$$

$$G(R_{hypo}) = \begin{cases} c_{1A} log(R_{hypo}) if R_{hypo} \leqslant R_H \\ c_{1A} log(R_H) + c_{1B} log\left(\frac{R_{hypo}}{R_H}\right) + c_3 R_{hypo} otherwise \end{cases} \quad (5)$$

where in equation (1) the random effect (RE) are

$$RE = \delta B_e + \delta S2S_s + \delta C2C_c \quad (6)$$

In equations (4) and (5), the reference and hinge magnitude and distances are set to M_{ref}=4.5, M_H=6, R_H=15km, respectively. The country-to-country random effect ($\delta C2C_c$), where *c* indicates country, is applied either on the offset term (e_1) or to c_3 (that is, two different regressions are performed). The other random-effects applied to the offset are the customary between-event (δB_e) and between-station ($\delta S2S_s$) residuals, where subscripts *e* and *s* are event and station indices. Residual ε in equation (3) represents the event- and station-corrected record-to-record variability. The standard deviations of ε, ~~$\delta S2S$~~$\delta S2S_s$, $\delta C2C_c$, and δB_e random distributions are indicated with ϕ_0, ϕ_{S2S}, ϕ_{C2C} and τ, respectively. The

Formatted: Subscript

Formatted: Subscript

regressions for SA and SD are performed over the orientation independent (rotD50) ~~component~~ground-motion measure (Boore, 2010), while the geometric ~~-al~~ mean of the horizontal components is used for FAS GMPEs.

For the GMPEs regressed on SA considering $\delta C2C_c$ on the offset e_1, Figure 2 shows the period-dependent coefficients along with their standard ~~-~~errors. Regarding the scaling with distance, coefficient c_{1B} increases from about -2 at T<0.1s to about -1.3 for T>0.5s while coefficient c_{1A} is maximum around T=0.1s and decreases below -2 for T>4s. The term c_3 is negative for short periods and turns positive for T $\geq$ 1.5s. The positive values of c_3 suggest to investigate ~~-a~~ further segmentation of the distance scaling component to ~~allow~~ captur~~inge~~ the effects of late~~r~~ arrivals on long period ground ~~-~~motion at distances beyond 80km (e.g., Lanzano et al., 2016; Kotha et al., 2018).

Regarding the scaling with magnitude, although large uncertainties affect b_3 (describing the magnitude scaling above the hinge magnitude), its values are ~~which is~~ always smaller than b_1, suggesting weaker scaling at larger magnitudes. The trend of the coefficients suggest that regressions above 8s are unstable due to a strong reduction in the number of data ~~used~~usable in ~~for~~ GMPE regressions (Lanzano et al., 2018a). For example, the number of recordings used for the regressions at 0.01, 4, 8, and 10 s are 18859, 16222, 9664 and 2767, respectively. The ~~statistical uncertainty~~asymptotic standard error (σ_μ) on the median predictions, quantifying the within-model statistical estimate of epistemic uncertainty ~~affecting the median predictions~~ (e.g., Al Atik and Youngs, 2014 ; Bindi et al., 2017a), due to the combined effect~~s~~ of limited data availability and complexity of implemented functional form (in terms of extent of GMPE parametrization) are shown in Figure 3. For Mw $\leq$6.5, σ_μ is smaller than 0.2 and it shows a weak dependency on magnitude for Mw<6. For Mw<6.5, σ_μ reaches the minimum values around T=2s while the sharp increase above T=8s indicates that regressions ~~at~~ are not well ~~-~~constrained for these periods. ~~For~~At Mw > 6.5, σ_μ increases ~~significantly~~rapidly, exceeding 0.25 at Mw=7.5. ~~-~~Above magnitude 7, σ_μ shows a bump around T=5s, suggesting that at periods longer than 4s GMPE median predictions for Mw >7 are ~~not well constrained~~ uncertain. The dependence of σ_μ on distance (here not shown) is negligible because the data appears to constrain very well the distance scaling component of the GMPE (small errors on c_{1A}, c_{1B}, c_3 in Figure 2).

Results

The results are presented in terms of event- and station--corrected residuals (ε), between--event (δB_e) and between--station ($\delta S2S_s$) distributions evaluated for SA with regional attributes $\delta C2C_c$ on e_1.

Event- and station--corrected residuals ε

The event- and station--corrected residuals for T=0.01s and T=8s are shown in Figure 4 as function of the recording Id (top), of magnitude (middle) and of distance (bottom). In the middle and bottom frames, the average residuals computed over magnitude and distance bins (red error bars) are shown as well. Please note that the number of recordings used in GMPE regression decrease towards long periods. The residuals do not show any significant trend ~~n~~either with distance ~~n~~or with magnitude. The number of residuals (in absolute value) above 3 times ϕ_0 is 60 at T=0.01 s and 20 at T=4s, as shown in Tables 1 and 2, respectively. A further check of the waveforms does not reveal any peculiarities in the recordings except for the records [IT-2016-0001, IV.INTR], which is characterized by ~~signal-to-noise ratiosmalla~~ high noise level at high frequencies, as shown in Figure 5.

Between--event random -effects δB_e

The between--event random-effects for T=0.01s and T=8s are shown in Figures 6a and 6b ~~for T=0.01s and T=8s~~, respectively. δB_e does not show significant trend for M> 4.5, while δB_e increases with decreasing magnitude for M<4.5. Unlike the functional form implemented in this study, an additional segmentation of F(Mw) at Mw = 4.5 (e.g, Campbell and Bozorgnia 2014, Kotha et al. 2018) would allow different scaling for small, moderate and large magnitude bins, removing the trend observed for small magnitudes. We note here that moment magnitudes below 4 are ~~mostly generated~~not converted ~~by conversions~~ from local magnitude.

Differences in observed SAs based on Style--of--Faulting (SoF) are widely debated in GMPE development. For the dataset analyzed in this study, the box and whisker plots corresponding to δB_e dependence on SoF do not show any systematic differences. Since the model included a random--effect ~~on~~at the ~~C~~country level (Italy-IT, Turkey-TK, Romania-RO, Greece-GR, other countries-Others), a dependence of δB_e on the country grouping is not expected, as shown by the box and whisker plots in Figures 6a and 6b. Events with δB_e larger than 3 times τ, in absolute value, are listed in Table 3 and 4 for T=0.01 and 8 s, respectively.

Formatted: Subscript

Between-station random-effects δS2S.

The between-station ~~residuals~~random-effects for T=0.01s and T=8s are shown in Figures 7a and 7b ~~for T=0.01s and T=8s~~, respectively. At short periods, $\delta S2S_s$ shows a trend with Vs30, as expected ($\log_{10}$Vs30) is commonly used as proxy for site amplifications). The best-fit model obtained through a break-point regression (Muggeo, 2003) is described by a piece-wise linear function with break-point at $\log_{10}$(V~~ɣ~~s30)=(2.78 ± 0.09) (i.e., V~~ɣ~~s30=600 m/s) and slopes (-0.6 ±0.2), (-1.7±0.40). A weaker relation exists between $\delta S2S_s$ and V~~ɣ~~s30WA (which is an estimate of V~~ɣ~~s30 derived from the topographic gradient, see Lanzano et al., 2018b). At long periods, the trend with measured V~~ɣ~~s30 is strong only for V~~ɣ~~s30<600 m/s, while considering V~~ɣ~~s30WA a correlation over a broader velocity range is observed. For T=8s, the break-point is located at $\log_{10}$Vs30=(2.77±0.09), and the slopes are (-1.6 ±0.2) and (-0.3 ±0.2).

When split among the different countries (box and whiskers plots in Figure 7), no residual offset is detected. Often the station V~~ɣ~~s30 distributions (mean and range) differ from region-to-region depending on the local geology over which the networks reside. As a result, systematic region-dependent offsets could be expected. In our analysis, since all the country-to-country variability is accounted by the $\delta C2C_c$ random-effects on e_1, we observe no differences in mean $\delta S2S_s$ among the countries. Nevertheless, it is worth of further investigation whether the scaling of $\delta S2S_s$ with V~~ɣ~~s30 is identical across various regions or not. Kotha et al. (2016) for instance, observed a similar (regionally) unbiased $\delta S2S_s$ for Italy and Turkey, but with significantly different scaling with ~~ɣ~~Vs30. Stations with $\delta S2S_s$ larger than 3 times ϕ_{S2S}, in absolute value, are listed in Table 5 and 6 for T=0.01 and 8 s, respectively.

Standard deviations of random-effects

The standard deviations of the random-effects are summarized in Figure 8. The between-station variability ϕ_{S2S} dominates the total aleatory variability of the dataset (σ): it shows a peak around 0.10 s and it decreases above 2s. The peak at T = 0.1s is observed in several other studies as well (e.g. Zhao et al. 2016, Cauzzi and Faccioli 2017; Kotha et al. 2018), indicating a high variability of short period site-response in response spectra domain. The between-event standard deviation τ has a peak located around 2s, otherwise almost constant with a small bump around 0.1 s; τ shows a suspicious drop at T=9s

and 10s, probably due to the strong reduction in the usable recordings. The event- and station--corrected variability (ϕ_0) is close to τ for T<0.5s, while for longer periods it is smaller than the between-event variability. Finally, the between-country variability ϕ_{C2C} (region-to-region variability of offset e_1) is small (of the order of 0.25) with a trough around 2s. The anti-symmetry of τ and ϕ_{C2C} around T = 2s suggests a trade-off between event-to-event and country-to-country variability in this GMPE regression, suggesting the need of further investigating the formulation of REs in GMPEs.

Regional differences on anelastic attenuation.

An additional regression is performed by applying the country-group random--effect not on the overall offset e_1 in equation 3 but on the anelastic coefficients c_3 of equation5. Figure 9 shows the obtained random effects δc_3 , which can be interpreted as regional adjustment to the generic region-independent "anelastic attenuation" term c_3. The~~e estimation errors on th~~ese country-specific adjustments (to c_3) are statistically significant with small standard errors (horizontal bar in the caterpillar plot of Figure 9), suggesting considerable differences in propagation effects across these regions. For instance, since the δc_3 value for Romania is not well-constrained with the very few data from that region, its standard error is larger than that for Italy, Greece, and Turkey. For periods between 0.01 and 2 s, Greek and Italy show stronger attenuation than Turkey; for longer periods we recall that c_3 becomes positive.

Since in this case the country-based RE is not applied to the overall offset, the between-- events show some variability ~~with~~across the countries~~y class~~ (Figure 10). Note that since $\delta C2C_c$ is no longer applied on e_1, systematic country-to-country variability in δB_e and $\delta S2S_s$ distributions may now appear, as seen in the between-event distributions --of Figure 10. For example, on average the events from Greece (GR) show a higher δB_e compared to those from Turkey (TR) and Italy (IT) (Figure 9).

Formatted: Subscript

Formatted: Subscript

δB_e obtained using either Mw^{EMEC} or Mw^{ESM}.

Several studies (e.g., Rohades, 1997; Moss, 2011; Kuehn and Abrahamson, 2018) evaluated the impact of the uncertainty in the ~~explanatory~~predictor variables on the aleatory variability. The effect of ~~the~~

magnitude uncertainty on δB_e variability ~~have~~has been observed in Europe for the RESORCE flatfile by comparing Mw values computed following different procedures (Kotha et al., 2016). For the ESM flatfile, we assess the impact of using Mw^{EMEC} or Mw^{ESM} by repeating the regression for each magnitude type, with $\delta C2C_c$ applied on e_1. The results are compared in Figure 11 in terms of δB_e for T=0.1 and 2.5s. The red dots in Figure 11 indicate ~~-~~recordings for which both ~~-~~Mw^{EMEC} and Mw^{ESM} ~~-~~are available, while the other colors indicate earthquakes for which at least one of the two Mw values is derived from ML~~l~~. In particular, the green dots are for earthquakes without Mw (neither EMEC nor ESM). Figure 11 shows that at short periods δB_e for the two regressions show a good correlation although with a large scatter. On the other hand, at longer periods (T=2.5s in Figure 11 but results for longer periods are similar) the correlation is very weak (being trivial the correlation shown by the green points). Figure 12 shows the standard deviation τ of the δB_e residuals for the two regressions, considering both actual Mw values and values derived from M~~l~~L. τ for Mw^{EMS} is higher than the one obtained using Mw^{EMEC} for periods T<2.0s whereas, from T=2.0 to 8s, similar values are obtained for the two regressions.

Formatted: Subscript

Formatted: Subscript

Summary for SD and for FAS

The ESM flatfile includes several intensity measure of engineering interest~~s~~, such as spectral acceleration (SA) and displacement (SD); Fourier amplitude spectra (FAS); peak ground velocity (PGV) acceleration (PGA) and displacement (PGD); Arias intensity, Housner intensity, effective duration, cumulative absolute velocity (CAV). While the check of all those parameters is beyond the aim of this work, the residual analysis have been extended to SD and FAS as well (Bindi et al., 2017b). Figure 13 summarizes the results by comparing the standard deviations of the residual distributions. The main conclusions drawn from SD and FAS analysis agree with those from SA. Regarding the uncertainty on the median model for FAS, the dependency of σ_μ on frequency is shown in Figure 14. The models are robust up to 20 Hz, for higher frequencies the number of available recording decreases due the condition applied to the usable frequency range (see Lanzano et al., 2018b, for the distribution of the low-pass corner frequencies of the filter applied during the processing phase) and the uncertainty becomes very large. Towards low frequencies, σ_μ for $Mw \geq 6.5$ shows a bump around 0.2 Hz. In general, in agreement with the results obtained for SA, σ_μ significantly increases above magnitude 7.

Conclusion

We performed residuals analysis over the ESM flat-file (Lanzano et al., 2018b) as complementary information to the flatfile description (Lanzano et al., 2018a), for supporting those studies aiming at developing ground-motion models. The main outcomes of the ~~sanity~~consistency check performed on recordings from shallow earthquakes (hypocentral depth less than 40 km) can be summarized as follow:

1) considering the data population and the selection criteria applied to define the usable period range, robust results for response spectra have been obtained for periods up to 8s, although the median models for Mw≥7 are affected by large within-model epistemic uncertainty (σ_μ) above 5s. Regarding the model for Fourier, σ_μ increases significantly above 20Hz and shows a bump around 0.2 Hz for magnitudes above 7. The epistemic uncertainty for both response and Fourier spectra increases significantly for Mw≥ 6.5;

2) the sanity check has been performed considering a simple functional form for describing the magnitude scaling and the scaling with hypocentral distance, although for larger finite-source events the Joyner-Boore or the rupture distances should be considered. Regarding the distance scaling, a piece-wise linear function of logarithm of distance was needed to avoid trends in the residuals at distances shorter than 15km (see also Kotha et al., 2018). Dependence of the hinge distance on magnitude and period is worth ~~of~~ future investigation in order to model the magnitude dependence of geometric spreading as well;

3) regional dependencies of the attenuation with distance (so-called apparent anelastic attenuation parameter) already observed in previous studies (e.g., Kotha et al., 2016) are confirmed. In addition, it is now possible to refine the regionalization scheme given the significantly larger amount of data from a few more regions;

4) regarding the scaling of SA with magnitude, it is worth considering a second hinge magnitude at small magnitudes, around 4 (see also Kotha et al 2018);

5) the scaling of the between-station ~~residuals~~random-effects $\delta S2S_s$ with the logarithm of V~~v~~s30 is well described by a piece-wise linear function with period-dependent slopes and hinge velocity around 580 m/s;

6) magnitude is one of the two fundamental explanatory variables used to develop ground motion models, and the flat-files includes information about different magnitude scales ~~as~~ extracted from the ESM data-base. In ~~particular~~addition, two different moment magnitude values are provided (when

available): Mw^{ESM} is the moment magnitude as included in ESM without any conversion (a reference for the source is provided); Mw^{EMEC} is the moment magnitude according to the EMEC catalog (Grünthal et al., 2012). ~~Being based on~~Derived through different procedures, the two magnitudes show an overall agreement (linear correlation) but with large scatter around the mean trend. The between-event residuals δB_e at long periods absorb the uncertainties in the seismic moment estimates: therefore, the weak correlation shown by the δB_e distribution obtained considering the two different moment magnitudes warns about the critical role played by the magnitude selection;

7) ~~tables with r~~Records, events, and stations showing standardized residuals larger, in absolute value, than 3 are listed in tables to allow for a comparison with other models.

8) among different intensity measures, Fourier amplitude spectra allow an easier and direct comparison between components of the observed ground-motion variability and seismological interpretations.

Data and Resources

In this study, we analyzed the ESM flatfile (Lanzano et al, 2018b) with persistent identifier PID: 11099/ESM_flatfile_2018. The flatfile can be downloaded from the ESM webpage: http://esm.mi.ingv.it.

Acknowledgments

This work has been developed within the European Plate Observing System (EPOS; https://epos-ip.org/) and the Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe (SERA; http://www.sera-eu.org) projects. Comments from B. Edwards and one anonymous reviewer helped us to improve the manuscript.

References

Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BS-J, Wooddell KE, Graves RW, Kottke AR, Boore DM, Kishida T, Donahue JL (2014) NGA-West2 database. Earthquake Spectra 30(3):989–1005

Akkar S, Sandıkkaya M A, Şenyurt M, Azari Sisi A, Ay B Ö, Traversa P, Douglas J, Cotton F, Luzi L, Hernandez B and Godey S (2014) Reference database for seismic ground-motion in Europe (RESORCE), Bull. Earthquake Eng 12:311–339, DOI 10.1007/s10518-013-9506-8

Al Atik L, Youngs B (2014) Epistemic uncertainty for NGA-West2 models. Earthquake Spectra 30(3):1301–1318

Ambraseys NN (1990) Uniform magnitude re-evaluation of European earthquake associated with strong-motion records. Earthq Eng Struct Dyn 19:1–20

Ambraseys NN, Bommer JJ (1991) Database of European strong-motion records. Eur Earthq Eng 5:18–37

Ambraseys N, Smit P, Berardi R, Rinaldis D, Cotton F, Berge-Thierry C (2000) Dissemination of European strong-motion data. CD-ROM collection. European Commission, Directorate General XII, Science, Research and Development, Environment and Climate Programme, Bruxelles

Ambraseys NN, Smit P, Douglas J, Margaris B, Sigbjörnsson R, Olafsson S, Suhadolc P, Costa G (2004) Internet site for European strong-motion data. Bollettino di Geofisica Teorica ed Applicata 45:113–129

Bastías, N, Montalva GA (2016) Chile strong ground motion flatfile, Earthq. Spectra 32, 2549–2566, doi: 10.1193/102715EQS158DP.

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

Bindi D, Cotton F, Kotha S, Bosse C, Stromeyer D, Grünthal G (2017a) Application-driven ground motion prediction equation for seismic hazard assessments in non-cratonic moderate-seismicity areas, Journal of Seismology 21, 5, 1201-1218.

Bindi D, Spallarossa D, Pacor F (2017b) Between-event and between-station variability observed in the Fourier and response spectra domains: comparison with seismological models, Geophysical Journal International 210, 1092–1104, https://doi.org/10.1093/gji/ggx217

Boore DM (2010) Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. Bull Seim Soc Am 100(4):1830-1835

Brillinger DR, Preisler HK (1985) Further analysis of the Joyner-Boore attenuation data. Bull. Seism. Soc. Am. 75, no. 2, 611–614.

Campbell KW, Bozorgnia Y (2014) NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. Earthquake Spectra 30,1087–115.

Cauzzi C, Faccioli E (2017) Anatomy of sigma of a global predictive model for ground motions and response spectra. Bull Earthq Eng, 1–19. doi: 10.1007/s10518-017-0278-4

Dawood HM, Rodriguez-Marek A, Bayless J, Goulet C, Thompson E, (2016) Flatfile for the KiK-net database processed using an automated protocol. Earthq Spectra 32,1281–1302.

Grünthal G, Wahlström R (2012) The European-Mediterranean Earthquake Catalogue (EMEC) for the last millennium. J Seismol 16(3):535–570

Kotha SR, Bindi D, Cotton F (2016) Partially non-ergodic region specific GMPE for Europe and Middle-East. Bull Earthq Eng,14:1245–1263.

Kotha SR, Cotton F, Bindi D (2018) A new approach to site classification: Mixed-effects Ground Motion Prediction Equation with spectral clustering of site amplification functions, Soil Dynamics Earth Eng, doi: 10.1016/j.soildyn.2018.01.051

Kuehn NM, Abrahamson NA (2018) Uncertainty in predictor variables on the estimation of Ground-Motion Prediction Equations, Bull. Seism. Soc. Am 108, doi: 10.1785/0120170166

Lanzano G, D'Amico M, Felicetta C, Puglia R, Luzi L, Pacor F, Bindi D (2016) Ground motion prediction equations for region specific probabilistic seismic hazard analysis. Bull Seim Soc Am 106(1):73–92

Lanzano G, Sgobba S, Luzi L, Puglia R, Pacor F, Felicetta C, D'Amico M, Cotton F, Bindi D (2018a) The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and data statistics. Bull Earthquake Eng (submitted).

Lanzano G, Puglia R, Russo E, Luzi L, Bindi D, Cotton F, D'Amico M, Felicetta C, Pacor F & ORFEUS WG5 (2018b) ESM strong-motion flat-file 2018. Istituto Nazionale di Geofisica e Vulcanologia (INGV), Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum (GFZ), Observatories & Research Facilities for European Seismology (ORFEUS). PID: 11099/ESM_flatfile_2018

Luzi L, Puglia R, Russo E, D'Amico M, Felicetta C, Pacor F, Lanzano G, Çeken U, Clinton J, Costa G, Duni L, et al (2016) The Engineering Strong-Motion Database: A Platform to Access Pan-European Accelerometric Data, Seismological Research Letters 87, 987-997. doi: 10.1785/0220150278

Moss RES (2011) Reduced sigma of ground-motion prediction equations through uncertainty propagation, Bull. Seismol. Soc. Am. 101, 250–257, doi: 10.1785/0120090325

Muggeo VMR (2003) Estimating regression models with unknown break-points. Statistics in Medicine, 22, 3055-3071

Puglia R, Russo E, Luzi L, D'Amico M, Felicetta C, Pacor F, Lanzano G (2018) Strong motion Processing Service: a tool to access and analyse earthquakes strong motion waveforms. Bull Earthquake Eng., doi: 10.1007/s10518-017-0299-z.

Rhoades DA (1997) Estimation of attenuation relations for strong-motion data allowing for individual earthquake magnitude uncertainties, Bull. Seismol. Soc. Am. 87,1674 – 1678.

Zhao JX, Zhou SL, Gao PJ, Zhang YB, Zhou J, Lu M, and Rhoades DA (2016) Ground-motion prediction equations for shallow crustal and upper mantle earthquakes in Japan using site class and simple geometric attenuation functions, Bull. Seism. Soc. Am. 106, 1552-1569, doi**:** 10.1785/0120150063

Table 1. Event- and station-corrected residuals greater than 3 times ϕ_0 in absolute value ($|\varepsilon|>3\phi_0$) at T=0.01s

Event ID	Station	ε(0.01s)	Event ID	Station	ε(0.01s)
EMSC-20101109_0000039	FR.REVF	-1.943414	EMSC-20161129_0000079	IT.AQG	-1.870889
EMSC-20110707_0000048	RA.IRPV	2.168072	EMSC-20161129_0000079	IT.AQV	-1.864027
EMSC-20130125_0000039	IV.BDI	2.24683	EMSC-20161129_0000079	IT.GNU	-1.864027
EMSC-20130829_0000020	KO.GONE	-1.996238	EMSC-20161129_0000079	IT.PSC	-1.912796
EMSC-20140907_0000025	IT.BRM	1.882042	EMSC-20161129_0000079	IT.SPD	-2.056146
EMSC-20140907_0000025	IT.PRT	3.342786	EMSC-20161129_0000079	IT.SULA	-1.839875
EMSC-20141216_0000014	KO.GONE	-1.881109	EMSC-20161129_0000079	IT.SULC	-1.954815
EMSC-20141219_0000039	IT.FOC	-1.853693	EMSC-20161129_0000079	IV.BSSO	-1.938496
EMSC-20141224_0000053	IT.SULC	1.87239	EMSC-20161129_0000079	IV.CERA	-2.019819
EMSC-20150130_0000003	IV.OPPE	1.919609	EMSC-20161129_0000079	IV.MOMA	2.029745
EMSC-20150130_0000003	NI.POLC	-2.251522	EMSC-20161129_0000079	IV.RNI2	-2.023737
EMSC-20150130_0000003	RF.GEPF	-1.924863	EMSC-20161129_0000079	IV.VAGA	-1.965676
EMSC-20150205_0000061	HL.ZKR	2.124032	EMSC-20161201_0000064	IT.CNE	1.943368
EMSC-20150726_0000076	HL.PLG	2.325277	GR-1999-0001	HL.PREA	3.271508
EMSC-20150913_0000012	KO.CAME	2.257234	IT-1987-0003	IT.PNN	-2.084439
EMSC-20151101_0000023	MN.PDG	1.961717	IT-2007-0208	IT.FMG	2.013387
EMSC-20151118_0000047	KO.CAME	-2.911697	IT-2009-0186	IV.SACR	-2.090624
EMSC-20151118_0000094	KO.CAME	-2.740185	IT-2009-0189	IV.SACR	-2.126487
EMSC-20160530_0000085	IV.SACS	2.118076	IT-2010-0003	IV.NRCA	2.243845
EMSC-20160824_0000006	IT.GRD	-1.950702	IT-2011-0029	IT.GAI	-2.075673
EMSC-20160824_0000295	IV.FIAM	1.86564	IT-2012-0002	FR.REVF	2.025614
EMSC-20160824_0000295	IV.RM33	2.219246	IT-2013-0003	IT.PSC	1.974364
EMSC-20160903_0000009	IT.ACC	-1.866856	IT-2013-0003	IV.POFI	2.157508
EMSC-20160903_0000009	IT.CIT	-1.983131	IT-2013-0005	IT.MDN	2.272414
EMSC-20160903_0000009	IT.FOC	2.200381	IT-2013-0010	IV.SNTG	-2.148043
EMSC-20160903_0000009	IV.NRCA	1.919352	IT-2016-0001	IV.INTR	-3.470851
EMSC-20160903_0000009	IV.T1201	-1.887857	ME-1979-0003	EU.BAR	-2.051513

EMSC-20160903_0000009	IV.T1218	-2.022831	TK-1999-0077	TK.8101	1.865905
EMSC-20161027_0000072	IV.SNTG	2.005795	TK-2003-0180	TK.1701	-1.970150
EMSC-20161129_0000079	IT.AQF	-1.928834	TK-2006-0098	TK.1001	-2.058961

Table 2. Event- and station-corrected residuals greater than 3 times ϕ_0 in absolute value ($|\varepsilon|>3\phi_0$) at T=8s

Event ID	Station	ε(8.0s)	Event ID	Station	ε(8.0s)
EMSC-20110707_0000048	RA.IRPV	1.996037	IT-2010-0003	IV.NRCA	1.48384
EMSC-20150205_0000061	HL.ZKR	1.78805	IT-2010-0006	IV.NRCA	1.627814
EMSC-20151106_0000016	FR.REVF	-2.236079	IT-2012-0010	TV.MIR01	-1.484999
EMSC-20151106_0000016	FR.SURF	1.709347	IT-2012-0025	IT.FIN0	1.660915
EMSC-20151118_0000094	KO.CAME	-1.846812	IT-2012-0025	IV.T0800	1.737367
EMSC-20161030_0000033	IT.ACC	1.99265	IT-2012-0025	IV.T0811	2.127796
EMSC-20161129_0000079	IT.AQV	-1.492742	IT-2012-0035	IV.CPGN	-1.493063
IT-2009-0009	IT.ANT	-1.557013	IT-2012-0063	IT.SBT	1.630845
IT-2010-0001	IV.NRCA	1.59095	IT-2013-0013	IT.NAS	1.569184
IT-2010-0001	IV.SENI	-1.600929	TK-1999-0077	TK.8101	1.501051

395

Table 3. Between-event greater than 3 times τ, in absolute value ($|\delta B_e|>3\tau$), at T=0.01 s

Event ID	δB_e(0.01s)	N records
EMSC-20121203_0000066	1.73	18
EMSC-20150411_0000019	-1.69	10
EMSC-20161102_0000149	-1.66	72
IT-2009-0316	1.97	6

Table 4. Between-event greater than 3 times τ, in absolute value ($|\delta B_e|>3\tau$), at T=8.0 s

Event ID	δB_e(8.0s)	N records
EMSC-20121203_0000066	1.65	11
EMSC-20141124_0000030	1.84	5
IT-2010-0003	1.95	3

Table 5. Between-station greater than 3 times ϕ_{s2s} , in absolute value ($|\delta S2S| > 3\phi_{s2s}$), at T=0.01 s

Station	δS2S(0.01s)	N records
IV.T0203	-2.35	1
HL.KSL	-2.04	42
CR.VELS	2.01	1
RO.JURR	2.04	2
FR.REVF	2.05	5

Table 6. Between-station greater than 3 times ϕ_{s2s} , in absolute value ($|\delta S2S| > 3\phi_{s2s}$), at T=8.0 s

Station	δS2S(8.0s)	N records
IV.T0501	-1.86	3
FR.REVF	2.29	4
IV.SNTG	-1.74	15
KO.ENZZ	1.70	7

Figures

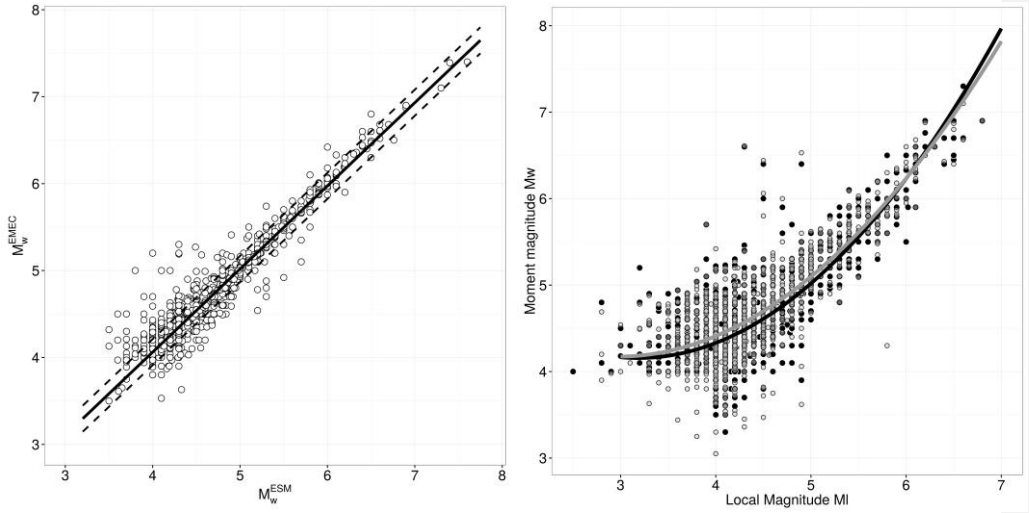


Figure 1. Left: regression between Mw^{ESM} and $M\sout{W}w^{EMEC}$; right: quadratic regressions between Mw^{ESM} and $M\underline{L}\sout{l}$ (black line and points) and between Mw^{EMEC} and $M\underline{L}\sout{l}$ (gray line and points) (see equations 1 and 2).

415

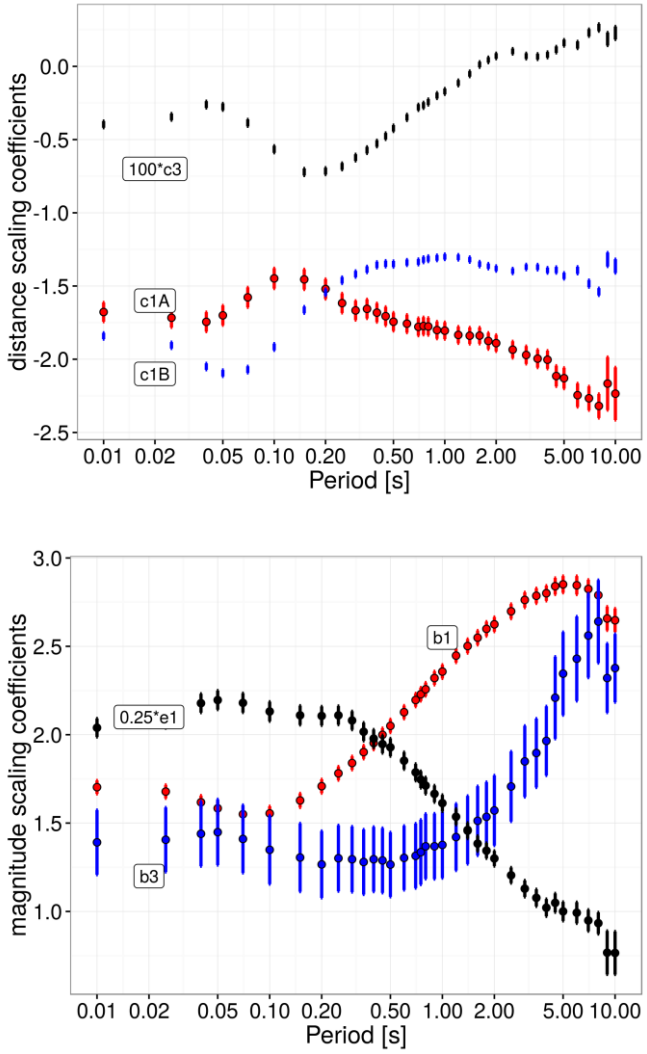
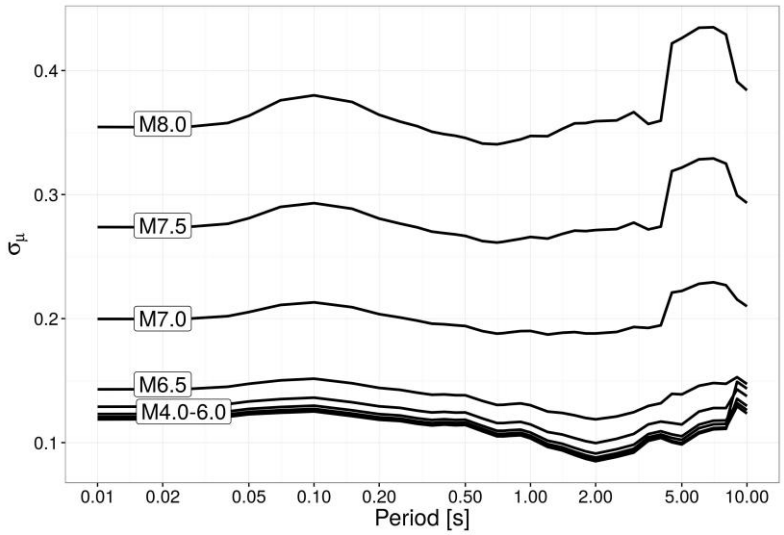
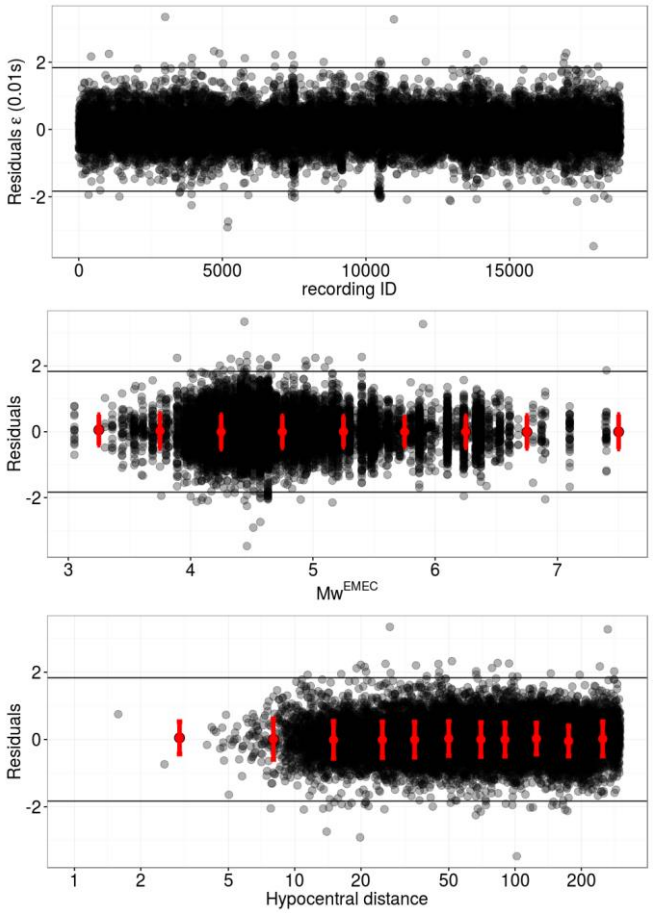


Figure 2. Coefficients (mean and standard error) of the model in equation (4) and (5) for Spectral Acceleration (SA).



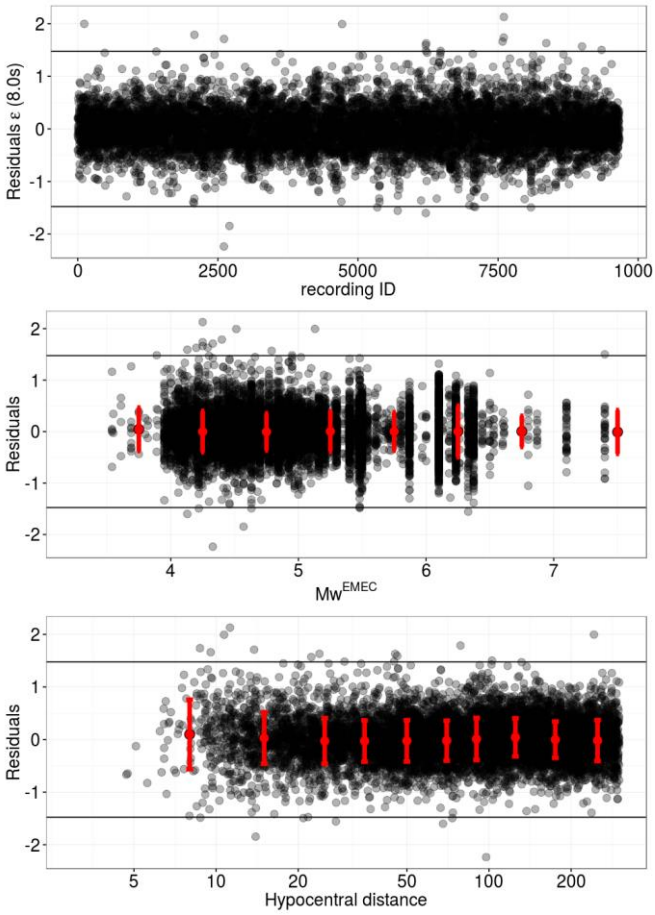
420

Figure 3. Period dependence of the ~~epistemic uncertainty on~~ asymptotic standard error on the median (σ_μ) for SA considering different magnitudes and a fixed distance of 30 km.



425

Figure 4a. Residuals ε computed for SA at T=0.01s. Horizontal lines correspond to $\pm 3\phi_0$. The average residuals ±1 standard deviation computed over magnitude and distance bins are shown as red bars.



430

Figure 4b. The same as Figure 4a but for T=8s.

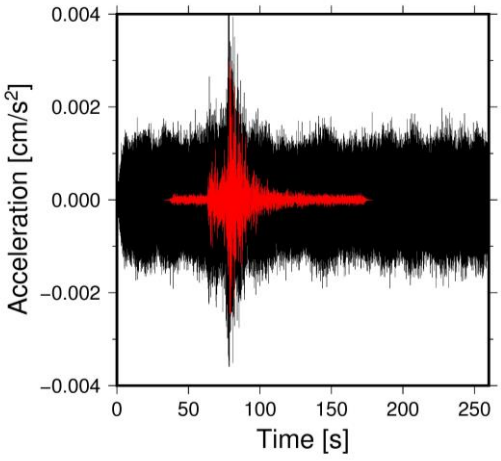


Figure 5 Record (EW component) of event IT-2016-0001 at station IV.INTR: comparison between raw data (black) and processed acceleration (red). The processed acceleration has been filtered over the bandwidth 0.15-30Hz using a second order a-causal Butterworth filter.

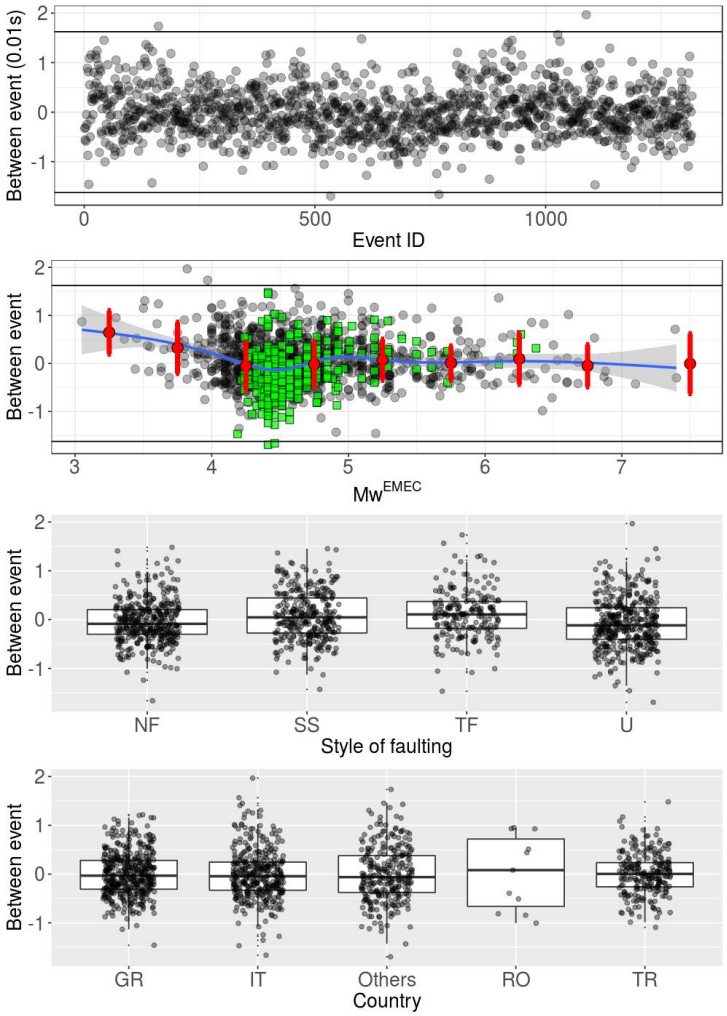


Figure 6a. Between-event residuals δB_e at T=0.01 s for SA. Green squares indicate earthquakes with Mw estimated from ML using equation (2). Horizontal lines correspond to ±3τ-.- The following abbreviations are used for the style of faulting: NF- normal faulting; SS – strike slip; TF – thrust faulting; U – unknown style of faulting. The country codes are: GR- Greece; IT – Italy; RO – Romania; TR- Turkey; Others – any other country different from the previous ones.

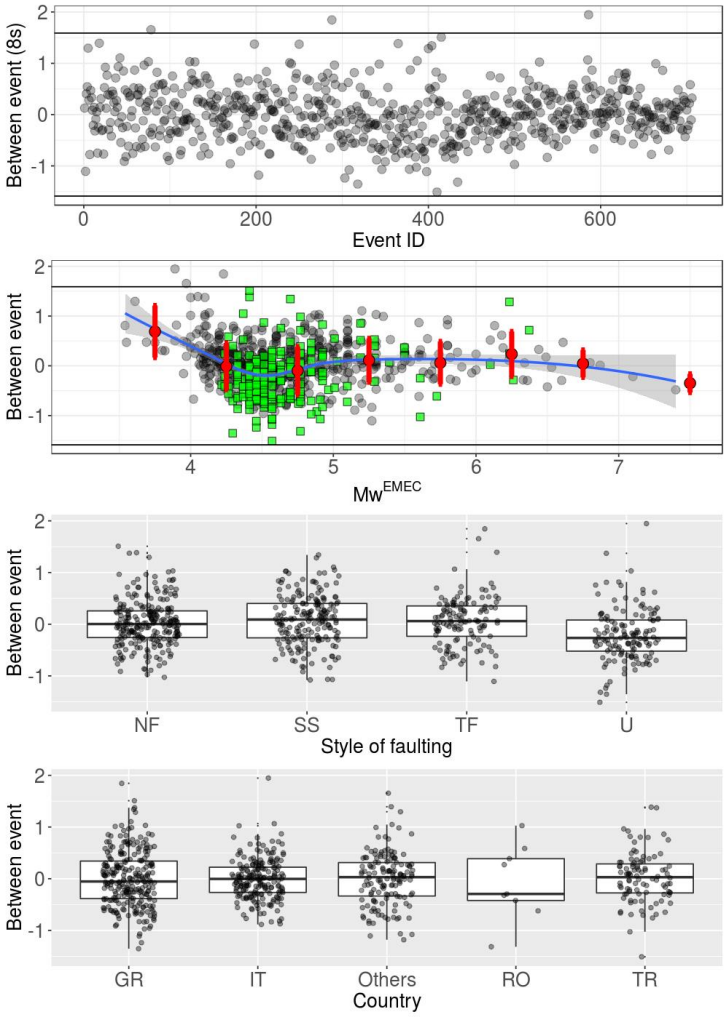


Figure 6b. The same as Figure 6a but for T=8 s

455

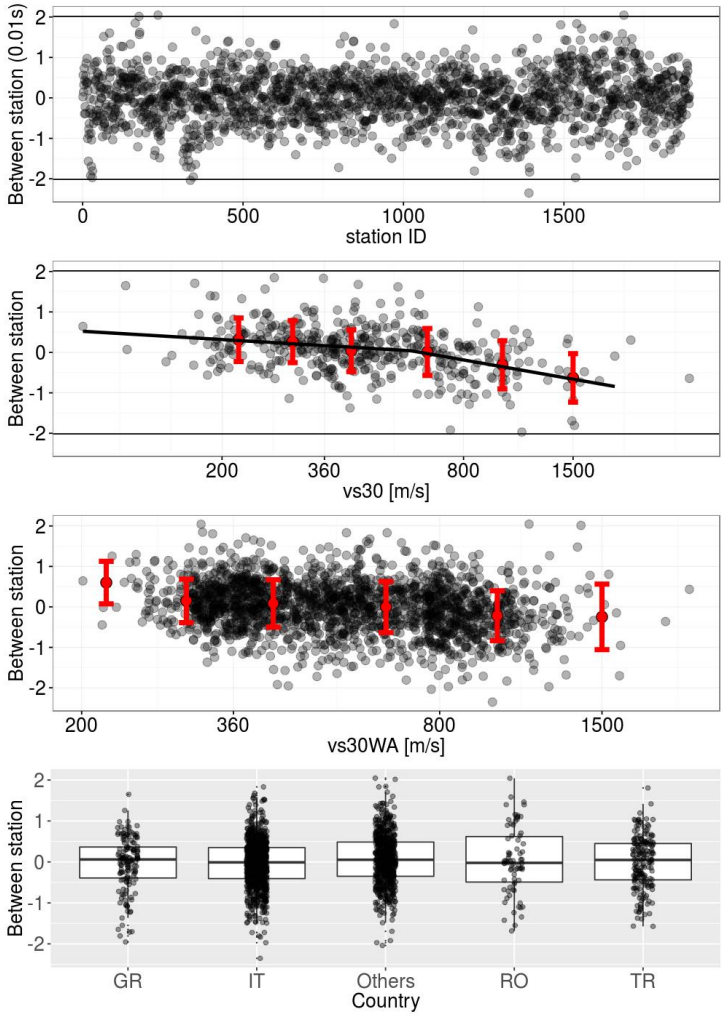


Figure 7a. Between-station residuals $\delta S2S_s$ at T=0.01 s for SA. Horizontal lines correspond to $\pm 3\phi_{S2S}$.

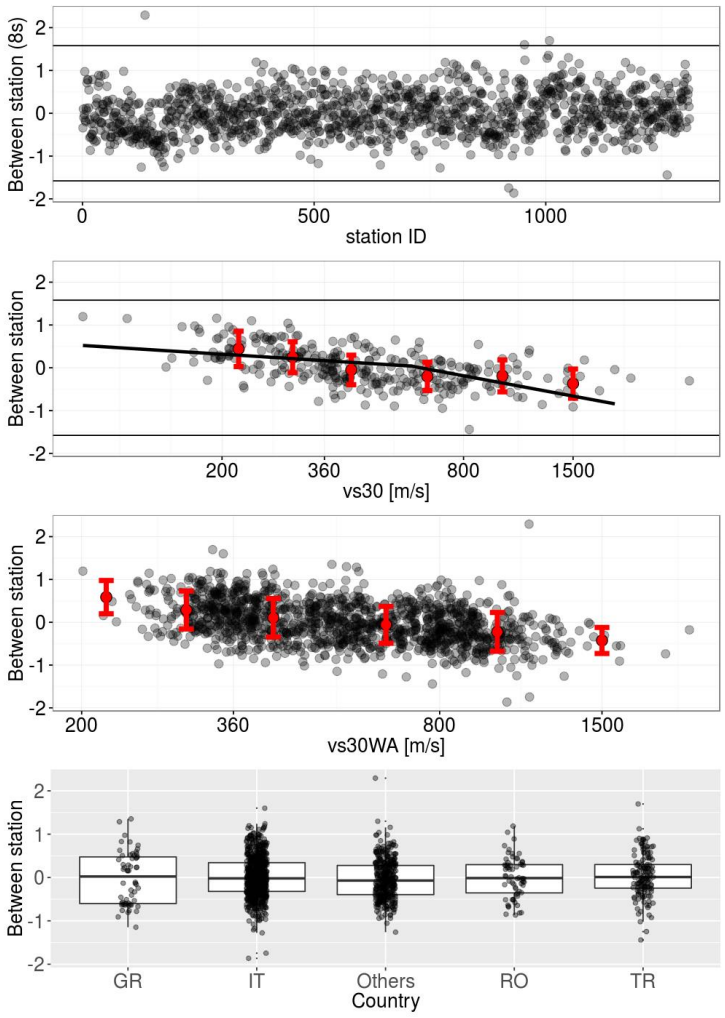


Figure 7b. The same as Figure 7a but for T= 8 s.

465

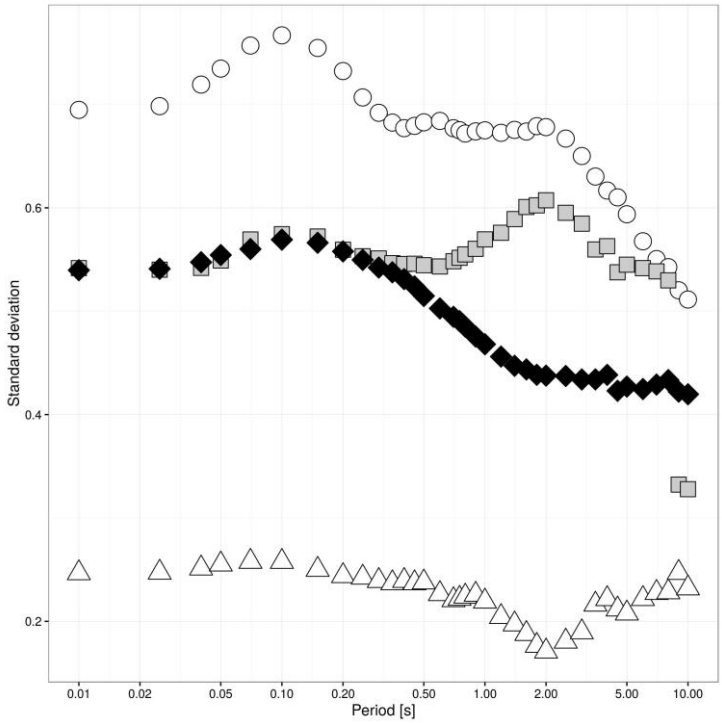


Figure 8. Standard deviation of different residual distributions for SA ($\delta C2C_c$ applied to the offset coefficient e_1). White circle: standard deviation of the between-station distribution (i.e., ϕ_{S2S}); gray square: standard deviation of the between-event distribution (i.e., τ); white triangle: standard deviation of the between-country distribution (ϕ_{C2C}); black diamond: standard deviation of the residual distribution (ϕ_0).

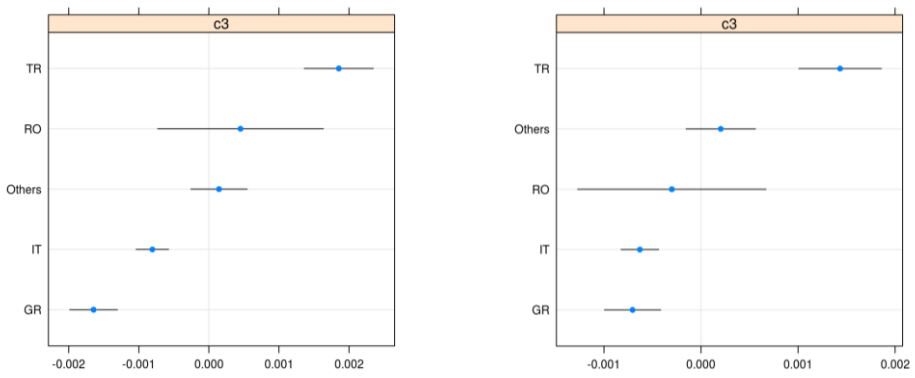


Figure 9. Random–effect on c_3 (anelestic parameter) depending on the country grouping level for T=0.01s (left) and T=2s (right), considering SA. Please, note the different order for the country classes and the different horizontal scale.

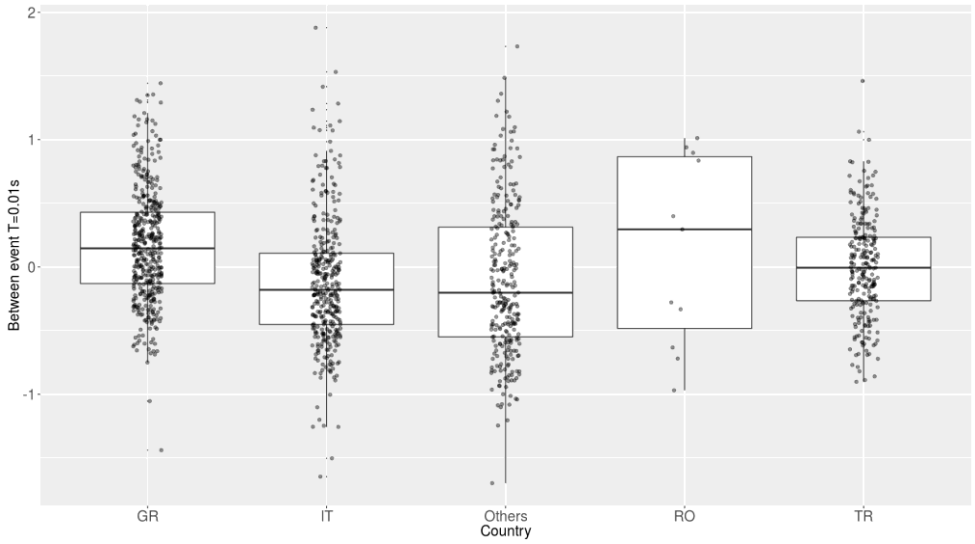


Figure 10. Between-events δB_e for SA at T=0.01s applying the country-based random-effect $\delta C2C_c$ to the anelastic attenuation parameter c_3.

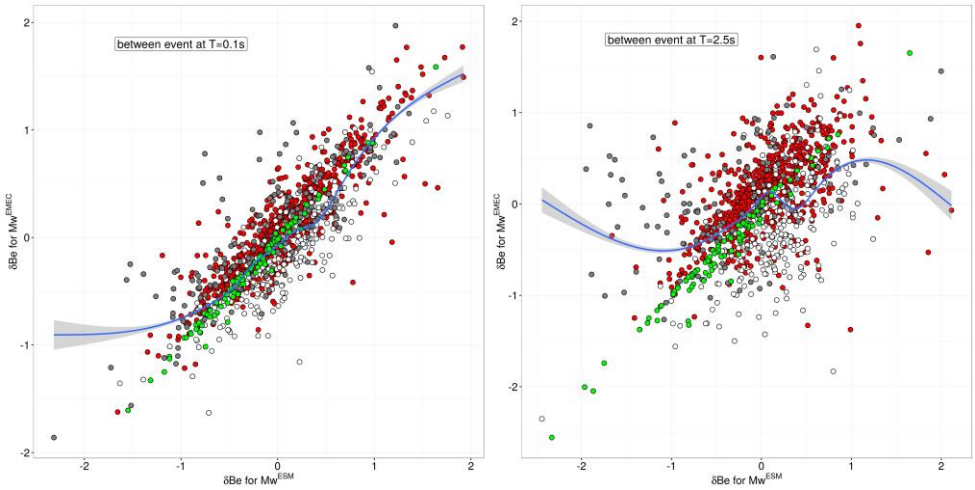
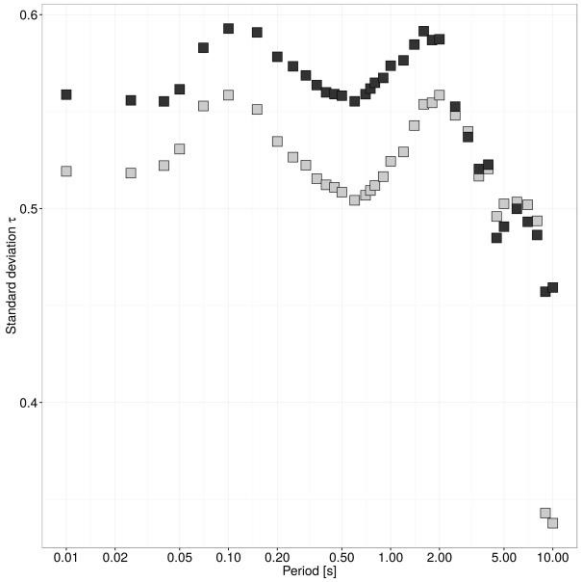


Figure 11. Comparison between δBe obtained using Mw from EMEC and ESM, considering T=0.1s (left) and 2.5s (right). The colors represent different selections on magnitude. In particular, only red represents recordings with both Mw^{EMEC} and Mw^{ESM} available, whereas the other colors ~~imply~~indicate conversions from ML~~l~~ to Mw. In particular, the green dots represent~~s~~ earthquakes for which both Mw^{ESM} and Mw^{EMEC} are derived from the ~~local magnitude~~ML (and therefore showing high correlation). The blue curves are the results of ~~the trend analysis (LOESS) with high order~~ a local polynomial regression (LOESS)~~s~~, giving an idea about the degree of correlation between the two δBe estimates.



550

Figure 12. Between-event standard deviation τ for the regression using Mw^{EMEC} (gray) and Mw^{ESM} (black)

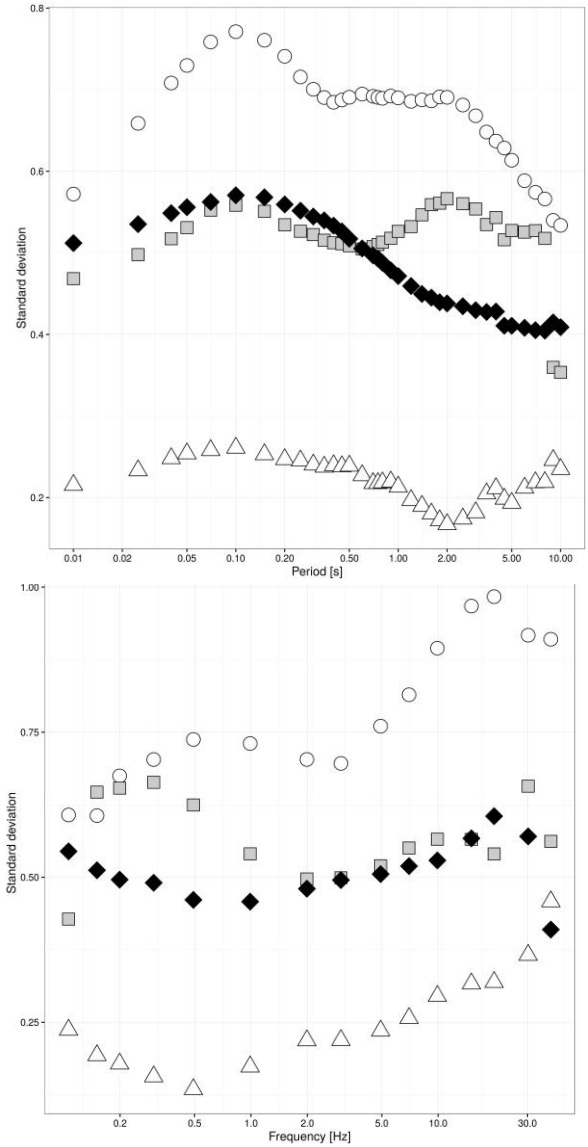
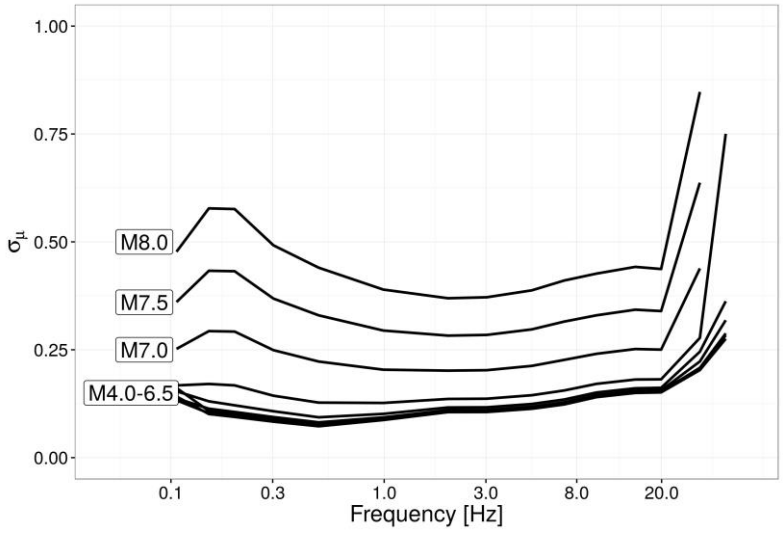


Figure 13. Standard deviation of residual distributions for spectral displacement-SD (top), ~~as function of~~according to period, and for Fourier Amplitude Spectra-FAS (bottom), according to~~as function of~~ frequency. White circle: standard deviation of the between-station distribution (ϕ_{S2S}); gray square: standard deviation of the between-event distribution (τ); white triangle: standard deviation of the between country distribution (ϕ_{C2C}); black diamond: standard deviation of the residual distribution (ϕ_0).



560

Figure 14. Frequency dependence of the ~~epistemic uncertainty~~asymptotic standard error on the median (σ_μ) for FAS considering different magnitudes and a distance of 30km .

Replies to Rewiewer 1

Click here to view linked References

Replies to Rewiewer 2

Click here to view linked References