

Near-Source Magnitude Scaling of Spectral Accelerations: Analysis and Update of Kotha et al. (2020) Model

Sreeram Reddy Kotha (✉ sreeram-reddy.kotha@univ-grenoble-alpes.fr)

Institut des Sciences de la Terre <https://orcid.org/0000-0002-4874-3730>

Graeme Weatherill

Helmholtz-Zentrum Potsdam GFZ: Deutsches Geoforschungszentrum Potsdam

Dino Bindi

Helmholtz-Zentrum Potsdam GFZ: Deutsches Geoforschungszentrum Potsdam

Fabrice Cotton

University of Potsdam: Universitat Potsdam

Research Article

Keywords: Ground-motion model, spectral accelerations, magnitude scaling, near-source saturation, within-model uncertainty, heteroskedastic variability

Posted Date: July 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-679233/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Bulletin of Earthquake Engineering on January 28th, 2022. See the published version at <https://doi.org/10.1007/s10518-021-01308-5>.

Near-Source Magnitude Scaling of Spectral Accelerations: Analysis and Update of Kotha et al. (2020) Model

Sreeram Reddy Kotha¹, Graeme Weatherill², Dino Bindi², Fabrice Cotton^{2,3}

ABSTRACT

Ground-motion models (GMMs) are often used to predict the random distribution of spectral accelerations (SAs) at a site due to an earthquake at a distance. In probabilistic seismic hazard and risk assessment, large earthquakes occurring close to a site are considered as critical scenarios. GMMs are expected to perform well for such rare scenarios i.e., to predict realistic SAs with low prediction uncertainty. However, the datasets used to regress GMMs are usually deficient of data from rare/critical scenarios. The Kotha et al. (2020) GMM developed from the Engineering Strong Motion (ESM) dataset was found to predict decreasing short-period SAs with increasing $M_W \geq M_h = 6.2$, and with large within-model uncertainty at near-source distances $R_{JB} \leq 30km$. In this study, we analysed and updated the parametrisation of the GMM based on non-parametric and parametric analyses of ESM and the NEar Source Strong motion (NESS) datasets. By reducing M_h to 5.7, we could rectify the M_W scaling issue, while also reducing the within-model uncertainty on predictions at $M_W \geq 6.2$. We then evaluated the updated GMM against NESS data, and found that the SAs from a few large, thrust-faulting events in California, New Zealand, Japan, and Mexico are significantly higher than GMM median predictions. However, near-source recordings of these events were mostly made on soft-soil geology and contain anisotropic pulse-like effects. A more thorough non-ergodic treatment of NESS was not possible because most sites sampled unique events in very diverse tectonic environments. Therefore, for now, we provide an updated set of GMM coefficients, within-model uncertainty, and heteroskedastic variance models.

Keywords: Ground-motion model, spectral accelerations, magnitude scaling, near-source saturation, within-model uncertainty, heteroskedastic variability

¹Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ. Gustave Eiffel, ISTERre, 38000 Grenoble, France

²Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, 14467 Potsdam, Germany

³University of Potsdam, Institute of Geosciences, Potsdam, Germany

INTRODUCTION

For engineering purposes, a ground-motion observation at a site from an earthquake is often reported as the peak acceleration response of a 5% damped single-degree-of-freedom (SDOF) oscillator with fundamental period T when excited by the recorded accelerogram at the site. For a variety of SDOF fundamental periods, the vector of spectral accelerations $[SA(T)]$ is the so-called response spectrum. Thousands of response spectra collected from several hundreds of earthquakes at several dozens of recording sites are compiled into engineering ground-motion datasets, such as the recent pan-European Engineering Strong-Motion (ESM by Bindi et al. 2018; Lanzano et al. 2018), the global Next-Generation-Attenuation West-2 (NGA-W2 by Ancheta et al. 2014), the Japanese KiK-net (Okada et al. 2004) ground-motion datasets (Bahrapouri et al. 2021; Dawood et al. 2016), and a few others. Such datasets are essential to studying the physical characteristics of $SA(T)$ from past earthquakes and predict $SA(T)$ of future earthquakes. For this purpose, a common practice has been to regress empirical attenuation relations connecting these observed $SA(T)$ to the properties of earthquake ruptures (e.g. moment magnitude M_W), the local geological conditions at the effected site (e.g. time-averaged shear-wave velocity V_s), and the rupture-to-site distances (e.g. Joyner-Boore distance R_{JB}). Such empirical relations are commonly referred to as ground-motion models (GMMs), and are widely used in probabilistic seismic hazard assessments (PSHA). Therefore, the quality and quantity of ground-motion datasets, and the physical and statistical robustness of GMMs – especially beyond the (M_W, R_{JB}, V_s) ranges of the regressed dataset – are of utmost importance in PSHA.

The most recent, and the largest yet, pan-European Engineering Strong-Motion dataset (ESM) was meticulously developed for the next update of European Seismic Hazard Map of the region - from ESHM13 (Woessner et al. 2015) to ESHM20 (Danciu et al. 2021 – in-preparation). This dataset is 10-fold larger than its predecessor RESORCE (Akkar et al. 2014b), and facilitated the development of the Kotha et al. (2020) GMM, which superseded its RESORCE based predecessor Kotha et al. (2016). The new GMM also made use of the latest geological and tectonic regionalisation models (e.g. Basili et al. 2019) to predict more accurate (e.g. region-, locality- and site-specific) ground-motions in the extremely heterogeneous (active) shallow crustal regions of pan-Europe. Therefore, this regionally-adaptable GMM was adapted into a scaled-backbone GMM logic tree implementation (e.g. Bommer 2012 and Douglas 2018) in the new ESHM20 by Weatherill et al. (2020b).

Current ground-motion datasets, including the ESM, are a few thousands of recordings larger than a decade ago. However, a typical PSHA requires ground-motion predictions for rare scenarios (e.g. $M_W \geq 5.5$ & $R_{JB} \leq 30km$) that are very sparsely or never sampled before, but are also deemed the most critical for built environments in the region. For instance, only 2% of the ESM data qualifies the *hazard critical* criterion above. Of course, the other 98% data is quite useful in robustly calibrating the model for its region-, locality-, and site-specific adaptability. But in regions such as Italy, Turkey, Greece it is perhaps more important to maintain GMM's physical and statistical robustness in the hazard critical scenarios.

Thanks to the timely feedback from the community of seismic hazard and risk modellers involved in ESHM20, we were motivated to evaluate the performance of the Kotha et al. (2020) GMM in hazard critical scenarios of Mediterranean active shallow crustal regions. Especially at short distances ($R_{JB} \leq 20km$), the short period ($SA(T \leq 0.2s)$) ground-motions predicted by the GMM for $M \geq 6.5$ are lower than those for $5.7 \leq M_W \leq 6.2$, and similar to those for $M_W \approx 5.5$. Although such behaviour reflects the modelling choice and is data-driven,

39 it is discordant with the behaviour of other contemporary GMMs at near-source distances (at short periods and for
 40 large magnitudes); and more importantly, has a rather high within-model uncertainty due to the lack of calibrating
 41 data. To resolve this issue, we propose an essential revision to the model, while also evaluating its behaviour
 42 against the recently published Near Source Strong-motion datasets (NESS and NESS2 by Pacor et al. 2018; Sgobba
 43 et al. 2021b). NESS2 is an update of the NESS dataset. Therefore, we simply refer to them as NESS from hereon.

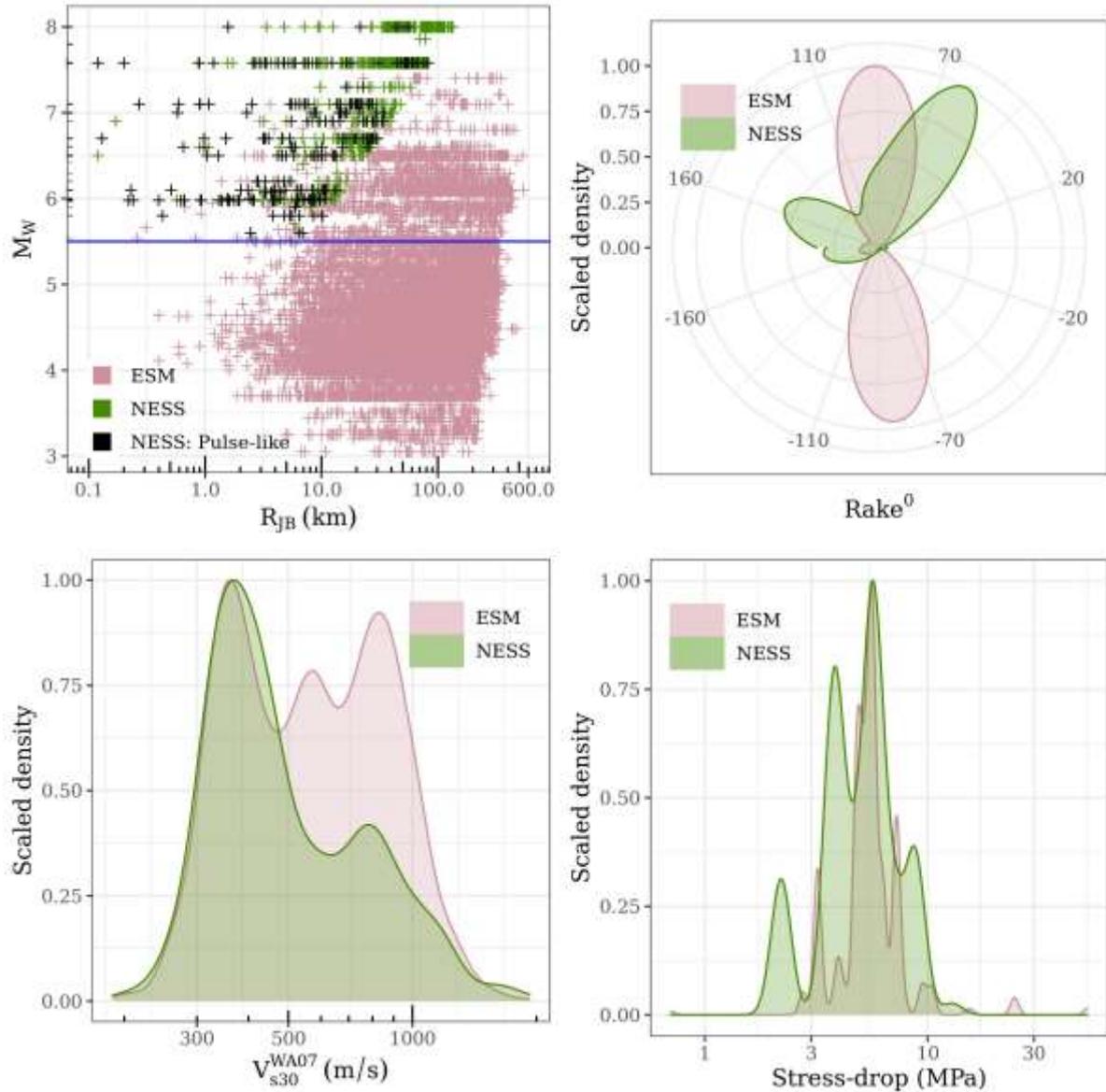


Fig.1 Comparison of ESM and NESS datasets: Top-left panel shows the (M_W, R_{JB}) distribution of the two datasets, with the horizontal (blue) at $M_W = 5.5$. Top-right panel shows the scaled kernel density of recordings from $M_W \geq 5.5$ events with various *rake* angles, with slope-inferred V_{s30}^{WA07} in the bottom-left panel, and stress-drop in bottom-right panel.

44

45 Fig.1 compares the ESM and NESS datasets. Since both datasets are compiled with identical data
 46 qualification procedures, ESM and NESS are seamlessly compatible. However, it is to be noted that while ESM
 47 contains data exclusively from shallow crustal earthquakes in the Mediterranean region, NESS is primarily
 48 composed of data from global large earthquakes from Western United States, Japan, New Zealand, Mexico, and
 49 Taiwan. The top-left panel of Fig.1, showing the distribution of recordings in (M_W, R_{JB}) ranges of the two datasets,
 50 clearly shows that NESS complements ESM with several recordings from hazard critical scenarios ($M_W \geq$

51 5.5 & $R_{JB} \leq 30\text{km}$) – albeit from global earthquakes originating in very diverse tectonic environments, recorded
 52 by stations in very particular geological settings. In addition, this plot also shows that 30% of the NESS recordings
 53 are characterised by pulse-like characteristics, either due to geological or finite-fault effects. In the other panels,
 54 we only compare the two datasets at $M_W \geq 5.5$ (data above the blue line). The top-right panel of Fig.1 shows the
 55 scaled kernel density of recordings versus *rake* angles (fault mechanism) of $M_W \geq 5.5$ events in the two datasets.
 56 Scaled (to 1) kernel density plots are smoothed version of frequency histograms, and are more intuitive when
 57 comparing the sampling rate of continuous data between two datasets. Evidently, the ESM dataset is dominated
 58 by recordings from normal faulting ($-110^\circ \leq \textit{rake} < -70^\circ$) and reverse faulting ($70^\circ \leq \textit{rake} < 110^\circ$) events
 59 from Italy, Turkey, and Greece. NESS on the other hand contains data from strike-slip ($160^\circ \leq |\textit{rake}|$) and
 60 reverse-oblique ($20^\circ \leq \textit{rake} < 70^\circ$) faulting events from elsewhere. Similarly, the bottom-left panel shows that
 61 the NESS data (from $M_W \geq 5.5$) is predominantly recorded on soft-soils characterised by very low V_{s30}^{WA07} – time-
 62 averaged shear-wave velocity in top 30m of soil inferred from topographic *slope* using Wald and Allen (2007) –
 63 when compared to the more uniform distribution in ESM. Finally, the bottom-right panel shows that the
 64 distribution of reported stress-drops – by various studies cited in Sgobba et al. (2021b) and Bindi and Kotha (2020)
 65 – of the sampled $M_W \geq 5.5$ events are comparable.

66 Although the two datasets are quite complementary, the NESS2 dataset of Sgobba et al. (2021b) was not
 67 available during the development of Kotha et al. (2020) GMM. Even if the smaller NESS dataset of Pacor et al.
 68 (2018) was available in a limited way, it was decided against integrating it into ESM. The reasons being:

- 69 1) Kotha et al. (2020) GMM was intended to be partially non-ergodic, in the sense that; events were localised
 70 into tectonic localities to capture systematic differences between earthquake generating sources, and
 71 stations were regionalised into attenuating regions to capture differences in apparent anelastic attenuation
 72 of $SA(T)$. The tectonic localisation and attenuation regionalisation models used therein could not be
 73 extended to contain global events. The possibility of simply assigning NESS (or ESM) data to larger
 74 political boundaries was decided against because the region-to-region and locality-to-locality random
 75 variances estimated from ESM dataset were to be used in constructing the backbone GMM logic tree for
 76 ESHM20 (Weatherill et al. 2020b), and these estimates were intended to capture the diversity of active
 77 shallow crustal regions in Mediterranean region alone; and not partially global.
- 78 2) Kotha et al. (2020) GMM was also required to function as a site-specific GMM for sites with sufficient
 79 number of recordings or otherwise, use the site-specific random-effects to develop a large scale site-
 80 response model for use in European Seismic Risk Map (ESRM20 by Crowley et al. 2019 and Weatherill
 81 et al. 2020a). For this purpose, it was necessary that each site recorded a variety of magnitudes, and event-
 82 to-site distances and azimuths. As well that events are recorded by stations with a variety of site
 83 conditions. Most stations in NESS dataset have recorded a single large event at a close distance, and most
 84 events are recorded by sites located in very similar geological settings. In this regard, as opposed to the
 85 *crossed and nested* nature of ESM dataset, NESS is exclusively *nested*. In order to avoid any trade-offs
 86 between site-specific and event-specific (between-event) random-effects, we chose not to integrate the
 87 two datasets.
- 88 3) As seen in top-left panel of Fig.1, around 30% of NESS recordings show pulse-like characteristics, which
 89 are often characterised by higher long-period *SAs* and lower short-period *SAs* than average, depending

90 on the event-to-site distance and azimuth. Instead of biasing the (isotropic) GMM median (at short
 91 distances and long periods) with pulse-like effects, we chose to reserve NESS datasets for investigation
 92 of complex near-fault phenomenon, as done by Sgobba et al. (2021a).

93 In this study, we discuss the re-parametrization of the Kotha et al. (2020) GMM, consequent impact on
 94 its within-model and aleatory uncertainties, its limitation in capturing the complex near-source finite-fault
 95 phenomenon, and the apparent (or actual) global diversity of $M_W \geq 5.5$ shallow crustal earthquakes.

96 FUNCTIONAL FORM

97 We briefly discuss the choice of parametrization (or functional form) of the Kotha et al. (2020) GMM. The number
 98 and combination of parameters in this GMM is far fewer than the more sophisticated NGA-W2 GMMs, such as
 99 Abrahamson et al. (2014); Campbell and Bozorgnia (2014); Chiou and Youngs (2014), and comparable to Boore
 100 et al. (2014). Of course, the sophistication of NGA-W2 GMMs comes from a more complex characterisation of
 101 the large magnitude near-fault effects, and complex site-effects, such as nonlinear soil response and basin effects.
 102 While in application driven GMMs for Europe (as also in Bindi et al. 2017), the decision for a more modest
 103 parametrization has been made to respect the limited metadata on earthquake rupture (event, from hereon),
 104 propagation path (or simply, path), and local geological conditions at the receiving site (site properties, in short);
 105 in seismic hazard models (e.g. ESHM13 and ESHM20), and also in the strong-motion datasets (e.g. RESORCE
 106 and ESM). In fact, the functional form had only minor changes since Bindi et al. (2014) and Kotha et al. (2016),
 107 despite the 10-fold increase in data since RESORCE. A more detailed description is available in the Kotha et al.
 108 (2020) study. Here we reintroduce the functional form through equations (1) through (4):

$$\ln(\mu) = e_1 + f_{R,g}(M_W, R_{JB}) + f_{R,a}(R_{JB}) + f_M(M_W) + \delta L2L_l + \delta B_{e,l}^0 + \delta S2S_s + \varepsilon \quad (1)$$

$$f_{R,g} = (c_1 + c_2 \cdot (M_W - M_{ref})) \cdot \ln^2 \sqrt{(R_{JB}^2 + h_D^2) / (R_{ref}^2 + h_D^2)} \quad (2)$$

$$f_{R,a} = \frac{c_3 + \delta c_{3,r}}{100} \cdot (\sqrt{R_{JB}^2 + h_D^2} - \sqrt{R_{ref}^2 + h_D^2}) \quad (3)$$

$$f_M = \begin{cases} b_1 \cdot (M_W - M_h) + b_2 \cdot (M_W - M_h)^2 & M_W \leq M_h \\ b_3 \cdot (M_W - M_h) & M_h < M_W \end{cases} \quad (4)$$

109 The coefficients $e_1, b_1, b_2, b_3, c_1, c_2, c_3$ are the so-called fixed-effect coefficients of a mixed-effect model;
 110 while $\delta c_{3,r}, \delta L2L_l, \delta B_{e,l}^0, \delta S2S_s$ are the random-effects, and ε are the residuals. In addition, the h_D (pseudo-
 111 depth, h-parameter or near-source saturation parameter) was a fixed-effect regression coefficient in Bindi et al.
 112 (2014) and Kotha et al. (2016) – making them nonlinear mixed-effect models – while Kotha et al. (2020) assigned
 113 it a priori values to keep the regression linear. The random-effects' description and their purposes are elaborated
 114 in Kotha et al. (2020), and will not be repeated here for brevity. Since this study is about revising the fixed-effect
 115 coefficients, we explain their original purpose here:

- 116 • e_1 is the generic offset, which is always recommended in a multivariate regression – irrespective of the choice
 117 of independent predictor variables (here, M_W, R_{JB}). Absence of this parameter may destabilize the mixed-
 118 effects regression, and redistribute the offset among the random-effects
- 119 • b_1 captures the linear increase of ground-motion with M_W and is perhaps the most prevalent parameter among
 120 50years of GMMs (see Douglas 2010)
- 121 • b_2 captures the nonlinear increase of ground-motion with M_W , as in equation (4). Fukushima (1996) has
 122 demonstrated eloquently the necessity of this term in a GMM, especially when the Fourier source spectrum is
 123 approximated to be an ω^{-2} source model proposed by Brune (1970). Of course, if b_2 should take a positive
 124 or negative value depends on magnitude scale. Fukushima (1996) showed that it should be negative if the
 125 choice of magnitude scale is M_W , and positive if the scale were to be M_L or M_{JMA} . In accordance, both the
 126 pan-European Bindi et al. (2014) and Kotha et al. (2016) estimates of their respective b_2 (at all periods) were
 127 negative, when regressed over the data from $4 \leq M_W \leq 7.6$ events in RESORCE dataset. The Akkar et al.
 128 (2014a) model, also derived from the RESORCE dataset, had negative scaling coefficient for their M_W^2 term.
 129 Among the NGA-W2 GMMs, the Abrahamson et al. (2014) GMM also reports a negative scaling with
 130 their M_W^2 term(s); while Campbell and Bozorgnia (2014) GMM has no M_W^2 term, and the Chiou and Youngs
 131 (2014) model uses hyperbolic magnitude scaling functions which are hard to compare with the simpler
 132 quadratic forms (in equation 4). To our knowledge, the M_W^2 scaling coefficient of all the GMMs mentioned
 133 here are data-driven and are not pre-sets based on theoretical expectations or numerical simulations.

134 In contrast, the Kotha et al. (2020) has a positive M_W^2 scaling coefficient; i.e. b_2 in equation (4) is
 135 positive, unlike Fukushima (1996) and other GMMs mentioned above. The only other model based on ESM
 136 is the recent Lanzano et al. (2019), which does not feature an M_W^2 term for comparison. It is however
 137 interesting that the Boore et al. (2014), which shares its quadratic form with pan-European models mentioned
 138 here, while being derived from the NGA-W2 dataset also has small but positive M_W^2 scaling coefficient. The
 139 Kotha et al. (2020) and Boore et al. (2014) both show positive M_W^2 scaling at short periods, and their respective
 140 predecessors Kotha et al. (2016) and Boore and Atkinson (2008) show the contrary.

141 A strong decrease in short-period ground-motions towards lower magnitudes has been demonstrated
 142 both theoretically and empirically by Douglas and Jousset (2011), and was recognised to be a κ_{source} effect –
 143 a parameter that attenuates high frequency ground-motions in the ω^{-2} source model of Brune (1970).
 144 Theoretically, b_2 should be negative to accommodate this. However, the non-parametric regressions on ESM
 145 (see Figure 4 of Kotha et al. 2020) showed instead a strong positive trend of b_2 . In our opinion, the main
 146 reason for this reversal is the introduction of more than 10000 new records from $M_W \leq 4.5$ events in ESM
 147 (details in Kotha et al. 2020) with respect to RESORCE. We suspect poorly constrained $M_W \leq 4.5$ is a factor,
 148 but probably not the only one, as also discussed in Kotha et al. (2016).

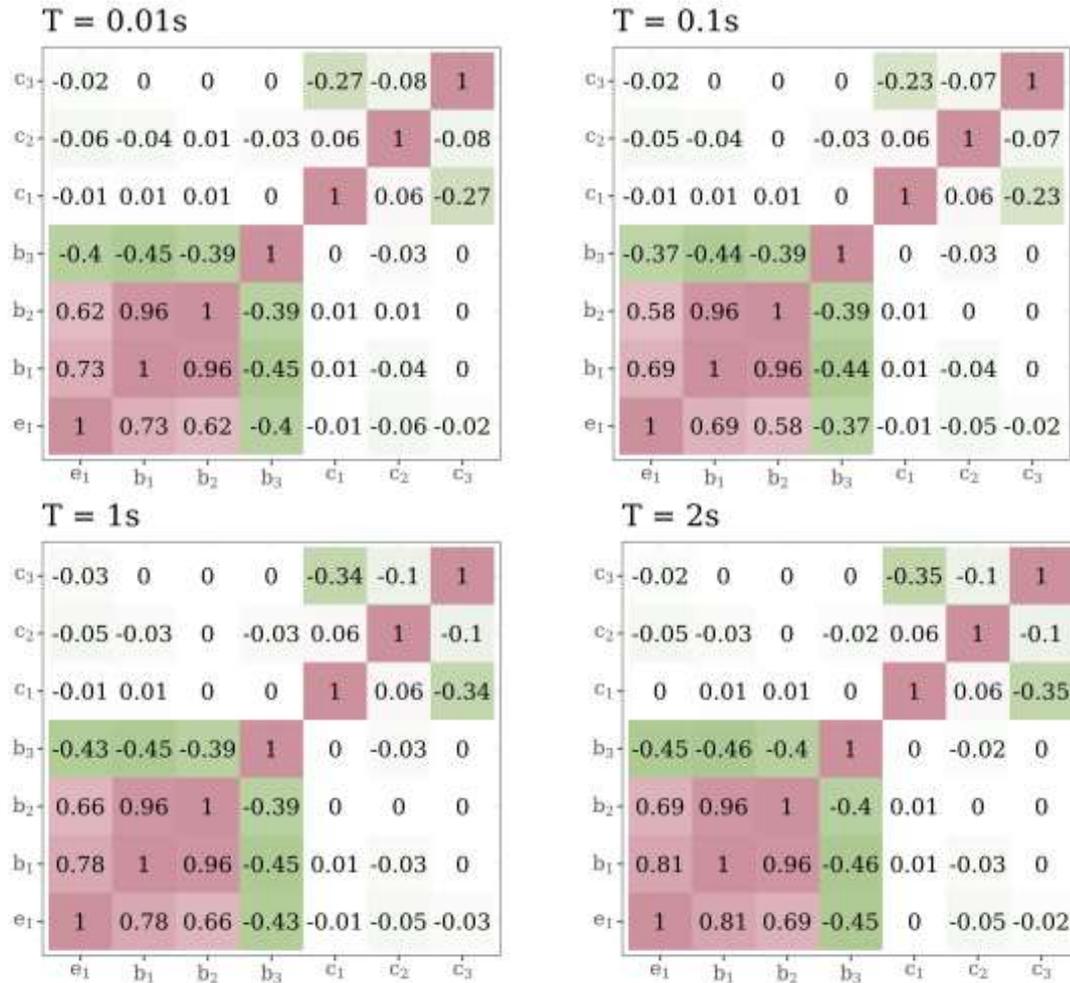


Fig.2 Correlation matrix of fixed-effects at $T = 0.01, 0.1, 1, 2s$ of the Kotha et al. (2020) GMM

149

150

151

152

153

154

155

156

157

158

159

Fig.2 shows the correlation matrix of the fixed-effects coefficients of Kotha et al. (2020) GMM. It is clear that at all periods, the coefficients b_1 and b_2 are almost perfectly correlated, which suggests one of them could be dropped out of the regression. GMM regression trials removing b_1 rendered b_2 estimates to be negative – as expected by Fukushima (1996) – but lead to under prediction of short period ground-motions at small magnitudes. Trials without b_2 , imitating Campbell and Bozorgnia (2014), would mean abandoning the Fukushima (1996) recommended nonlinear scaling with M_W^2 . And finally, removing either b_1 or b_2 led to a substantial increase in AIC, BIC, Log-likelihood, and deviance estimates (typical metrics of the analyses of variance) with respect to the model containing both. Meaning, presence of both these fixed-effects coefficients is beneficial to model fit. Therefore, we chose not to modify this component of the functional form for time being.

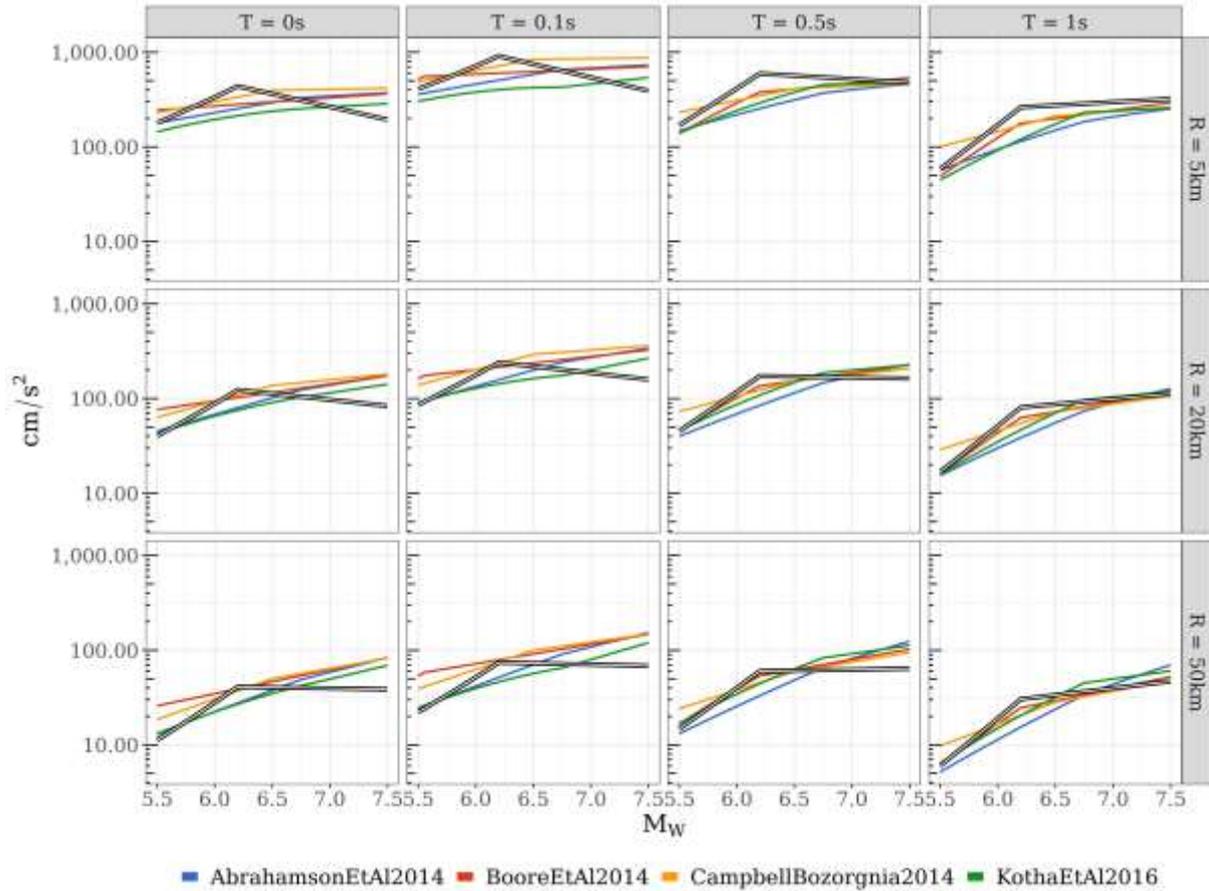


Fig.3 Trellis plots comparing SA predictions by contemporary GMMs at $T = 0.01, 0.1, 1, 2s$. The color legend shows the various GMMs, while the double-line in black is the prediction by Kotha et al. (2020) GMM

160

- 161 • b_3 is the coefficient controlling the linear scaling of ground-motion at large magnitudes $M_W > M_h$. In Kotha
 162 et al. (2020) GMM, M_h is placed at $M_W = 6.2$ based on the non-parametric analyses presented therein. The
 163 choice of M_h is largely ambiguous and is chosen iteratively. For example, Boore et al. (2014) chose a period-
 164 dependent $M_h \in [5.5, 6.2]$, while Abrahamson et al. (2014) and Campbell and Bozorgnia (2014) chose to
 165 introduce multiple hinges with multiple period-independent M_h . Several GMMs use the so-called hinge-
 166 magnitude (M_h) to disassociate the scaling at large magnitudes from that at smaller magnitudes. Fig.2 shows
 167 that b_3 is poorly correlated to b_1 and b_2 , as intended.

168

169 Intuitively, removing b_3 (and the hinge at M_h) while maintaining both b_1 and b_2 in the functional
 170 form would strongly bias their estimates towards the relatively numerous $M_W \leq M_h$ events. This would lead
 171 to a severe over-prediction at larger magnitudes, which are often of greater relevance (hazard critical) than
 172 smaller events in PSHA. In fact, the choice $M_h = 6.2$ has led to strong over-prediction for events in the
 range $6.2 \leq M_W \leq 6.5$, which is one factor motivating us to revise this part of the GMM in this study.

173

174 Fig.3 compares the $SA(T)$ magnitude-scaling of a few recent GMMs for $T = 0.01, 0.1, 1, 2s$ at $R =$
 175 5, 20, 50km. Evidently, the choice of $M_h = 6.2$ led to predictions by Kotha et al. (2020) to be *oddly* higher
 176 for M6.2 events than M6.5 events. Moreover, the *kink* at $M_h = 6.2$, although data-driven, doesn't seem to
 177 reflect any physical phenomenon that would suggest M6.2 events to be often stronger than any other
 magnitudes. This oddity is more prevalent at short distances and short periods. In this study, we rectify this

178 issue by revising our choice of M_h . Accepting that b_3 is necessary in the GMM, we will demonstrate the
 179 impact of (apparently ambiguous) choice of M_h on the GMM predictions for large magnitudes and short
 180 distances.

181 • c_1 is the fixed-effect coefficient capturing the linear decay of $SA(T)$ with R_{JB} , i.e. the geometrical spreading
 182 of ground-motion as a spherical wave front. In distance range $10 < R_{JB} < 80km$, analytical values of decay
 183 rate are about -1.1 to -1.3, but can also be strongly region dependent. For example, in the spectral
 184 decomposition of this ESM dataset using Generalized Inversion Technique (GIT, Castro et al. 1990, Edwards
 185 et al. 2008, and Oth et al. 2011], Bindi and Kotha (2020) showed that the slopes of non-parametric 1Hz Fourier
 186 ground-motion amplitude decay curves at $R_{JB} < 80km$ are: -1.36 in Italy, -1.34 in regions encompassing
 187 Balkans and Romania, and -1.03 in regions covering Aegean and East Anatolian regions. Since the c_1
 188 estimates of our model range between from -1.34 and -1.12 for $SA(T = 0.1s - 8s)$, we are not concerned
 189 with revising this part of functional form.

190 • c_2 complements c_1 in capturing the magnitude dependence of geometric spreading due to finiteness of large
 191 magnitude ruptures. This parameter is only necessary in GMMs predicting spectral accelerations, and not as
 192 much in those predicting Fourier amplitudes (Cotton et al. 2008). This parameter usually takes positive values
 193 to simulate the more gradual decay of $SA(T)$ in near-source distances from large ruptures, and increases
 194 towards longer periods. However, this is simply a consequence of the functional form. For our concerns, Fig.2
 195 shows that c_1 and c_2 are almost uncorrelated, which means it deserves place in the functional form.

196 • c_3 captures the apparent anelastic attenuation of $SA(T)$ with distance. Unlike the geometric spreading of c_1
 197 and c_2 , c_3 is intended to simulate the exponential decay of $SA(T)$ with distance due to intrinsic scattering and
 198 absorption of radiated seismic energy. The effect of this parameter is more pronounced at $R_{JB} > 80km$, and
 199 is known to be strongly region dependent. We often find that c_1 and c_3 fixed-effects are inversely correlated
 200 because both capture the decay of ground-motion with distance; similar to correlation of $\ln(R)$ and R .
 201 However, since we have used a dataset of recordings with $0 \leq R_{JB} \leq 545km$, the negative correlation
 202 between c_1 and c_3 is reasonably low at all periods. Inverse correlation of c_1 and c_3 is lower at short periods
 203 because anelastic attenuation effects high frequency ground-motions more than low frequency ground-
 204 motions. While at longer periods, c_3 values become close to zero and slightly more inversely correlated to c_1 .
 205 For instance, a few GMMs force $c_3 = 0$ at $T \geq 1s$. But as long as they are non-positive – indicating an
 206 unphysical exponential increase with distance – we are not too concerned about revising this part of the model.

207 Essentially, we mean to say that all the fixed-effects parameters of our model are necessary, and most
 208 behave as expected. The only exception being b_3 , which in a way is also the most critical parameter controlling
 209 the GMM predictions for large magnitudes at short distances, and therefore the hazard and risk assessments. In the
 210 following sections, without changing the mixed-effects configuration of the GMM, and using the exact same
 211 dataset used in Kotha et al. (2020), we will choose a variety of M_h values to understand how the b_3 estimates vary;
 212 and of course, its impact on median, aleatory variability, and within-model uncertainty of the GMM.

213

APPROACH

214 The published fixed-effect coefficients of Kotha et al. (2020) GMM were estimates using $M_h = 6.2$, based on the
 215 non-parametric trends of $SA(T)$ scaling with M_w . Here we repeat the regressions and estimate the fixed- and
 216 random-effects for the choices of $M_h = 5.5, 5.7, 6.2, 6.5$. While the Kotha et al. (2020) GMM was derived using a
 217 robust linear mixed-effects regression algorithm (*robustlmm* by Koller 2016), in this exercise involving several
 218 trials, we use a less computationally intense ordinary least-square mixed-effects regression algorithm (*lme4* by
 219 Bates et al. 2014). Both packages are developed for use with R software (Team 2013). We have already verified
 220 (in Kotha et al. 2020) that both the algorithms yield nearly identical estimates of fixed-effect coefficients. In fact,
 221 even the fixed-effect variance-covariance matrices differ ever so slightly between robust and ordinary least-square
 222 estimates – at least in this exercise. The biggest differences to be expected are in the random-effect and residual
 223 variances, but this will be handled in the following sections. Therefore, the inferences we make in this study – at
 224 least about the fixed-effects – will not change if we preferred either robust or ordinary least-square regressions.

225

OUTCOME

226 Essentially, we re-derived the GMM for PGA , PGV , and $SA(T)$ at 34 values of $T = 0.01 - 8s$, while choosing
 227 $M_h = 5.5, 5.7, 6.2, 6.5$ in Eq. (4) and keeping all other pre-sets unchanged; i.e. $M_{ref} = 4.5$, $R_{ref} = 30km$ and
 228 $h_D = 4, 8, 12km$ depending on the hypocentral depth of the event. We discuss the fixed-effect (median)
 229 predictions, and the (M_w, R_{JB}) dependent within-model uncertainty σ_μ estimated using their variance-covariance
 230 matrices. σ_μ^2 is the asymptotic variance defined in Atik and Youngs (2014).

231

FIXED-EFFECTS

232 Fig.4 shows the magnitude and distance scaling of the four M_h -trial GMMs. Note that the predictions
 233 corresponding to $M_h = 6.2$ in this figure are identical to those presented in the Kotha et al. (2020). The four GMMs
 234 have identical magnitude and distance scaling in the best sampled range $4 \leq M_w \leq 5.5$, irrespective of location
 235 of the hinge magnitude. But beyond these ranges, particularly at $M_w \geq 5.5$, the median predictions deviate
 236 significantly. The apparent oversaturation of short period SAs for large magnitudes and short distances disappears
 237 completely when $M_h = 5.5, 5.7$; while for $M_h = 6.2, 6.5$ the GMM predicts a greater degree of oversaturation than
 238 that could be constrained by the data, and in doing so diverged from the general trend found in other GMMs (see
 239 Fig.3).

240

241 We consider the apparent oversaturation towards larger magnitudes as indefensible because there seems
 242 to be no clear physical explanation as to why predictions for an M7 event would be lower than M6.5. Despite,
 243 several independent datasets we have reviewed – for example, those used in the development of Cauzzi et al.
 244 (2015) and Boore et al. (2021), near-source datasets of Pacor et al. (2018) and Sgobba et al. (2021b), etc. – showed
 245 recorded short-period SAs of $M_w \geq 7$ to be lower than smaller magnitude events. In fact, Cauzzi et al. (2015)
 246 provided an explicit adjustment to the magnitude scaling of their GMM if one prefers to ignore the apparently
 247 data-driven oversaturation at higher magnitudes. Even the recently published dataset of synthetic near-source
 248 ground-motions BB-SPEEDset by Paolucci et al. (2021) also shows an on-average lower short-period SAs for
 249 $M_w \geq 7$. Boore et al. (2014) refer to the simulations of Schmedes and Archuleta (2008) to support the apparent
 oversaturation of short-periods SA , which was avoided in their earlier Boore and Atkinson (2008) GMM.

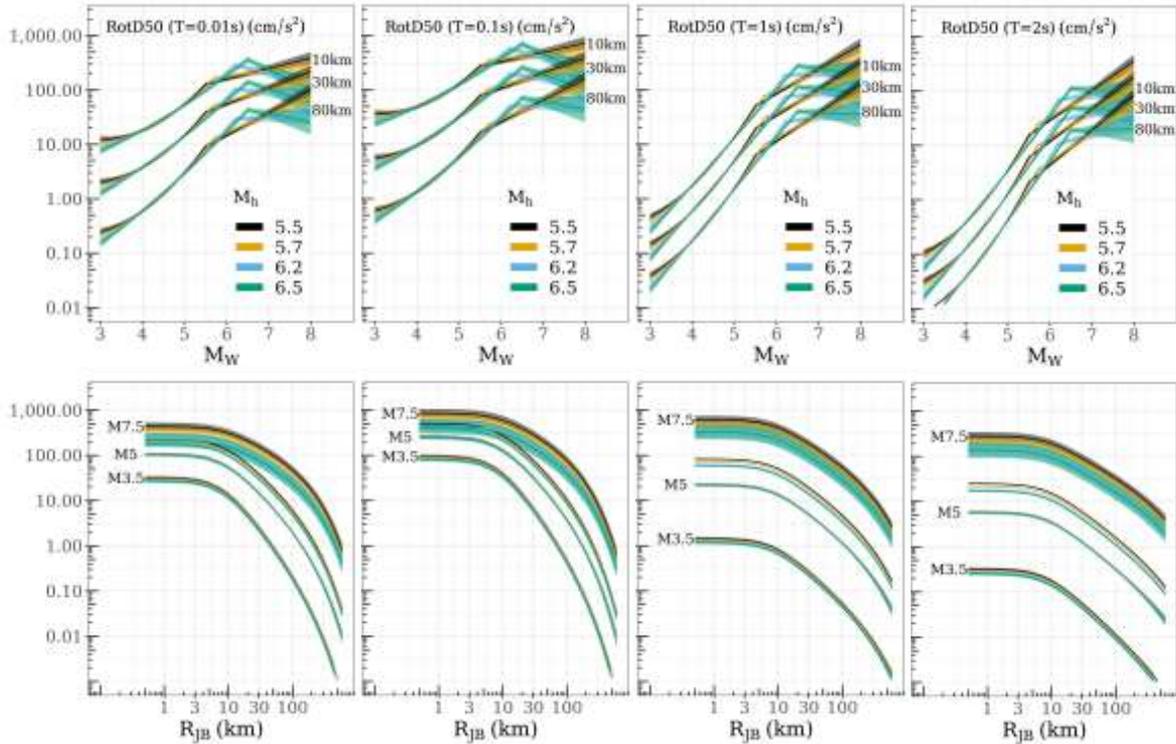


Fig.4 Magnitude and distance scaling of $SA(T = 0.01, 0.1, 1, 2s)$ along with the within-model uncertainty σ_μ when $M_h = 5.5, 5.7, 6.2, 6.5$

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

Here we mean to say that, oversaturation is apparent in several datasets, but the reasons as to why could be from: 1) very poor sampling at near-source distances from the rare large magnitude events, 2) finiteness of large ruptures and location of recording stations relative to the largest asperity, i.e. stations with $R_{JB}, R_{rup} = 0km$ but $R_{epi}, R_{hypo} \neq 0km$, 3) location of recording stations in the areas receiving lower shear-wave radiation energy (due to anisotropy) or directivity effects (e.g. Dujardin et al. 2018; Kotha et al. 2019), 4) regional variability of large ruptures due to tectonic diversity, i.e. stress loading and release rates, maturity of fault systems, etc., as reported in Radiguet et al. (2009) and Chounet et al. (2018). Of all these hypothesis, the most favourable one could be that large *characteristic* ruptures of $M_W \geq 7$ are likely to occur on the same areas of the fault plane with lower fault friction, i.e. the mature sections of the fault plane, and thereby releasing lower amount of high frequency shear-wave energy than smaller ruptures occurring in the less mature sections. In addition, rupture directivity and Doppler's effect could have caused a stacking of high-frequency pulses into a single low frequency pulse-like ground-motions, leading to higher low frequency ground-motions and lower high frequency ground-motions. These effects need to be investigated thoroughly using the recently published near-source strong motion datasets (e.g. NESS1 and NESS2) and broadband simulations (e.g. BB-SPEEDset). In-lieu of developing a semi-empirical near-source adjustments to our GMM, as in the NGA-West2 GMMs (Donahue and Abrahamson 2014), here we will first focus on revising the magnitude scaling component.

267

268

269

270

It is interesting to see (in Fig.4) that changing the M_h to lower values has curtailed the issue of apparent oversaturation; although non-parametric plots in Kotha et al. (2020) suggested it to be close $M_W = 6.2$. The fact that the fixed-effect b_3 is poorly correlated (in Fig.3) to other coefficients also helps in maintaining the distinction between scaling on either sides of M_h . In Fig.4, we also observe that the within-model uncertainty of predictions

271 (σ_μ) at $M_h \leq M_W$ is lower when considering $M_h = 5.5, 5.7$, which could be from the increased amount of data in
 272 calibrating b_3 than when $M_h = 6.2, 6.5$.

273 Fig.5 shows the response spectra predicted by the four alternative GMMs for different scenarios.
 274 Although the shape of the response spectra does not change much, one striking feature is the clear reduction in the
 275 median within-model uncertainty. In this figure, as in the previous, the $\pm\sigma_\mu$ is shown as the coloured ribbon around
 276 the median prediction. Apparently the reduction in uncertainty on b_3 has led to a substantial decrease in σ_μ , to the
 277 point that the predicted median and uncertainty of response spectra for (M7.5,10km) when setting $M_h = 5.5, 5.7$
 278 has no overlap with that of (M5.5, 10km); unlike when using $M_h = 6.2, 6.5$.

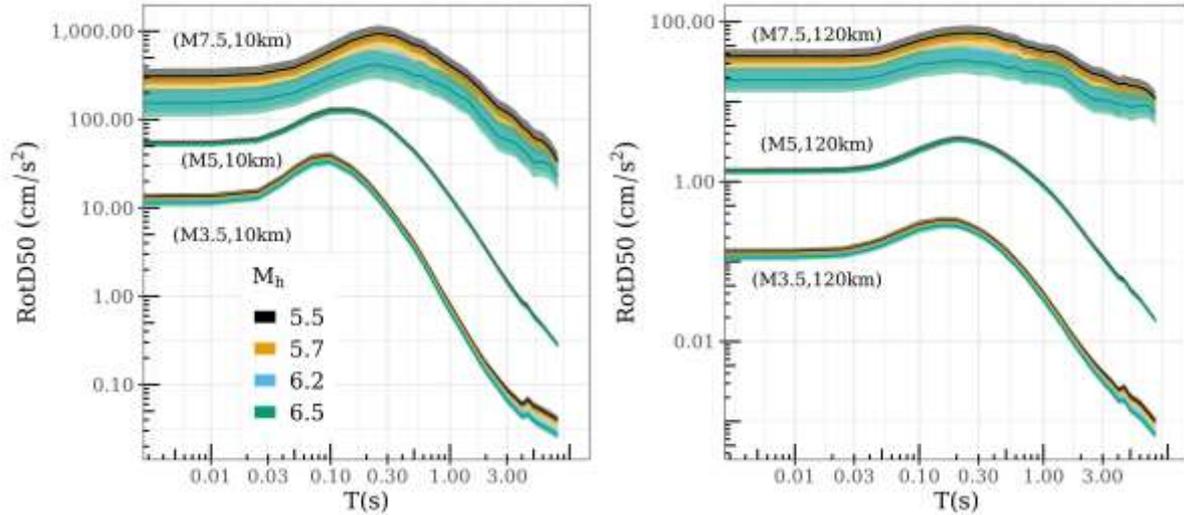


Fig.5 Response spectra predictions including within-model epistemic uncertainty ($\pm\sigma_\mu$) when $M_h = 5.5, 5.7, 6.2, 6.5$

279
 280 Fig.6 is more illustrative of the change in within-model epistemic uncertainty of the GMM depending on
 281 the choice of M_h . It is evident that σ_μ is lower for hazard critical scenarios ($M_W \geq 5.5$ & $R_{JB} \leq 30km$) when M_h
 282 is lowered to 5.5 or 5.7, due to the substantial increase in number of events and recordings used to constrain b_3 . In
 283 fact, such reduction was unexpected because it is not often the case that σ_μ is queried when choosing the functional
 284 form. Therefore, we have learned, and suggest, that reduction in σ_μ could be an interesting criterion when
 285 formulating a mixed-effects GMM; alongside the traditional non-parametric analyses, random-effect and residual
 286 analyses, and theoretical expectations. Based on all the figures referred to so far, we consider a reduction of M_h to
 287 $M_W = 5.7$ is the most reasonable revision to the fixed-effects of the GMM.

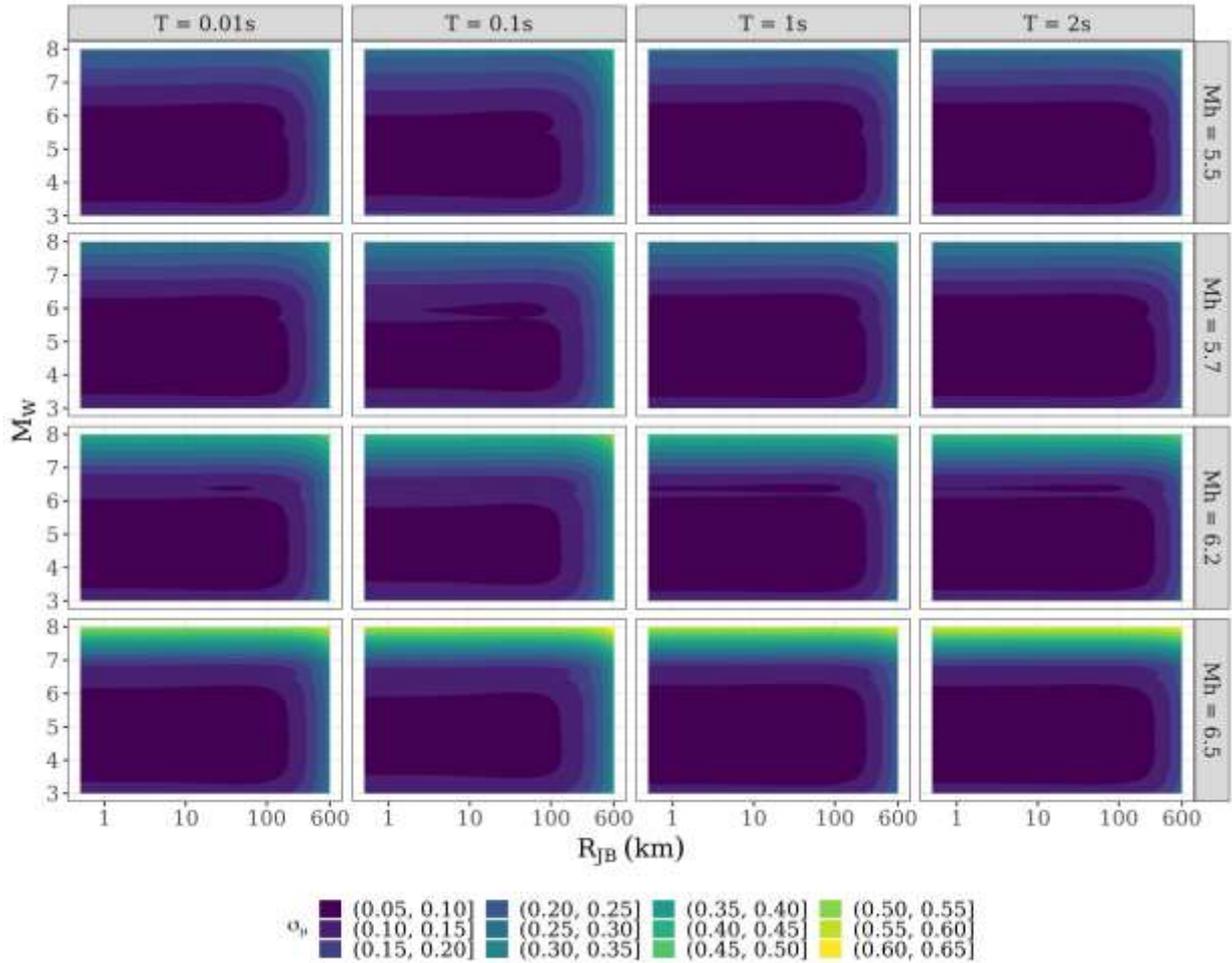


Fig.6 Within-model epistemic uncertainty over the magnitude and distance range of the ESM dataset at $T = 0.01, 0.1, 1, 2s$ when $M_h = 5.5, 5.7, 6.2, 6.5$

288

289 RANDOM-EFFECTS AND RESIDUALS

290 Assured that the fixed-effects estimates and predictions improve when shifting to $M_h = 5.7$, it is necessary to
 291 verify the impact on random-effect and residual variances. A major change in these quantities would require us to
 292 re-evaluate the physical meaning of random-effects, which would challenge the understanding that random-effects
 293 are poorly correlated to fixed-effects, as shown in von Specht and Cotton (2020).

294 Firstly, in Fig.7 we present the comparison of random-effect and residual standard-deviations (aleatory
 295 variability) of the four M_h -trial GMMs. Note that the curves associated to different choices of M_h overlap in this
 296 figure, indicating that the estimates are almost identical – with even a small reduction in between-locality (τ_{L2L})
 297 and between-event (τ_0) variabilities. Since our revision is primarily concerned with magnitude scaling, we see
 298 only minor changes in these event related aleatory variabilities. Note also that, the values shown here are not
 299 comparable to those in Kotha et al. (2020) because, the (lower) estimates therein are from robust mixed-effect
 300 regressions, while those (higher) here are from ordinary least-square mixed-effect regressions.

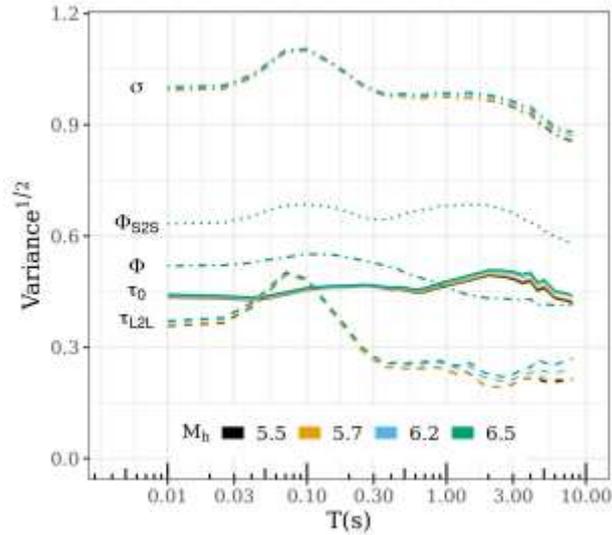


Fig.7 Residual and random-effect variances when $M_h = 5.5, 5.7, 6.2, 6.5$. Note that the curves associated to different choices of M_h overlap in this figure, indicating that the estimates are almost identical

301

302

303 We refer to the supplementary figures, Fig.S1, Fig.S2, and Fig.S3, to compare the random-effect estimates
 304 of the four M_h -trial GMMs. Therein we make a one-to-one comparison of random-effect values for *levels* of each
 305 random-effect *groups* – levels being regions, localities, and sites, when groups are between-region (attenuation,
 306 $\delta c_{3,r}$), between-locality (source, $\delta L2L_l$), and between-site ($\delta S2S_s$), respectively. Except for some very minor
 307 changes in $\delta L2L_l$ of a couple of tectonic localities l – at long periods where τ_{L2L} is smaller than at short periods
 308 (Fig.7) – there is no difference between the four M_h -trial GMMs. In addition, Fig.S4 and Fig.S5 show the trends
 309 of aleatory residuals $\varepsilon \approx N(0, \Phi_0)$ against M_W and R_{JB} , respectively. Neither plot show any striking differences
 310 among the four M_h -trial GMMs, and are not discussed further. This is a good outcome because this will require us
 311 to neither revise the discussions in Kotha et al. (2020) nor the GMM backbone logic tree implementation presented
 in the companion article of Weatherill et al. (2020b).

312

313 The most relevant changes in random-effect estimates are those associated to individual events, i.e. the
 314 inter-event or the between-event random-effect group $\Delta B_{e,l}^0 \approx N(0, \tau_0)$. Fig.7 shows that between-event standard-
 315 deviation τ_0 is almost identical among the four M_h -trial GMMs, with a minor decrease at long periods. However,
 316 this does not necessarily mean that the trends of $\Delta B_{e,l}^0$ against M_W will remain unchanged. In fact, they should
 317 necessarily change to reflect the change in the fixed-effects median predictions in Fig.4. Fig.8 shows the
 318 $\Delta B_{e,l}^0 \sim M_W$ trends for the four M_h -trial GMMs at $T = 0.01, 0.1, 1, 2s$. Here we see that the $\delta B_{e,l}^0$ in the range $4 \leq$
 319 $M_W \leq 5.5$ do not show any remarkable differences between the four regressions, reflecting the unchanged median
 prediction in Fig.4.

320

321 The most striking differences are seen at $M_W > 5.5$ – highlighted in red for clarity. In the lower two rows
 322 corresponding to the estimates when $M_h = 6.2, 6.5$, we see that most of the $\delta B_{e,l}^0$ values are negative, i.e. below
 323 the zero-baseline. This indicates that these two models – among which is also the Kotha et al. (2020) – are strongly
 324 overestimating the median predictions at $M_W > 5.5$. This aspect was overlooked in Kotha et al. (2020) because:
 325 1) that model was derived using a robust linear mixed-effects regression, and 2) the $\delta B_{e,l}^0$ of the largest events were
 326 more evenly distributed around zero-baseline. About point 1, robust random-effect variances are (generally)
 327 smaller than ordinary least-square estimates because the apparent outlier events are down-weighted, pushing their
 328 $\delta B_{e,l}^0$ further from zero and bringing the rest of $\delta B_{e,l}^0$ closer to zero. This led us to believe that the model was
 performing reasonably well. Regarding point 2, we still notice that when using $M_h = 6.2, 6.5$ the $\delta B_{e,l}^0$ of $M_W \geq 7$

329 events are almost symmetrically distributed around zero. The strong oversaturation at large magnitudes seen for
 330 these two models is of course data-driven and supported by non-parametric trends, but at the price of over
 331 predicting ground-motions of the more frequent (and relevant to Europe) $6 \leq M_W \leq 7$ events. Revising the model
 332 with $M_h = 5.5, 5.7$ easily resolves this issue. There are two events with $M_W \geq 6$ with abnormally large $\delta B_{e,l}^0$
 333 values: one of them is the M6.1 event of 10th June, 2012 in the Dodecanese Islands, Greece, and the other is the
 334 famous M6.45 Friuli earthquake of 6th May, 1976. The Friuli event is known to be one of the strongest ever in
 335 Italy with a very high stress-drop, and therefore, a strongly positive $\delta B_{e,l}^0$.

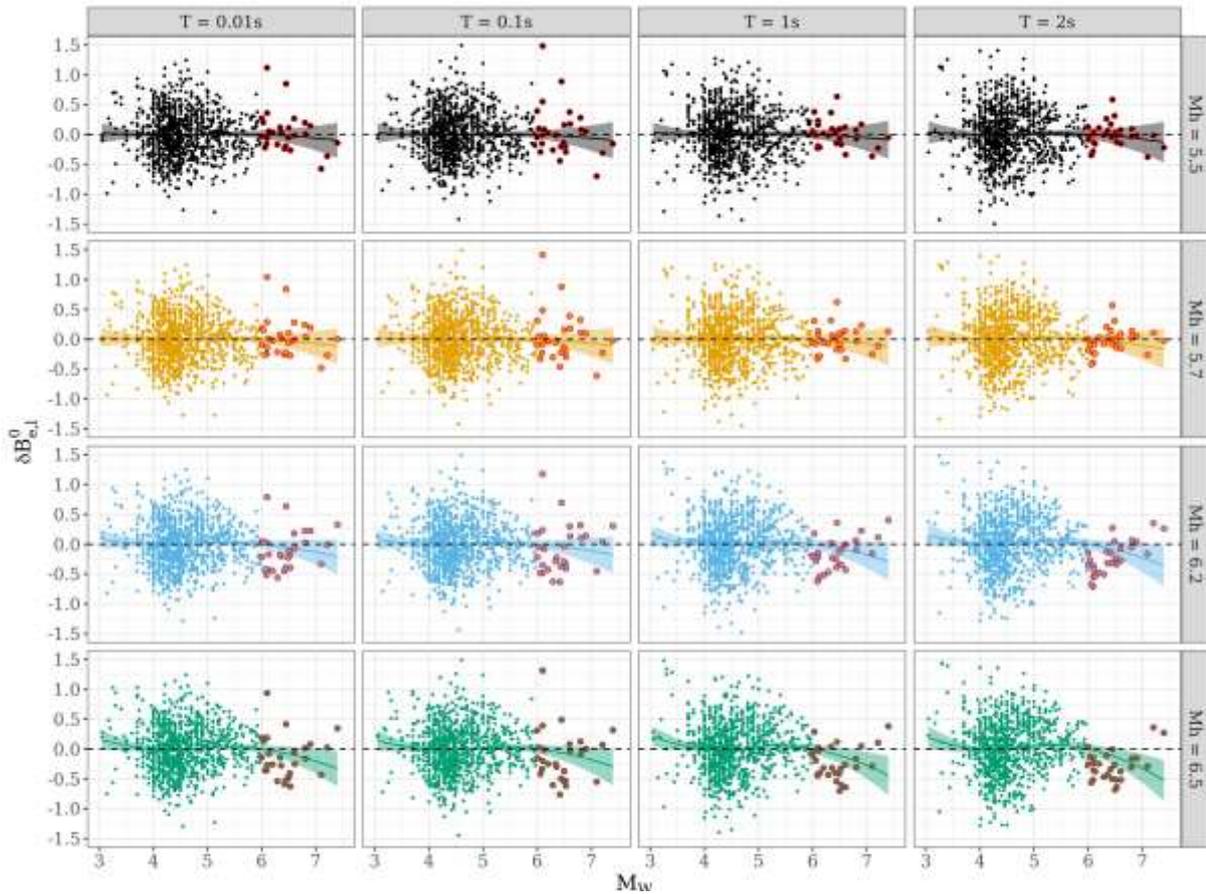


Fig.8 Between-event random-effects versus M_W of the events in ESM dataset at $T = 0.01, 0.1, 1, 2s$ when $M_h = 5.5, 5.7, 6.2, 6.5$ in the four M_h -trial GMMs

336

337 Based on the analyses shown here, we are assured that the most optimal revision to the Kotha et al. (2020)
 338 GMM would be to choose $M_h = 5.7$, while keeping the mixed-effects formulation the same. Across the trials,
 339 $M_h = 5.7$ trial GMM has the lowest within-model uncertainty on its median, and relatively unchanged random-
 340 effect and residual standard-deviations compared to the published estimates. Therefore, we move ahead by re-
 341 deriving the GMM with $M_h = 5.7$. In order to derive an updated set of coefficients to replace those provided by
 342 Kotha et al. (2020), we repeat the regressions shown previously but this time adopting the same robust linear
 343 mixed-effects regression approach used by Kotha et al. (2020).

344

EVALUATION WITH NESS DATASET

345 So far in this study, we have re-derived and revised the Kotha et al. (2020) GMM using the ESM dataset. The
 346 revised model is a clear improvement with regards to explaining the scaling of $SA(T)$ with M_W of pan-European

347 shallow crustal events; most of which originate in Italy, Turkey, and Greece. However, it still remains to be seen
 348 how this revision performs against a complementary dataset with high quality ground-motion recordings in near-
 349 source distances of large earthquakes from across the globe. For this purpose, we make use of the recently
 350 published Near Source Strong-motion datasets (NESS1 by Pacor et al. 2018 and NESS2 by Sgobba et al. 2021b).
 351 NESS2 is an update of NESS1 with several more near-source recordings and high quality event metadata. We will
 352 use and refer to NESS2 dataset as NESS from hereon.

353 DATASET

354 NESS dataset is an impressive collection of 1189 high-quality ground-motion recordings from global large
 355 magnitude events with $5.5 \leq M_W \leq 8.1$ recorded in near-source distances $0 \leq R_{JB} \leq 138km$. Along with various
 356 ground-motion intensity measures of engineering interest, NESS also provides advanced event metadata (e.g.
 357 centroid moment tensor solutions, fault geometry, stress-drop, etc.) and distance metrics (e.g. nucleation depth,
 358 fault parallel and perpendicular distances, etc.). These metrics are quite useful in exploring the near-source finite
 359 fault effects, such as directivity, shear-wave radiation pattern, and pulse-like ground-motions as in Sgobba et al.
 360 (2021a). Moreover, the NESS2 dataset is compiled using the same processing methodologies as those of ESM;
 361 and hence, complement each other seamlessly. We refer the readers to Pacor et al. (2018) and Sgobba et al. (2021b)
 362 for elaboration on the dataset.

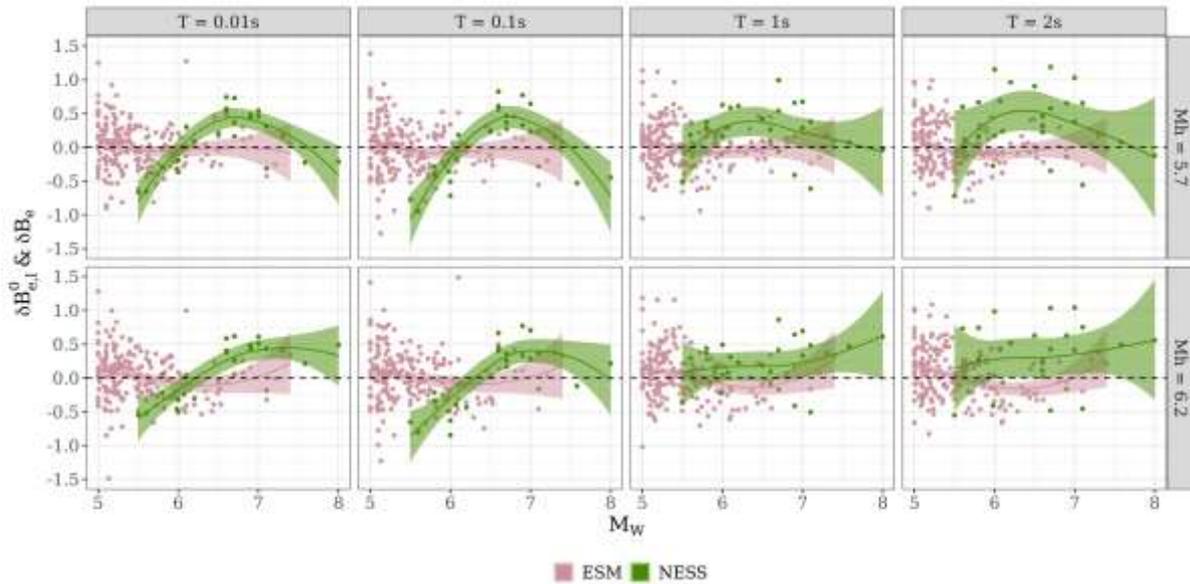
363 Fig.1 compares the $M_W - R_{JB}$ distribution of data between NESS and ESM datasets. It is evident that
 364 NESS has much more data from hazard critical scenarios. However, we note that there are around 177 records,
 365 and 39 events common to both the datasets. Of the 34 recordings exclusive to NESS but from events present also
 366 in ESM, 15 are from the M6.5 Norcia, Italy earthquake of 30th October, 2016; 10 from the M7.6 Izmit, Turkey
 367 earthquake of 17th August, 1999; and one each from 9 other events. In this study, we chose to discard all NESS
 368 records or events that are already present in ESM; amounting to 211 recordings. In addition, 8 recordings
 369 associated to the M8.1 Iquique, Chile earthquake of 1st April, 2014 have also been removed because it might have
 370 been wrongly classified in NESS as shallow crustal event while being an interface event. Following a further data
 371 screening with the same selection criteria as in Kotha et al. (2020), out of the 1189 recordings we are left to evaluate
 372 the performance of our revised GMM against 947 recordings from earthquakes originating in Taiwan (382), USA
 373 (242), New Zealand (204), Japan (85), Mexico (27), and Italy (7). Note that the 7 records from Italy that did not
 374 feature in ESM – which contains only data up to 23rd December, 2016 – are from an M5.5 aftershock of the 2016
 375 central Italy sequence occurring on 18th January, 2017.

376 RANDOM-EFFECTS AND RESIDUALS

377 To evaluate the performance of the Kotha et al. (2020) GMM and its revision presented here, we perform a random-
 378 effect and residual analyses of the NESS dataset. Instead of the customary plots showing GMM fixed-effects
 379 (median) prediction against observations, we choose to present and analyse the trends of random-effects and
 380 residuals against predictor variables M_W and R_{JB} . For this purpose, we perform an ordinary least-square random-
 381 effect decomposition as shown in equation below:

$$\ln(Obs_{e,s}) - \ln(\mu) = offset(0) + \delta B_e + \delta W_{e,s} \quad (5)$$

382 In Eq. (5), $\ln(Obs_{e,s})$ are the natural-log of $SA(T)$ in the NESS dataset recorded for event e by site s ,
 383 and $\ln(\mu)$ is the fixed-effects GMM prediction for that event at that event-to-site distance (in R_{JB}). Since there are
 384 two GMMs, $\ln(\mu)$ will also be different – one with $M_h = 5.7$ and other with $M_h = 6.2$. No region-, locality-, and
 385 site-specific adjustments are made to $\ln(\mu)$. Residuals from $\ln(Obs_{e,s}) - \ln(\mu)$ are then split into between-event
 386 and within-event terms. On the right-hand side of Eq. (5), $offset(0)$ means the random-effects decomposition
 387 does not estimate an offset or bias term. Unlike in the GMM, here we estimate the traditional between-event
 388 random-effect δB_e , with subscript e as the event index. These are different from the $\delta B_{e,l}^0$ of the GMM, which are
 389 corrected for tectonic locality-specific $\delta L2L_l$. Technically, $\delta B_e \approx \delta B_{e,l}^0 + \delta L2L_l$, but since we could not *localize*
 390 the events in NESS, we could only estimate the δB_e . Out of the 1049 sites in NESS, 826 do not have a
 391 measured V_{s30} . As most of the sites are singly sampled, we did not introduce a between-site random-effect $\delta S2S_s$,
 392 where subscript s would be the site index. This means that, all the site-effects are redistributed among between-
 393 event δB_e and within-event residuals $\delta W_{e,s}$.



394 **Fig.9** Between-event random-effects, δB_e for NESS and $\delta B_{e,l}^0$ for ESM, versus M_W of the events in ESM and NESS
 395 dataset at $T=0.01, 0.1, 1, 2s$ when $M_h = 5.7, 6.2$ (rlmm)

396 Now that there are two GMMs we are evaluating – revised GMM with $M_h = 5.7$ and original GMM with
 397 $M_h = 6.2$ – against the NESS datasets, we have two sets of δB_e and $\delta W_{e,s}$ for PGA , PGV , and $SA(T = 0.01 -$
 398 $8s)$. Similarly, for the ESM dataset, we have two sets of $\delta B_{e,l}^0$ and ε . Note again that the $\delta B_{e,l}^0$ and ε are filtered
 399 for tectonic locality-, attenuation region-, and site-specific effects, while δB_e and $\delta W_{e,s}$ are not. Therefore,
 although these are not technically the same quantities, they are still comparable.

400 Fig.9 plots the between-event random-effects against M_W of the $M_W \geq 5$ events in ESM and NESS
 401 dataset, estimated with respect to the GMMs with $M_h = 5.7, 6.2$. Of course, the localized non-parametric trend
 402 (*loess* by Jacoby 2000) always stays close to zero for the ESM data because the models were calibrated on it. The
 403 NESS data, however, is significantly off the zero baseline, and often positively biased. This indicates that, despite
 404 the choice of M_h , both the GMMs under-predict the ground-motions of events in NESS dataset – although the
 405 unrevised Kotha et al. (2020) GMM with $M_h = 6.2$ appears to perform slightly better with a smaller bias. The

406 between-event standard-deviation τ_0 of the two GMMs is about 0.45 at all periods (see Fig.7); which means,
 407 although most NESS events are within $\pm\tau_0$ at short periods, at long periods most are significantly above $+\tau_0$.

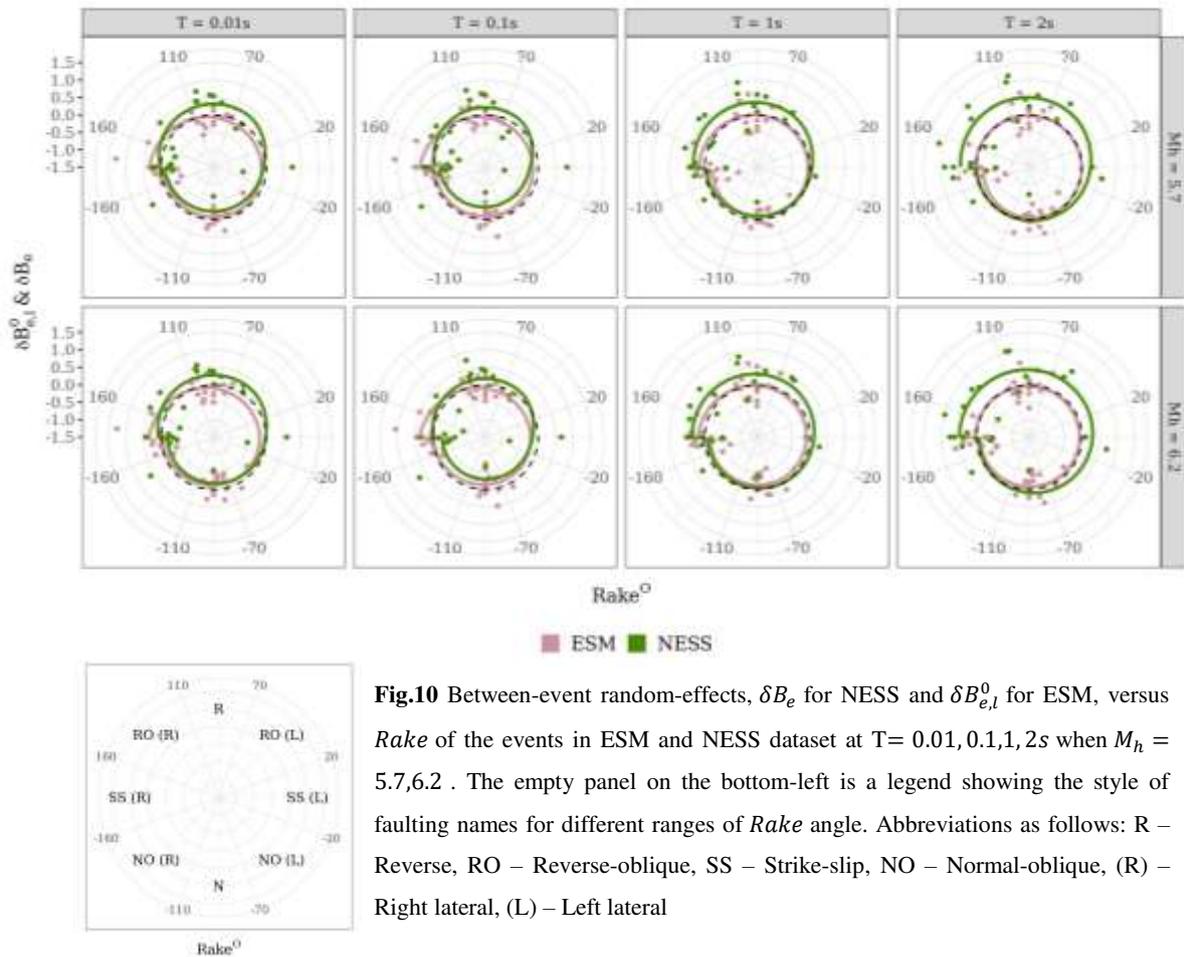


Fig.10 Between-event random-effects, δB_e for NESS and $\delta B_{e,l}^0$ for ESM, versus *Rake* of the events in ESM and NESS dataset at $T=0.01, 0.1, 1, 2s$ when $M_h = 5.7, 6.2$. The empty panel on the bottom-left is a legend showing the style of faulting names for different ranges of *Rake* angle. Abbreviations as follows: R – Reverse, RO – Reverse-oblique, SS – Strike-slip, NO – Normal-oblique, (R) – Right lateral, (L) – Left lateral

408

409 The misfit between both the GMMs and the NESS events is a concern. But it is important to note a few
 410 assumptions made in this comparison: 1) the $\delta B_{e,l}^0$ of the ESM events are corrected for specificities of the tectonic
 411 locality in which they originate using the $\delta L2L_l$, and of attenuating regions through which their shear-waves
 412 propagate using the $\delta c_{3,r}$ (more details in Kotha et al. 2020). While the δB_e of the NESS events are not, and may
 413 include some of these effects, 2) the $\delta B_{e,l}^0$ of the ESM events are corrected for site-response of their recording
 414 stations, while the δB_e of the NESS events are not. This means that, if these events were preferentially recorded
 415 by soft-soil sites, then the their site-amplification is likely to have crept into δB_e estimates, and possibly inflating
 416 them, and 3) ESM events are primarily recorded at $R_{JB} \geq 30km$, while the NESS events at $R_{JB} \leq 30km$. Any
 417 near-source finite-fault effects, such as directivity and shear-wave radiation pattern, may have aggravated their
 418 recorded ground-motions and in-turn, when unfiltered, may have inflated their δB_e .

419 Fig.10 is a polar coordinate diagram showing the δB_e of NESS events and $\delta B_{e,l}^0$ of ESM events against
 420 their *rake* angle. The radius of the plot measures the δB_e and $\delta B_{e,l}^0$ values, and the azimuth measures the *rake* of
 421 the event. In this figure, we see that the loess fit for ESM events coincides with the zero baseline – indicating no
 422 bias towards any focal mechanism. While for the NESS events, although the mean trend (of loess fit) coincides
 423 with zero baseline in the range $-180^\circ \leq rake \leq 20^\circ$, elsewhere there's a strong positive bias. This range $20^\circ <$
 424 $rake < 180^\circ$ corresponds to events often identified as reverse-slip or thrust faulting mechanism. Reverse-slip

425 ruptures often generate stronger ground-motions on their hanging-wall side than on foot-wall side (at the same R_{JB})
 426 due to *wedge* geometry. If the recording stations were concentrated on this side of the fault trace, then it is likely
 427 that the finite-fault effects and anisotropic radiation (along with site-effects) might have aggravated their recorded
 428 ground-motions (e.g. Spudich 2013, Donahue and Abrahamson 2014, Kamai et al. 2014, Dujardin et al. 2018,
 429 Kotha et al. 2019, Sgobba et al. 2021a, etc.)

430 **Table.1** NESS events with between-event random-effect δB_e larger than τ_0 of the Kotha et al. (2020) GMM. Abbreviations as
 431 follows: US – United States of America, MX – Mexico, JP – Japan, NZ – New Zealand; LL – Left lateral, RL – Right lateral,
 432 Rev – Reverse

Date	Region	M_W	Style of faulting	R_{JB} (km)	V_{s30} (m/s)	Records (Pulse-like)	Stress-drop (MPa)
				[0 th , 100 th]	[0 th , 50 th , 100 th]		
09/02/1971	US	6.7	Reverse	[3, 30]	[1116, 1116, 1116]	27 (7)	-
13/08/1978	US	5.8	LL Rev-Oblique	[0, 6]	-	4 (2)	-
15/10/1979	MX	6.5	RL Strike-slip	[0, 23]	[163, 202, 250]	27 (16)	-
18/10/1989	US	6.9	RL Rev-Oblique	[0, 35]	[222, 312, 730]	29 (12)	-
25/04/1992	US	7	Reverse	[8, 33]	[219, 219, 219]	7 (3)	-
17/01/1994	US	6.7	Reverse	[0, 29]	[257, 343, 1116]	44 (9)	-
23/10/2004	JP	6.6	Reverse	[0, 27]	[129, 372, 686]	15 (3)	3.7
03/09/2010	NZ	7.1	RL Strike-slip	[0, 44]	[154, 207, 435]	33 (17)	5.7
21/02/2011	NZ	6.2	RL Rev-Oblique	[0, 17]	[154, 208, 435]	14 (6)	8.8
22/02/2011	NZ	5.6	LL Strike-slip	[2, 7]	[187, 197, 207]	3 (3)	4.2
13/06/2011	NZ	6	RL Strike-slip	[0, 10]	[154, 194, 221]	9 (6)	7
24/08/2014	US	6.07	RL Strike-slip	[3, 14]	-	13 (5)	-

433 Table.1 lists the events whose between-event random-effect δB_e is often larger than τ_0 of the Kotha et al.
 434 (2020) GMM at $0.01s \leq T \leq 8s$. These events show all the characteristics that we hypothesized might have
 435 aggravated their recorded ground-motions. For instance, most events are of reverse-slip rupture type, recorded at
 436 very close distances, by sites whose V_{s30} characterises them as soft-soils with strong amplification of long period
 437 ground-motions, and most interestingly, a good proportion of these recordings are identified with pulse-like (finite-
 438 fault effect) features. Here we mean to say that, it is clear that our GMMs are not able to predict the recorded
 439 ground-motions of these events, and the reason is two-fold: lack of site information to correct for site-effects, and
 440 lack of GMMs ability to predict evident finite-fault effects. With this experience, we anticipate further
 441 sophisticating our GMM with empirical or semi-empirical finite-fault adjustments to near-source predictions. For
 442 this purpose, the NESS2 dataset of Sgobba et al. (2021b), BB-SPEEDset of Paolucci et al. (2021) and Donahue
 443 and Abrahamson (2014) provide exciting opportunities.

444 Supplementary figure Fig.S6 shows the trend of within-event residuals $\delta W_{e,s}$ of the NESS events plotted
 445 against M_W , along with the aleatory residuals ε of the ESM events. Since these plots show no discernible trends,
 446 we have excluded it from discussion. Fig.11 shows the same residuals against R_{JB} , with the pulse-like ground-

447 motion residuals in black. At longer periods, i.e. $T = 1, 2s$ here, residuals from pulse-like recordings are
 448 preferentially above zero baseline, indicating that these ground-motions are systematically stronger than the GMM
 449 predictions. Without dwelling further, we refer to a similar finding by Sgobba et al. (2021a); wherein, pulse-like
 450 recordings are characterised by a higher than average long-period SAs and lower than average short-period SAs .
 451 We recognize the need to upgrade our GMMs to reflect these finite-fault features, but perhaps not in this study.

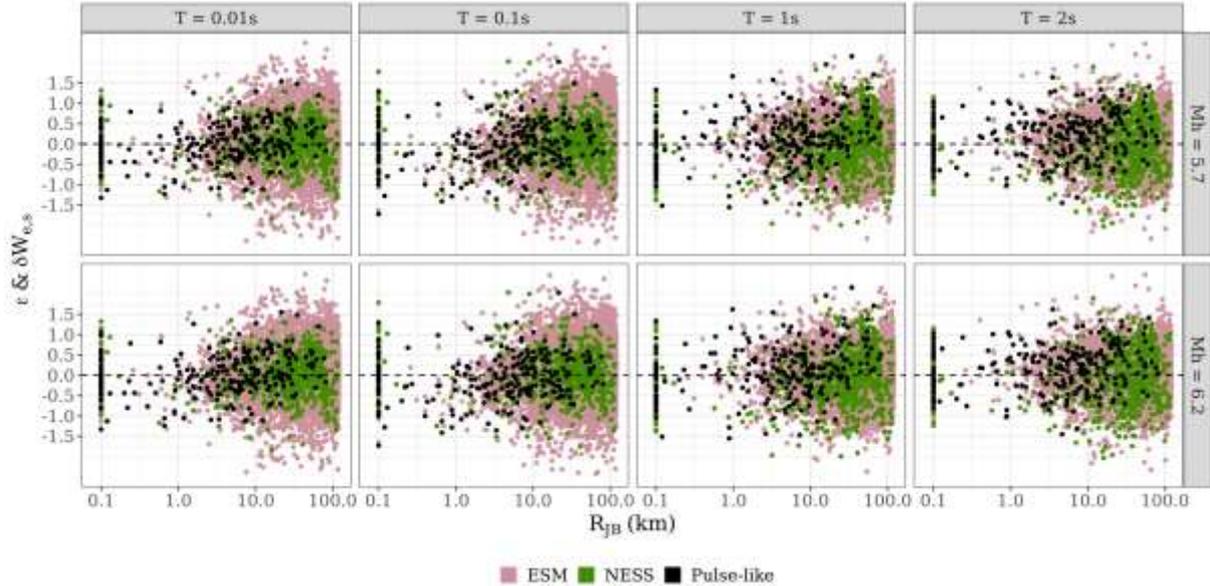


Fig.11 Within-event residuals versus R_{jB} of the records in ESM and NESS dataset at $T = 0.01, 0.1, 1, 2s$ when $M_h = 5.7, 6.2$. Black markers indicate those records with reported pulse-like effects after removing the shared events

452

453

HETEROSCEDASTICITY

454 Weatherill et al. (2020b) provided an M_W dependent heteroskedastic model of τ_0 for use with the implementation
 455 of (revised) Kotha et al. (2020) GMM in the regionally adaptable ESHM20 backbone logic-tree. While that
 456 heteroskedastic model of τ_0 is adapted from Al Atik (2015), which was developed from a global dataset of events,
 457 here we provide one using only the ESM events. Alongside a heteroskedastic $\tau_0 \sim M_W$ for the ESM events, we
 458 also examine the impact of adding NESS events to ESM events.

459 Fig.12 shows the investigation of $\tau_0 \sim M_W$ for two cases: $\delta B_{e,l}^0$ of only the ESM events, $\delta B_{e,l}^0$ of ESM
 460 events along with δB_e of NESS events. We perform a simple test for heteroscedasticity – the Breusch-Pagan test
 461 (Breusch and Pagan 1979) – programmed into an R package *lmtest* by Zeileis and Hothorn (2002). The Breusch-
 462 Pagan test fits a linear regression model to the residuals of a linear regression model (here $\delta B_{e,l}^0 \sim M_W$) and rejects
 463 the null hypothesis of homoscedasticity if too much of the variance (here τ_0) is explained by the additional
 464 explanatory variables (here M_W). H_0 is the test statistic of the Breusch-Pagan test, and follows a chi-squared
 465 distribution with the number of regression parameters (here M_W) as degrees of freedom (here 1). Therefore, if H_0
 466 has a p-value below an appropriate threshold (e.g. $p < 0.05$) then the null hypothesis of homoscedasticity is
 467 rejected and heteroscedasticity is assumed.

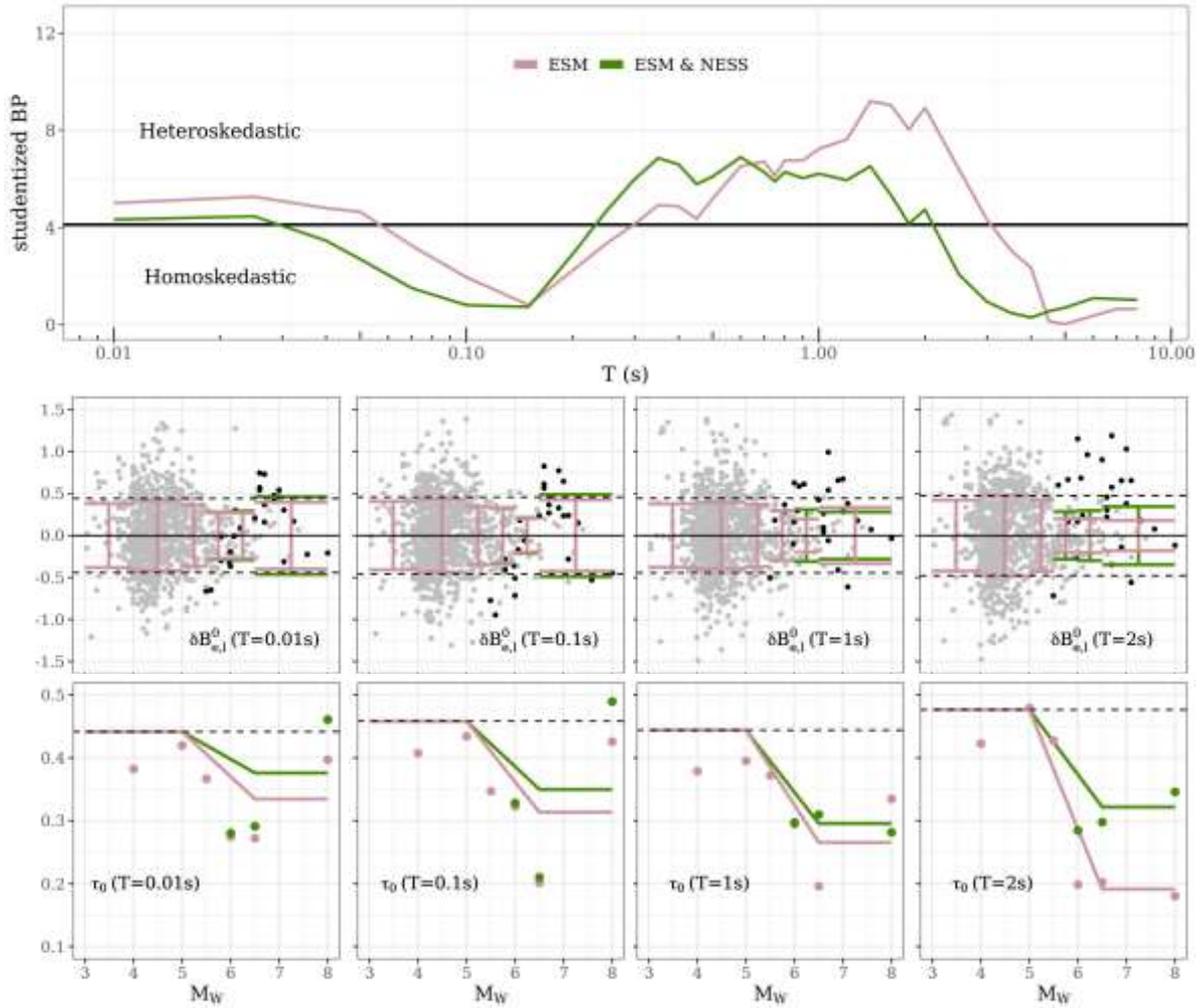


Fig.12 Heteroscedasticity of between-event ($\delta B_{e,l}^0, \delta B_e$) variability with M_W of the events in ESM and ESM&NESS datasets: Top row panel shows the studentised Breusch-Pagan statistic, which is a measure of heteroscedasticity of between-event variability with M_W for the $T = 0.01 - 8s$. Middle row panels show the $\delta B_{e,l}^0(T)$ and $\delta B_e(T)$ versus M_W for $T = 0.01, 0.1, 1, 2s$. The error-bars show the Median Absolute Deviance (sample MAD) of $\delta B_{e,l}^0$ and $(\delta B_{e,l}^0, \delta B_e)$ combination within bins $M_W \in (3, 4], (4, 5], (5, 5.5], (5.5, 6], (6, 6.5], (6.5, 8]$. Bottom row panels show the sample MAD (markers) within the M_W bins and the proposed heteroskedastic τ_0 models

468

469

470

471

472

473

474

475

476

477

478

The top panel of Fig.12 shows the studentised Breusch-Pagan statistic of $\delta B_{e,l}^0, \delta B_e \sim M_W$ at $T = 0.01s - 8s$. The horizontal black line is the estimated value of the H_0 statistic to be exceeded to reject the null hypothesis of homoscedasticity with 95% statistical significance. Lower values do not necessarily mean homoscedasticity, but simply a lower significance of heteroscedasticity. In this case, we observe that the heteroscedasticity is most significant for both ESM and ESM&NESS datasets in the range $0.2 < T < 2s$. Further, heteroscedasticity of ESM events' $\delta B_{e,l}^0$ is stronger than when including also the δB_e of global, tectonically diverse, events of NESS. As seen in Table.1, one of the reasons a few NESS $M_W \geq 6$ events have large δB_e – thereby increasing the scatter – is likely because of their origin in a tectonically different environment (region), and style of faulting compared to those in Italy, Greece, and Turkey. Nevertheless, this could also be due to the finite-fault and site effects mentioned earlier.

479 In the middle panels of Fig.12, we show the $\delta B_{e,l}^0 \sim M_W$ of ESM events in grey, overlain by $\delta B_e \sim M_W$
 480 of NESS events in black. The error-bars indicate the $\pm MAD$ (median absolute deviation) of $\delta B_{e,l}^0$ and $(\delta B_{e,l}^0, \delta B_e)$
 481 in magnitude bins $M_W \in [3.4, 4), [4,5), [5,5.5), [5.5,6), [6,6.5), [6.5,7.4]$. At $M_W < 5$, the MAD is comparable to
 482 the τ_0 , while at $M_W \geq 6.5$ the MAD is on average 35% smaller than τ_0 for ESM events depending on the period.
 483 For the combination ESM&NESS events however, although MAD for $M_W \geq 6.5$ is 20% lower than that of $M_W <$
 484 5 at $T \leq 3s$, at $T > 3s$ the MAD for $M_W \geq 6.5$ is in fact 15% larger. At such long periods, we hypothesis that
 485 many more complexities are responsible for the apparently increased diversity of large event ground-motions.
 486 Therefore, we consider our heteroskedastic model for the combination ESM&NESS as purely experimental.

487 In the lower panels of Fig.12, we show the estimated M_W binwise MAD (markers), and the proposed
 488 piece-wise linear model of heteroskedastic between-event variability is shown in equation (6). For the moment,
 489 we provide a readily applicable $\tau_0(M_W)$ for the ESM events, and an experimental one for ESM&NESS data:

$$\tau_0(M_W) = \begin{cases} \tau_{0,M_1} = \tau_0 & M_W < M_1 = 5 \\ \tau_{0,M_1} + (M_W - M_1) \left(\frac{\tau_{0,M_2} - \tau_{0,M_1}}{M_2 - M_1} \right) & M_1 \leq M_W < M_2 = 6.5 \\ \tau_{0,M_2} & M_W \geq M_2 \end{cases} \quad (6)$$

490 In equation (6), at each period (T , not shown in equation), τ_{0,M_1} is the $\tau_0(M_W)$ for events with $M_W <$
 491 $M_1 = 5$, kept equal to τ_0 . For $M_W \geq M_2 = 6.5$, $\tau_0(M_W)$ is a smaller τ_{0,M_2} . For $M_1 \leq M_W < M_2$, $\tau_0(M_W)$
 492 decreases linearly from τ_{0,M_1} to τ_{0,M_2} . Values of τ_{0,M_1} are the same for ESM and ESM&NESS datasets, but τ_{0,M_2}
 493 values are different, as seen in the lower-panels of Fig.12.

494 In addition to the heteroscedasticity of between-event variability, we have examined the magnitude
 495 dependent heteroscedasticity of residual variability; i.e. the so-called within-event $\delta W_{e,s}$ residuals of NESS data
 496 containing the various regional and site-specific effects, and the aleatory residuals ε of ESM data corrected for
 497 region-, locality- and site-specific effects by Kotha et al. (2020). Once again, although $\delta W_{e,s}$ and ε are not
 498 technically the same, they are both the *left-overs* after filtering out all possible the fixed- and random-effects from
 499 observations. Supplementary Fig.S7 is similar to Fig.12 shown above for between-event random-effects. Therein
 500 we observed that the $\delta W_{e,s}$ of NESS data show strong heteroscedasticity at long periods. However, unlike the
 501 decreasing variability of δB_e with increasing M_W , long period $\delta W_{e,s}(T \geq 1s)$ variability increases rapidly with
 502 increasing M_W . We understand this to be due to the anisotropy of shear-wave radiation, pulse-like effects, etc. in
 503 near-source ground-motions. We believe it is necessary to investigate these effects in more detail than that can be
 504 presented here. Therefore, for the moment, we will not further discuss heteroscedasticity of within-event $\delta W_{e,s}$
 505 residuals of the NESS dataset.

506

DISCUSSION

507 In this study, we have revised the fixed-effect regression coefficients of the Kotha et al. (2020) GMM. We first
 508 discussed the adequacy of various fixed-effects in the GMM in capturing the primary physical phenomenon
 509 controlling the seismic ground-motion observations. An outstanding issue with the GMM was the oversaturation
 510 of short-period SAs with increasing magnitude at near-source distances. Alongside, the selection of a hinge-
 511 magnitude $M_h = 6.2$ led to large within-model uncertainty of median predictions at $M_W \geq 6.2$. Since large

512 magnitude earthquakes occurring at short distances from the site are often the hazard critical scenarios, the GMM
 513 within-model uncertainty has a large impact on seismic hazard results. Therefore, the most optimal revision, i.e.
 514 without effecting the random-effects and residual variances, was to revise the GMM with a different M_h . We have
 515 carefully analysed the within-model uncertainties associated to the choice of M_h and finally chose $M_h = 5.7$ to
 516 curtail the oversaturation issue. A new set of fixed-effect coefficients are generated using robust linear mixed-
 517 effect regressions to replace the one provided with Kotha et al. (2020).

518 In fact, there are several alternative magnitude-scaling formulations already implemented in various
 519 GMMs. The one in Kotha et al. (2020) is among the simplest one with a linear and quadratic scaling terms for
 520 small-moderate magnitudes ($M_W < M_h$), and linear-only term for $M_W \geq M_h$. Even with such simple formulation,
 521 the choice of M_h is still ambiguous, and may render the GMM behaviour unexplainable at $M_W \geq M_h$. Therefore,
 522 in addition to non-parametric analyses, which was the sole justification for $M_h = 6.2$ in Kotha et al. (2020), we
 523 have learned that querying the within-model uncertainty can be helpful in choosing an appropriate M_h .

524 Another interesting observation in this study is the scaling of *SAs* for small magnitudes (e.g. $M_W \leq 4$).
 525 While empirical and theoretical expectations suggest that ground-motions decay rapidly towards smaller
 526 magnitudes, our data-driven model coefficients suggest it need not be so. Newer models with lower magnitude
 527 limit of applicability up to $M_W = 3$, despite very low within-model uncertainty, are perhaps biased by uncertainty
 528 in M_W estimates. Therefore, we suggest caution when applying this GMM to low-moderate seismicity regions (e.g.
 529 France and Germany), where the M_W estimates of small earthquakes may be quite uncertain, if not inferred from
 530 another magnitude scale (e.g. local magnitude M_l , body-wave magnitude m_b , etc.)

531 Once the GMM coefficients were revised to have lower within-model uncertainty on large magnitude
 532 predictions at near-source distances, we evaluated its performance against a high-quality near-source strong
 533 ground-motion dataset. Unlike the regressed ESM dataset however, the NESS dataset contains ground-motion data
 534 collected from tectonically diverse, global, large magnitude earthquakes recorded by dense network of stations
 535 located on particularly soft-soil geology. In addition, a large fraction of NESS data are characterised by pulse-like
 536 characteristics. Owing to its peculiar sampling, NESS was decided against integration into ESM; in order to not
 537 bias the isotropic Kotha et al. (2020) predictions with anisotropic pulse-like ground-motion observations. We
 538 noticed that most of the large magnitude events in NESS have produced stronger ground-motions than that the
 539 GMM predicts. The misfit between NESS observations and GMM median predictions is larger at longer periods,
 540 where pulse-like effects and soil conditions play a major role in aggravating ground-motions. At short periods, the
 541 misfit was within one standard-deviation, and that's an improvement we could achieve by revising the model with
 542 $M_h = 5.7$ instead of $M_h = 6.2$ of Kotha et al. (2020). For the sake of brevity however, we have only demonstrated
 543 that between-event random-effect and aleatory residual variabilities become larger if NESS data is integrated into
 544 ESM. To evaluate these estimates would require us to impose the same level of non-ergodicity (event, site, tectonic
 545 locality, and attenuation region) on NESS data as that was possible on ESM dataset – but was not possible. NESS
 546 dataset can instead be used to develop empirical or semi-empirical models to capture near-source finite-fault
 547 effects. We anticipate developing and appending such models to our revised GMM in the near future.

548

DATA AND RESOURCES

549 The European Strong Motion flatfile is available at <https://esm.mi.ingv.it/flatfile-2018/> with persistent identifier
550 PID: 11099/ESM_flatfile_2018. The NESS flatfiles are available at <http://ness.mi.ingv.it/>. The analyses in this
551 study have been performed in R software (Team 2013). In addition to those cited in the maintext, we used the
552 libraries *dplyr* (Wickham et al. 2019b), *ggplot2* (Wickham et al. 2019a), *ggmap* (Kahle et al. 2019), *viridis* (Garnier
553 2019), etc.

554

ACKNOWLEDGEMENTS

555 The model revision has benefitted immensely from feedback provided by the collaborators in Horizon 2020
556 “Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe (SERA)” project with Grant
557 Agreement No. 730900. A special thanks to Prof. Roberto Paolucci and his team for a thorough test of the model
558 against BB-SPEEDset numerical simulations.

559

DECLARATIONS

560 FUNDING

561 Publication costs of the manuscript are supported by the Géophysique des Risques sismiques Et gravitaires (GRE)
562 team at ISTerre, Grenoble 38000, France

563 CONFLICTS OF INTEREST/COMPETING INTERESTS

564 The authors declare no conflicting or competing interest

565 AVAILABILITY OF DATA AND MATERIAL

566 Availability of data is declared under Data and Resources section

567 CODE AVAILABILITY

568 All the open-source scripts can be made available upon request

569

REFERENCES

- Abrahamson NA, Silva WJ, Kamai R (2014) Summary of the ASK14 Ground Motion Relation for Active Crustal Regions Earthquake Spectra 30:1025-1055 doi:10.1193/070913eqs198m
- Akkar S, Sandıkkaya M, Bommer J (2014a) Empirical ground-motion models for point-and extended-source crustal earthquake scenarios in Europe and the Middle East Bulletin of earthquake engineering 12:359-387
- Akkar S et al. (2014b) Reference database for seismic ground-motion in Europe (RESORCE) Bulletin of earthquake engineering 12:311-339 doi:10.1007/s10518-013-9506-8
- Al Atik L (2015) NGA-East: Ground-motion standard deviation models for central and eastern North America PEER Report 2015 7
- Ancheta TD et al. (2014) NGA-West2 database Earthquake Spectra 30:989-1005
- Atik LA, Youngs RR (2014) Epistemic uncertainty for NGA-West2 models Earthquake Spectra 30:1301-1318
- Bahrampouri M, Rodriguez-Marek A, Shahi S, Dawood H (2021) An updated database for ground motion parameters for KiK-net records Earthquake Spectra 37:505-522
- Basili R et al. (2019) NEAMTHM18 Documentation: the making of the TSUMAPS-NEAM Tsunami Hazard Model 2018
- Bates D, Mächler M, Bolker B, Walker S (2014) Fitting linear mixed-effects models using lme4 arXiv preprint arXiv:14065823
- Bindi D, Cotton F, Kotha SR, Bosse C, Stromeyer D, Grünthal G (2017) Application-driven ground motion prediction equation for seismic hazard assessments in non-cratonic moderate-seismicity areas Journal of Seismology 21:1201-1218 doi:10.1007/s10950-017-9661-5
- Bindi D, Kotha S (2020) Spectral decomposition of the Engineering Strong Motion (ESM) flat file: regional attenuation, source scaling and Arias stress drop Bulletin of Earthquake Engineering:1-26
- Bindi D, Kotha SR, Weatherill G, Lanzano G, Luzi L, Cotton F (2018) The pan-European engineering strong motion (ESM) flatfile: consistency check via residual analysis Bulletin of Earthquake Engineering:1-20
- Bindi D, Massa M, Luzi L, Ameri G, Pacor F, Puglia R, Augliera P (2014) Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset Bulletin of earthquake engineering 12:391-430
- Bommer JJ (2012) Challenges of building logic trees for probabilistic seismic hazard analysis Earthquake Spectra 28:1723-1735
- Boore DM, Atkinson GM (2008) Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s Earthquake Spectra 24:99-138
- Boore DM, Stewart JP, Seyhan E, Atkinson GM (2014) NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes Earthquake Spectra 30:1057-1085
- Boore DM et al. (2021) A ground-motion prediction model for shallow crustal earthquakes in Greece Bulletin of the Seismological Society of America 111:857-874
- Breusch TS, Pagan AR (1979) A simple test for heteroscedasticity and random coefficient variation Econometrica: Journal of the econometric society:1287-1294

- Brune JN (1970) Tectonic stress and the spectra of seismic shear waves from earthquakes *Journal of geophysical research* 75:4997-5009
- 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* 30:1087-1115 doi:10.1193/062913eqs175m
- Castro R, Anderson J, Singh S (1990) Site response, attenuation and source spectra of S waves along the Guerrero, Mexico, subduction zone *Bulletin of the Seismological Society of America* 80:1481-1503
- Cauzzi C, Faccioli E, Vanini M, Bianchini A (2015) Updated predictive equations for broadband (0.01–10 s) horizontal response spectra and peak ground motions, based on a global dataset of digital acceleration records *Bulletin of Earthquake Engineering* 13:1587-1612
- Chiou BS-J, Youngs RR (2014) Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra *Earthquake Spectra* 30:1117-1153
- Chounet A, Vallée M, Causse M, Courboux F (2018) Global catalog of earthquake rupture velocities shows anticorrelation between stress drop and rupture velocity *Tectonophysics* 733:148-158
- Cotton F, Pousse G, Bonilla F, Scherbaum F (2008) On the discrepancy of recent European ground-motion observations and predictions from empirical models: Analysis of KiK-net accelerometric data and point-sources stochastic simulations *Bulletin of the Seismological Society of America* 98:2244-2261
- Crowley H et al. (2019) *Methods for Estimating Site Effects in Risk Assessments* vol Deliverable 26.4, Final edn. Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe (SERA),
- Dawood HM, Rodriguez-Marek A, Bayless J, Goulet C, Thompson E (2016) A Flatfile for the KiK-net Database Processed Using an Automated Protocol *Earthquake Spectra* 32:1281-1302
- Donahue JL, Abrahamson NA (2014) Simulation-based hanging wall effects *Earthquake Spectra* 30:1269-1284
- Douglas J (2010) Consistency of ground-motion predictions from the past four decades *Bulletin of Earthquake Engineering* 8:1515-1526
- Douglas J (2018) Calibrating the backbone approach for the development of earthquake ground motion models *Best Practice in Physics-based Fault Rupture Models for Seismic Hazard Assessment of Nuclear Installations: Issues and Challenges Towards Full Seismic Risk Analysis*
- Douglas J, Jousset P (2011) Modeling the difference in ground-motion magnitude-scaling in small and large earthquakes *Seismological Research Letters* 82:504-508
- Dujardin A, Causse M, Berge-Thierry C, Hollender F (2018) Radiation Patterns Control the Near-Source Ground-Motion Saturation Effect *Bulletin of the Seismological Society of America* doi:10.1785/0120180076
- Edwards B, Rietbrock A, Bommer JJ, Baptie B (2008) The acquisition of source, path, and site effects from microearthquake recordings using Q tomography: Application to the United Kingdom *Bulletin of the Seismological Society of America* 98:1915-1935
- Fukushima Y (1996) Scaling relations for strong ground motion prediction models with M 2 terms *Bulletin of the Seismological Society of America* 86:329-336
- Garnier S (2019) Viridis: Default color maps from “matplotlib”. 2018 URL <https://github.com/sjmgarnier/viridis> R package version 03 4:27
- Jacoby WG (2000) Loess:: a nonparametric, graphical tool for depicting relationships between variables *Electoral Studies* 19:577-613
- Kahle D, Wickham H, Kahle MD (2019) Package ‘ggmap’

- Kamai R, Abrahamson N, Graves R (2014) Adding fling effects to processed ground-motion time histories *Bulletin of the Seismological Society of America* 104:1914-1929
- Koller M (2016) *robustlmm*: An R package for robust estimation of linear mixed-effects models *Journal of statistical software* 75:1-24
- Kotha SR, Bindi D, Cotton F (2016) Partially non-ergodic region specific GMPE for Europe and Middle-East *Bulletin of Earthquake Engineering* 14:1245-1263
- Kotha SR, Cotton F, Bindi D (2019) Empirical models of shear-wave radiation pattern derived from large datasets of ground-shaking observations *Scientific reports* 9
- Kotha SR, Weatherill G, Bindi D, Cotton F (2020) A Regionally Adaptable Ground-Motion Model for Shallow Crustal Earthquakes in Europe *Bulletin of Earthquake Engineering* 18:4091–4125 doi: 10.1007/s10518-020-00869-1
- Lanzano G, Luzi L, Pacor F, Felicetta C, Puglia R, Sgobba S, D'Amico M (2019) A Revised Ground-Motion Prediction Model for Shallow Crustal Earthquakes in Italy *Bulletin of the Seismological Society of America* 109:525-540
- Lanzano G et al. (2018) The pan-European engineering strong motion (ESM) flatfile: compilation criteria and data statistics *Bulletin of Earthquake Engineering*:1-22
- Okada Y, Kasahara K, Hori S, Obara K, Sekiguchi S, Fujiwara H, Yamamoto A (2004) Recent progress of seismic observation networks in Japan—Hi-net, F-net, K-NET and KiK-net— *Earth, Planets and Space* 56:xv-xxviii
- Oth A, Bindi D, Parolai S, Di Giacomo D (2011) Spectral analysis of K-NET and KiK-net data in Japan, Part II: On attenuation characteristics, source spectra, and site response of borehole and surface stations *Bulletin of the Seismological Society of America* 101:667-687
- Pacor F et al. (2018) NESS1: A worldwide collection of strong-motion data to investigate near-source effects *Seismological Research Letters* 89:2299-2313
- Paolucci R, Smerzini C, Vanini M (2021) BB-SPEEDset: a dataset of broadband near-source earthquake ground motions from physics-based simulations *Bulletin of the Seismological Society of America*
- Radiguet M, Cotton F, Manighetti I, Campillo M, Douglas J (2009) Dependency of near-field ground motions on the structural maturity of the ruptured faults *Bulletin of the Seismological Society of America* 99:2572-2581
- Schmedes J, Archuleta RJ (2008) Near-source ground motion along strike-slip faults: Insights into magnitude saturation of PGV and PGA *Bulletin of the Seismological Society of America* 98:2278-2290
- Sgobba S, Lanzano G, Pacor F, Felicetta C (2021a) An Empirical Model to Account for Spectral Amplification of Pulse-Like Ground Motion Records *Geosciences* 11:15
- Sgobba S, Pacor F, Felicetta C, Lanzano G, D'Amico M, Russo E, Luzi L (2021b) NEar-Source Strong-motion flatfile (NESS), version 2.0 (Version 2.0). doi:<https://doi.org/10.13127/NESS.2.0>
- Spudich P, Bayless, J.R., Baker, J.W., Chiou, B. S.-J., Rowshandel, B., Shahi, S. K., and Somerville, P., (2013) Final Report of the NGA-West2 Directivity Working Group. Pacific Earthquake Engineering Research Center, University of California, Berkley
- Team RC (2013) R foundation for statistical computing Vienna, Austria 3

- von Specht S, Cotton F (2020) A Link between Machine Learning and Optimization in Ground-Motion Model Development: Weighted Mixed-Effects Regression with Data-Driven Probabilistic Earthquake Classification Bulletin of the Seismological Society of America 110:2777-2800
- Wald DJ, Allen TI (2007) Topographic slope as a proxy for seismic site conditions and amplification Bulletin of the Seismological Society of America 97:1379-1395
- Weatherill G, Crowley H, Lemoine A, Roullé A, Tourlière B, Kotha SR, Cotton F (2020a) Modelling Seismic Site Response at Regional Scale for the 2020 European Seismic Risk Model (ESRM20) Bulletin of Earthquake Engineering *in-preparation*
- Weatherill G, Kotha SR, Cotton F (2020b) A Regionally-Adaptable "Scaled-Backbone" Ground Motion Logic Tree for Shallow Seismicity in Europe: Application in the 2020 European Seismic Hazard Model Bulletin of Earthquake Engineering
- Wickham H, Chang W, Henry L, Pedersen T, Takahashi K, Wilke C, Woo K (2019a) R Package 'ggplot2' v. 3.1.1 Cran R
- Wickham H, François R, Henry L, Müller K (2019b) dplyr: A Grammar of Data Manipulation. R package version 0.8.0.1. ed,
- Woessner J et al. (2015) The 2013 European Seismic Hazard Model: key components and results Bulletin of Earthquake Engineering 13:3553-3596 doi:10.1007/s10518-015-9795-1
- Zeileis A, Hothorn T (2002) Diagnostic checking in regression relationships

SUPPLEMENTARY FIGURES

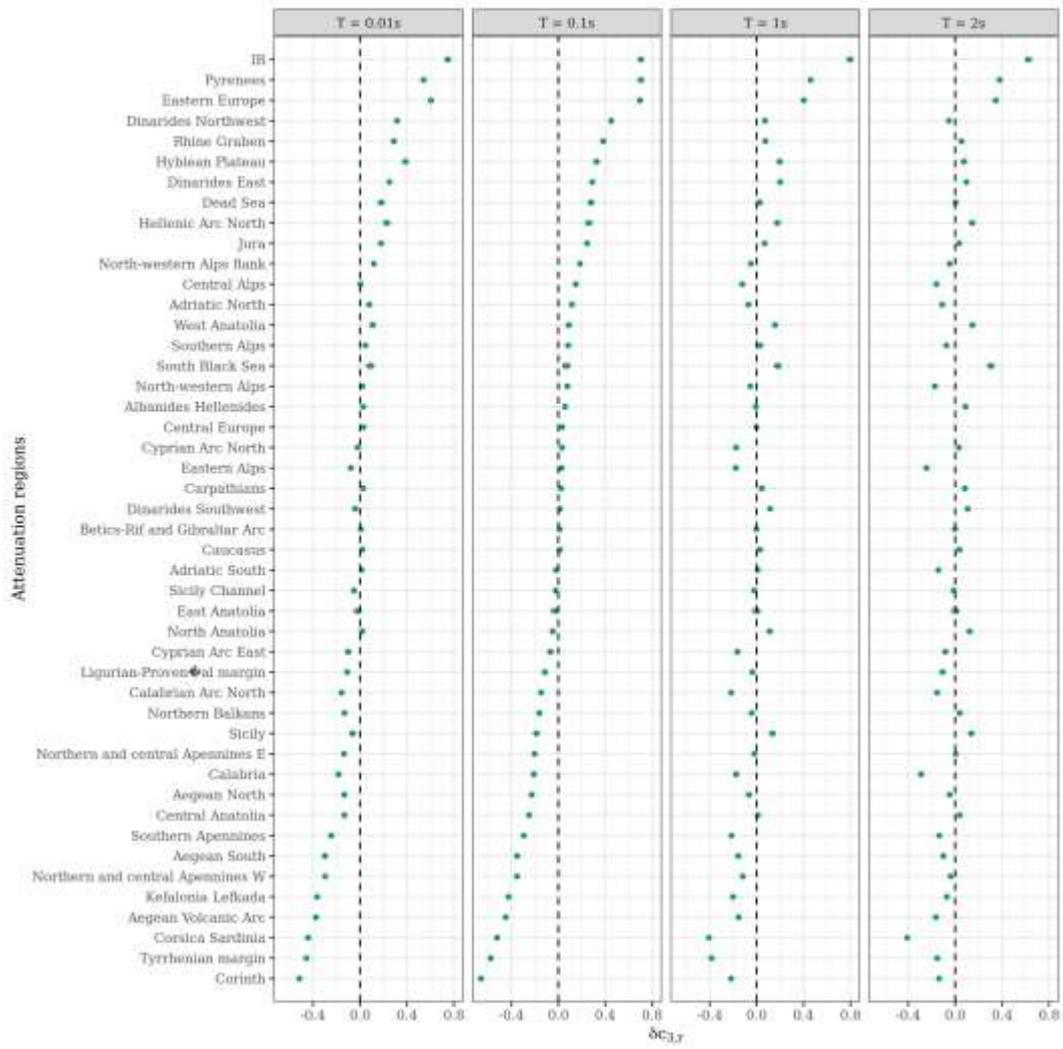


Fig.S1 Between-region attenuation random-effects $\delta c_{3,r}$ at $T=0.01, 0.1, 1, 2s$ when using $M_h = 5.5, 5.7, 6.2, 6.5$ in GMM regressions. Note that the estimates from the four trial GMMs are identical, and thus the markers overlap and appear to be estimates only from a single model

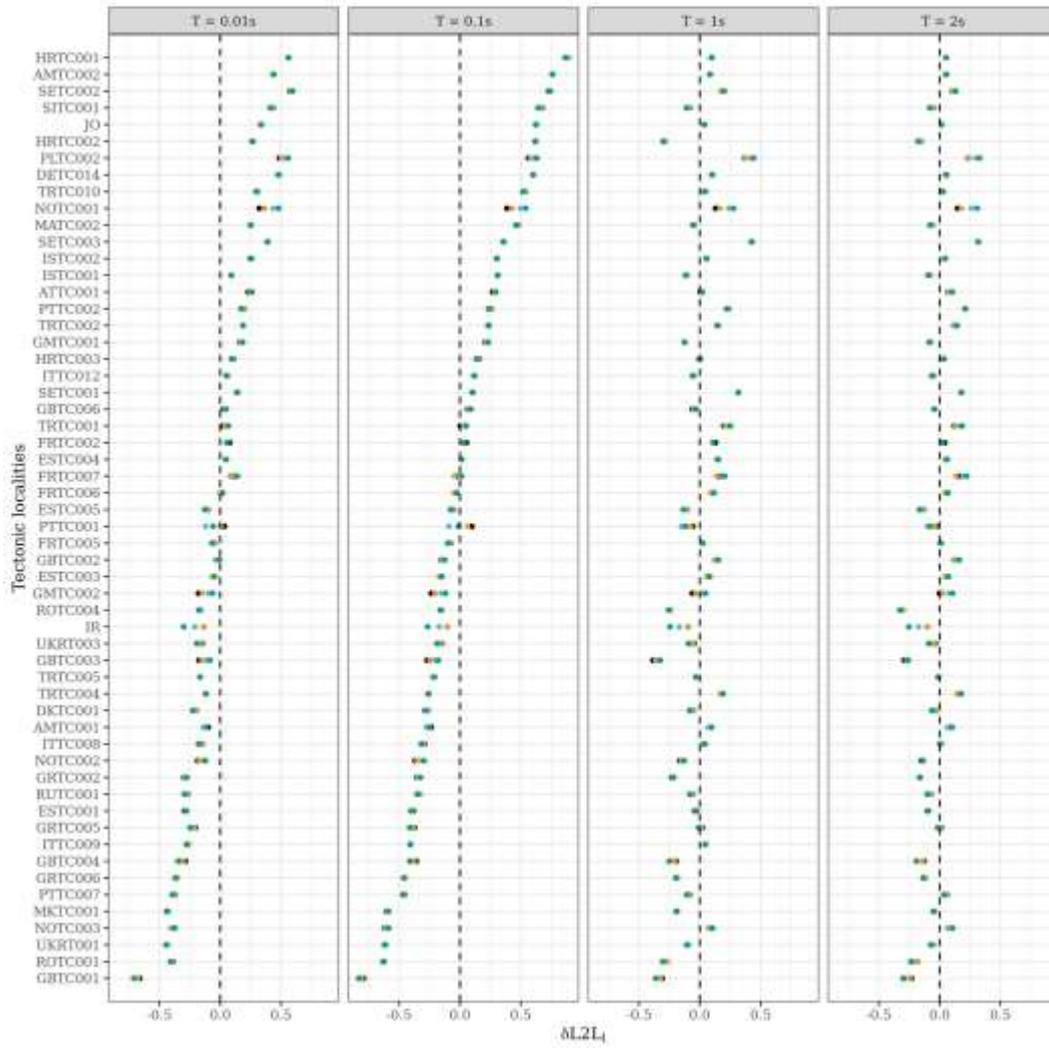


Fig.S2 Between-locality source random-effects $\delta L2L_l$ at $T = 0.01, 0.1, 1, 2s$ when using $M_h = 5.5, 5.7, 6.2, 6.5$ in GMM regressions. Note that the estimates from the four trial GMMs are identical, and thus the markers overlap and appear to be estimates only from a single model

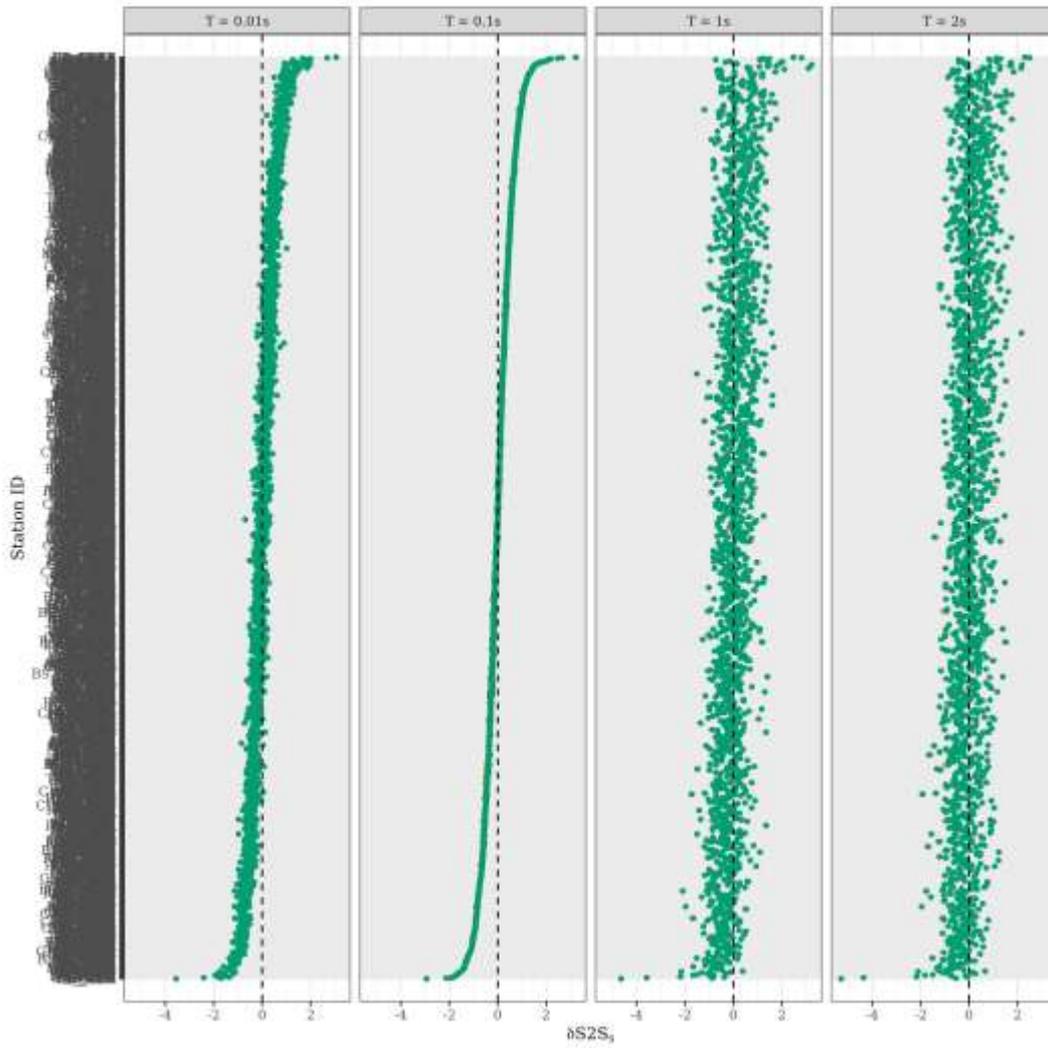


Fig.S3 Between-site random-effects $\delta S2 S_s$ at $T= 0.01, 0.1, 1, 2s$ when using $M_h = 5.5, 5.7, 6.2, 6.5$ in GMM regressions. Note that the estimates from the four trial GMMs are identical, and thus the markers overlap and appear to be estimates only from a single model

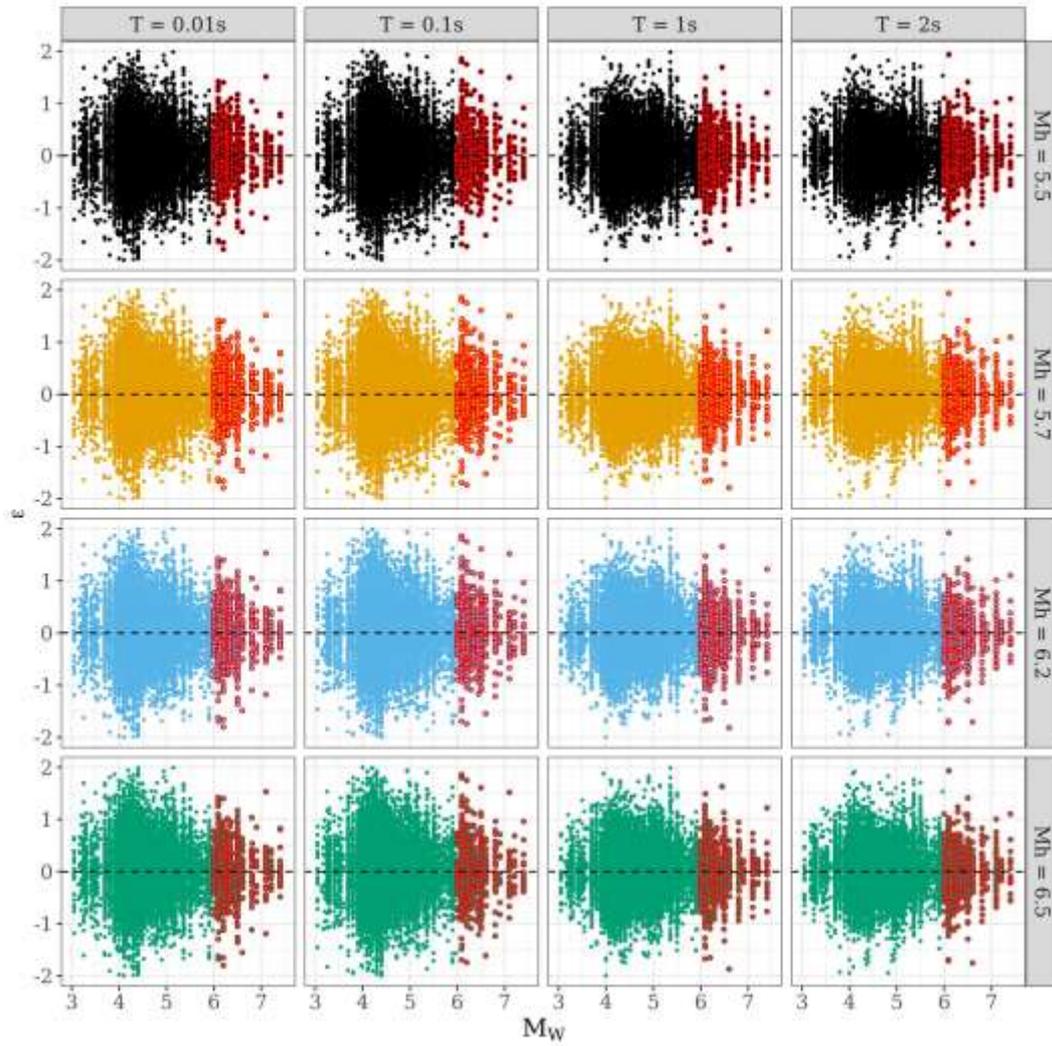


Fig.S4 Trends of aleatory residuals ε against M_W at $T = 0.01, 0.1, 1, 2s$ when using $M_h = 5.5, 5.7, 6.2, 6.5$ in GMM regressions. Residuals corresponding to events with $M_W \geq 6$ are highlighted in red for visual clarity

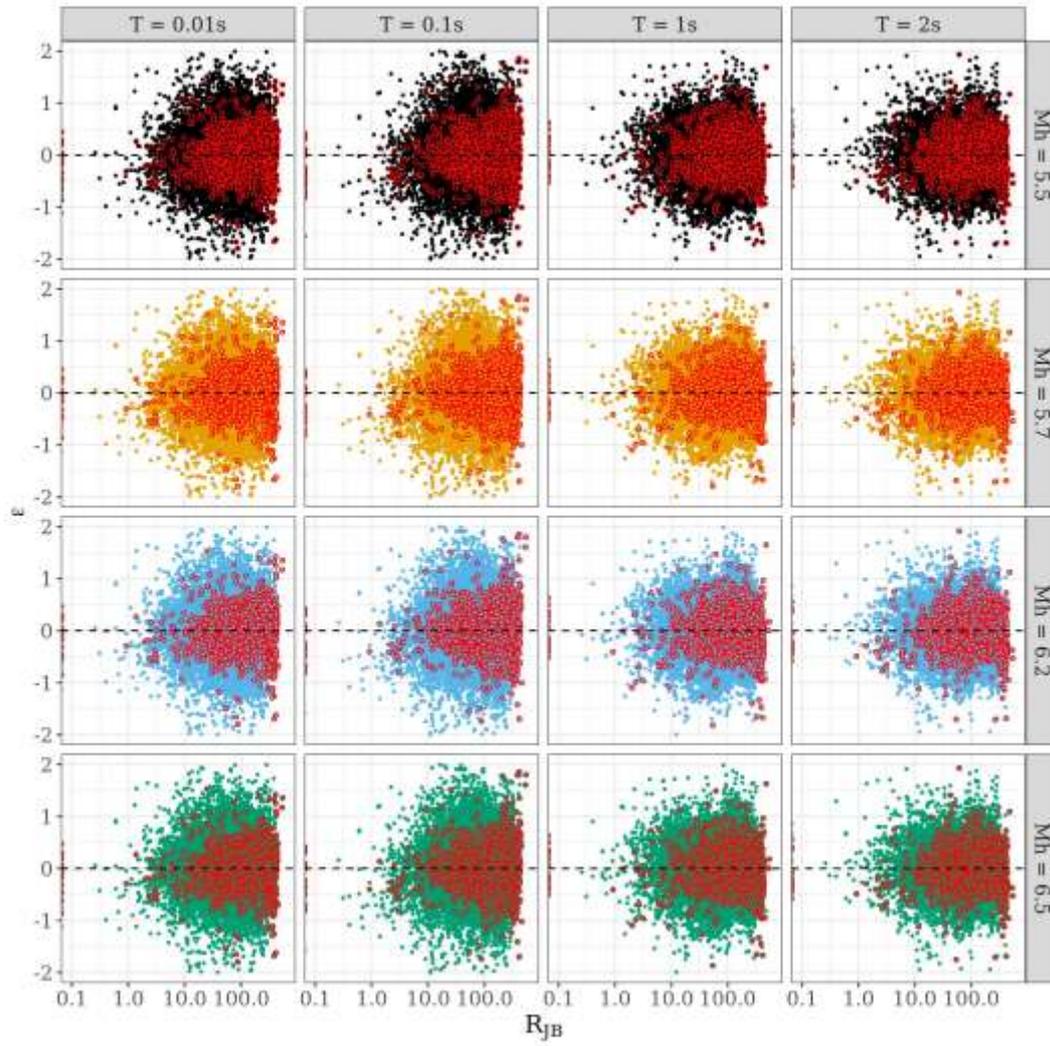


Fig.S5 Trends of aleatory residuals ϵ against R_{JB} at $T = 0.01, 0.1, 1, 2s$ when using $M_h = 5.5, 5.7, 6.2, 6.5$ in GMM regressions. Residuals corresponding to events with $M_w \geq 6$ are highlighted in red for visual clarity

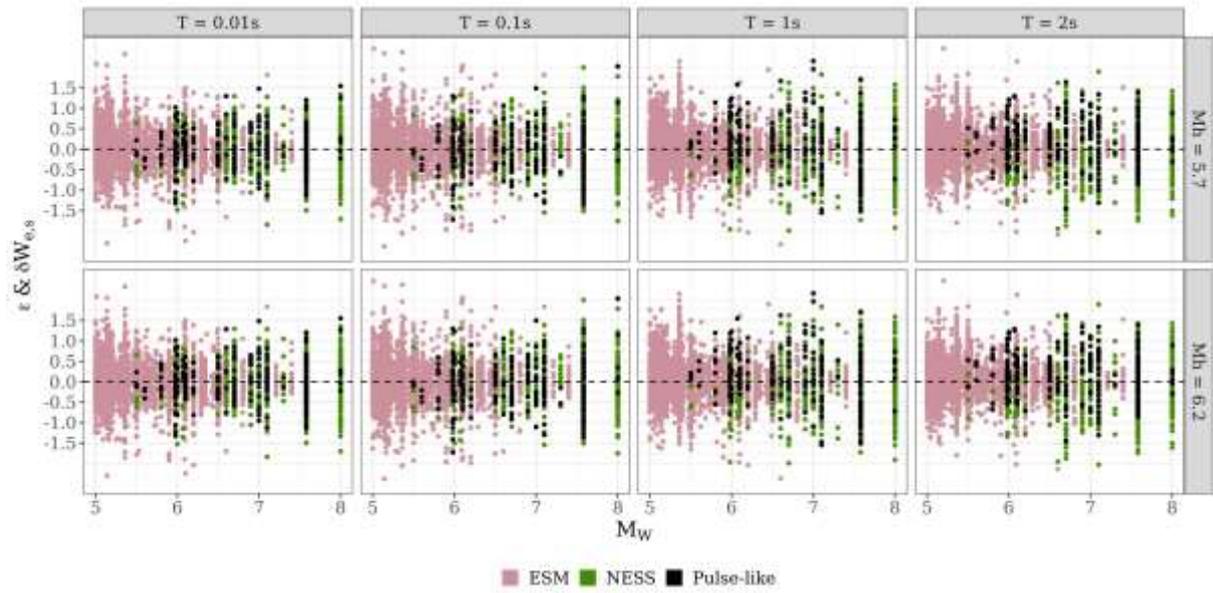


Fig.S6 Within-event residuals versus M_w of the records in ESM and NESS dataset at $T= 0.01, 0.1, 1s$ when $M_h = 5.7, 6.2$. Black markers indicate those records with reported pulse-like effects

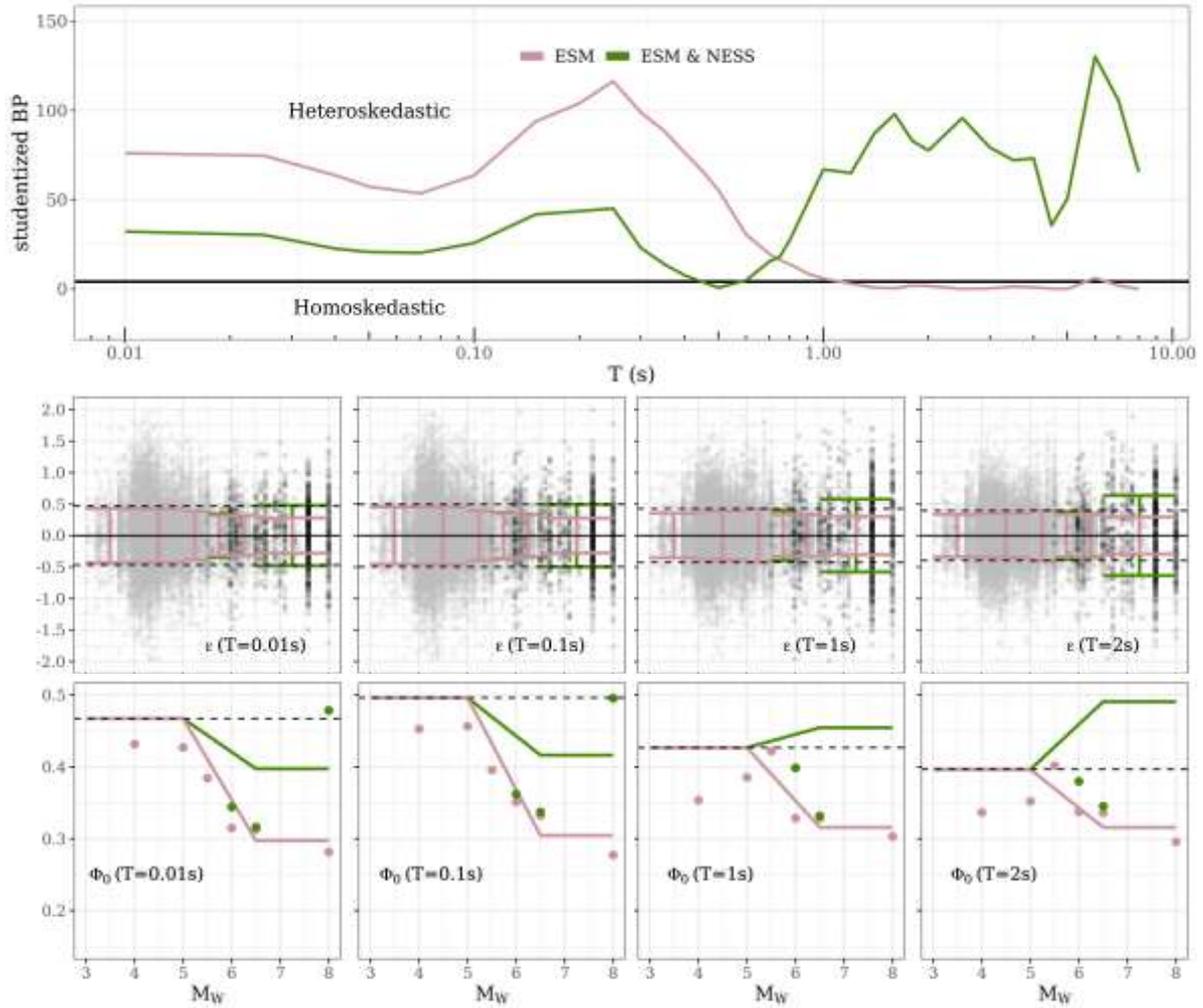


Fig.S7 Heteroscedasticity of residual (ϵ) variability with M_W of the events in ESM and ESM&NESS datasets: Top row panel shows the studentised Breush-Pagan statistic, which is a measure of heteroscedasticity of residual variability with M_W for the $T = 0.01 - 8s$. Middle row panels show the $\epsilon(T)$ versus M_W for $T = 0.01, 0.1, 1, 2s$ when $M_h = 5.7, 6.2$. The error-bars show the Median Absolute Deviance (sample MAD) of ϵ within bins $M_W \in (3, 4], (4, 5], (5, 5.5], (5.5, 6], (6, 6.5], (6.5, 8]$. Bottom row panels show the sample MAD (markers) within the M_W bins and the proposed heteroskedastic Φ_0 models