Bindi, D., Cotton, F., Kotha, S., 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*, 5, pp. 1201—1218.

DOI: http://doi.org/10.1007/s10950-017-9661-5

# Application-driven Ground Motion Prediction Equation for Seismic Hazard Assessments in non-Cratonic Moderate-Seismicity Areas

**D. Bindi, F. Cotton, S. R. Kotha, C. Bosse, D. Stromeyer and G. Grünthal**

*German Research Center for Geosciences GFZ, Potsdam, Germany*
*bindi@gfz-potsdam.de*

## Abstract

We present a Ground Motion Prediction Equation (GMPE) for probabilistic seismic hazard assessments (PSHA) in low-to-moderate seismicity areas, such as Germany. Starting from the NGA-west 2 flat-file (Ancheta et al., 2014), we develop a model tailored to the hazard application in terms of data selection and implemented functional form. In light of such hazard application, the GMPE is derived for hypocentral distance (along with the Joyner-Boore one), selecting recordings at sites with vs30≥360 m/s, distances within 300 km, magnitudes in the range 3 to 8 (being 7.4 the maximum magnitude for the PSHA in the target area). Moreover, the complexity of the considered functional form is reflecting the availability of information in the target area. The median predictions are compared with those from the NGA-west 2 models and with one recent European model, using the Sammon's map constructed for different scenarios. Despite the simplification in the functional form, the assessed epistemic uncertainty in the GMPE median is of the order of those affecting the NGA-west2 models for the magnitude range of interest of the hazard application. On the other hand, the simplification of the functional form led to an increment of the apparent aleatory variability. In conclusion, the GMPE developed in this study is tailored to the needs for applications in low-to-moderate seismic areas and for short return periods (e.g., 475 years); its application in studies where the hazard is involving magnitudes above 7.4 and for long return periods is not advised.

**Key words**: Ground Motion Prediction Equations; moderate seismicity region; NGA-West2

## Introduction

Since 2003, the Pacific Earthquake Engineering Research Center (PEER) is conducting a large research program to develop the next generation of ground motion prediction equations (GMPEs) for shallow crustal earthquakes in active tectonic regions. The second phase of this project (called NGA-West2) concluded in 2014 and provided important results, including a strong motion database of recorded ground motions (Ancheta et al. 2014) and a set of peer-reviewed GMPEs (Abrahamson et al. 2014).

Several recent hazard projects have shown that the models developed by NGA-west projects may be of interest not only for active regions but also for non-cratonic and lower seismicity regions. NGA-west ground-motion models have been selected as part of ground-motion logic tree to compute recent probabilistic seismic hazard assessments (PSHA) in Europe (Delavaud et al., 2012), Switzerland (Edwards et al., 2016) and Germany (Grünthal et al., 2017). These models are also widely used in regions where active faults have not been identified for most seismic sources , which means that rock hazard computations have to be conducted for distributed seismicity (area sources or zoneless approach) without taking into account the refinements introduced by recently developed GMPEs (directivity and hanging wall effects, non-linear site effects, basin effects).

These projects show that existing NGA-west2 GMPEs, despite their high quality, are then not fully fitting the needs of regional hazard computations in moderate seismicity areas. The main problems encountered in the application of the NGA-west2 models in low-to-moderate seismic areas are the following:

1) Modern GMPEs use definitions of the source-to-site distance that reflect the dimensions of the fault rupture for larger earthquakes better than point-source measures relative to the epicenter or hypocenter. This is a positive development since it more realistically reflects the fact that energy is released from the crust around the entire fault rupture during a large earthquake. However, seismic source configurations defined for PSHA in low-to-moderate seismicity areas almost invariably include areas of distributed point-source seismicity. Point-source simulations can be enhanced to include simulations of virtual extended ruptures. These adaptations are computationally demanding and not easily implemented. As suggested by Bommer and Akkar (2012), there is a need to compute pairs of equations, one using an extended-source distance metric, the other a point-source measure. To our knowledge, such pairs of equations have not been performed by the NGA-west 2 project.

2) A problem often encountered in the application of the NGA-west 2 GMPEs based on complex functional forms is related to the availability of suitable metadata in the target

region. In low to moderate seismicity regions the source and site characterizations are generally not as detailed as in the data set used to derive the GMPE (host region). In such cases, the GMPEs are applied in simplified forms, where one or several variables (e.g., basin depth, hanging wall foot wall effects) are constrained to default values. This operation should be accompanied by either a proper handling of the epistemic uncertainty introduced when fixing some variables, or by propagating the uncertainty to the aleatory component. Both choices imply some additional work and expert decisions.

3) The hazard computed at a given location depends on both the seismic source model and on the ground motion model. Hazard computations in low-to-moderate seismicity areas are particularly dependent on the GMPE magnitude scaling around magnitude 5.5-6. Some NGA-west 2 models have chosen functional forms with a magnitude hinge around M=5.5. Such a choice has a low impact on hazard computations in high seismicity regions but a larger one in moderate seismicity regions. This application-driven practical issue motivates the development of functional forms adapted to moderate seismicity areas.

To overcome these problems, we derived a new GMPE using the high quality PEER flat file but tuning the complexity of the model to the information available in moderate seismicity regions. This development is also motivated by the needs of hazard computations in low-to-moderate seismicity areas: pairs of equations (one using an extended-source distance metric, the other a point- source measure), focus on stiff soil and rock site conditions, specific magnitude-scaling analysis in the magnitude range 5.5-6.

The use of simpler GMPEs (point source distance metric, lower number of input parameters) has however two main drawbacks, which will be analyzed in the following:

- The aleatory variability of GMPEs ($\sigma$) has a strong impact on the results of PSHA at long return periods. The $\sigma$ values are indeed an estimate of the apparent aleatory variability since they are evaluated with respect to the chosen model. The use of simpler models implies larger $\sigma$ and the impact of such increase has to be carefully evaluated.
- GMPEs are used as part of a logic tree or selected as a backbone equation. In both cases, it is important to evaluate the GMPEs epistemic uncertainty of the median (particularly in the magnitude range of interest) and the proximity of the model with other published models, e.g., using Sammon's maps (Scherbaum et al., 2010).

This article is organized as follow. First, we discuss the motivations that led us to derive a GMPE tailored to our specific hazard application (Grünthal et al., 2017). Then, after describing the functional form and the data considered for the GMPE development, we discuss the epistemic uncertainty in the median and the aleatory variability. Finally, the

comparisons of the median predictions with NGA-West2 GMPEs are presented in terms of Sammon's map and Trellis charts.

### Towards an application driven GMPE: Hazard assessment in a moderate seismicity area (Germany) and associated needs

This study is part of the German Hazard map project accomplished on behalf of the Deutsches Institut für Bautechnik (DIBt; The Centre of Competence in Civil Engineering). The new version of the national PSHA should predict uniform hazard spectra (UHS) for any site within Germany, hazard maps for spectral accelerations, peak ground accelerations, and deaggregations for the hazard levels of 10%, 5% and 2% exceedance probability within 50 years (Grünthal et al., 2017). All hazard calculations had to be performed for vs30 =800 m/s, where vs30 is the time-averaged shear-wave velocity of the top 30 m. Induced events are excluded from the study. Epistemic uncertainties have been explored both for the seismicity and ground-motion model part. The tectonic context of Germany (Grünthal et al., 2017) is complex with active structural elements mainly along the chain of the Rhine Graben up to rather stable parts towards the north and northeast. Because of this complexity, GMPEs logic tree used in past seismic hazard studies for this part of the West European Platform (e.g. Delavaud et al., 2012) included equations for active crustal regions (ACR).

The use of ACR models calibrated to the NGA-west database was also motivated by recent stochastic models and GMPE testing performed in Western Europe. In France, Beauval et al. (2012) tested several GMPEs: the NGA-west 1 Abrahamson and Silva (2008) model was ranked as one of the best models. Drouet and Cotton (2015) developed and tested a new stochastic model based on data recorded in the French Alps and their resulting model is consistent with GMPEs derived for active crustal regions (e.g. Boore and Atkinson 2008). In Switzerland, Edwards and Fäh (2013) and Cauzzi et al. (2015a) proposed stochastic ground-motion models of the Swiss Foreland and the Swiss Alpine region. They also showed that Swiss stochastic ground-motion models are broadly consistent with the NGA-west-1 Chou and Young (2008) model.

The GMPEs logic tree implemented for updating the seismic hazard in Germany is composed by three main branches, each of them including one or more models derived from different data sets. In particular, while the first branch includes models derived for Europe using the RESORCE data set (Akkar et al., 2014), the other two branches are relevant to GMPEs calibrated considering global ACR data sets (Grünthal et al., 2017). The model of Cauzzi et al. (2014b), calibrated over a data set mostly populated by

Japanese earthquakes, is considered for one of the two branches while the other was reserved to a GMPE derived from the NGA2-west flat file (Ancheta et al, 2014). The seismicity model used to derive the German hazard map is based mainly on area sources. Disaggregation analyses have been performed over magnitude and distance scenarios for preliminary hazard assessments for several representatives sites (Figure 1). The hazard (for the return period RP=475 years) is controlled mainly by earthquakes of moderate magnitude (M<5.5) at distance below 25 km. The disaggregation results suggest a significant impact of the functional form chosen to define the magnitude scaling between M=5 and M=6 (Figure 1). For example, the kink in the magnitude scaling of the BSSA14 model (Boore et al., 2014) around the hinge magnitude Mw=5.5 increases the relative contribute to hazard of scenarios for magnitude between 5 and 6 and distances around 20 km. In the following, we use the PEER-NGA2 flat file to derive a new GMPE whose functional form is selected for the specific hazard study of interest.

### NGA-west2 data and GMPE development

To develop a global GMPE for logic tree implemented in the hazard assessment for Germany, we consider the PEER-NGA2 flat file (Ancheta et al., 2014). In particular, starting from the Campbell and Bozorgnia (2014) data selection, further selection criteria related to the specific application are applied to the data set. We selected only recordings for distances less than 300 km and stations with vs30 computed or estimated from shear wave measurements (i.e., vs30 code either 0 or 1, Ancheta et al., 2014). Moreover, since the hazard application is performed for rock condition, only vs30 ≥360 m/s are selected in order to limit possible bias in the median due to not properly modelled site effects for soft sites (i.e., neglecting non-linear effects). Figure 2 shows the scatterplot for the distribution of magnitude with hypocentral distance. The selected data set is composed by 4692 recordings from 242 earthquakes and 1025 stations. The [16th, 50th, 84th] percentiles of the magnitude, Joyner-Boore and hypocentral distributions are [3.7, 4.3, 6.7], [43.8, 111.9, 209.5] km, and [49.7, 117.2, 217.8] km, respectively. In particular, for magnitudes above 4.5, hypocentral distances below 10 km are almost not sampled. Regarding the selected stations, the percentiles for vs30 are [393, 511, 786] m/s, while the number of recordings for Eurocode8 class A and B are 689 and 4013, respectively (Figure 2).

Regarding the functional form, the involved explanatory variables should reflect the metadata availability in the target region and the requirements of the specific hazard assessment. For example, since the hazard is computed for rock site conditions (vs30=800 m/s) and in a low-to-moderate seismicity area, only the linear site effect is

implemented, without correction for basin effects. Also, since the hazard assessment is based on source area models, the hypocentral distance is preferred, although the model is also derived for the Joyner-Boore distance, following (Bommer and Akkar, 2012). Moreover, extended source effects such as hanging-wall/foot-wall terms are not modeled. For the aforementioned reasons, the following functional form is considered:

$$\ln Y = e_1 + F_D(R, M) + F_M(M) + F_S \quad (1)$$

Where the distance $F_D$ and magnitude $F_M$ functions are given by:

$$F_D(R_{JB}, M) = \left[c_1 + c_2(M - M_{ref})\right] ln\left(\sqrt{{R_{JB}}^2 + h^2}/R_{ref}\right) + c_3\left(\sqrt{{R_{JB}}^2 + h^2} - R_{ref}\right) \quad (2)$$

$$F_D(R_{hypo}, M) = \left[c_1 + c_2(M - M_{ref})\right] ln(R_{hypo}/R_{ref}) - c_3(R_{hypo} - R_{ref}) \quad (3)$$

$$F_M(M) = \begin{cases} b_1(M - M_{ref}) + b_2(M - M_{ref})^2, & M < M_h \\ b_3(M - M_h) + b_1(M_h - M_{ref}) + b_2(M_h - M_{ref})^2, & otherwise \end{cases} \quad (4)$$

$$F_S = sA\ ln(V_{S30}/800) \quad (5)$$

In equations (2) and (3), the Joyner-Boore ($R_{JB}$) and hypocentral ($R_{hypo}$) distances are considered, respectively. The reference distance $R_{ref}$ have be set equal to 1 km, the reference magnitude $M_{ref}$ to 4.5 (i.e., close to the 50$^{th}$ percentile of the cumulative number of recordings versus magnitude). The hinge magnitude $M_h$ introduced to handle the saturation in the magnitude scaling, is set equal to 6.5, that is, slightly above the values suggested by data (6-6.2), to move the kink in the magnitude scaling at a magnitude larger than those controlling the hazard (Figure 1). After preliminary tests, the style of faulting term is not considered because not justified in term of bias-variance trade-off, using the AIC parameter (Akaike, 1973).

The regression is performed using a mixed effect approach (Abrahamson and Youngs, 1992; Bates et al., 2014), accounting for the between-event residuals as random effect on the offset depending on the earthquake grouping level. The models are calibrated for 5%-damped pseudo-acceleration response spectra, considering 90 periods ranging between 0.01 and 4 s.

## Results

In the following paragraphs, we discuss the results in terms of: fixed effects, epistemic uncertainty in the median and aleatory variability.

### Fixed effects (median model)

The obtained coefficients are listed in Tables 1 and 2 included in the Electronic Supplements. Figure 3 shows the variability with period of the coefficients obtained for the GMPE implementing the Joyner-Boore distance. The trends of the parameters with periods determine the scaling of the model with respect to the explanatory variables. For example, the decrease (in absolute value) of coefficient $c_3$ with period reflects the decrease of the attenuation proportional to distance, sometimes referred to as anelastic attenuation term, although this interpretation is strictly valid only in the Fourier domain, e.g. Bora et al., (2016), which almost vanishes above 2 s. The effect of this term is largest at 0.1 s and, for periods smaller than 0.03s, it is constant. The coefficient $c_2$ in equations (2) and (3) controls the magnitude dependence of the attenuation with the logarithm of distance (sometimes referred to as geometrical spreading, in analogy with a model for Fourier). As shown in Figure 3, $c_2$ is positive and almost constant. Its effect on lnY depends on the sign of $(M-M_{ref})$: for $M>M_{ref}$, the $c_2$ term reduces the attenuation with distance while, for $M<M_{ref}$, it increases the attenuation with distance. Then, the distance attenuation is more significant for small magnitudes. This effect is more evident for short periods (below 0.05 s) and almost vanishes between 0.1 and 0.6 s. Another example is the site coefficient sA (equation 5). For velocity lower than 800 m/s, the term ln(vs30/800) is negative. Then, the trend of sA in Figure 3 implies that the site amplification effects are larger between 0.03 and 0.1 s for velocities larger than 800 m/s and between 0.2 and 1 s for velocities smaller than 800 m/s. Since the different terms in equations (2) through (5) could be affected by mutual trade-offs, Figure 4 shows the overall dependencies of the predictions on magnitude and distance, i.e., by grouping all terms depending on these explanatory variables. In the top panel, the period dependence of the term $F_M$ given by equation 4 is shown for two different magnitudes (i.e., 4 and 6). Since also the $c_2$ term includes the magnitude, the plot is repeated for two different distances (continuous line for R=30 km; dashed lines for R= 100 km). The dependence on the source recalls the shape of the response spectra, with a sharper increase with decreasing period for smaller magnitude. For short periods, the curve flattens and it is almost independent on the frequency of the oscillator. The dependence on distance is weak, that is the role of the $c_2$ term with respect the terms with $b_1$, $b_2$ and $b_3$, and only appreciable for short periods (see the discussion above about $c_2$). The

period dependence of $F_D$ (equation 3) is shown in the bottom panel of Figure 4, for four scenarios defined by M6 and M4 at 30 and 100 km. By comparing the curves for R=30 and 100 km, we see that the overall effect of $c_1$ is to scale the prediction. For a given distance, the impact of the magnitude (through the term with $c_2$) becomes significant for periods below 0.15 s, where large magnitude are less attenuated than the small ones, as previously discussed for the period dependence of $c_2$.

## Epistemic variability in the median

The variance-covariance matrix of the model quantifies the uncertainties of coefficients (diagonal elements) and their mutual trade-offs (off-diagonal elements). Following (Al-Atik and Youngs, 2014), the variance-covariance matrix and the matrix of the gradients of the model with respect to the coefficients can be used to assess the epistemic uncertainty in the median model (see Al-Atik and Youngs, 2014, for a detailed discussion of the methodology):

$$var[\overline{lnY}]_{x_0} = J_0^T \left[ varCov_{x_i} \right] J_0 \qquad (6)$$

where the Jacobian matrix *J* is evaluated in the predictive location $x_0$ and the variance-Covariance matricx *varCov* is evaluated in the data points $x_i$ used to develop the model. Figure 5 shows the ingredients to assess the variance of ln(SA) at 0.1s. For graphical reasons, the variances of the model coefficients and the correlation matrix are shown instead of *varCov*. The standard deviation $\sigma_\mu$ of the median (i.e., the square root of the left hand side term in equation 6) quantifies the epistemic uncertainty in the median due to the combined effects of limited data availability and implemented functional form.

The largest variances are those of $e_1$ and $b_3$ (Figure 5, panel a), while the largest trade-off occurs between $c_1$ and $c_3$, between $e_1$ and $c_1$, between $b_1$ and $c_2$, and between $e_1$ and $c_3$ (Figure 5, panel b), reflecting of the trade-off between the source and attenuation terms. Following equation (6), the uncertainty on the median is controlled by the product of these terms with the values of the gradient of the model with respect to the coefficients, evaluated for the predictive scenarios. Figure 5, panel (c), shows the gradients for different magnitude and distance combinations, and for vs30=600 m/s. It is worth noting that, in the case of hypocentral model, the model is linear with respect to the coefficients and therefore the gradients are period independent. The dependence of $\sigma_\mu$ on period eventually arises from the variance-covariance matrix. On the contrary, for

the Joyner-Boore model, the derivative of $F_D$ in equation 2 with respect to the coefficients $c_1$, $c_2$, $c_3$, and $h$ depends on the model coefficients, making the gradients period dependent.

Figure 6 shows $\sigma_\mu$ for different magnitude and distance scenarios. Panels (a) and (b) show that $\sigma_\mu$ for magnitude 6 and vs30=800 m/s (i.e., a typical scenario of interest for our application) is weakly dependent on distance and is of the order of $\sigma_\mu$ modelled for NGA2 GMPEs (Al-Atik and Youngs, 2014). For distances shorter than 10 km, $\sigma_\mu$ slightly increases with decreasing distance, reflecting the fact that short distances are weakly constrained by data. In the magnitude range from 4 to 6 (and for a distance of 25 km and vs30=800 m/s), $\sigma_\mu$ is weakly dependent on magnitude, while it increases outside this range, in particular above magnitude 7. The bump around magnitude 6.5 is a consequence of introducing the hinge magnitude for handling the saturation with magnitude. The overall dependence on period is weak, as shown in panel (c). For magnitude 6, $\sigma_\mu$ is between the models for normal and reverse faulting derived for NGA2 while, as already shown in panel (b), larger values are obtained for magnitude 8. The increases of $\sigma_\mu$ for periods longer than 2s is stronger for small magnitudes, which less constrain the ground motion at low frequencies. Regarding the dependence of $\sigma_\mu$ on vs30 (here not shown), it is negligible. Finally, the overall contribution of $\sigma_\mu$ to the mean response spectra uncertainty are shown in panel (d), for different magnitudes, at a distance of 25 km and for vs30=800 m/s.

## Aleatory variability

Figure 7 shows the aleatory variability in terms of period dependent $\tau$ (between-event), $\phi$ (within-event) and $\sigma$ (total aleatory variability). To provide a term of comparison, Figure 7 also reports the models for BSSA14 (Boore et al., 2014) and RES14 (Bindi et al., 2014). Since BSSA14 is heteroscedastic, its standard deviations are evaluated for magnitudes 4 and 7 (representing the range of main interest for the hazard application in Germany), at a distance of 40 km, and for vs30=800 m/s.

Considering the simplification applied to the functional form, $\sigma$ values larger than those of BSSA14 were expected. Indeed, for short periods, $\sigma$ is very close to the values of BSSA14 for M4 while, for longer periods, it approaches the values of RES14 (which was calibrated for Europe). The main differences among the models are observed for $\tau$. For periods shorter than 0.3 s, $\tau$ is larger than the values of the other models while for longer periods it overlaps to BSSA14 evaluated for magnitude 4. The values for RES14

are smaller, probably as a consequence of a more regional composition, since it is manly composed by earthquakes occurred in Italy and Turkey. Regional differences are also present in the European data but they are mainly affecting the distance scaling and site effects (e.g., Kotha et al., 2016). The largest contribution to $\sigma$ is coming from $\phi$. Below 0.2 s, $\phi$ from this study is similar to RES14 while, for longer periods, it is smaller than RES14 and close to the BSSA14 one evaluated for magnitude 7. In conclusion, the aleatory variability of the derived model is close to the variability of BSSA14 for low magnitude and, since we do not allowed $\sigma$ to be heteroscedastic, larger than the BSSA14 one for large magnitude. The simplifications applied to the functional form mainly affect the source variability for periods below 0.4 s and the record-to-record variability for the longer periods.

## Discussions

In the following, the derived model is discussed in terms of residuals analysis and by comparing the predictions with those from previous models.

### Analysis of residuals

The explanatory power of the models is evaluated through the analysis of the residual distributions. Figure 8 shows the prediction versus distance for PGA and PGV, considering magnitudes 4 and 6.7, and vs30=800 m/s. The predictions are compared to observations selected in ±0.2 range with respect the magnitude used for the prediction, and considering all available stations (circles for stations with vs30 ≥800 m/s, triangles for vs30 <800 m/s). The model captures well the trend and the variability in the data, with perhaps the tendency of overestimating the ground motion for large magnitude at large distances. To quantify the overall agreement between data and predictions, the within and between residuals are shown in Figure 9 against distance and magnitude, respectively. Generally, the within event residuals (left panels) for both selected periods (i.e., 0.1 and 1.0 s) do not show any significant trend with distance, except for a slight tendency to underestimate the spectral acceleration at distance smaller than 20 km for T=0.1 s. Regarding the between event residuals (right panels), they are shown accordingly to their focal mechanism. The data set is dominated by strike-slip (SS) events (181 earthquakes), shown as black circles, while the number of events with normal (NF) and reverse (RF) mechanisms are 16 and 45, respectively. While the between event distribution is unbiased when considered as a whole, the average residuals computed separately for the three style of faulting classes [SS, NF, RF] are [0.007, -0.116, 0.013] at 0.1 s, and [0.005, -0.096, 0.009] at 1s. Therefore, the model

tends to slightly overestimate the spectral acceleration for normal events, although the large standard deviations (of the order of 0.4) make all these values not significantly different from zero.

### Comparison with NGA2 and RES14 models

The median predictions of the model calibrated in this study are compared to four different NGA2 models and to one European model. The models considered are: BSSA14 (Boore et al., 2014), CY14 (Chiou and Youngs, 2014), ASK14 (Abrahamson et al., 2014), CB14 (Campbell and Bozorgnia, 2014), RES14 (Bindi et al., 2014). The Idriss (2014) model is not used here because its application is suggested above magnitude 5 which does not fit with the needs of hazard computations in moderate seismicity areas. Both the hypocentral and Joyner-Boore versions of the model derived in the present study are discussed here. The comparison is performed in terms of Sammon's map (Scherbaum et al., 2010) and Trellis plot. Since the implemented GMPEs use different distance definitions, the comparison is performed for a set of a-priori defined source scenarios (Figure 10). In particular, 7 different fault configurations are adopted varying the dip $\delta$ and the rake $\lambda$ angles (Figure 10 shows the list of the 7 considered combinations). For each fault model, 4 different magnitudes are selected to generate the fault extension (i.e., 4, 5, 6, and 7). Regarding the station locations, they are arranged along a line orthogonal to the strike, located over the hanging wall at distances equal to $\Delta$=[0.01, 0.1, 0.2, 0.5, 1, 2] degrees. For all stations, vs30=800 m/s is used. For each models, those parameters of NGA2 GMPEs like $Z_{tor}$, $Z_{2.5}$, $Z_{10}$, etc., are set equal to default values suggested by the GMPE's authors and no regional attributes are considered. For each source-station combination, the distances required by the GMPEs are computed (i.e., rupture distance $R_{rup}$, Joyner-Boore distance $R_{JB}$, and hypocentral distance $R_{hypo}$, being the latter computed locating the hypocenter in the middle of the fault). In total, 168 source-station combinations are generated, which are used to compile the multi-dimensional vectors for the Sammon's maps. In order to provide a reference in the Sammon's map, the mixture of the 4 considered NGA2 GMPEs is computed with equal weights, and indicated with MIX in Figure 11. Moreover, artificial scaling with distance and magnitude are applied to MIX, to add further reference points in the maps. In particular, M+ and M++ in Figure 11 refer to MIX with added the term $0.25(M-M_{ref})$ and $0.5(M-M_{ref})$, respectively. Similar definitions apply to M- and M--, but the artificial scaling is in this case subtracted. Regarding the distance scaling, R+ and R- correspond to adding $0.25\ln(R_{rup})$ or subtracting $0.25\ln(R_{rup})$ to MIX, respectively. To make it easier the comparison between different maps, we have applied translation and rotation to the

Sammon's maps in order to locate MIX always in the origin of the coordinate system and R+ along the positive x-axis (Figure 11). If necessary, a reflection with respect to the x-axis is finally applied.

The inter-point distances in Figure 11 are an estimate of the GMPEs proximity in predicting similar ground motion levels (Scherbaum et al., 2010), assessed from the feature vectors constructed considering the scenarios in Figure 10. Considering the distances among the NGA-west2 GMPEs shown in Figure 11, the model derived in this study for the Joyner-Boore distance (GERjb) and for periods up to 1s, is close to the NGA-west2 group. On the contrary, GERjb for longer periods (i.e., 4s in Figure 11) and the hypocentral model (GERhypo) for all periods predict significant different ground motion. In particular, GERhypo and GERjb show a stronger magnitude scaling with respect to MIX, being the differences more evident for GERhypo. This is confirmed by the magnitude scaling shown in Figure 12, in particular for short distances and T=0.1s. Regarding the scaling with distance, the Sammon's maps show a stronger decay of Rhypo than MIX (i.e., Rhypo is along the R- direction) at T=0.02 and 0.1 s while the attenuation is weaker (R+ direction) at 0.1 and 4s. The Trellis plots in Figure 13 confirm these overall trends, although with differences depending on specific scenarios (e.g., at T= 0.1s for M6.5, Rhypo shows a stronger attenuation than MIX). The Sammon's maps also provide information for the other models. For example, RES14 shows a weaker magnitude scaling than MIX for 0.02 and 0.1s (see also Figure 11) and a general weaker attenuation with distance than MIX (see also Figure 13). It is worth noting that 4s is beyond the range of applicability of RES14 suggested by the authors.

Figure 14 summarizes the results of the Sammon's map analysis, showing the distance between each considered model and the reference one (i.e., the mixture of the NGA2 models), as function of periods. The predictions from the model derived in this study for the Joyner-Boore distance are close to the NGA-west2 ones, in particular for short periods. For long periods, the Joyner-Boore model is close to CY14 and BSSA14 while between 0.3 and 2 s, the predictions are closer to the pan European model (RES14). Figure 14 confirms that the hypocentral model derived, where a point source is considered for computing the distances, shows larger differences in the prediction of the ground shaking for the considered scenarios.

## Conclusions

Motivated by its application in the update of the seismic hazard assessment for Germany, we developed a ground motion prediction equation (GMPE) in this study

tailored to such specific needs. Starting from the high quality NGA-west2 flat file, we constructed our model taking into consideration the requirements from the specific hazard application in a low-to-moderate seismicity area, being the following the main ones: a model implementing a point source measure of distance (i.e. hypocentral) along with an extended source metric (Joyner-Boore in our case); develop a model for a reference rock condition of vs30=800 m/s, avoiding possible bias due to low velocities (see also the discussions in Idriss, 2014); a GMPE with a smooth magnitude scaling around magnitude 5.5, which control the hazard at short return period in the target area; a complexity of the model suitable for its application in a low-to-moderate seismic area, that is, a functional form not requiring a-priori assumptions of variables not known in the target area, that would imply additional assumptions for refining the aleatory variability model.

The simplification in the functional form with respect to the NGA-west2 GMPEs had the effect of increasing $\sigma$. Indeed, the obtained values of $\sigma$ are close to the NGA-west2 values for small magnitudes and periods shorter than about 0.6s whereas, for longer periods, $\sigma$ increases to the values observed for the Bindi et al. (2014) Pan European model (RES14). This increased variability level for large magnitudes hamper the applicability of the model derived in this study for those applications where long mean return periods are of concern, such as site-specific hazard assessments. On the other hand, the followed approach can be of interest for many other applications, such as shake maps or earthquake early warning, as well as for the development of GMPEs for new intensity measure (e.g. Koufoudi et al., 2015). The comparison between the median predictions with those from the NGA-west2 and RES14 models in terms of Sammon's map shows that the predictions from our GMPE derived for the Joyner-Boore distance are closer to the NGA-west2 ones than the RES14 model whereas, for the GMPE implementing the hypocentral distance, larger differences are observed in the magnitude and distance scaling. Furthermore, the analysis of the covariance matrix shows that the epistemic uncertainty in the median of the model calibrated for the hypocentral distance, controlled by both the functional form and the data availability, is of the order of those affecting the NGA-west 2 models for magnitudes smaller than 7.5.

In conclusion, the suggested ranges of applicability for the GMPE derived in this study for the Joyner-Boore distance are between magnitude 3 and 8, distances shorter than 300 km and vs30 larger than 360 m/s (i.e., class A and B of Eurocode 8); for the hypocentral distance GMPE, the suggested ranges are from 10 to 300 km and magnitudes between 3 and 7.

## Acknowledgments

The need for the study summarized in this paper was born in the frame of the research project partly financed by the DIBt (https://www.dibt.de/) for the reappraisal of the seismic hazard of Germany for the National Annex to the updated Eurocode 8. We thank two anonymous Reviewers and the Editorial board who helped us to improve our work. Scripts downloaded from the Baker Research Group in Stanford (http://stanford.edu/~bakerjw/GMPEs.html, last accessed May 2016) were used to compute the predictions from the NGA2 models; we thanks the authors for sharing their programs. The R software (R Development Core Team, 2008; http://www.r-project.org, last accessed May 2016) has been used in this study to perform the regressions. In particular, the package lme4 (Bates et al., 2015; https://cran.r-project.org/web/packages/lme4/news.html, last accessed May 2016) has been used for mixed-effect regressions. The NGA2-West flat file is available at the PEER web page (http://peer.berkeley.edu/ngawest2/final-products/ , last accessed May 2016).

## References

Abrahamson N, Youngs R R (1992). A stable algorithm for regression analyses using the random effects model. Bull Seismol Soc Am 82(1):505–510

Abrahamson N, Silva W (2008) Summary of the Abrahamson and Silva NGA ground-motion relations. Earthquake Spectra 24(1): 67–97

Abrahamson NA, Silva WJ, Kamai R (2014) Summary of the ASK14 ground motion relation for active crustal regions. Earthquake Spectra 30(3): 1025-1055

Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csáki F (eds) 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, pp 267-281

Akkar S, Sandıkkaya MA, Şenyurt M, Sisi AA, Ay BÖ, Traversa P, Douglas J, Cotton F, Luzi L, Hernandez B, Godey S (2014a) Reference database for seismic ground-motion in Europe (RESORCE). Bulletin of Earthquake Engeneering 12(1):311–339

Al Atik L, Youngs RR (2014) Epistemic Uncertainty for NGA-West2 Models. Earthquake Spectra 30(3): 1301-1318

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

Bates D, Mächler M, Bolker B, Walker S (2014) Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823

Beauval C, Tasan H, Laurendeau A, Delavaud E, Cotton F, Gueguen P, Kuehn N (2012) On the Testing of Ground-Motion Prediction Equations against Small-Magnitude Data. Bull Seismol Soc Am 102(5): 1994-2007

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(1): 391-430

Bommer JJ, Akkar S (2012) Consistent Source-to-Site Distance Metrics in Ground-Motion Prediction Equations and Seismic Source Models for PSHA. Earthquake Spectra 28(1): 1-15

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 and 10 s, Earthquake Spectra 24(1): 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(3): 1057-1085

Bora SS, Scherbaum F, Kuehn N, Stafford P (2016) On the Relationship between Fourier and Response Spectra: Implications for the Adjustment of Empirical Ground-Motion Prediction Equations (GMPEs). Bulletin of the Seismological Society of America 106(3): 1235-1253

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(3): 1087-1115

Cauzzi C, Edwards B, Fäh D, Clinton J, Wiemer S, Kästli K, Cua G, Giardini D (2015a). New predictive equations and site amplification estimates for the next-generation Swiss ShakeMaps, Geophysical Journal International 200, 421-438. DOI: https://doi.org/10.1093/gji/ggu404

Cauzzi C., Faccioli E, Vanini M, Bianchini A (2015b). 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 (2008) An NGA model for the average horizontal component of peak ground motion and response spectra. Earthquake Spectra 24(1): 173-215

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(3): 1117-1153

Delavaud E, Cotton F, Akkar S, Scherbaum F, Danciu L, Beauval C, Drouet S, Douglas J, Basili R, Sandikkaya MA, Segou M, Faccioli E, Theodoulidis N (2012) Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe. Journal of Seismology 16(3): 451-473

Drouet S, Cotton F (2015) Regional Stochastic GMPEs in Low Seismicity Areas: Scaling and Aleatory Variability Analysis - Application to the French Alps. Bulletin of the Seismological Society of America 105(4): 1883-1902

Edwards B, Fäh D (2013) A stochastic ground-motion model for Switzerland. Bulletin of the Seismological Society of America 103(1): 78–98

Edwards B, Cauzzi C, Danciu L, Fäh D (2016). Region‐Specific Assessment, Adjustment, and Weighting of Ground‐Motion Prediction Models: Application to the 2015 Swiss

Seismic ‐ Hazard Map, Bulletin of the Seismological Society of America 106, 1840-1857, DOI: 10.1785/0120150367

Gruenthal G, Bosse C, Stromeyer D, Cotton F, Bindi D (2017) The Probabilistic Seismic Hazard Assessment of Germany, Version 2016 - Considering the Ranges of Epistemic Uncertainties and Aleatory Variabilities, submitted

Koufoudi E, Ktenidou O-J, Cotton F, Dufour F, Grange S (2015). Empirical ground-motion models adapted to the intensity measure $ASA_{40}$, Bull Earthquake Eng 13, 3625-3643 doi:10.1007/s10518-015-9797-z

Kotha SR, Bindi D, Cotton F (2016) Relaxing the ergodic assumption in Europe and Middle East: from ground motion regional attributes to site-specific hazard curves. Paper presented at 35th General Assembly of the European Seismological Commission, Trieste/Italy, 4-11 September 2016

Idriss I M, (2014). An NGA-West2 Empirical Model for Estimating the Horizontal Spectral Values Generated by Shallow Crustal Earthquakes. Earthquake Spectra. Vol 30(3): 1155-1177.

Scherbaum F, Kuehn NM, Ohrnberger M, Koehler A (2010). Exploring the Proximity of Ground-Motion Models Using High-Dimensional Visualization Techniques. Earthquake Spectra 26(4): 1117-1138

# Figures

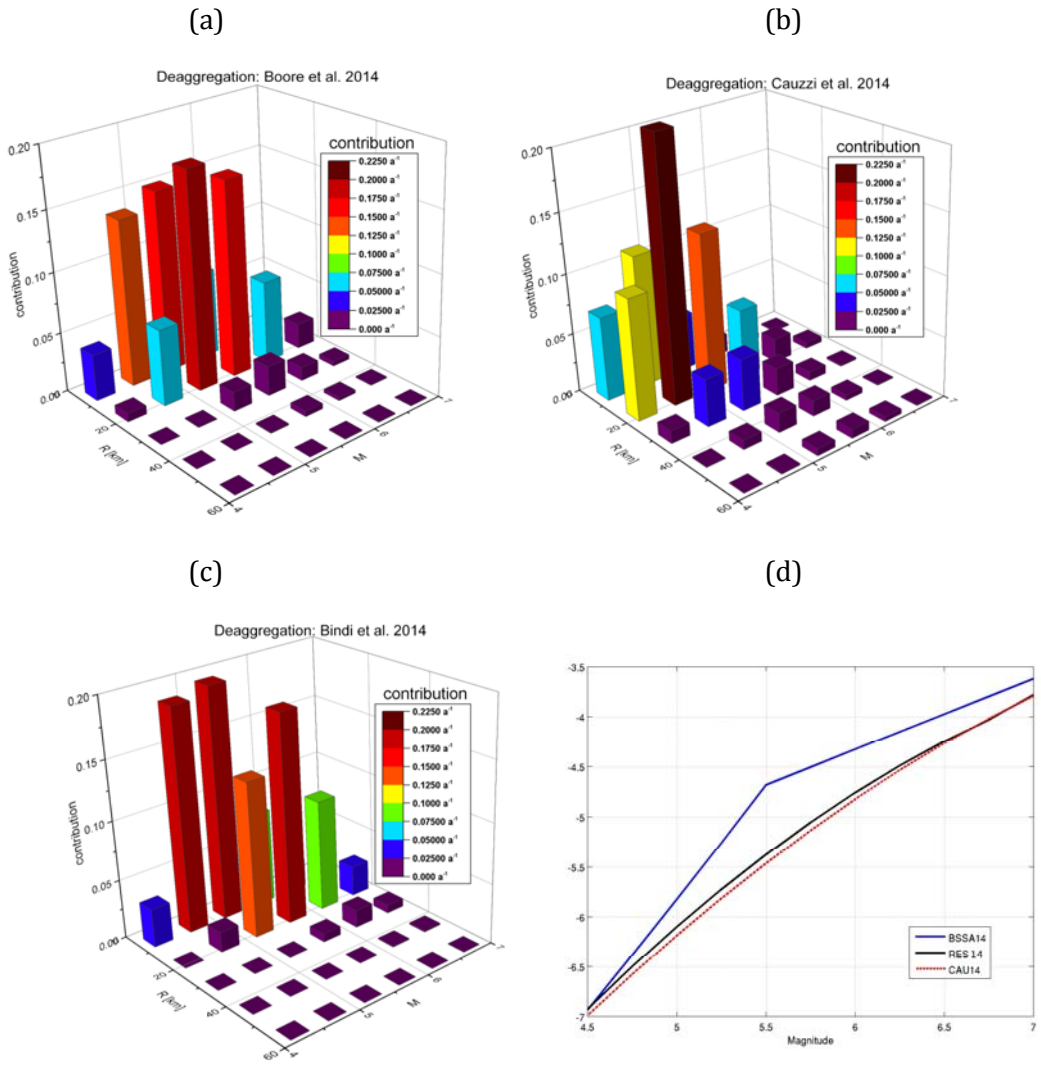



**Figure 1**. Disaggregation (i.e., normalized contribution to the total annual rate equal to 1/475, for PGA) over magnitude and distance scenarios for preliminary hazard assessment at a representative site in Germany: (a) BSSA14 (Boore et al., 2014) model; (b) Cauzzi et al. (2015b) model (CAU14); (c) Bindi et al. (2014) model (RES14). In panel (d), the PGA magnitude scaling for the three GMPEs is compared considering a vertical strike slip earthquakes at 30 km (vs30=800m/s).

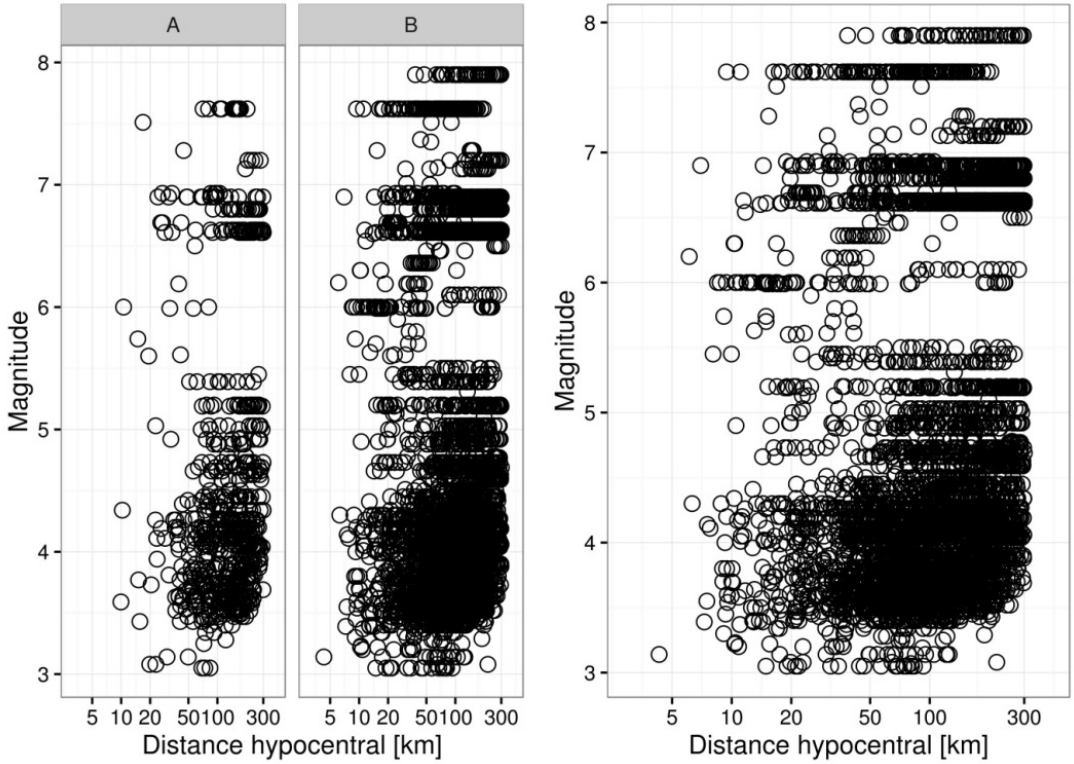



**Figure 2**. Distribution of magnitude versus hypocentral distance for the considered recordings. In the left panels, the distributions for classes A and B of Eurocode 8 are shown separately.

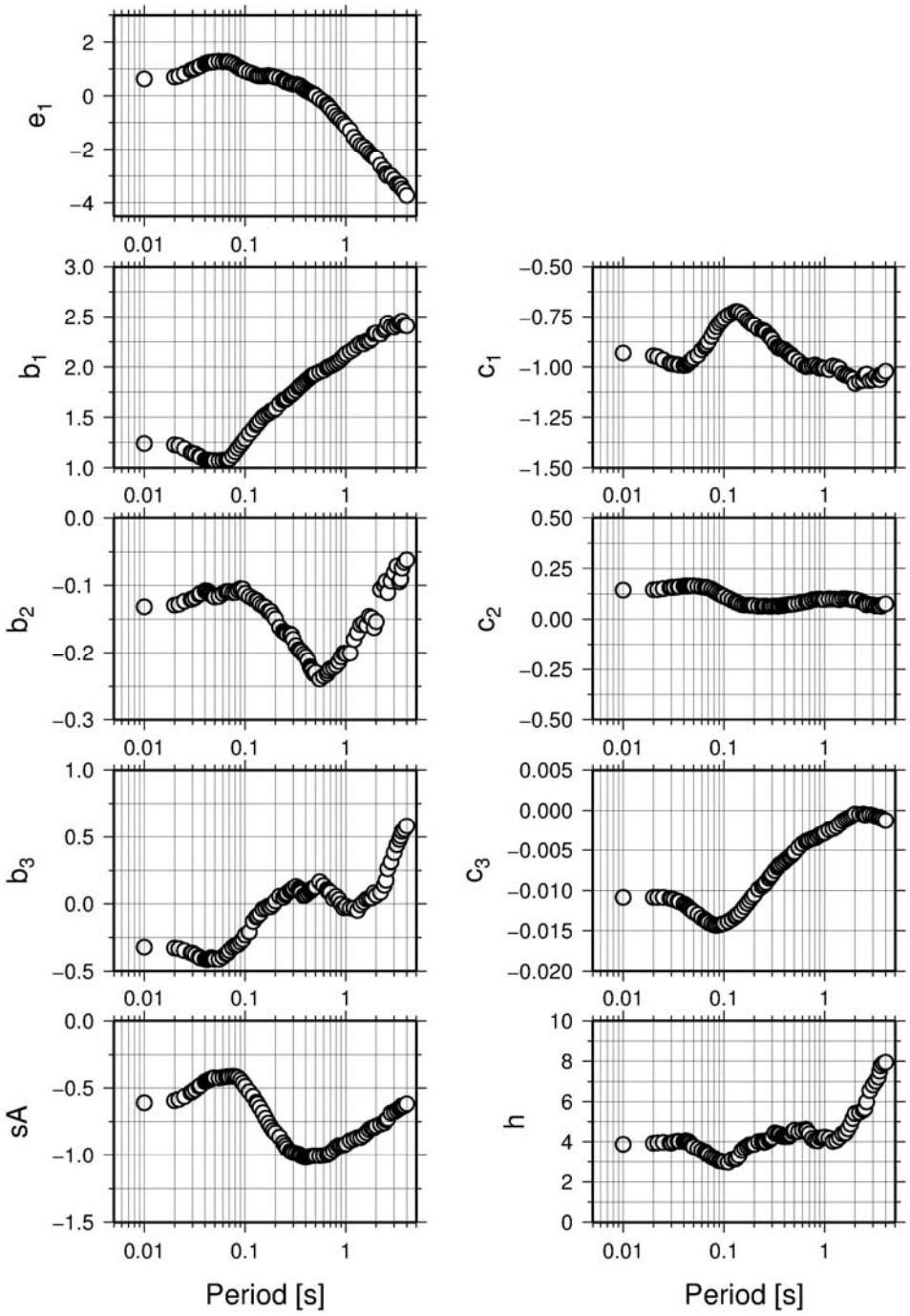


Figure 3. Coefficients of the GMPE calibrated for the Joyner-Boore distance.

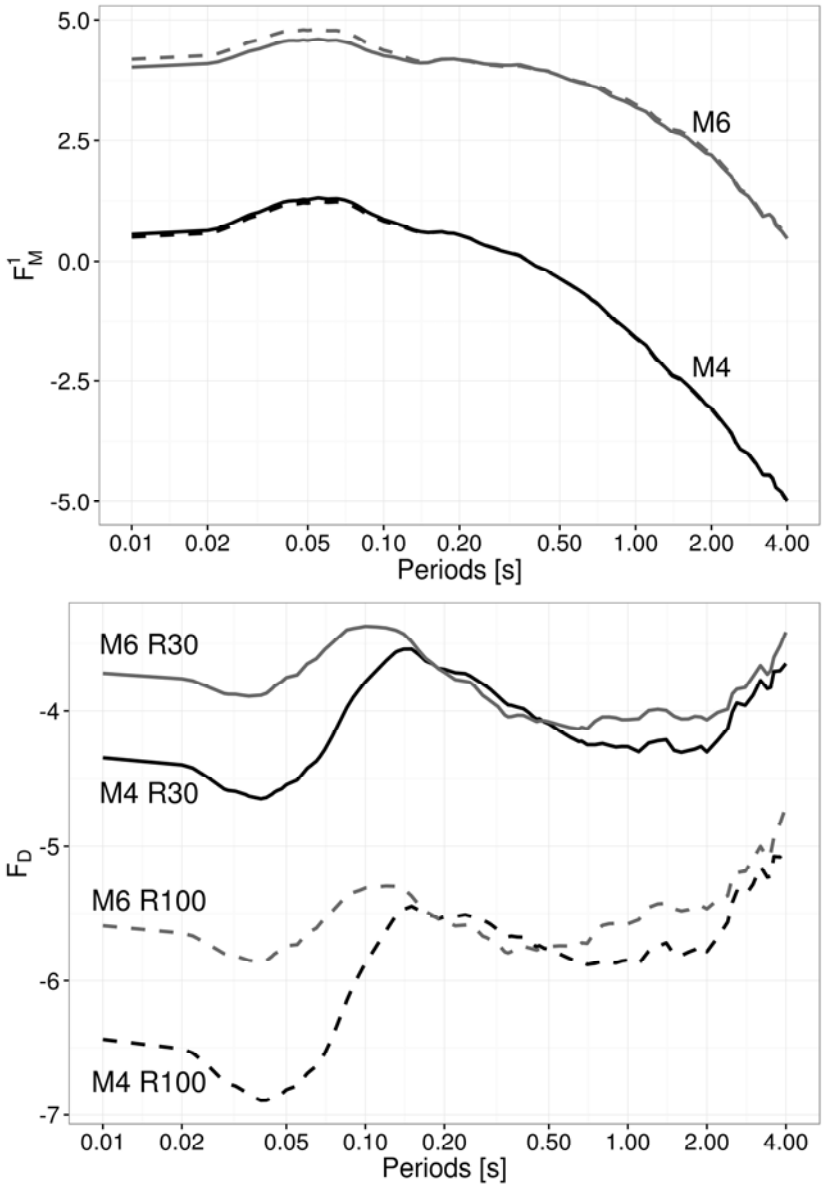



**Figure 4**. Top: Period dependence of the magnitude terms controlled by parameters $b_1$, $b_2$, $b_3$ and $c_2$ ($F^1_M$ is given by the $F_M$ term in equation 4 plus the $c_2$ term of equation 3), for magnitude 6 (gray) and 4 (black), and two distances (R=30 km, continuous line; R=100 km, dashed lines). Bottom: Period dependence of the distance terms controlled by $c_1$, $c_2$ and $c_3$ ($F_D$ term, see equation 3), for magnitude 6 (gray) and 4 (black) and two distances (R=30 km, continuous lines; R=100 km, dashed lines).

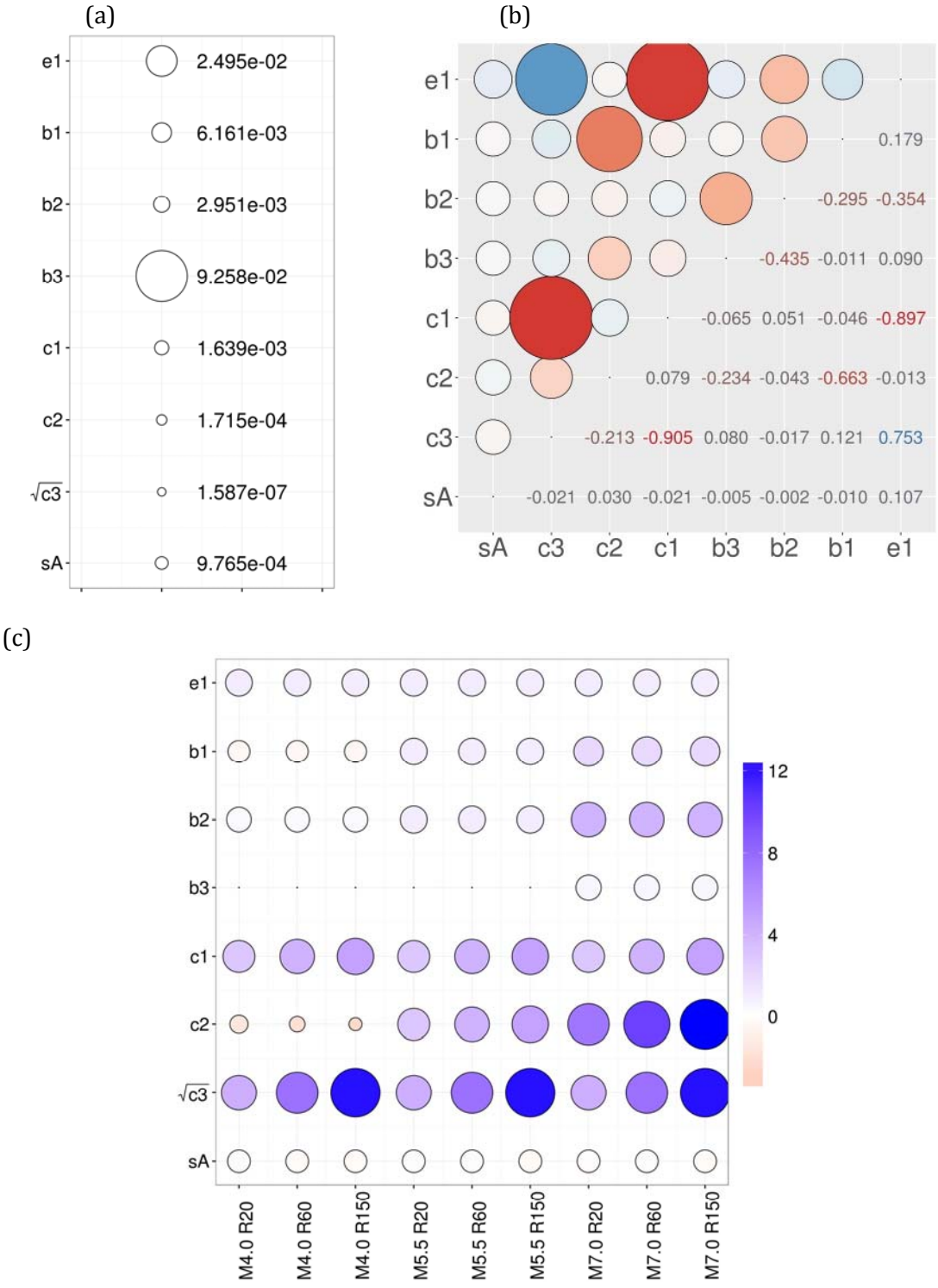
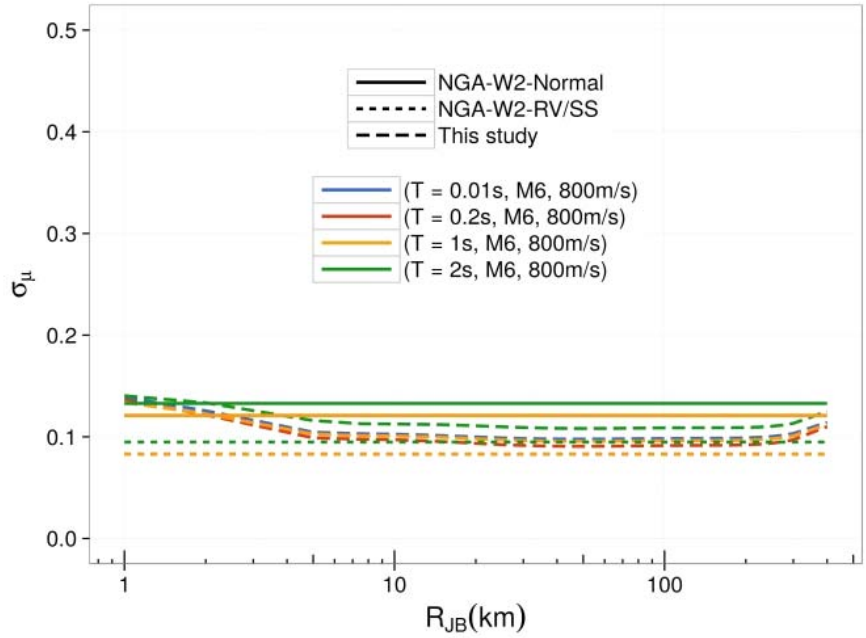



**Figure 5.** (a) Variance of the coefficients for T=01.s, considering the model for hypocentral distance. (b) correlation of the coefficients at 0.1 s, evaluated at the data points used to develop the model. (c) each column includes the gradient of the model with respect to one coefficient (from $e_1$ to sA, see equations 2 through 5), evaluated at a given magnitude and distance (for example, the column M4.0 R20 is evaluate for magnitude 4 at 20km, M7.0 R60 for magnitude 7 at 60, and so forth). The vs30 velocity is fixed to 600 m/s. Please note that the square root of variance for $c_3$ (panel a) and the square root of the derivative with respect to $c_3$ (panel c) are considered.

(a)




(b)

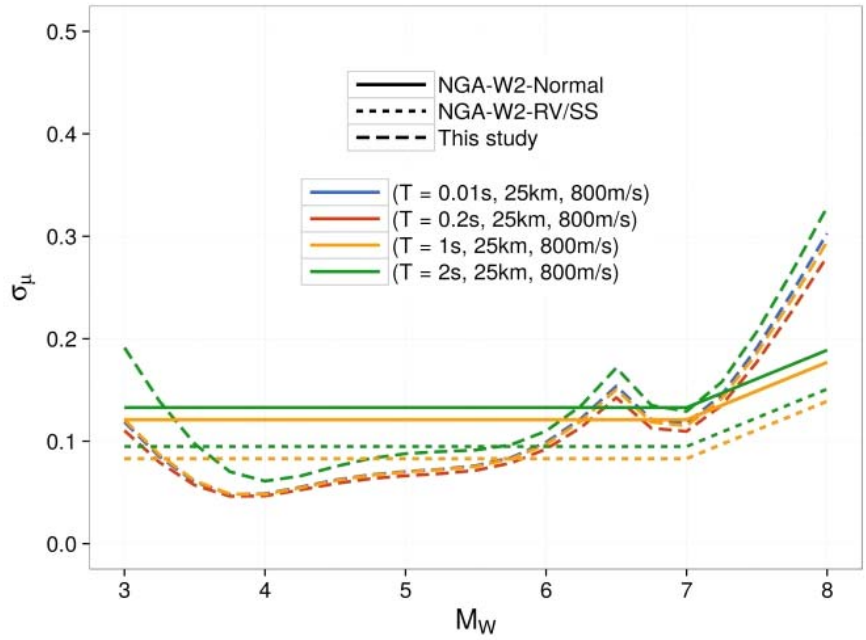
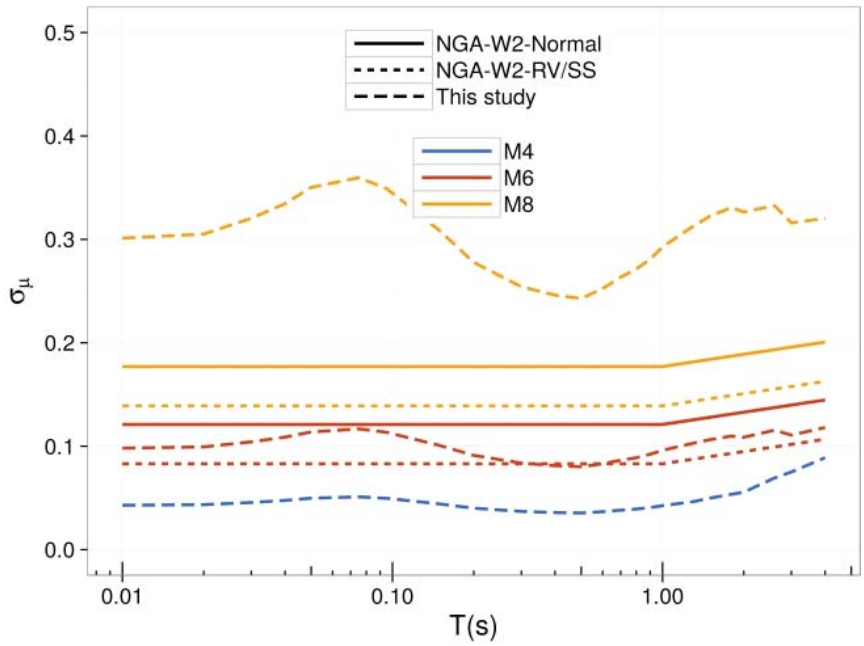



Figure 6- Continue.

(c)




(d)

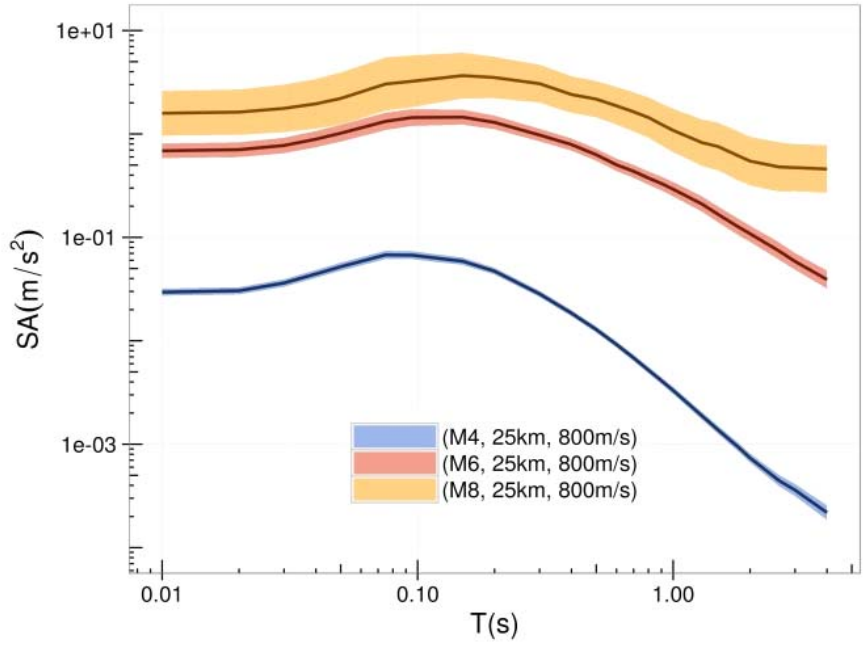



**Figure 6** Uncertainty in the mean computed for the Joyner-Boore model. (a) Dependence of $\sigma_\mu$ on distance, for different periods (colors); (b) Dependence of $\sigma_\mu$ on magnitude, for different periods; (c) Dependence of $\sigma_\mu$ on periods for different magnitude (colors) at a distance of 25 km; (d) overall effect of $\sigma_\mu$ on the mean response spectra, for three different magnitudes at a distance of 25 km (vs30=800 m/s). In panels (a), (b), and (c) the continuous and dotted lines represent the models proposed by Al-Atik and Youngs (2014) for NGA2 for normal and reverse or strike faults, respectively.

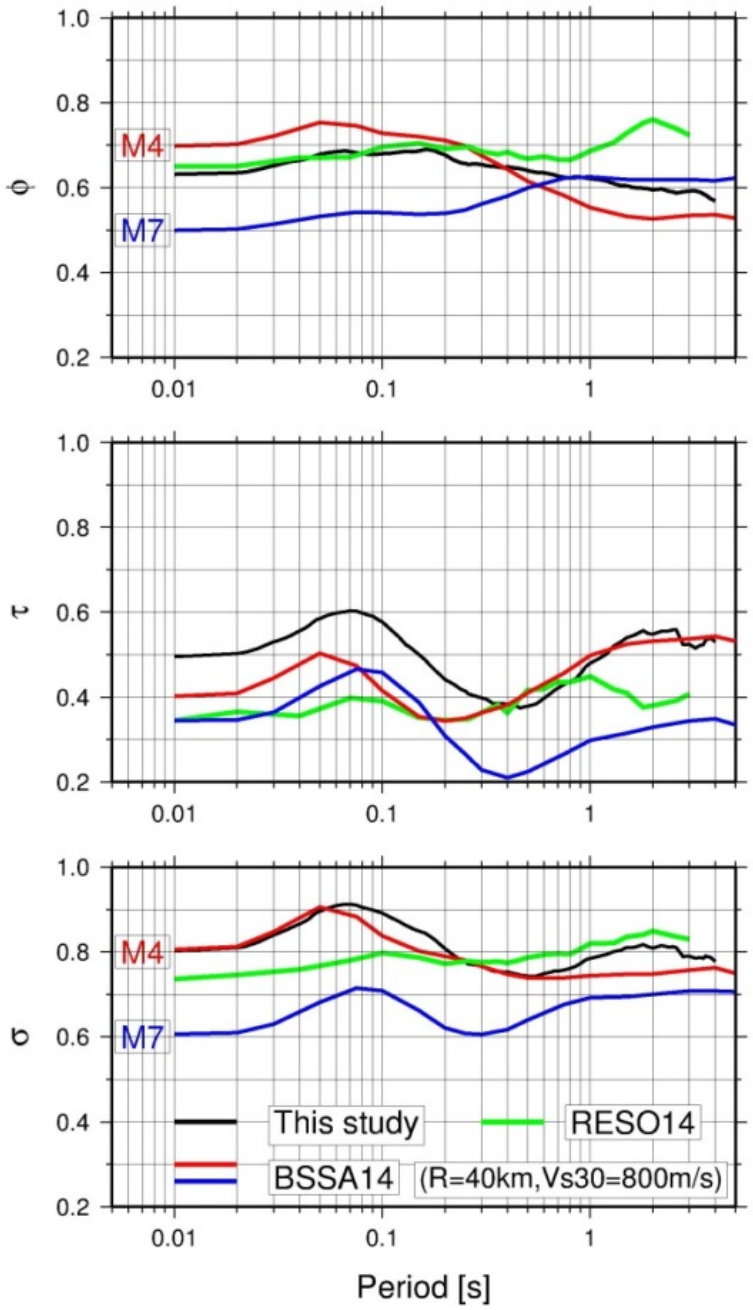
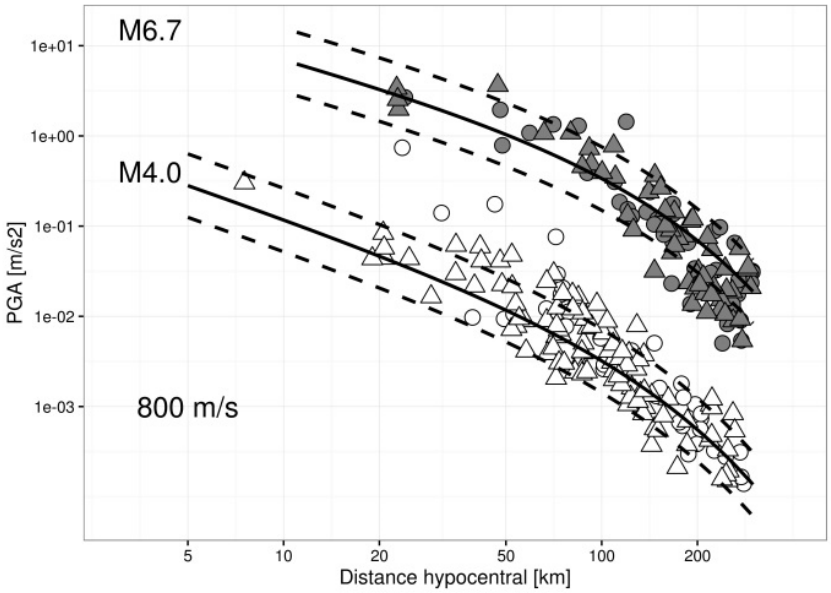



**Figure 7**. Within-event ($\phi$), between-event ($\tau$) and total ($\sigma$) standard deviations versus periods for the Joyner-Boore model (black lines) and comparison with BSSA14 (Boore et al., 2014) and RES14 (Bindi et al., 2014) GMPEs.

(a)




(b)

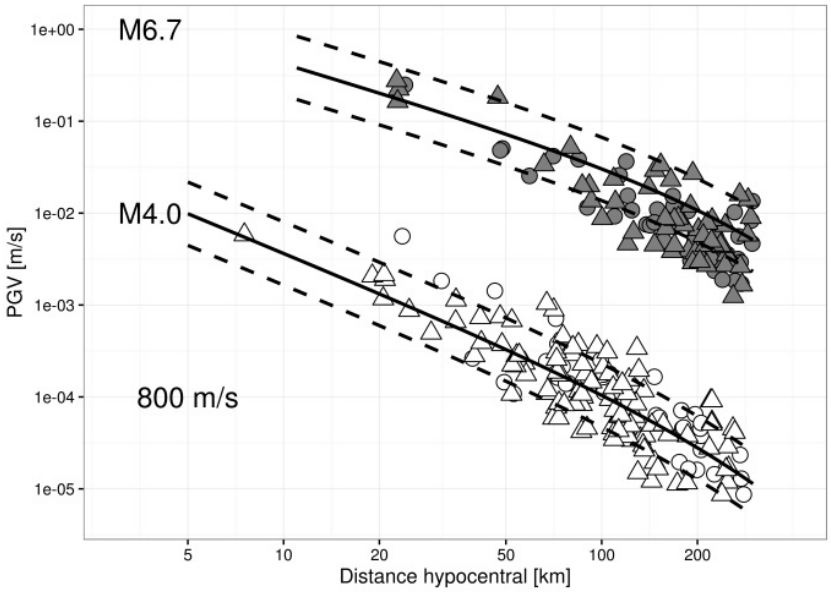



**Figure 8**. Prediction versus data for PGA (a) and PGV (b), for the hypocentral distance model. The median (continuous lines) ± 1σ (dashed lines) are computed for magnitude 4 and 6.7, and vs30=800 m/s. Symbols are observations for M=4.0±0.25 (white) and 6.7±0.25 (gray), considering stations with vs30≥800 m/s (circles) and <800 m/s (triangles).

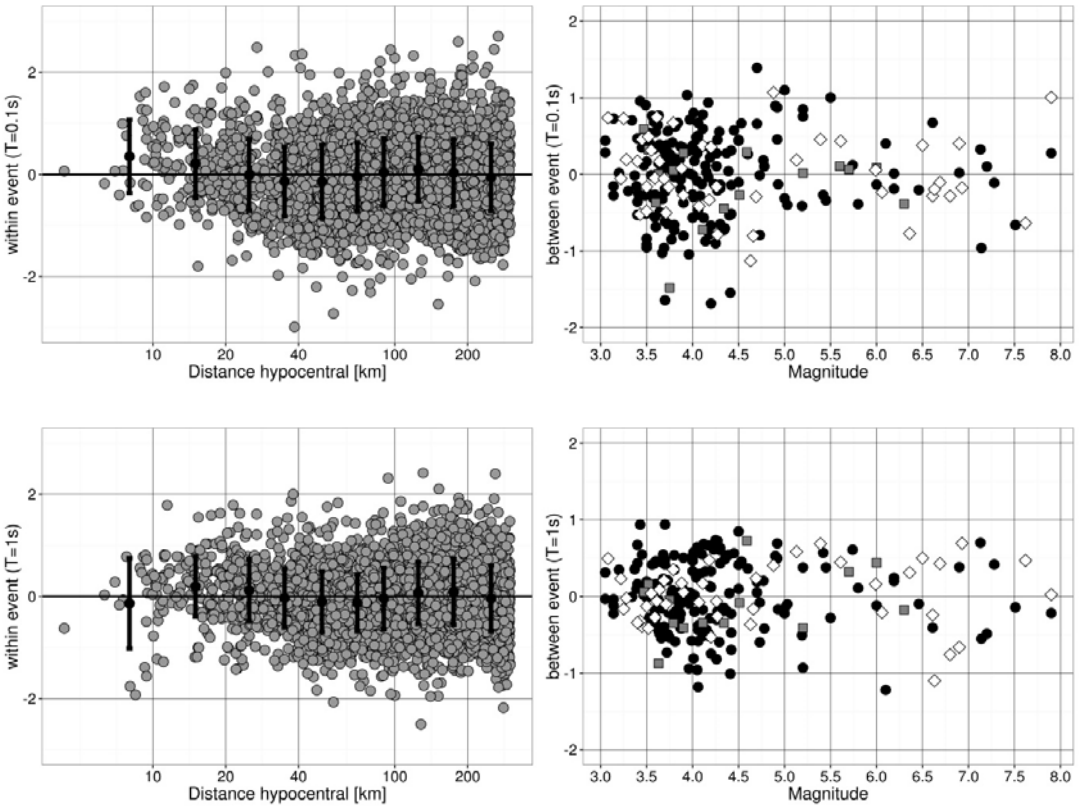
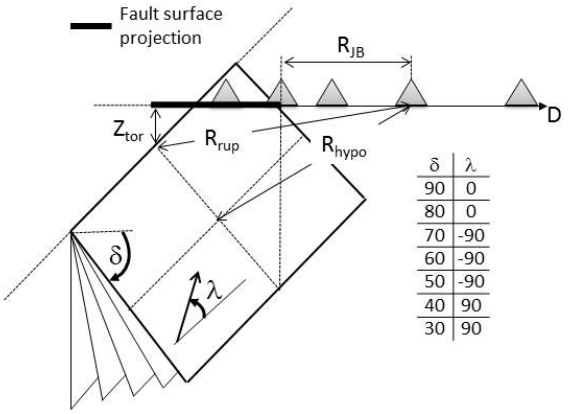



**Figure 9.** Top. Within event residuals versus distance (left) and between event residuals versus magnitude (right), for spectral acceleration at 0.1 s. The symbols in the between event plot indicate different focal mechanisms (circle: strike slip; square: normal; diamonds: reverse). In the bottom panels, the same distributions are shown but for 1 s.




| δ | λ |
|---|---|
| 90 | 0 |
| 80 | 0 |
| 70 | -90 |
| 60 | -90 |
| 50 | -90 |
| 40 | 90 |
| 30 | 90 |

**Figure 10.** Geometries used to generate the scenarios for the Sammon's map. For each source, the distances required by the considered GMPEs are computed for 6 stations located at different distances along a line perpendicular to the strike direction, being all stations located on the hanging wall. The combinations between the dip δ and rake λ angles are given in the Figure. For each combination, 4 different magnitudes are generated (i.e., 4, 5, 6 and 7). vs30 is fixed to 800 m/s.

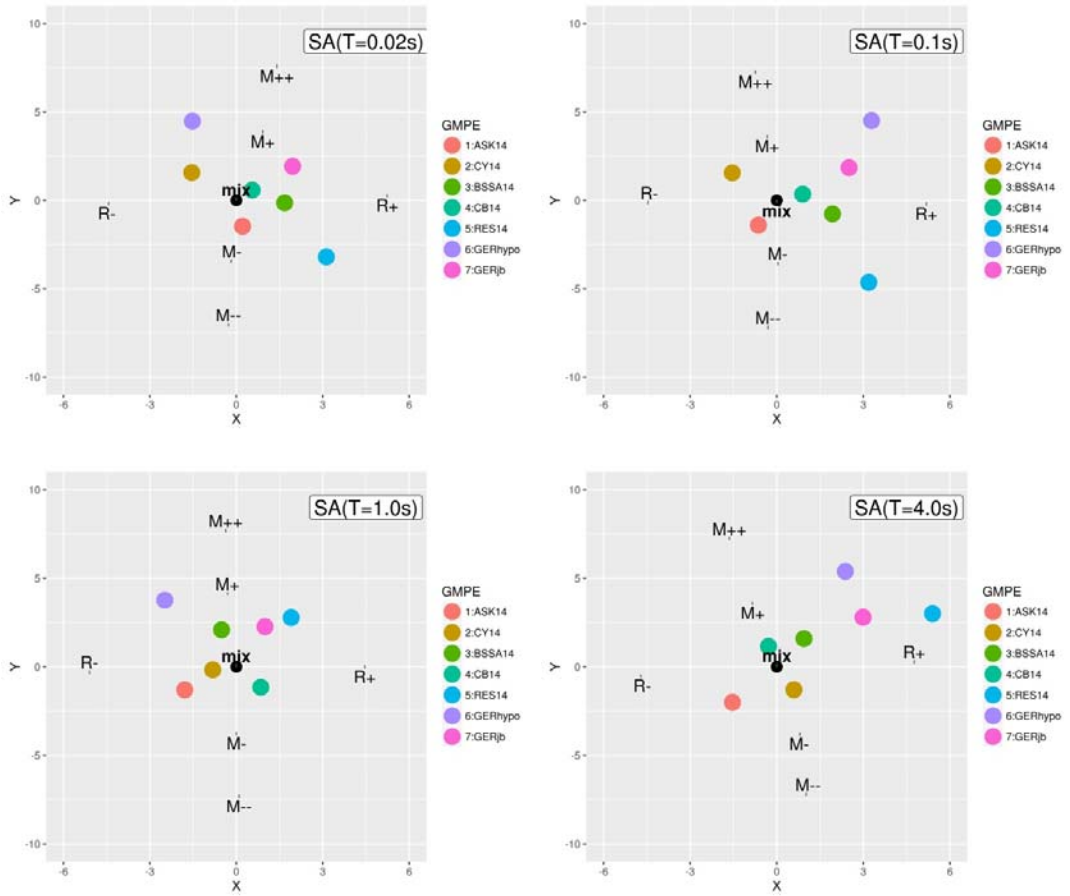



**Figure 11**. Sammon's maps for T=0.02, 0.1, 1, and 4 s, considering the source scenarios described in Figure 10. The considered GMPEs are represented by colored circles while MIX is the reference model computed as mixture of the four considered NAGA2 GMPEs with equal weights. M+, M++, M-, M--, R+, and R- are GMPEs where artificial either magnitude or distance scaling are added to MIX (see text for details).

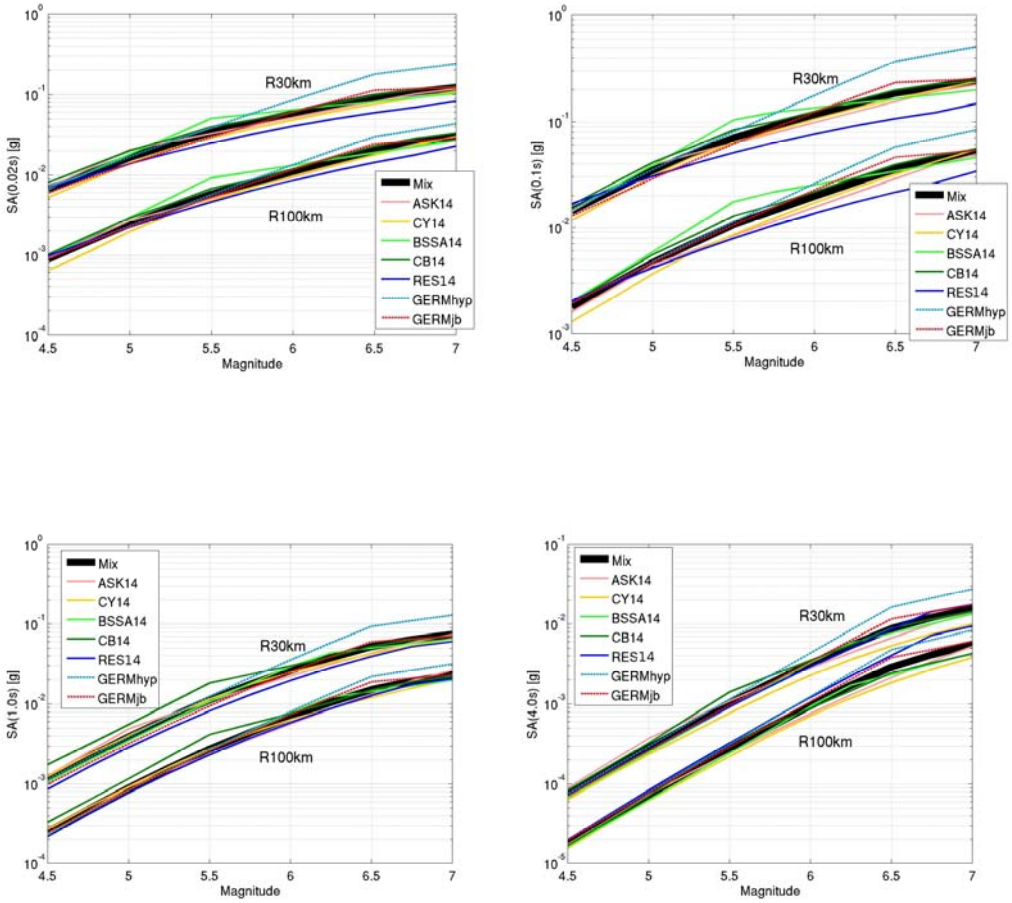



**Figure 12**. Magnitude scaling for different GMPEs a T=0.02, 0.1, 1.0 and 4.0 s, considering two different distances (i.e., 30 and 100km), vs30=800 m/s, and a vertical strike slip earthquake. The model MIX is the mixture of the four considered NGA2 models considered with equal weights (see text for details).

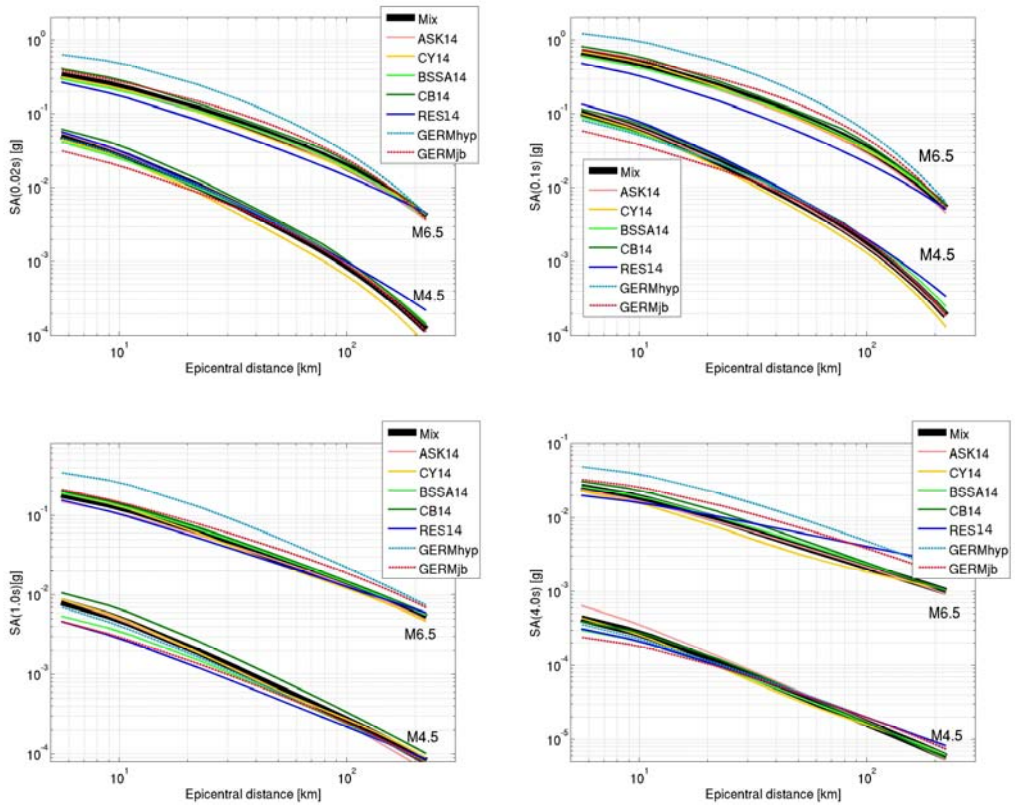



**Figure 13**. Distance scaling for different GMPEs a T=0.02, 0.1, 1.0 and 4.0 s, considering two different magnitudes (i.e., 4.5 and 6.5), vs30=800 m/s, and a vertical strike slip earthquake. The model MIX is the mixture of the four considered NGA2 models considered with equal weights (see text for details).

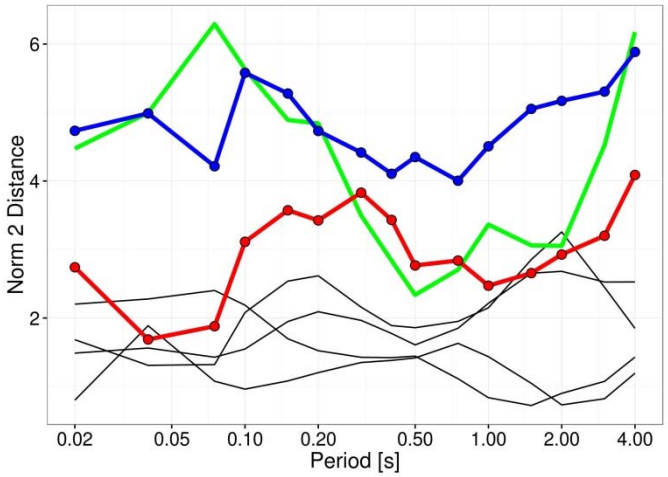



**Figure 14**. Period dependence of the distance in the Sammon's plane between each considered model in Figure 11 and the reference one (MIX). The results for the NGA2 models are shown in black, RES14 in green, while the models derived in study for the Joyner-Boore and hypocentral distances are shown in red and blue, respectively.