A Partially Non-Ergodic Ground-Motion Prediction Equation for Europe and the Middle East

Nicolas M. Kuehn
Pacific Earthquake Engineering Research Center
University of California, Berkeley

Frank Scherbaum
Institute of Earth and Environmental Science
University of Potsdam

March 30, 2016

Abstract

A partially non-ergodic ground-motion prediction equation is estimated for Europe and the Middle East. Therefore, a hierarchical model is presented that accounts for regional differences. For this purpose, the scaling of ground-motion intensity measures is assumed to be similar, but not identical in different regions. This is achieved by assuming a hierarchical model, where some coefficients are treated as random variables which are sampled from an underlying global distribution. The coefficients are estimated by Bayesian inference. This allows one to estimate the epistemic uncertainty in the coefficients, and consequently in model predictions, in a rigorous way. The model is estimated based on peak ground acceleration data from nine different European/Middle Eastern regions. There are large differences in the amount of earthquakes and records in the different regions. However, due to the hierarchical nature of the model, regions with only few data points borrow strength from other regions with more data. This makes it possible to estimate a separate set of coefficients for all regions. Different regionalized models are compared, for which different coefficients are assumed to be regionally dependent. Results show that regionalizing the coefficients for magnitude and distance scaling leads to better performance of the models. The models for all regions are physically sound, even if only very few earthquakes comprise one region.

1 Introduction

An important question in engineering seismology is whether ground-motion predictions equations (GMPEs) are regionally dependent [24]. For example, in Europe both regional GMPEs (e.g. [4] for Turkey; [11] for Italy; [16] for the Eastern Alps; [21] for Greece) as well as pan-European GMPEs (e.g. [5, 9]) have been developed. The recent pan-European GMPEs described in [25], [5, 14, 23, 31] do
not include regional differences. From the models developed recent NGA-West 2 project [15], four [11, 13, 17, 20] include regional adjustment terms. Generally, these models start out from a global model, which is then regionally adjusted based on residual analysis. For California, differences between Northern and Southern California have been found [10, 19].

Incorporating regional differences in ground-motion scaling into PSHA is a first step to remove the ergodic assumption from seismic hazard estimation [9, 37, 32]. Under the ergodic assumption states the distribution of ground motions at a given site is the same as the spatial distribution over all sites given the same magnitude, distance and site conditions [9]. Regionally dependent GMPEs relax this assumption by breaking a data set from a large region into smaller units.

Generally, it makes sense to assume that there are regional differences in ground-motion scaling between different regions, as geological conditions differ. However, one problem that one faces when developing regional GMPEs is lack of data for smaller regions. Recently, [30] showed a way to constrain coefficients to be similar across regions, and also provided evidence of better predictive performance as opposed to a global model. In a similar vein, [37] shows how regional difference between countries (regions) can be included into a GMPE.

Here, we investigate differences in GMPEs between regions in Europe and the Middle East, based on data from the RESORCE data base [6]. We develop a multi-level model for peak ground acceleration (PGA) that accounts for regional differences from the start. Therefore, the coefficients of the model are regionalized but assumed to be sampled from a global distribution – this allows the different regions to borrow statistical strength from each other. We investigate the predictive performance of several models, which differ in the amount of coefficients that are regionalized. The model presented in this work shares similarities with the work of [32, 33, 37]. We use the same data as [30], but constrain the coefficients for different regions in a different way. The model is cast in the formulation of [33]. The integration of regionalized coefficients is similar to [37], but the model presented here is estimated on a different data set. In addition, we show the effect of different regionalized models on median predictions, and how predictive performance can be used for model selection.

In general, regionalization of coefficients improves predictive performance and decreases the associated aleatory variability (see section Results), which has consequences for seismic hazard calculations [12].

2 Data

We use the same data set that was used by [30]. It is a subset of the RESORCE data base [6], which was compiled for the SIGMA project [1]. We provide only a brief description here. More information can be found in [30] and [6]. In total, there are 1261 records from 362 earthquakes, recorded at 359 stations. The events are divided into 9 regions, mainly based on geographical considerations, but informed by tectonic units defined for the SHARE project (Seismic Hazard Harmonization in Europe) [22]. The events and their regions are shown in Figure [1]. The magnitude/distance distribution of the data is shown in Figure

\[\text{http://projet-sigma.com/}\]
\[\text{http://www.share-eu.org/}\]
Table 1 summarizes the data for the different regions. As one can see, some regions have only a very limited magnitude/distance range. However, as shown later, we can still estimate a physically plausible model for those regions.

3 Model Formulation

We use a slightly adjusted multi-level formulation of [33] to formulate the model:

\[ Y_{es} \sim N(f(x) + \eta_e + \lambda_s, \phi_{SS}) \]  
\[ \eta_e \sim N(0, \tau) \]  
\[ \lambda_s \sim N(0, \phi_{S2S}) \]

Equation (1) reads as: the ground-motion value \( Y \) from event \( e \) recorded at station \( s \) is distributed according to a normal distribution with a mean that is a
function of the predictors $x$ plus the event term $\eta_e$ and the station term $\lambda_s$, and standard deviation $\phi_{SS}$, which denotes single-station within-event variability. The event and station terms are distributed normally with standard deviation $\tau$ and $\phi_{S2S}$, which describe between-event and station-to-station variability [7]. The ground-motion parameter of interest $Y$ is logarithmic PGA.

We use a slightly different model than [30]. The predictor variables are the moment magnitude $M$, the Joyner-Boore distance $R_{JB}$, the shear wave velocity averaged over the top 30m $V_{S30}$, and the focal mechanism. The dependence of the median predictions on the predictors is modeled by the following functional form:

$$f(x) = c_0 + c_1 M_e + c_2 M_e^2 + c_3 F_R + c_4 F_N + (c_5 + c_6 M_e) \ln \sqrt{R_{es}^2 + c_7^2} + c_8 R_{es} + c_9 \ln \frac{V_{S30}}{760}$$  \hspace{1cm} (4)

where $F_R$ and $F_N$ are dummy variables taking the value 1 for reverse and normal faulting events, respectively, and 0 otherwise. The coefficients $c_i$ vary by region, which is indicated by index $r$. However, to allow data from one region influence the ground-motion scaling in other regions, the coefficients are connected

$$c_i \sim N(\mu_{c_i}, \sigma_{c_i})$$  \hspace{1cm} (5)

Hence, the regional coefficients are sampled from a (global) distribution. In terms of a multi-level model [27, 33] the regional coefficients would be another level, a hierarchy higher than the event/station level. In terms of a mixed-effects model, the global coefficients $\mu_{c_i}$ would be called the fixed effects, while the difference $c_i - \mu_{c_i}$ are the regional random-effects. We prefer the multi-level formulation, because we believe that it is easier to understand.

The parameters of the model comprise the coefficients $c_{i,r}$, the mean and standard deviations of their global distribution, $\mu_{c_i}$ and $\sigma_{c_i}$, the components of the aleatory variability $\phi_{SS}$, $\phi_{S2S}$ and $\tau$, as well as the event and station terms $\eta_e$ and $\lambda_s$.

The parameters of the model are estimated via Bayesian inference using the program Stan [33], which performs Hamiltonian Monte Carlo sampling. Bayesian inference relies on Bayes theorem to estimate the posterior distribution of the parameters, which is proportional to the product of the likelihood, determined by the model of equation (4), times the prior distribution of the

| Region       | $N_{eq}$ | $N_{rec}$ | $M$-range | $R_{JB}$-range |
|--------------|----------|-----------|-----------|----------------|
| Alps         | 29       | 91        | 4.1 – 6.4 | 2. – 100.      |
| Appenines    | 78       | 303       | 4.1 – 6.9 | 0. – 200.      |
| South Greece | 41       | 87        | 4.1 – 6.9 | 0. – 154.      |
| North Greece | 18       | 58        | 4.7 – 6.9 | 2. – 102.      |
| North Turkey | 53       | 324       | 4.1 – 7.6 | 0. – 191.      |
| South Turkey | 70       | 250       | 4.2 – 6.5 | 0. – 199.      |
| East Turkey  | 62       | 122       | 4.4 – 7.1 | 0. – 198.      |
| Iran         | 6        | 19        | 5.9 – 7.4 | 4. – 175.      |
| Sicily       | 5        | 7         | 4.2 – 5.9 | 20. – 151.     |
parameters. The prior distribution describes the uncertainty of the parameters before data is taken into account. Hence, the posterior distribution describes the uncertainty of the model parameters given the observed data. For a more detailed explanation of Bayesian inference see e.g. [36]. For the parameters describing standard deviations we use Half-Cauchy distributions as prior distributions [20, 39]. The priors for the global parameters $\mu_c$ are normal distributions, based on the model of [1]. This model has been shown to extrapolate well to large magnitudes and short distances, so using it as a basis for setting the prior distributions makes these more informative than just setting wide normal distributions. The model of [1] is chosen because it is predominantly based on a global data set and thus has only partial overlap with RESORCE database. Hence, data is not used twice in both the prior and the likelihood.

We evaluate the model of [1] at different settings of the predictor variables ($M = 4.5, 5, \ldots, 8; R_{JB} = 1, 5, 10, 20, 30, 50, 75, 100, 150, 200; V_{S30} = 300, 400, \ldots, 1100; F_N = F_R = 0, 1$). Values for the depth of the top of the rupture are calculated from the model presented in [20]. All evaluations are carried out on the footwall. We then fit equation (1) to the predictions of the model of [1]. The fitted coefficients provide the mean values for the prior distributions. Ideally, the width of the prior distributions would be set by taking into account the uncertainty of the model of [1], together with the uncertainty of the fitted coefficients. Since the former uncertainty is not readily available, we fit equation (4) to the data and use the standard error of the coefficients, together with the standard error of the fit to the model of [1], for the standard deviations of the prior distributions for the coefficients.

To avoid unphysical behavior of the expected models, we constrain the parameters $c_i$ and $\mu_c$ to be either positive or negative, except for the constant term $c_0$.

The parameters are estimated by sampling from the posterior distribution using the program Stan [38]. We run 4 chains of samples, with different starting values. To avoid influence of the starting values, we discard a burn-in of the first 1000 samples. After the burn-in period, each chain is run for another 1000 samples. We keep every fifth sample, leading to 800 draws from the posterior distribution. Convergence of the chains is assessed using the Gelman-Rubin statistic [29].

4 Model Selection

In the previous section we have formulated a multi-level ground-motion model that can take regional differences into account by constraining the parameters via equation (5). This can be done for all coefficients, or just for a subset. It is also possible to estimate a model that pools all data together into one region. Furthermore, we can estimate a separate model for each region, where the model coefficients are not connected via equation (5). It is important to test whether a regionalized model performs better than a global model or individual models. One often used measure to compare different models is the Akaike Information Criterion (AIC) [2]. However, AIC cannot be easily generalized to Bayesian models. Furthermore, the calculation of AIC requires the number of parameters in the model, which is not straightforward for multi-level models, since shared parameters (such as an event term) cannot be treated like a fixed
coefficient in the model (see [28] for a discussion on these issues).

[30] performed experiments splitting the data set into training and test data and found that a regionalized, connected model performs generally better on the test data than the pooled model. This already provides a compelling argument for splitting a data set into subregions, for which individual models can be estimated.

Splitting the data set into a training and test set and evaluating the model on the test set provides a measure how well the model can predict new, future data. This is called the generalization error. There are different possibilities how the data set could be split into training/test data. One could select the test data such that it contains only records from events that are not in the training data. Another way would be to have records from one event in both the training and test data set. In the latter case, one can make use of an estimated event term from the training data, and hence get better residuals (in terms of a smaller overall residual). Ultimately, the two strategies answer two different questions – how well does the model predict data from a new event, and how well does it perform on an event from which data already exists. Both of these measures are important.

It is computationally very demanding to repeatedly fit different models to training and test data. Therefore, in this work estimate the generalization error using a measure that is based on draws from the posterior distribution of the parameters. This measure is the widely applicable Information Criterion (WAIC) [40, 42]. The estimated quantity is $\hat{\text{elpd}}_{\text{WAIC}}$, which is an approximation to the expected log pointwise predictive density for a new data set. The predictive density describes the full uncertainty distribution of a prediction from a model, taking into account the uncertainty of the parameters (characterized by the posterior distribution) and the aleatory variability. For details on theory, derivation and implementation of $\hat{\text{elpd}}_{\text{WAIC}}$, see [40, 41, 28].

The measure $\hat{\text{elpd}}_{\text{WAIC}}$ is calculated based on the log-likelihood of the data given draws from the posterior distribution of the parameters $\theta$. The log-likelihood can be calculated in different ways, based on the question one might want to ask. Here, we calculate two different likelihoods, and consequently get two different values for $\hat{\text{elpd}}_{\text{WAIC}}$:

\begin{equation}
\label{eq:elpd1}
 p(y_{es}\mid \theta^S) = N(f(x) + \eta_e + \lambda_s, \phi_{SS})
\end{equation}

\begin{equation}
\label{eq:elpd2}
 p(y_{es}\mid \theta^S) = N(f(x), \sigma_T)
\end{equation}

where $\theta^S$ is a draw from the posterior distribution of the parameters, and $\sigma_T^2 = \tau^2 + \phi_{SS}^2 + \phi_{SS}^2$ is the total variability.

The $\hat{\text{elpd}}_{\text{WAIC}}$ based on the likelihood calculated using equation (6) is a measure of how well the model can predict a new record, given knowledge of the event term $\eta$ and station term $\lambda$. On the other hand, the second likelihood describes how well the model can predict a new record, given only the median prediction $f(x)$. Hence, equation (6) measures how well a model can predict a new record for an already observed event, while equation (7) provides a measure how well a model can predict a new record from an unknown event.

We use WAIC (or more precisely $\hat{\text{elpd}}_{\text{WAIC}}$) to compare different models:

- A model estimated on the whole, pooled data set, which we term “global”. 
A model with independent, separate coefficients for each region, which we call “individual” (I).

- A model with coefficients that are connected via equation (5). The model where all coefficients vary by region is called “regional 1” (R1).

- A model with regionalized coefficients, where the near-source saturation term $c_6$ does not vary by region. This model is called “regional 2” (R2).

- A model where coefficients $c_3, c_4, c_7$ do not vary by region, called “regional 3” (R3).

- A model where only coefficients $c_0, c_8, c_9$ vary by region, called “regional 4” (R4).

- A model where only coefficient $c_0$ varies by region, called “regional 5” (R5).

In all cases, we assume the standard deviations to be the same for all regions. This avoids trade-offs between differences in the coefficients and the standard deviations. Implicitly, this means that the variability of stress drop and geological conditions within the regions is assumed to be similar.

5 Results

Figure 3 shows the difference in $\hat{\text{elpd} \text{WAIC}}$ between the regionalized models and the global model, where a positive difference indicates a better fit of the regional model.

We can see a positive $\Delta \hat{\text{elpd} \text{WAIC}}$ for all the regionalized models compared to the global model. There is not a big difference between the different regionalized models, except for the “regional 5” model, which is close to the global model. This indicates that there are differences in scaling with the predictor variables between the different regions, and that these differences are important for predicting new data points. The separate model also has better predictive capability (according to $\text{elpd}_W \text{AIC}$) than the global model, and is comparable to the regional models when the event and station term are included (equation
Figure 4: Between-event, within-event and between-station standard deviations. Error bars show the standard deviation of the posterior distribution.

This is not really surprising, since knowing these terms provides a lot of information in the case of a small data set. Furthermore, the informative prior provides some constraint on the parameters, so that even with a small number of data the estimated coefficients are not completely physically unreasonable.

Figure 4 shows the estimated between-event, within-event and station-to-station standard deviations. The values decrease for the regional models, with the main contribution to the lower total standard deviation coming from $\phi_{S2S}$. This indicates that there are mainly differences in site scaling between the different regions.

The values of the $\hat{\text{elpd}}_{WAIC}$ parameters provide some indication that a hierarchical, regionalized model might be beneficial in terms of ground-motion modeling. However, $\hat{\text{elpd}}_{WAIC}$ is an approximation to the expected log point-wise predictive density for a new data set, where the new data set may include new events and stations, but does not go beyond the ranges of the predictor variables. For an application of a ground-motion model in PSHA, however, it is important that the model extrapolates, e.g. to large magnitudes and short distances, in a physically meaningful way. In particular, some of the regions have a pretty limited coverage of magnitudes and distances (cf. Figure 2), so it is important to ensure that they extrapolate well to $M/R$-ranges of interest. On the one hand, using an informative prior on the parameters (based on the model of [1]) constraints the variability of the parameters and should lead to reasonable behavior. However, since the coefficients for the individual regions are only indirectly constrained by the prior via equation (5), it is worthwhile to check their scaling. In the following, we use the “regional 3” model, for which all coefficients except $c_3$, $c_4$ and $c_7$ vary by region. According to Figure 3 is has good generalization performance. It makes sense to not regionalize the scaling with focal mechanism since the distribution of focal mechanisms across regions is very uneven.

In Figure 5, the scaling of PGA with magnitude and distance is shown, using the regionalized coefficients. The curves are based on the mean of the posterior distributions for the individual parameters. As one can see, the models for all regions have a physically plausible shape – even for regions such as Sicily or Iran, which have a very limited number of indigenous data. In these cases, their model borrows strength from the other regions via the hierarchical model formulation of equation (6). One can see in Figure 5 that the regional models for Turkey produce relatively low median predictions. This is concordant with the
models of [4, 14], which both produce relatively low ground-motion predictions and which contain a large percentage of Turkish data.

The residuals of the model are shown in Figure 6. They are unbiased for all regions, indicating that the model formulation works well in fitting the data, and also in accounting for regional differences. Since the main reduction in the standard deviations is for the site-to-site variability $\phi_{S2S}$, the scaling of PGA with $V_{S30}$ is shown in Figure 7. The slopes are quite different for some regions, which explains the lower standard deviation.

6 Discussion and Conclusions

We have shown a way to estimate regionalized GMPEs, using a multi-level approach, as previously explored by [37]. Using this approach, it is possible to estimate individual, physically sound models even for regions that do not comprise a lot of data, by constraining their coefficients to be close to the global coefficients. The amount by which they can differ is determined by the data.

We have seen in Figure 3, that the hierarchical regionalized model gives higher values of $elpd_{WAIC}$ than a global model that pools all data, which means that the regionalized model has higher generalization capability. Furthermore, the models for the individual regions are physically plausible.

In Figure 8, we illustrate the effect of the constraint imposed by equation (5). The magnitude scaling is shown for Sicily and Northern Greece, which
Figure 7: Scaling of regional models with $V_{S30}$.

Figure 8: Scaling of regionalized models for two regions, calculated using the “regional 3” model (solid lines) and a model with independent coefficients (“individual”, dashed lines). Black line is the global model. The predictions of the regionalized models are drawn towards the global model in the case of only few data points in the regions.

both have relatively few data points, for the regionalized model (R3) and the “individual” model. We can see that in the regionalized model, the curves are pulled towards the global model – the models for these regions are constrained by data from the other regions. This effect makes it possible to extrapolate the regionalized models to data ranges (for example higher magnitudes or shorter distances) that are not covered by data for the individual regions in these ranges, the fact the regional coefficients are connected via equation (5) ensures that the regional models behave similarly to the global model. Hence, we believe that the proposed methodology is a good way to estimate regional ground-motion models.

The parameters of the model in this work are estimated using Bayesian inference (e.g. [36]). This means that all parameters of the model are associated with a (posterior) distribution, which is a measure of (epistemic) uncertainty associated with the parameters. This uncertainty translates into uncertainty regarding the median predictions, and should not be neglected. Figure 9 shows the posterior distribution of the magnitude scaling for two regions, by plotting the the 5%, 50% and 95% quantiles of the predicted PGA values based on the posterior distribution of the coefficients. We only show two regions because it is difficult to make out differences in the posterior predictions if more than two scaling distributions are shown. It is obvious from Figure 9 that there is some
overlap in the posterior distributions for Northern Greece and Northern Turkey, but this overlap is small. This indicates that the median predictions for both regions are indeed different. However, based on Figure 9 alone, we can only make statements for the two regions for $R_{JB} = 20$km.

To investigate the differences in predictions of the different regions over a broader magnitude/distance range, we look at a Sammon’s map [34, 35]. Therefore, for each region we draw 100 samples from the posterior distribution of the respective coefficients, and calculate predictions at magnitudes $M = 4.5, 5.0, \ldots, 7.5$ and distances $R_{JB} = 5.0, 10.0, 20.0, 50.0, 75.0, 100.0, 150.0, 200.0$. For each sample, this gives a point in a 56-dimensional ground-motion space, which is projected into two dimensions by the Sammon’s map algorithm – in short, a two-dimensional projection is sought that preserves the differences between all points in the high-dimensional space (see [34, 35] for details). In addition to the regional samples, we draw 100 samples from the posterior distribution of the global model. The final map is shown in Figure 10. This map is a representation of the mean difference between any two models over the selected magnitude/distance range. The distance between two points on the map represents the average difference between their corresponding models over the whole magnitude/distance range in ln units. As one can see, the predictions of some regions overlap quite a bit (e.g. North and South Greece), but many regions have distinct predictions, which are also relatively far away from the global model. In particular, the Appenines and Northern Turkey occupy distinct sections on the map, indicating that their predictions are quite different from the other regions. We can also see that the posterior distribution of median predictions for Sicily are relatively broadly distributed, which suggest a larger uncertainty. This is a reflection of the small amount of data for Sicily. Overall, the Sammon’s map shows that the median predictions (both the average and the posterior distribution) are quite different for different regions, which will have effects on hazard calculations, since these average over a wide range of $M/R$-scenarios. In addition, the regionalized model has a lower standard deviation than the global model, which also effects hazard calculations [12].

We assessed the different models using approximations of predictive error (WAIC). This measure estimates the error made for predictions on a new data set. We believe that it is beneficial to look at measures such as these to assess the quality of a model and not only look at residuals and estimated standard de-
Figure 10: Sammon’s map of regionalized predictions, based on samples from the posterior distribution of the coefficients for each region, together with the global model (black). Distance between two points on the map correspond to the average difference in median predictions for two GMPEs (which differ in the values of the coefficients). The axes do not have a physical meaning, only relative differences are of interest. For details on GMPEs and Sammons maps, see [35].
viations. On the other hand, one should not rely on WAIC (or similar measures) alone, since they do not capture the extrapolation capability of a model. This is demonstrated in the model with independent regional coefficients, which performs well in terms of WAIC, but does not extrapolate well to large magnitudes for Sicily (cf. Figure 8).

Data and Resources

Data comes from the Resorce database [6]. Figures are prepared using the program Mathematica (https://www.wolfram.com/mathematica/). The values of $\Delta \hat{elpd}_{WAIC}$ are calculated using the R-package loo, version 0.1.3. The Sammon’s map is calculated using the R-package MASS, version 7.3. Inference is carried out using the program STAN, version 2.5.0 [8].

Acknowledgements

We would like to thank Norm Abrahamson, Christine Goulet and Justin Hollenback for discussions on the topic of regionally varying GMPEs. We would also like to thank the reviewers Sinan Akkar and John Douglas for their thoughtful comments, which greatly helped to improve the manuscript.

References

[1] Abrahamson, N.A., Silva, W.J., Kamai, R.: Summary of the ask14 ground motion relation for active crustal regions. Earthquake Spectra 30(3), 1025–1055 (2014). DOI 10.1193/070913EQS198M.

[2] Akaike, H.: Information Theory And An Extension Of The Maximum Likelihood Principle. In: B. N. Dans and F. C. Petrov (Editors), 2nd International Symposium on Information Theory, Akademiai Kiado, Budapest (1973), Reprinted in: Kotz, S. (ed.) Breakthroughs in Statistics, pp. 610-624. Springer, New York (1992)

[3] Akkar, S., Bommer, J.J.: Empirical equations for the prediction of pga, pgv, and spectral accelerations in europe, the mediterranean region, and the middle east. Seismological Research Letters 81(2), 195–206 (2010). DOI 10.1785/gssrl.81.2.195.

[4] Akkar, S., Cagnan, Z.: A local ground-motion predictive model for turkey, and its comparison with other regional and global ground-motion models. Bulletin of the Seismological Society of America 100(6), 2978–2995 (2010). DOI 10.1785/0120090367.

[5] Akkar, S., Sandikkaya, M.A., Bommer, J.J.: Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in europe and the middle east. Bulletin of Earthquake Engineering 12(1), 359–387 (2014). DOI 10.1007/s10518-013-9461-4.
[6] Akkar, S., Sandikkaya, M.A., Şenyurt, M., Azari Sisi, A., Ay, B.Ö., Traversa, P., Douglas, J., Cotton, F., Luzi, L., Hernandez, B., Godey, S.: Reference database for seismic ground-motion in europe (resorce). Bulletin of Earthquake Engineering 12(1), 311–339 (2014). DOI 10.1007/s10518-013-9506-8.

[7] Al-Atik, L., Abrahamson, N., Bonner, J.J., Scherbaum, F., Cotton, F., Kuehn, N.: The variability of ground-motion prediction models and its components. Seismological Research Letters 81(5), 794–801 (2010). DOI 10.1785/gssrl.81.5.794.

[8] Ambraseys, N.N., Douglas, J., Sarma, S.K., Smit, P.M.: Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from europe and the middle east: Horizontal peak ground acceleration and spectral acceleration. Bulletin of Earthquake Engineering 3(1), 1–53 (2005). DOI 10.1007/s10518-005-0183-0.

[9] Anderson, J.G., Brune, J.N.: Probabilistic seismic hazard analysis without the ergotic assumption. Seismological research letters 70(1), 19–28 (1999).

[10] Atkinson, G.M., Morrison, M.: Observations on regional variability in ground-motion amplitudes for small-to-moderate earthquakes in north america. Bulletin of the Seismological Society of America 99(4), 2393–2409 (2009). DOI 10.1785/0120080223.

[11] Bindi, D., Pacor, F., Luzi, L., Puglia, R., Massa, M., Ameri, G., Paolucci, R.: Ground motion prediction equations derived from the italian strong motion database. Bulletin of Earthquake Engineering pp. 1899–1920 (2011). DOI 10.1007/s10518-011-9313-z.

[12] Bommer, J.J., Crowley, H.: The influence of ground-motion variability in earthquake loss modelling. Bulletin of Earthquake Engineering 4(3), 231248 (2006). DOI 10.1007/s10518-006-9008-z.

[13] Boore, D.M., Stewart, J.P., Seyhan, E., Atkinson, G.M.: Nga-west2 equations for predicting pga, pgv, and 5% damped psa for shallow crustal earthquakes. Earthquake Spectra 30(3), 1057–1085 (2014). DOI 10.1193/070113EQS184M.

[14] Bora, S.S., Scherbaum, F., Kuehn, N., Stafford, P.: Fourier spectral- and duration models for the generation of response spectra adjustable to different source-, propagation-, and site conditions. Bulletin of Earthquake Engineering 12(1), 467–493 (2014). DOI 10.1007/s10518-013-9482-z.

[15] Bozorgnia, Y., Abrahamson, N.A., Atik, L.A., Ancheta, T.D., Atkinson, G.M., Baker, J.W., Baltay, A., Boore, D.M., Campbell, K.W., Chiou, B.S.J., Darragh, R., Day, S., Donahue, J., Graves, R.W., Gregor, N., Hanks, T., Idriss, I.M., Kamai, R., Kishida, T., Kottke, A., Mahin, S.A., Rezaeian, S., Rowshandel, B., Seyhan, E., Shahi, S., Shantz, T., Silva, W., Spudich, P., Stewart, J.P., Watson-Lamprey, J., Wooldell, K., Youngs, R.: Nga-west2 research project. Earthquake Spectra 30(3), 973–987 (2014). DOI 10.1193/072113EQS209M.
[16] Bragato, P.L., Slejko, D.: Empirical ground-motion attenuation relations for the eastern alps in the magnitude range 2.5-6.3. Bulletin of the Seismological Society of America 95(1), 252276 (2005). DOI 10.1785/0120030231.

[17] Campbell, K.W., Bozorgnia, Y.: 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 (2014). DOI 10.1193/062913EQS175M.

[18] Chiou, B., Darragh, R., Gregor, N., Silva, W.: Nga project strong-motion database. Earthquake Spectra 24(1), 23 (2008). DOI 10.1193/1.2894831.

[19] Chiou, B., Youngs, R., Abrahamson, N., Addo, K.: Ground-motion attenuation model for small-to-moderate shallow crustal earthquakes in california and its implications on regionalization of ground-motion prediction models. Earthquake Spectra 26(4), 907 (2010). DOI 10.1193/1.3479930.

[20] Chiou, B.S.J., Youngs, R.R.: 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 (2014). DOI 10.1193/072813EQS219M.

[21] Danciu, L., Tselentis, G.A.: Engineering ground-motion parameters attenuation relationships for greece. Bulletin of the Seismological Society of America 97(1), 162183 (2007). DOI 10.1785/0120040087.

[22] Delavaud, E., Cotton, F., Akkar, S., Scherbaum, F., Danciu, L., Beauval, C., Drouet, S., Douglas, J., Basili, R., Sandikkaya, M.A., Segou, M., Faccioli, E., Theodoulidis, N.: Toward a ground-motion logic tree for probabilistic seismic hazard assessment in europe. Journal of Seismology pp. 451–473 (2012). DOI 10.1007/s10950-012-9281-z.

[23] Derras, B., Bard, P.Y., Cotton, F.: Towards fully data driven ground-motion prediction models for europe. Bulletin of Earthquake Engineering 12(1), 495–516 (2014). DOI 10.1007/s10518-013-9481-0.

[24] Douglas, J.: On the regional dependence of earthquake response spectra. ISET Journal of Earthquake Technology 44(1), 71–99 (2007).

[25] Douglas, J., Akkar, S., Ameri, G., Bard, P.Y., Bindi, D., Bommer, J.J., Bora, S.S., Cotton, F., Derras, B., Hermkes, M., Kuehn, N.M., Luzi, L., Massa, M., Pacor, F., Riggelsen, C., Sandikkaya, M.A., Scherbaum, F., Stafford, P.J., Traversa, P.: Comparisons among the five ground-motion models developed using resorce for the prediction of response spectral accelerations due to earthquakes in europe and the middle east. Bulletin of Earthquake Engineering 12(1), 341–358 (2014). DOI 10.1007/s10518-013-9522-8.

[26] Gelman, A.: Prior distributions for variance parameters in hierarchical models. Bayesian analysis 3, 515–533 (2006).

[27] Gelman, A., Hill, J.: Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press (2006)
[28] Gelman, A., Hwang, J., Vehtari, A.: Understanding predictive information criteria for bayesian models. Statistics and Computing 24(6), 997–1016 (2013). DOI 10.1007/s11222-013-9416-2

[29] Gelman, A., Rubin, D.: Inference from iterative simulation using multiple sequences. Statistical science 7, 457–511 (1992).

[30] Gianniotis, N., Kuehn, N., Scherbaum, F.: Manifold aligned ground motion prediction equations for regional datasets. Computers and Geosciences 69, 72–77 (2014). DOI 10.1016/j.cageo.2014.04.014.

[31] Hermkes, M., Kuehn, N.M., Riggelsen, C.: Simultaneous quantification of epistemic and aleatory uncertainty in gmpes using gaussian process regression. Bulletin of Earthquake Engineering 12(1), 449–466 (2014). DOI 10.1007/s10518-013-9507-7.

[32] Kuehn, N.M., Abrahamson, N.A.: Non-ergodic seismic hazard: Using bayesian updating for site-specific and path-specific effects for ground-motion models. In: CSNI Workshop on Testing PSHA Results and Benefit of Bayesian Techniques for Seismic Hazard Assessment, pp. 84–98 (2015)

[33] Kuehn, N.M., Scherbaum, F.: Ground-motion prediction model building: a multilevel approach. Bulletin of Earthquake Engineering 13(9), 2481–2491 (2015). DOI 10.1007/s10518-015-9732-3.

[34] Sammon, J.W.: A nonlinear mapping for data structure analysis. IEEE Transactions on computers C-18, 401–409 (1969).

[35] Scherbaum, F., Kuehn, N.M., Ohrnberger, M., Kochler, A.: Exploring the proximity of ground-motion models using high-dimensional visualization techniques. Earthquake Spectra 26(4), 1117–1138 (2010). DOI 10.1193/1.3478697.

[36] Spiegelhalter, D., Rice, K.: Bayesian statistics. Scholarpedia, 4(8):5230, 2009. revision #91036. DOI 10.4249/scholarpedia.5230

[37] Stafford, P.J.: Crossed and nested mixed-effects approaches for enhanced model development and removal of the ergodic assumption in empirical ground-motion models. Bulletin of the Seismological Society of America 104(2), 702–719 (2014). DOI 10.1785/0120130145.

[38] Stan Development Team: Stan: A c++ library for probability and sampling, version 2.5.0 (2015). URL http://mc-stan.org

[39] Stan Development Team(b): Stan modeling language users guide and reference manual, version 2.5.0 (2015)

[40] Vehtari, A., Gelman, A.: Waic and cross-validation in stan (May), 1–15 (2014). URL http://www.stat.columbia.edu/~gelman/research/unpublished/waic Stan.pdf

[41] Vehtari, A., Gelman, A., Gabry, J.: Efficient implementation of leave-one-out cross-validation and waic for evaluating fitted bayesian models (2015). URL http://arxiv.org/abs/1507.04544
[42] Watanabe, S.: Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory, Journal of Machine Learning Research (2010) 3571-3594 11, 3571–3594 (2010).