PSHA Workshop
5-7 September 2017 Lenzburg, Switzerland

EARTHQUAKE ENGINEERING RESEARCH CENTRE
UNIVERSITY OF ICELAND

Recalibration and Selection of Ground Motion Models for Seismic Hazard Analyses in Iceland

Schweizerischer Erdbebendienst
Service Sismologique Suisse
Servizio Sismico Svizzero
Swiss Seismological Service

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Milad Kowsari[1], Benedikt Halldorsson[1,2] and Birgir Hrafnkelsson[3]

ABSTRACT

In this study, a Bayesian random effects model that uses a Markov Chain Monte Carlo algorithm for inference is presented to recalibrate several different ground-motion models (GMMs) to the Icelandic dataset. This method is able to partition the aleatory variability into inter-event and intra-event components to account for uneven sampling of the different earthquakes and also correlations of the recorded ground-motions from a single event. The recalibrated GMMs fit the recorded data very well in the distance range where data is available and the residuals of the recalibrated GMMs versus magnitude, distance and frequency are unbiased. The epistemic uncertainties which arise from shortcomings of our models and/or our understanding of the physical causes of data characteristics can be handled by data-driven selection methods which reduce subjectivity and guide the selection process in a quantitative way. Therefore, we review the likelihood-based and the Euclidean distance-based ranking (EDR) methods, before introducing a new procedure on the basis of the deviance information criterion (DIC), for identifying the GMMs that best fit the data. Moreover, the DIC method is not only shown to optimize the selection of GMMs, e.g., for the assesment of earthquake hazard using logic trees for a given region in an unbiased way through the Bayesian statistics.

BAYESIAN STATISTICS USING A MCMC SIMULATION

Finding a GMM which appropriately describes the attenuation of peak-parameters of earthquake strong-motion of the area under study is problematic when performing a seismic hazard analysis. To overcome this problem, Bayesian random effects model that uses a Markov Chain Monte Carlo algorithm can be applied to recalibrate GMMs to the Icelandic dataset. Here, four GMMs including Akkar and Bommer (2010), AB10; Danciu and Tselentis (2007), DT07; Zhao et al. (2006), Zh06 and Lin and Lee (2008), LL08. The difference between the original and recalibrated models at different periods is shown in Figure 1.

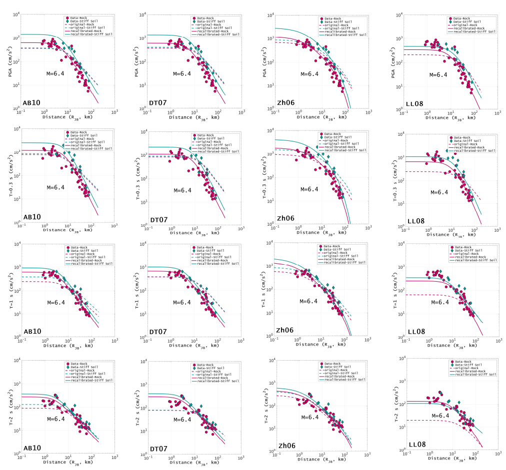

Figure 1.The original and recalibrated GMMs at PGA and Sa at T=0.3, 1 and 2s. The models are evaluated at data's mean magnitude (Mw=6.1). The circles show the recorded data at rock and the diamonds at stiff soil sites.

ALEATORY VARIABILITY

A Bayesian random effect model is applied as one of the motivations of this study to partition the aleatory variability into inter-event and intra-event components, shown in Figure 2.

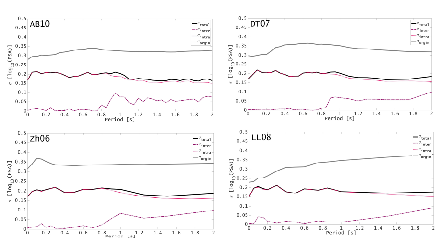

Figure 2.The decomposed variability into an inter-event and an intra-event component, and the total standard deviation for both recalibrated and original GMMs at different periods.

SELECTION OF GMMS

The selection of appropriate GMM is still a major challenge for regions where an appropriate GMM does not exist due to the low seismicity or limited observational data.

I. Scherbaum et al. (2004) pointed out that the likelihood of the normalized residuals is an appropriate measure for the goodness-of-fit of a GMM. Then, they categorized GMMs into four classes based on the obtained LH value and the median, mean and standard deviation of normalized residuals.
II. Scherbaum et al. (2009) suggested an information theoretic approach, called as LLH (log-likelihood), that overcomes several shortcomings of LH method. LLH was not sample size dependent and also it did not require any ad hoc assumptions regarding classification boundaries.
III. These likelihood-based approaches inspired Kale and Akkar (2013) to propose the Euclidean distance-based ranking (EDR) method.

However, when the observed data are accumulated away from the median estimations of the two GMMs, LLH prefers the predictive model with larger sigma, while EDR favors a model with smaller sigma, regardless of what the true uncertainty is (Mak et al. 2014). Therefore, we introduce a new data-driven method for the optimal selection of a GMMs that is not affected by these shortcomings.

DEVIANCE INFORMATION CRITERION

The deviance has an important role in statistical model comparison and is defined as:

$$D(y,\theta) = -2\log p(y|\theta) \tag{1}$$

where, θ is the unknown parameter and y is the observed data.
A summary based on D(y,θ) that does not depend on θ is given by:

$$D_{\hat{\theta}}(y) = D\{y, \hat{\theta}(y)\} \tag{2}$$

where $\hat{\theta}$ is a point estimate for θ such as the mean of the posterior simulations.
Another summary is the posterior mean of D(y,θ), which can be estimated with:

$$\hat{D}_{avg}(y) = \frac{1}{L}\sum_{l=1}^{L} D(y,\theta^l) \tag{3}$$

where θ^l is the l-th draw from posterior distribution.

$$DIC = 2\hat{D}_{avg} - D_{\hat{\theta}}(y) \tag{4}$$

APPLICATION TO GROUND MOTION MODELS

Assuming that the ground-motion data follow a normal distribution:

$$p(y|\beta,\sigma^2) = \prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y_i-\mu(\beta)_i)^2}{2\sigma^2}\right) \tag{5}$$

where, N is the number of observations, y is the observed ground-motions, β_i are the regression coefficients of GMM, $\mu(\beta)$ is the mean value predicted by GMM and σ is the standard deviation of GMM. Therefore, we assumeIt is not obvious to use that the ground motion uncertainty of several other regions should be imposed onto another single and tectonically different region. that σ is unknown and follows a scaled inverse chi-squared distribution because for the normal model, the conjugate prior distribution for σ^2 is scaled inverse-χ^2:

$$p(\sigma^2) \propto (\sigma^2)^{-\frac{\vartheta}{2}-1} \exp\left(-\frac{\vartheta}{2}\cdot\frac{s^2}{\sigma^2}\right) \tag{6}$$

Here, ϑ is the number of chi-squared degrees of freedom and s^2 is the scaling parameter. Therefore, the posterior distribution of σ^2 is given by:

$$p(\sigma^2|y) \propto p(\sigma^2)p(y|\beta,\sigma^2) \tag{7}$$

PERFORMANCE OF THE PROPOSED METHOD

To evaluate the performance of the proposed method, different sets of synthetic data and corresponding fictitious GMMs are simulated. Here, the strong-motion databases have been generated with predefined statistical properties in term of different mean and standard deviations by the median PGA estimates of Ambraseys et al. (2005) for strike-slip earthquakes over the rock site. The datasets are simulated for a M_w6 earthquake and uniformly distributed source-to-site distances up to 100 km. A simple functional form is fitted to the data of each database and finally nine corresponding GMMs are developed to cover all possible situations. Then, the EDR, LLH and DIC methods are applied to rank these models which the scores of the fictitious GMMs and their ranking in addition to the prior and posterior sigma are presented in Tanle 1.

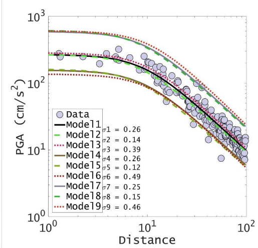

Table 1. Scores of the fictitious GMMs and their ranking based on the different methods.

Models	EDR		LLH		DIC		Sigma	
	Score	Rank	Score	Rank	Score	Rank	Prior	Post.
Model 1	0.316	2	0.096	1	28.479	1	0.260	0.260
Model 2	0.288	1	1.167	3	42.805	2	0.139	0.269
Model 3	0.413	3	0.333	2	55.558	3	0.392	0.278
Model 4	0.922	4	3.108	5	358.930	4	0.260	0.593
Model 5	1.055	5	19.823	6	400.385	6	0.120	0.658
Model 6	1.083	6	1.493	4	384.728	5	0.488	0.632
Model 7	1.550	8	7.279	8	501.065	8	0.254	0.846
Model 8	1.434	7	19.050	9	481.741	7	0.151	0.806
Model 9	1.858	9	3.287	7	547.865	9	0.458	0.951

Figure 3. GMMs developed from simulated datasets and their median comparisons.

CONCLUSIONS

- The GMMs considered in this study underestimate the Icelandic ground motions at short distances from the epicentre and overestimate at distances farther from the epicentre.
- A MCMC method which forms the backbone of modern Bayesian posterior inference is used to infer the GMM parameters and recalibrate the GMMs to the dataset. As a result, all the recalibrated models fit the recorded data very well in the magnitude and distance range and over a range of oscillator different periods.
- The Bayesian random effects model is applied to separate the intra-event and inter-event component which is important to reveal the sources of the variability.
- The DIC overcomes the problems associated with the LLH and EDR methods and can be used to quantify the the quality of fit to the dataset in question.
- This has direct implementation in seismic hazard assessment using different GMMs in a logic tree structure.
- The presented methods are of importance in future seismic hazard analyses for Iceland with implications on the seismic protection of residential buildings and urban infrastructures.

ACKNOWLEDGEMENTS

This study was funded by Grant of Excellence (No. 141261-051/052/053) from the Icelandic centre for research, and partly by the Icelandic Catastrophe Insurance. The support is gratefully acknowledged.

Affiliation: 1 Earthquake Engineering Research Centre & Faculty of Civil and Environmental Engineering, University of Iceland; 2 Icelandic Meteorological Office, Iceland, 3 Faculty of Physical Sciences, University of Iceland.

Earthquake Engineering Research Centre
Austurvegur 2A, IS-800 Selfoss, Iceland
Web page: www.eerc.hi.is · Tel.: +(354) 525 4141 · Fax: +(354) 525 4140