A methodological approach to update Ground Motion Prediction Models using Bayesian Inference

Saran Srikanth Bodda (✉ ssbodda@ncsu.edu)

North Carolina State University <https://orcid.org/0000-0001-5849-7468>

Merlin Keller

EDF Lab Saclay

Abhinav Gupta

NCSU: North Carolina State University

Gloria Senfaute

EDF Lab Saclay

Research Article

Keywords: GMPE, Bayesian model selection, MCMC, Recalibration, ESM database, PSHA

Posted Date: May 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-481786/v1>

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

[Read Full License](#)

Version of Record: A version of this preprint was published at Pure and Applied Geophysics on November 30th, 2021. See the published version at <https://doi.org/10.1007/s00024-021-02915-8>.

A methodological approach to update Ground Motion Prediction Models using Bayesian Inference

Saran Srikanth Bodda^{1*}, Merlin Keller², Abhinav Gupta³, Gloria Senfaute⁴

¹Postdoctoral Researcher, Department of CCEE, North Carolina State University, Raleigh, NC 27695, USA (ssbodda@ncsu.edu)

²EDF R&D, PRISME Dept., 6 quai Watier, 78401 Chatou, France (merlin.keller@edf.fr)

³Director, Center for Nuclear Energy Facilities and Structures, North Carolina State University, Raleigh, NC 27695, USA (agupta1@ncsu.edu)

⁴EDF R&D Paris-Saclay, 7 Boulevard Gaspard Monge, 91120 Palaiseau, France (gloria.senfaute@edf.fr)

*Corresponding author

Abstract

In recent decades, prediction of ground motion at a specific site or a region is of primary interest in probabilistic seismic hazard assessment (PSHA). Historically, several ground motion prediction equation (GMPE) models with different functional forms have been published using strong ground motion records available from NGA-West and European databases. However, low-to-moderate seismicity regions, such as Central & Eastern United States and western Europe, is characterized by limited strong-motion records in the magnitude-distance range of interest for PSHA. In these regions, the available data for the development of empirical GMPEs is very scarce and limited to small magnitude events. For these regions, the general practice in PSHA is to consider a set of GMPEs developed from data sets collected in other regions with high seismicity. This practice generates an overestimation of the seismic hazard for the low seismicity regions. There are two potential solutions to overcome this problem: (i) a new GMPE model can be developed; however, development of such a model can require significant amount of data which is not usually available, and (ii) the existing GMPE models can be recalibrated based on the data sets collected in the new region rather than developing a new GMPE model. In this paper, we propose a methodological approach to recalibrate the coefficients in a GMPE model using different algorithms to perform Bayesian inference. The coefficients are recalibrated for a subset of European Strong-Motion (ESM) database that corresponds to low-to-moderate seismicity records. In this study, different statistical models are compared based on the functional form given by the chosen GMPE, and the best model and algorithm are recommended using the concept of information criteria.

Keywords

GMPE; Bayesian model selection; MCMC; Recalibration; ESM database; PSHA

DECLARATIONS

Funding

This research was supported by Électricité de France (EDF) and Center for Nuclear Energy Facilities and Structures (CNEFS) at North Carolina State University.

Conflicts of interest/Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and material

All of the data and models that support the findings of this study are available.

Code availability

All the custom codes that support the findings of this study are available.

Authors' contributions

All authors contributed to the study conception and design in a significant manner. Material preparation, data collection and analysis was performed by Saran Bodda. The first draft of the manuscript was written by Saran Bodda and remaining authors reviewed the manuscript carefully and provided comments in multiple meetings to improve the quality of the manuscript. All authors read and approved the final manuscript.

1 Introduction

The expected ground motion at a site is represented by an attenuation relationship or a ground motion prediction equation (GMPE). GMPEs estimate the intensity of ground shaking and their underlying uncertainty based on ground motion characteristics such as earthquake magnitude, propagation path and soil conditions, style-of-faulting, and other seismological parameters. In recent years, many methodologies (Bertin et al., 2020; Kowsari et al., 2019, 2020; Wang and Takada, 2009) based on Bayesian inference have been used to update the model parameters by accounting for various sources of uncertainty. Wang and Takada (2009) updated an existing GMPE model for site-specific ground motion records based on K-NET and KiK-net databases. The functional form of the original GMPE is considered as fixed and a linear correction term function of magnitude and distance is added to it. The unknown parameters are estimated using analytical formulations and the predictions obtained from the calibrated model are found out to be unbiased. In a recent study, Bertin et al. (2020) calibrated nine existing GMPEs against RESORCE-2013 database. The GMPE models are considered as fixed and a bias term is added to the GMPE functional form. The bias term and the variance of GMPE model are estimated using Metropolis Hastings algorithm. Kowsari et al. (2019) updated seven GMPEs of different functional forms based on strong-motion PGA data from the South Iceland Seismic Zone. The posterior probability density function of all regression parameters are estimated using Metropolis within Gibbs Markov Chain Monte Carlo (MCMC) sampler and the priors are chosen as non-informative. More recently, Kowsari et al. (2020) further improved upon Kowsari et al. (2019) and updated four GMPEs at various periods ranging from 0.05s to 5s. In Kowsari

et al. (2020), informative priors are used to update higher order regression parameters of GMPE, and Delayed Rejection and Adaptive Metropolis (DRAM) algorithm is implemented for efficient sampling of highly correlated regression parameters.

In the aforementioned studies (Bertin et al., 2020; Kowsari et al., 2019, 2020; Wang and Takada, 2009), the original GMPEs are calibrated for new data set either by (i) fixing the GMPE functional form and estimating additional terms to account for bias correction, or (ii) estimating all the regression parameters in the GMPE model. The former approach is easy to implement and the bias correction terms can be estimated using analytical formulations. The second approach to estimate all the regression parameters in a non-linear GMPE functional form can be challenging, computationally demanding, and it requires careful selection of MCMC algorithms. However, none of these studies compared the performance of both these approaches against a new testing data set. Moreover, the non-linearity in a generic GMPE functional form which is a function of moment magnitude (M_W) and distance (R) terms usually comes from a regression coefficient (h) that represents the pseudo-depth parameter (see Eq. 1). Therefore, the performance of the recalibrated GMPE model with h being a fixed (linear functional form) parameter versus h not being a fixed (non-linear functional form) parameter must be studied.

$$\mathbb{E}[\log y] = b_0 + (b_1 + b_2 M_W) \log \left(\sqrt{R^2 + h^2} \right) + b_3 M_W + b_4 M_W^2 \quad (1)$$

where, $\log y$ is the logarithm of the ground motion intensity measure of interest, M_W is the moment magnitude, R is the distance. $b_0, b_1, b_2, b_3, b_4, h$ are the coefficients of the GMPE model.

In this research, we propose a Bayesian methodology to update GMPE internal parameters or coefficients for ground motion records corresponding to a new region. The parameters are estimated using different algorithms and the effect of prior information (weakly informative vs informative priors) on the GMPE predictions is evaluated. In this study, different statistical models are compared based on the functional form given by the chosen GMPE and the best model is selected using several information criteria. The proposed methodology is implemented for European Strong-Motion (ESM) database. This manuscript is organized as follows: Section 2 presents different GMPE functional forms and the selection criteria used for selecting the ground motion records from ESM database. Section 3 describes the proposed methodology for Bayesian update of GMPE models. Then, the Bayesian estimation methods and the information criteria used for model selection is described in Section 4. The application of the proposed methodology to a realistic GMPE model is shown in Section 5. Finally, summary and conclusions of this study are presented in Section 6.

2 GMPE Functional Form and Database

In this study, we have selected the GMPE by [Bindi et al. \(2014\)](#) (referred to as BI14 in the rest of the manuscript) and the functional form of the GMPE model is shown below:

$$\begin{aligned}\log y_i &= \log \left(10^{f_{BI14}(x_i)-2} / g \right) + \varepsilon_i \\ &= f_{BI14}(x_i) \log 10 - 2 \log 10 - \log g + \varepsilon_i\end{aligned}\tag{2}$$

where, y_i is the i -th ground motion intensity measure of interest (PGA or pseudo spectral acceleration (PSA) at a given period in terms of g) of the considered database, $i = 1, \dots, n$, and $f_{BI14}(x_i)$ is the intensity measure predicted by BI14, given a vector x_i of characteristics of the earthquake; $\varepsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$ is a stochastic error term, representing the variability of BI14 prediction with respect to observed records. The deterministic part of the GMPE model is given by:

$$f_{BI14}(x_i) = e_1 + f_D(R_i, M_i) + f_M(M_i) + f_S(V_{s,30,i}) + f_{SOF}(FC_i),\tag{3}$$

where $x = (M_i, R_i, V_{s,30,i}, FC_i)$, corresponds to:

- moment magnitude M_i (real values)
- distance R_i (real values)
- shear-wave velocity $V_{s,30,i}$ (real values)
- and fault class FC_i (binary values).

The distance function f_D includes a magnitude-dependent geometrical spreading term and is given by:

$$f_D(R_i, M_i) = [c_1 + c_2(M_i - M_{ref})] \log_{10} \left(\sqrt{R_i^2 + h^2} / R_{ref} \right) - c_3(\sqrt{R_i^2 + h^2} - R_{ref})\tag{4}$$

where, c_1, c_2, c_3, h are the coefficients to be determined. R_{ref} is the reference distance at which near-source predictions start decreasing. M_{ref} is the reference magnitude (close to 50th percentile of the cumulative number of recordings vs magnitude).

The magnitude scaling f_M is given by:

$$f_M(M_i) = \begin{cases} b_1(M_i - M_h) + b_2(M_i - M_h)^2 & \text{for } M_i \leq M_h \\ b_3(M_i - M_h) & \text{otherwise} \end{cases}\tag{5}$$

where, b_1, b_2, b_3 are the coefficients to be determined. M_h is the hinge magnitude for the shape of the magnitude scaling, is a coefficient to be set during the analysis.

The site amplification term f_S is given by:

$$f_S(V_{s,30,i}) = \gamma \log_{10}(V_{s,30,i}/V_{ref}) \quad (6)$$

where, γ is the coefficient to be determined. V_{ref} is the reference shear-wave velocity.

The style-of-faulting term f_{SOF} is given by:

$$f_{SOF}(FC_i) = f_1 F_{N,i} + f_2 F_{R,i} + f_3 F_{S,i}$$

$$F_{N,i}, F_{R,i}, F_{S,i} = \begin{cases} 1 & \text{for normal, reverse and strike-slip, respectively} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where, f_1, f_2, f_3 are the coefficients to be determined. $F_{N,i}, F_{R,i}, F_{S,i}$ are dummy variables used to represent fault mechanism classes of normal, reverse, and strike-slip, respectively.

Database

The coefficients in BI14 are originally calibrated using RESORCE database (Akkar et al., 2014). BI14 is valid for a moment magnitude range of [4; 7.6] and distances less than 300 km. Recently, Lanzano et al. (2019), developed an Engineering Strong-Motion (ESM) flatfile for pan-European region that consists of 23,014 recordings from 2179 earthquakes and 2080 stations. The comparison of data distribution between RESORCE and ESM databases is shown in Figure 1 (Kotha et al., 2020). As seen in Figure 1, the ESM database has more records compared to RESORCE database in the low-to-moderate seismicity region ($3 \leq M_W \leq 5.2$) and for distances less than 600 km. Therefore, in this study, we use a subset of ESM database to update the BI14 coefficients. The selection criteria for moment magnitude [3; 5.2], distance [1; 600 km], $V_{s,30}$ velocity [800; 1200 m/s], and fault mechanism [normal, strike-slip and reverse] is chosen such that the selected records correspond to low-to-moderate seismicity records such as French territory. This resulted in a subset of 3154 observations, out of which 2200 records are used for updating the BI14 coefficients and 954 records are used for testing the performance of the recalibrated BI14 model.

Figure 1: Comparison of RESORCE (yellow) and ESM (black) Databases (Kotha et al., 2020)

3 Proposed Methodology

The proposed methodology employs four key stages that are described below:

Stage 1: Assessment Base

- Develop an assessment base of ground motion records for a region or site-specific location of interest. Some of the widely used databases around the world are RESORCE (Akkar et al., 2014), ESM (Lanzano et al., 2019), NGA-West2 (Ancheta et al., 2014), NGA-East (Goulet et al., 2014), K-NET and KiK-net (NIED, 2019) databases.
- For a given application domain of interest, select a subset of records from the complete database based on earthquake characteristics.
- Divide the selected subset records into two data sets: (i) training data set – for updating the coefficients of GMPE, and (ii) testing data set – for testing the performance of the recalibrated GMPE.

Stage 2: Evaluation Models

- Select GMPE models from the existing literature that are applicable for the target seismotectonic context.

Stage 3: Model Updating

When a GMPE model calibrated on a particular data set predicts response on a new unseen data set, some kind of discrepancy always exists between the observed response and the predicted response. The process of estimating this discrepancy (or bias correction) and any unknown parameters in the model is known as inverse uncertainty quantification (Nagel, 2019; Wu et al., 2018). In the proposed methodology, the following types of statistical models are updated and subsequently compared in terms of information criteria:

- *Fixed parameter setting with added bias:* In this setting (Eq. 8), the functional form of original GMPE model is fixed and a bias correction term μ_k is added. Since $f_{BI14}(x_i)$ is fixed, the unknown parameters that need to be estimated are $\theta = (\mu, \sigma^2)$.

$$\log y_i = \log \left(10^{f_{BI14}(x_i) + \mu - 2} / g \right) + \varepsilon_i, \quad \varepsilon_i \sim N(\mu, \sigma^2) \quad (8)$$

- *Uncertain parameter setting:* In this setting (Eq. 9), the functional form of the GMPE model is not fixed and one or more parameters in the model are updated based on the new observations. The unknown parameters that needs to be estimated are $\theta = (\eta, \sigma^2)$.

$$\log y_i = \log \left(10^{f_{BI14}(x_i, \eta) - 2} / g \right) + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \quad (9)$$

where, η is a vector of regression coefficients (aka parameters) in BI14.

Stage 4: Assess Evaluation Model Adequacy

- Compare/rank the updated statistical models using information criteria (described in section 4.2).
- Test the performance of the recalibrated GMPE models using the testing data set.

4 Estimation Methods and Information Criteria

The concept of Bayesian inference has been widely used in many applications to estimate unknown parameters with uncertainty (Bodda et al., 2020a,b; Keller et al., 2014; Vaishnav et al., 2020; Viallet et al., 2017). Bayesian inference is a type of statistical inference method in which the probabilities or distribution of uncertain parameters are updated as new evidence or information becomes available. The posterior density is computed using Bayes theorem (Reich and Ghosh, 2019):

$$p(\theta|D) = \frac{f(D|\theta)\pi(\theta)}{m(D)} \propto f(D|\theta)\pi(\theta) \quad (10)$$

where,

$D = (x_1, x_2, \dots, x_n)$, vector of observed data

$\pi(\theta)$ = prior density of θ

$f(D|\theta)$ = Likelihood or sampling density of D given θ

$m(D) = \int f(D|\theta)\pi(\theta)d\theta$, marginal density of D

$p(\theta|D)$ = posterior density of θ given D

Except in very simple cases, $m(D)$ is not available in closed-form, making the posterior density intractable. That is why, in many cases widely used numerical methods such as Markov Chain Monte Carlo (MCMC) methods or Importance sampling (IS) methods are employed to generate many posterior samples which approximates the true posterior distribution. MCMC and IS are a family of methods, rather than one particular algorithm. Some of the well-known MCMC algorithms are Metropolis-Hastings algorithm, Gibbs sampling, Slice sampling etc. (Reich and Ghosh, 2019). The sampling importance resampling (SIR) algorithm is one of the most commonly applied IS methods to nonlinear Bayesian problems (Heine, 2005; Keller et al., 2014). The estimation methods used for Bayesian update of GMPE parameters are described in the following section.

4.1 Estimation Methods

Bayesian Linear Models (BLM) using Conjugate Priors: This approach is applicable for GMPE models with linear functional form and the likelihood function with a known conjugate prior. In this study, multivariate normal distribution is chosen for regression coefficients η , inverse gamma distribution is chosen for variance σ^2 , and the likelihood is that of a normal distribution. Due to conjugacy, the joint posterior distribution of $\theta = (\eta, \sigma^2)$ follows a normal-inverse-gamma distribution. The derivation of conjugate priors are shown in detail in Banerjee (2008) and Bodda et al. (2021). This method is the most convenient choice if it is applicable for the given scenario.

Sampling Importance Resampling with Laplace Approximation: In Laplace approximation method, the posterior distribution is asymptotically approximated using a multivariate normal distribution. Then, the Laplace approximated posterior is chosen as a proposal distribution $q(\theta|D)$ in importance sampling method. Next, a large number of samples are generated from the proposal distribution. Each sample is then weighted by calculating the relative likelihood of getting that sample from the target distribution to that of the proposal distribution. An approximated independent and identically distributed (iid) sample from the posterior density of θ can be obtained by resampling from the samples with probabilities proportional to their weights. This procedure is known as sampling importance resampling (SIR), and unlike other MCMC approaches, SIR avoids tuning and convergence issues. However, the quality of the weighted samples in SIR is measured by the effective sample size (ESS), with $1 < ESS < S \equiv$ the size of an equivalent iid posterior

sample (Keller et al., 2014).

$$\begin{aligned}
&\text{Proposal density: } q(\theta|D) \sim \mathcal{MVN}(\tilde{\theta}, -H^{-1}) \\
&\text{Sample weights: } w_i = \frac{p(\theta_i|D)/q(\theta_i|D)}{\sum_{\theta} p(\theta_i|D)/q(\theta_i|D)} \\
&\text{Effective sample size: } ESS = \frac{(\sum_i w_i)^2}{\sum_i w_i^2}
\end{aligned} \tag{11}$$

where, $\tilde{\theta}$ is the mode of $p(\theta|D)$, $H = \nabla^2 \log p(\theta|D)|_{\theta=\tilde{\theta}}$ is a Hessian matrix.

MCMC Algorithms: In addition to SIR, we tested two MCMC algorithms: Adaptive Metropolis (AM) and Automated factor slice sampling (AFSS) for Bayesian update of GMPE internal parameters. The algorithms are implemented in R using LaplacesDemon package (Statisticat and LLC., 2020). In this study, MCMC convergence diagnostics is assessed by visual inspection of trace plots and using formal diagnostic measures such as Autocorrelation plots.

4.2 Information Criteria

Information criteria provide scoring values to candidate models based on their performance on training data set and complexity of the model (Gelman et al., 2013, 2014). The scores are calculated using likelihood-based measures such as deviance. The deviance of the data given parameters is defined as twice the negative log likelihood, $Dev(D|\theta) = -2\log[f(D|\theta)]$. In this study, we use deviance information criteria (*DIC*) and Watanabe–Akaike (or widely applicable) information criteria (*WAIC*) for model comparison. *DIC* uses MCMC results directly for Bayesian model comparison. The model complexity is represented by effective number of parameters p_D and the model fit is represented by posterior mean of deviance $\overline{Dev}$. The models with small *DIC* are preferred as they fit well (small $\overline{Dev}$) and are simple (small p_D).

$$\begin{aligned}
DIC &= \overline{Dev} + p_D = \widehat{Dev} + 2p_D \\
p_D &= \overline{Dev} - \widehat{Dev}
\end{aligned} \tag{12}$$

where, $\widehat{Dev} = Dev(D|\hat{\theta})$ and $\hat{\theta}$ is the posterior mean of θ .

Like *DIC*, *WAIC* estimates the out-of-sample expectation by computing log pointwise posterior predictive density (*lppd*) and then adding a correction term p_W for the effective number of parameters to adjust for overfitting. Similar to *DIC*, models with small *WAIC* are preferred as they fit well and are simple.

$$\begin{aligned}
WAIC &= -2(lppd - p_W) \\
lppd &= \sum_{i=1} \log E[f(D_i|\theta)|D] \\
p_W &= \sum_{i=1} \text{Var}[\log(f(D_i|\theta))|D]
\end{aligned} \tag{13}$$

5 Application

In this study, three statistical models (M1, M2, M3) based on the functional form of the BI14 GMPE are considered (see Eq. 14). M1 refers to the fixed parameter setting with added bias term. M2 refers to the uncertain parameter setting with coefficient h being fixed in Eq. (4). Therefore, M2 becomes a linear model. M3 is same as M2 with h not being fixed, making it as a nonlinear model. In both M2 and M3, coefficient b_3 in Eq. (5) is not updated because the training data set does not contain data that has magnitude M_W more than hinge magnitude $M_h = 6.75$. Also, coefficient f_3 in Eq. (7) is not updated as the third fault class (F_S strike-slip) or categorical variable is redundant. Coefficients c_3 and h are constrained to be non-negative. In the following, we note $\eta = (e_1, c_1, c_2, c_3, b_1, b_2, \gamma, f_1, f_2)$. The three considered models are defined formally as follows, together with the corresponding set of uncertain parameters to be determined:

$$\begin{aligned} \text{M1: } \log y_i &= \log \left(10^{f_{BI14}(x_i) + \mu_{BI14} - 2} / g \right) + \varepsilon_{BI14,i}, \quad \theta_{M1} = (\mu_{BI14}, \sigma_{BI14}) \\ \text{M2: } \log y_i &= \log \left(10^{f_{BI14}(x_i, \eta) - 2} / g \right) + \varepsilon_{BI14,i}, \quad \theta_{M2} = (\eta, \sigma_{BI14}) \\ \text{M3: } \log y_i &= \log \left(10^{f_{BI14}(x_i, \eta, h) - 2} / g \right) + \varepsilon_{BI14,i}, \quad \theta_{M3} = (\eta, h, \sigma_{BI14}) \end{aligned} \quad (14)$$

5.1 Parameter Updating

The 5% damped peak ground acceleration (PGA) and pseudospectral acceleration (PSA) for periods $T = 0.2s, 1s$ are computed using OpenQuake Engine (Pagani et al., 2014) and GMPE-SMTK toolkit (Weatherill et al., 2014). The PGA and PSA observations are obtained by calculating the geometric mean of the two horizontal components. The parameters are estimated/recalibrated using four algorithms (BLM, SIR, AM, and AFSS) described in section 4.1. BLM algorithm is implemented only for models M1 and M2. For models M1, M2, and M3 the parameters are first estimated using weakly informative priors: $\beta_{BI14} \sim \mathcal{N}(0, 10^2)$, $\mu_{BI14} \sim \mathcal{N}(0, 10^2)$, $\sigma_{BI14}^2 \sim \mathcal{IV}(a = 0.001, b = 0.001)$.

For M2 and M3, the parameters are also recalibrated based on informative normal-inverse gamma priors. The prior for regression coefficients is a normal distribution with mean equal to their original value and variance equal to the total variance of BI14 (Table 1). The samples from MCMC (AM; AFSS) algorithms are obtained by considering two parallel chains with (100,000; 5000) iterations for each chain and a thinning interval of (50; 2). Then, a burn-in of 500 iterations is applied for the output MCMC samples after thinning. The first chain is initiated using the original values of regression coefficients and the second chain is initiated by specifying all the parameter values equal to 0.1 (random initial values can also be chosen). In Laplace approximation, the mode $\tilde{\theta}$ of posterior distribution $p(\theta|D)$ is estimated using Trust Region optimization method (Statisticat and LLC., 2020). Next, 10,000 random samples are generated using the Laplace approximated posterior distribution. In SIR, the posterior density of parameters is estimated from the above samples by resampling the samples with a size equal to the effective sample size (ESS) and probabilities proportional to the sample weights (Eq. 11).

The results of the recalibrated BI14 regression coefficients estimated using various algorithms for models M2 and M3 are tabulated in [Table 1](#) and [Table 2](#), respectively. For comparison, the original coefficients of the BI14 are also provided in this table along with the mean and standard deviation (given in brackets) obtained from the posterior distribution of regression coefficients. The BLM approach estimates the regression coefficients without any approximations where as the results obtained from the MCMC and SIR algorithms are based on approximations. As seen in [Table 1](#), most of the recalibrated regression coefficients estimated using different algorithms are close to each other. However, the accuracy of the algorithms are compared in [Section 5.3](#) in a formal way using the concept of information criteria presented in [Section 4.2](#). As observed in [Table 1](#), the standard deviation of coefficients is greater for uninformative priors, compared to the informative priors. Therefore, it is better to use informative priors for GMPE parameter recalibration.

Next, the coefficient h that is causing non-linearity in M3 is examined as shown in [Table 2](#). The recalibrated value of h is very close its original value indicating that the new data set does not have much effect on the estimation of coefficient h . Therefore, the predictions estimated using non-linear model M3 can be similar to the predictions estimated using linear model M2. We also see an increase in total sigma σ_T of the recalibrated GMPE compared to the original BI14 GMPE. This is because the BI14 GMPE is originally derived from the RESORCE database whereas the ESM database comprises of records from RESORCE database, and many other additional records from the pan-European region. Therefore, increasing amount of data implies increasing the spatio-temporal diversity of ground motion records, and thus an increase in total sigma σ_T for the recalibrated BI14 GMPE on the ESM training database ([Kotha et al., 2020](#)).

Table 1: Comparison of original BI14 coefficients with the M2 recalibrated coefficients for PSA ($T = 1s$)
(the standard deviation of coefficients are given in brackets)

Coefficients	Org.	<i>M2: Linear</i>			
		BLM	SIR	AM	AFSS
e_1	3.1247				
Uninformative		2.63 (0.224)	3.7095 (0.456)	3.3144 (0.407)	3.689 (0.441)
Informative		3.5769 (0.148)	3.2312 (0.252)	3.1479 (0.177)	3.226 (0.236)
c_1	-1.0527				
Uninformative		-1.0379 (0.095)	-1.0379 (0.095)	-0.9743 (0.103)	-1.0458 (0.097)
Informative		-0.7448 (0.082)	-0.9903 (0.077)	-1.0511 (0.069)	-0.9933 (0.078)
c_2	0.1035				
Uninformative		0.2357 (0.061)	0.2363 (0.062)	0.2623 (0.069)	0.2288 (0.065)
Informative		0.4395 (0.051)	0.2746 (0.048)	0.2429 (0.067)	0.2687 (0.053)
c_3	0				
Uninformative		7e-4 (2.8e-4)	6.6e-4 (2.8e-4)	8.7e-4 (5.9e-4)	6.7e-4 (2.9e-4)
Informative		0.001 (2.8e-4)	6.9e-4 (2.8e-4)	5.8e-4 (3.1e-4)	7.1e-4 (2.7e-4)
b_1	0.3066				
Uninformative		0.9674 (0.295)	0.9669 (0.3)	0.6894 (0.258)	0.9317 (0.319)
Informative		-0.1853 (0.194)	0.6336 (0.175)	0.4649 (0.186)	0.6188 (0.177)
b_2	-0.1476				
Uninformative		0.1012 (0.054)	0.1014 (0.054)	0.0481 (0.063)	0.0905 (0.066)
Informative		-0.0643 (0.039)	0.0448 (0.039)	-0.0052 (0.055)	0.0391 (0.046)
γ	-0.8266				
Uninformative		-1.2179 (0.191)	-1.2138 (0.188)	-1.1921 (0.219)	-1.211 (0.19)
Informative		-1.0783 (0.166)	-1.1217 (0.166)	-1.1093 (0.168)	-1.1282 (0.167)
f_1	0.0263				
Uninformative		-0.0034 (0.021)	-0.003 (0.022)	-0.005 (0.04)	-0.0032 (0.021)
Informative		0.0004 (0.022)	-0.0024 (0.021)	0.003 (0.04)	-0.0031 (0.022)
f_2	0.0186				
Uninformative		-0.081 (0.027)	-0.0808 (0.027)	-0.0721 (0.044)	-0.08 (0.027)
Informative		-0.0794 (0.027)	-0.0807 (0.027)	-0.0608 (0.091)	-0.0806 (0.028)
h	4.4161				
Uninformative		—	—	—	
Informative		—	—	—	
σ_T	0.3561				
Uninformative		0.8672 (0.151)	0.8698 (0.013)	0.8651 (0.049)	0.8701 (0.013)
Informative		0.8842 (0.154)	0.8699 (0.013)	0.8649 (0.046)	0.8702 (0.017)

Table 2: Comparison of original BI14 coefficients with the M3 recalibrated coefficients for PSA ($T = 1s$)
(the standard deviation of coefficients are given in brackets)

Coefficients	Org.	<i>M2: Linear</i>	<i>M3: Non-Linear</i>		
		BLM	SIR	AM	AFSS
e_1	3.1247				
Uninformative		2.63 (0.224)	3.7075 (0.469)	3.19 (0.277)	3.2189 (0.243)
Informative		3.5769 (0.148)	3.2299 (0.253)	3.1023 (0.14)	3.2383 (0.248)
c_1	-1.0527				
Uninformative		-1.0379 (0.095)	-1.039 (0.111)	-0.9946 (0.071)	-0.9877 (0.076)
Informative		-0.7448 (0.082)	-0.9902 (0.079)	-1.0104 (0.052)	-0.9904 (0.078)
c_2	0.1035				
Uninformative		0.2357 (0.061)	0.2377 (0.063)	0.2589 (0.083)	0.2764 (0.045)
Informative		0.4395 (0.051)	0.2751 (0.047)	0.2775 (0.044)	0.274 (0.048)
c_3	0				
Uninformative		7e-4 (2.8e-4)	6.5e-4 (3.3e-4)	7.5e-4 (4.2e-4)	7e-4 (2.7e-4)
Informative		0.001 (2.8e-4)	6.8e-4 (2.8e-4)	5.5e-4 (2e-4)	6.9e-4 (2.8e-4)
b_1	0.3066				
Uninformative		0.9674 (0.295)	0.9628 (0.302)	0.5825 (0.193)	0.6252 (0.174)
Informative		-0.1853 (0.194)	0.6328 (0.176)	0.5032 (0.063)	0.6385 (0.172)
b_2	-0.1476				
Uninformative		0.1012 (0.054)	0.1011 (0.055)	0.0271 (0.047)	0.0437 (0.039)
Informative		-0.0643 (0.039)	0.0449 (0.039)	0.0185 (0.019)	0.0455 (0.038)
γ	-0.8266				
Uninformative		-1.2179 (0.191)	-1.2146 (0.191)	-1.1872 (0.192)	-1.1275 (0.168)
Informative		-1.0783 (0.166)	-1.1209 (0.166)	-1.1216 (0.189)	-1.1236 (0.168)
f_1	0.0263				
Uninformative		-0.0034 (0.021)	-0.003 (0.021)	0.0067 (0.05)	-0.0037 (0.022)
Informative		0.0004 (0.022)	-0.0026 (0.021)	-0.0018 (0.024)	-0.0025 (0.022)
f_2	0.0186				
Uninformative		-0.081 (0.027)	-0.0808 (0.027)	-0.0622 (0.08)	-0.081 (0.027)
Informative		-0.0794 (0.027)	-0.0811 (0.027)	-0.0811 (0.03)	-0.0801 (0.027)
h	4.4161				
Uninformative		—	4.5108 (1.205)	4.5272 (0.999)	4.4313 (0.339)
Informative		—	4.4215 (0.337)	4.4644 (0.316)	4.4102 (0.332)
σ_T	0.3561				
Uninformative		0.8672 (0.151)	0.87 (0.013)	0.8644 (0.054)	0.8697 (0.013)
Informative		0.8842 (0.154)	0.8698 (0.013)	0.8699 (0.013)	0.87 (0.013)

(a) Marginal Density

(b) Trace Plots

(c) Autocorrelation Plots

Figure 2: Convergence diagnostics plots of Adaptive Metropolis chains for PSA ($T = 1s$)

5.2 Convergence Diagnostics

Figure 2 shows the AM convergence diagnostic plots of all the regression coefficients in M3. The trace plots shown in Figure 2 have inefficient chains with poor mixing especially for coefficients $b_1, b_2, c_1, c_2, c_3, e_1$ and the autocorrelation is high even for lag of 50 indicating a very slow mixing of AM samples. The autocorrelation is very high when all the initial values of AM algorithm are initialized to 0.1 compared to the initialization of parameters equal to the original values of coefficients. To further examine the poor mixing of regression coefficients, the Pearson's correlation coefficient matrix is plotted using the posterior samples obtained from AM algorithm.

Figure 3: Pearson's correlation matrix for posterior samples – PSA ($T = 1s$)

As seen in Figure 3 (a), there is a strong correlation between regression coefficients $b_1, b_2, c_1, c_2, c_3, e_1$. AM algorithm suffers under high correlations as the tuning of proposal covariance can be often too slow to reach a covariance that approximates the target posterior density very well (Turner et al., 2013). On the other hand, AFSS is an efficient MCMC sampler even when there is a strong correlation between parameters as shown in Figure 3 (b) (Tibbitts et al., 2014). Figure 4 shows the AFSS convergence diagnostic plots of all the regression coefficients in M3. The samples from AFSS gives similar estimates for all the chains, and the autocorrelation is near zero for lag of 5 and beyond indicating a good convergence of AFSS samples. Also, the AFSS algorithm performs well even when all the parameters are initialized to 0.1. The convergence diagnostics for SIR algorithm is measured by the effective sample size (ESS, Eq. 11). In this study, the ESS for all the cases is in the range $[8300 - 9600]$, indicating a very good convergence of results. The estimates obtained using SIR algorithm are similar to the AFSS estimates as shown in Table 1. For the sake of simplicity, only the recalibrated results obtained for the PSA ($T = 1s$) are given here. However, we see similar observations as PSA ($T = 1s$) for the results of PGA and PSA ($T = 0.2s$).

Figure 4: Convergence diagnostics plots of AFSS chains for PSA ($T = 1s$)

5.3 Model Selection

In this section, the statistical models M1, M2, and M3 are compared based on their goodness of fit and model complexity using information criteria. Table 3 shows the DIC and $WAIC$ scores for all the different types of models and estimation methods considered in this study. The $WAIC$ scores for M2 and M3 are greater than their DIC scores when the parameters are estimated using AM algorithm. $WAIC$ is evaluated

by averaging over the posterior distribution where as DIC is estimated by conditioning on a point estimate (Gelman et al., 2014). As there is a lot of uncertainty in the predictions estimated using AM algorithm due to bad convergence, $WAIC$ captures the overall uncertainty and gives a high score. The $WAIC$ and DIC scores are close to each other for M2 and M3 when the parameters are estimated using BLM, SIR, and AFSS algorithms. Therefore, it is better to select the linear model (M2) rather than non-linear model (M3) for Bayesian recalibration of parameters. Also, BLM algorithm is preferred over AFSS or SIR algorithms because BLM provides results without approximations and is computationally efficient for large data sets. If the non-linear functional form of the GMPE is considered for Bayesian recalibration of coefficients, then SIR algorithm can be employed for better computational efficiency.

Table 3: Comparison of DIC and $WAIC$ scores

Model	Estimation Method	PGA		PSA (0.2s)		PSA (1s)	
		DIC	$WAIC$	DIC	$WAIC$	DIC	$WAIC$
M1: Added-Bias	BLM	6712.30	6713.18	6742.08	6743.01	5987.16	5988.64
	SIR	6716.35	6716.25	6746.15	6746.08	5991.22	5991.71
	AM	6722.23	6721.19	6749.61	6750.59	5992.52	5993.08
	AFSS	6716.59	6716.43	6746.07	6746.00	5991.18	5991.69
M2: Linear	BLM	6139.22	6140.53	6322.09	6323.29	5633.46	5635.42
	SIR	6140.92	6141.92	6324.39	6325.11	5633.72	5635.73
	AM	6168.59	6171.24	6669.41	9611.83	5793.69	6608.39
	AFSS	6150.25	6144.02	6326.97	6325.25	5636.17	5630.25
M3: Non-Linear	SIR	6135.53	6136.56	6321.81	6322.55	5633.96	5636.01
	AM	6403.54	8584.04	6704.76	9917.72	5768.62	6378.53
	AFSS	6142.98	6136.03	6323.77	6323.94	5638.72	5608.87

5.4 Testing

In the final step of the proposed methodology, the performance of the recalibrated GMPE is tested against the testing data set $(\tilde{x}, \tilde{y})$. The testing data set is not used in the recalibration of GMPE parameters. Figure 5 shows the root mean squared error (RMSE, Eq. 15) estimates obtained from different functional forms of BI14: M1, M2, and M3 considered in this study along with the original functional form of BI14 (M0) for comparison.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\tilde{y}_i - E[p(\tilde{y}_i|\tilde{x}_i, y_{train})])^2}{N}} \quad (15)$$

where, $E[p(\tilde{y}_i|\tilde{x}_i, y_{train})]$ is the expected value of posterior predictive distribution of the recalibrated

statistical models.

As seen in Figure 5, the model predictions has improved compared to the original BI14 predictions when the added bias term is added. It has improved further when the coefficients of BI14 are recalibrated using the training data set.

Figure 5: Comparison of RMSE statistic for testing data set

6 Summary and Conclusions

In this manuscript, we have introduced a Bayesian methodology to recalibrate the coefficients/parameters of a GMPE model for ground motion records corresponding to a new region. Three statistical models (added bias, linear, and non-linear) based on different function forms of BI14 GMPE are considered for Bayesian update of parameters. In this research, a subset of ESM database that corresponds to low-to-moderate seismicity region is considered. The selection criteria is chosen as $[3; 5.2]$, $[1; 600 \text{ km}]$, $[800; 1200 \text{ m/s}]$, and $[\text{normal, strike-slip and reverse}]$ for moment magnitude, distance, $V_{s,30}$ velocity, and fault mechanism, respectively. The subset records are divided into training and testing data sets, where the training data set is used for recalibration of parameters and the testing data is used for testing the performance of different GMPE functional forms.

The parameters are updated using Bayesian linear models (BLM), Sampling importance resampling (SIR) with Laplace approximation as a proposal distribution, and in addition to SIR, two MCMC algorithms: Adaptive Metropolis (AM) and Automated factor slice sampling (AFSS) are considered. BLM algorithm is only applicable for models M1 and M2, and the parameters are estimated without approximations. AM algorithm suffered from convergence issues due to high correlation between parameters. AFSS and SIR algorithms gave similar parameter estimates with good convergence of samples. The effect of prior information on the parameter recalibration is evaluated and the models with informative

priors performed better compared to weakly informative priors. In the current study, the variance of the regression coefficients in informative priors is chosen as the variance of the GMPE model. However, in future studies, a sensitivity analysis can be carried out by varying the parameters' coefficient of variation in informative priors.

The coefficient h that is causing non-linearity in M3 is also examined. The recalibrated value of h is found out to be almost same as the original value. Therefore, the predictions from models M2 and M3 must be in a similar range. To confirm this, we compared the *DIC* and *WAIC* information criteria scores between models. The scores are similar to each other for M2 and M3 when the parameters are estimated using BLM, SIR, and AFSS algorithms. Therefore, the linear functional form (M2) of BI14 can be used for parameter recalibration with BLM algorithm for computational efficiency. If one uses the non-linear functional form of the GMPE, then SIR algorithm can be utilized for recalibration of regression coefficients.

Finally, the performance of the models is tested against a testing data set and root mean squared error (RMSE) is used as a measure for comparing the performance. The Bayesian recalibration of parameters (M2 and M3) have improved the RMSE compared to the original BI14 model and added bias model (M1). The reduction in RMSE is significant for PGA and lower time period ($T = 0.2s$) giving confidence in the proposed methodology of Bayesian update of GMPE parameters. In summary, the key contributions of this proposed methodology are:

- The linear statistical model (M2) based on the linear functional form of the GMPE can be considered for Bayesian update of parameters and to reduce the RMSE when the GMPE is tested against a new data set.
- The parameters in M2 can be recalibrated using conjugate priors (analytical formulation) approach for Bayesian linear models (BLM) for the best computational efficiency and accuracy.
- The parameters in the non-linear functional form (M3) of the GMPE model can be recalibrated using sampling importance resampling (SIR) with informative priors and Laplace approximation as a proposal distribution for a better computational efficiency and accuracy compared to MCMC algorithms such as Adaptive Metropolis (AM) and Automated factor slice sampling (AFSS) algorithms.

The conclusions presented above are from an initial exploratory study. While it shows a significant promise, we need to evaluate the applicability of this study across a wider set of GMPEs and databases in the future. Also, we intend to extend this methodology for updating the coefficients in GMPE models and then integrate the model uncertainty by averaging predictions over all the GMPE's using Bayesian model averaging (BMA) approach (Bodda et al., 2021).

Acknowledgments

This research was supported by Électricité de France (EDF) and Center for Nuclear Energy Facilities and Structures (CNEFS) at North Carolina State University. Resources for the Center come from the dues paid

by member organizations and from the Civil, Construction, and Environmental Engineering Department and College of Engineering in the University.

References

- Akkar, S., Sandikkaya, M., Şenyurt, M., Sisi, A. A., Ay, B., Traversa, P., Douglas, J., Cotton, F., Luzi, L., Hernandez, B. et al. (2014). “Reference database for seismic ground-motion in Europe (RESORCE).” *Bulletin of earthquake engineering*, 12(1), pp. 311–339.
- Ancheta, T. D., Darragh, R. B., Stewart, J. P., Seyhan, E., Silva, W. J., Chiou, B. S.-J., Wooddell, K. E., Graves, R. W., Kottke, A. R., Boore, D. M. et al. (2014). “NGA-West2 database.” *Earthquake Spectra*, 30(3), pp. 989–1005.
- Banerjee, S. (2008). “Bayesian linear model: Gory details.” *Downloaded from <http://www.biostat.umn.edu/~ph7440>*.
- Bertin, M., Marin, S., Millet, C. and Berge-Thierry, C. (2020). “Using Bayesian model averaging to improve ground motion predictions.” *Geophysical Journal International*, 220(2), pp. 1368–1378.
- Bindi, D., Massa, M., Luzi, L., Ameri, G., Pacor, F., Puglia, R. and 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), pp. 391–430.
- Bodda, S. S., Gupta, A. and Dinh, N. (2020a). “Enhancement of risk informed validation framework for external hazard scenario.” *Reliability Engineering & System Safety*, 204, p. 107140.
- Bodda, S. S., Gupta, A. and Dinh, N. (2020b). “Risk informed validation framework for external flooding scenario.” *Nuclear Engineering and Design*, 356, p. 110377.
- Bodda, S. S., Keller, M., Gupta, A. and Senfaute, G. (2021). “A Bayesian Approach to Estimate Weights for GMPE Models in a Logic Tree.”
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.
- Gelman, A., Hwang, J. and Vehtari, A. (2014). “Understanding predictive information criteria for Bayesian models.” *Statistics and computing*, 24(6), pp. 997–1016.
- Goulet, C. A., Kishida, T., Ancheta, T., Cramer, C., Darragh, R., Silva, W., Hashash, Y., Harmon, J., Stewart, J., Wooddell, K. et al. (2014). “PEER NGA-east database.” *PEER Report 2014*, 17.
- Heine, K. (2005). “Unified framework for sampling/importance resampling algorithms.” In *2005 7th International Conference on Information Fusion*, volume 2, pp. 6–pp, IEEE.

- Keller, M., Pasanisi, A., Marcilhac, M., Yalamas, T., Secanell, R. and Senfaute, G. (2014). “A Bayesian methodology applied to the estimation of earthquake recurrence parameters for seismic hazard assessment.” *Quality and Reliability Engineering International*, 30(7), pp. 921–933.
- Kotha, S. R., Weatherill, G., Bindi, D. and Cotton, F. (2020). “A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe.” *Bulletin of Earthquake Engineering*.
- Kowsari, M., Halldorsson, B., Hrafnkelsson, B., Snæbjörnsson, J. P. and Jónsson, S. (2019). “Calibration of ground motion models to Icelandic peak ground acceleration data using Bayesian Markov Chain Monte Carlo simulation.” *Bulletin of Earthquake Engineering*, 17(6), pp. 2841–2870.
- Kowsari, M., Sonnemann, T., Halldorsson, B., Hrafnkelsson, B., Snæbjörnsson, J. P. and Jónsson, S. (2020). “Bayesian inference of empirical ground motion models to pseudo-spectral accelerations of south Iceland seismic zone earthquakes based on informative priors.” *Soil Dynamics and Earthquake Engineering*, 132, p. 106075.
- Lanzano, G., Sgobba, S., Luzi, L., Puglia, R., Pacor, F., Felicetta, C., D’Amico, M., Cotton, F. and Bindi, D. (2019). “The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and data statistics.” *Bulletin of Earthquake Engineering*, 17(2), pp. 561–582.
- Nagel, J. B. (2019). “Bayesian techniques for inverse uncertainty quantification.” *IBK Bericht*, 504.
- NIED (2019). “NIED K-NET, KiK-net, National Research Institute for Earth Science and Disaster Resilience.”
- Pagani, M., Monelli, D., Weatherill, G., Danciu, L., Crowley, H., Silva, V., Henshaw, P., Butler, L., Nastasi, M., Panzeri, L. et al. (2014). “OpenQuake engine: An open hazard (and risk) software for the global earthquake model.” *Seismological Research Letters*, 85(3), pp. 692–702.
- Reich, B. J. and Ghosh, S. K. (2019). *Bayesian statistical methods*. CRC Press.
- Statisticat and LLC. (2020). *LaplacesDemon: Complete Environment for Bayesian Inference*. R package version 16.1.4.
- Tibbits, M. M., Groendyke, C., Haran, M. and Liechty, J. C. (2014). “Automated factor slice sampling.” *Journal of Computational and Graphical Statistics*, 23(2), pp. 543–563.
- Turner, B. M., Sederberg, P. B., Brown, S. D. and Steyvers, M. (2013). “A method for efficiently sampling from distributions with correlated dimensions.” *Psychological methods*, 18(3), p. 368.
- Vaishnav, P., Gupta, A. and Bodda, S. S. (2020). “Limitations of traditional tools for beyond design basis external hazard PRA.” *Nuclear Engineering and Design*, 370, p. 110899.

- 1
2
3
4 Viallet, E., Humbert, N. and Mottier, P. (2017). “Updating of a PSHA based on Bayesian inference with
5 historical macroseismic intensities.” In *16th world conference on earthquake engineering, 16WCEE*.
6
7
8 Wang, M. and Takada, T. (2009). “A Bayesian framework for prediction of seismic ground motion.” *Bulletin*
9 *of the Seismological Society of America*, 99(4), pp. 2348–2364.
10
11
12 Weatherill, G., Pagani, M. and Garcia, J. (2014). “OpenQuake ground motion toolkit – User guide, Global
13 Earthquake Model (GEM).” *Technical Report*.
14
15
16 Wu, X., Kozłowski, T., Meidani, H. and Shirvan, K. (2018). “Inverse uncertainty quantification using the
17 modular Bayesian approach based on Gaussian process, part 1: theory.” *Nuclear Engineering and Design*,
18 335, pp. 339–355.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65