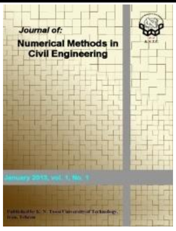
Numerical Methods in Civil Engineering, 6-2 (2021) 49-65



Numerical Methods in Civil Engineering

Journal Homepage: https://nmce.kntu.ac.ir/

A case study on the practicability of using linear analysis results in a Bayesian inference model to predict nonlinear responses in performance-based design methods

Shima Taheri* and Reza Karami Mohammadi**

ARTICLE INFO

RESEARCH PAPER

Article history:
Received:
June 2021.
Revised:
August 2021.
Accepted:
August 2021.

Keywords:
Bayesian inference,
Bayesian model comparison,
Demand parameters,
Prediction of responses,
Linear response spectrum analysis,
Likelihood function

Abstract:

Compared to traditional methods based on mean response evaluation of seismic parameters with significant confidence margin, the growing use of the new generation of performance-based design methods, which are based on loss and financial assessment, necessitates an increase in accuracy and reliability in probabilistic evaluation of structural response for all values of seismic parameters. Even with the same limited number of common nonlinear analyses, utilizing the Bayesian approach, which allows the use of diverse and even inaccurate data to form beliefs, is a powerful method to predict and enhance seismic response results. In this paper, the practicability of using linear analysis data in a Bayesian inference model to predict nonlinear responses is evaluated. A 20-story reinforced concrete special moment resisting frame is being considered, and a Bayesian model for prediction of the maximum story drift and the peak floor acceleration has been investigated. The Bayesian model was developed on linear results and finally updated with a limited number of nonlinear results. The predictability power of predictors, Bayesian model comparison among different likelihood functions, and common diagnostics tools in numerical solution of the Bayesian model developed on linear results, have all been examined. The results demonstrate a significant improvement in the outcomes, while proving the practicability of developing a stable and reliable model based on linear analysis data.

1. Introduction

With the introduction of the new generation of performance-based design methods in recent years [1], assessing the financial damage caused by earthquakes has consistently been one of the concerns for engineers. However, the combination of multiple epistemic and aleatory uncertainties strongly affects the accuracy of the results. In addition, to appropriately assess loss and damage, there is a need to obtain the Probability Density Function (PDF) curves, and not just the mean of responses. As a result, one of the appropriate solutions is to incorporate statistical inference approaches and probabilistic methodologies to improve the performance evaluation of buildings and consequently enhance corresponding loss assessment.

* *Ph.D. Candidate, Faculty of Civil Engineering, K.N. Toosi University of Technology, Tehran, Iran.*
** *Corresponding author: Associate Professor, Faculty of Civil Engineering, K. N. Toosi University of Technology, Tehran, Iran. E-mail: rkarami@kntu.ac.ir.*

The introduction of two distinct philosophies in the field of statistical inference has led to the development of two different problem-solving paradigms:

- The traditional Frequentist inference approach;
- The Bayesian inference approach.

Despite the fact that the Bayesian paradigm is as old as the frequentist one and has even simpler roots, the challenges of tackling its problems mathematically and explicitly, were not well-known until recently. In recent years, rapid and continuous advances in processing power have a significant impact on computational statistics practice [2]. Increased processing power has also contributed to the increased prevalence of computationally more complex approaches based on resampling, such as Gibbs sampling via the Markov Chain Monte Carlo (MCMC) algorithm [3], which makes use of Bayesian inference methods more practical [4, 5]. A Markov chain is a series of events in which the probability of the next event is solely determined by the current event's state [6].

After acquiring fresh data, Bayesian inference methods use Bayes' theorem to compute and update probabilities. The Bayes' theorem can directly assign a probability distribution that quantifies the belief to a parameter or set of parameters. Therefore, it can be said that Bayesian statistics interprets probability as a degree of belief. In contrast to frequentist statistics, Bayesian philosophy is based on the assumption that there may have been more evidence about a physical situation than was included in a single experiment's data, since, Bayesian methods can be used to combine results from different experiments or even scarce, sparse, noisy or biased sources, or unexpectedly all of them [7, 8].

In recent years, several studies have concentrated on the use of the Bayesian approach to assess buildings and structures, especially under seismic loads.

Esmaeili, Ludwig and Zareian used the results of Linear Response Spectrum Analyses (LRSAs) in a Bayesian inference model based on joint probability distribution to form a prior damage distribution and update the damage function by analyzing the Nonlinear Time History Analyses (NTHAs) with a limited number of Ground Motions (GMs) to enhance results [9]. On nuclear research reactors, Kwag, Lee and Ryu suggested a Bayesian-based seismic margin quantitative evaluation using seismic fragility data and fault tree analysis [10]. Erazo, Moaveni and Nagarajaiah used Bayesian method in accordance with a full-scale piece of a seven-story shear wall building with an array of sensors to measure the dynamic reactions. Following that, the estimated demands are utilized to compute damage measures in order to undertake a quantitative assessment of the structural integrity subjected to strong earthquake GMs [11]. On the basis of Bayesian networking, Gholami, Asgarian and Asil Gharebaghi provide a probabilistic approach for inspecting, maintaining, and repairing jacket platforms [12].

In contrast to frequentist approaches, Bayesian inference replaces the quantity of data with the following two fundamental assumptions:

- The likelihood function which is used as a proper distribution for the observations;
- Prior distribution for model or unknown parameters.

The authors of this paper, in another study, introduced a Bayesian-based inference method in which LRSA results are employed as belief to form informative prior distributions in a generalized linear regression model [13]. Artificially generated responses could then be generated by the Monte Carlo method to build a significantly more accurate and reliable PDF of demand parameters by updating the Bayesian model based on a limited number of NTHAs as new fresh evidence. The PDF obtained using this method can lead to a significant increase in accuracy and a sharp decrease in uncertainty associated with GM selection.

The practicability of just using LRSA results in determining the likelihood function and developing the model in a Bayesian generalized regression model is investigated in the present study, particularly concentrating on the technicalities. Using conventional diagnostic tools, the predictability power of the predictors, as well as the stability and reliability of the generated model based on the LRSAs, were also examined.

2. Methodology

There is no necessity for the correctness of the introduced fundamental assumptions in Bayesian inference, and depending on circumstances, accurate results can be achieved even though the wrong assumptions about the prior distributions or likelihood function are chosen. In most cases, the penalty for choosing the incorrect belief presents itself as uncertainty in the subsequent distribution of unknown parameters in Bayesian inference which leads to an increase in variation in predicted values [14]. Thus, one of the concerns in this research is concentrated on improving the accuracy of results by picking the appropriate likelihood function and examining the reliability of the model.

In the proposed Bayesian inference method, two distinct sets of observations are used as evidence:

- 994 LRSAs results to form an initial belief and model developing;
- 11 NTHAs results as proposed by most codes for Bayesian updating of the model [15].

However, in this study, 994 NTHAs were performed only to validate the final outcomes. The posterior distribution of the unknown parameters is estimated in the first step, utilizing the results of the 994 LRSAs as evidence. Given the large amount of data in this step, the suitable distribution for the likelihood function can also be determined directly from the results. In the second step, we do a Bayesian update on the posterior distribution obtained in the previous step, based on the results of 11 NTHAs, and we get a new posterior distribution of unknown parameters.

However, according to the following points, it seems that the determination of the likelihood function should either be determined by default and without using any evidence, or only the results of LRSAs can be employed:

- The distribution of structural responses under LRSAs and NTHAs will not always be the same, as will be shown later.
- Full NTHAs data can not be utilized to determine the appropriate distribution for the likelihood function, despite the fact that 994 NTHAs were utilized in the research process to validate the outcomes because it is assumed that this information isn't readily available in practice.
- A set of eleven NTHAs can not be utilized to choose the appropriate likelihood function distribution because, while this number of data is

reasonably accurate in predicting the mean of responses, it is not only accurate, but also error-prone in determining the appropriate distribution.

- Even if the appropriate distribution is chosen, an eleven NTHAs set is only sufficient to estimate the parameters of maximum three-parameter distributions and causes errors in distributions with more parameters.

As a result, in practice, the likelihood function must always be determined using the LRSAs results. In this paper, the likelihood function is evaluated using a Bayesian approach and LRSA observations and the correctness of the assumption will next be assessed due to the availability of 994 NTHA results.

The generated model is then presented in greater detail, followed by a discussion of common diagnostics tools for conforming the predictability power of the model and validating the Bayesian model's numerical solution.

3. Bayes' theorem

3.1 Bayesian inference

The Bayesian approach to statistical problem solving is founded on a mindset that is completely different from the frequentist methods. Existing information or beliefs are merged with new information in this way to generate a new degree of knowledge. One of the most important advantages of this method is the ability to update and use different and even distinct layers of information. In statistical analysis, the Bayesian method can produce different outcomes than the classical method. Even when using the Bayesian technique, the sequence in which evidence is entered to update the initial belief might have an impact on the ultimate outcome. The steps of Bayesian inference can be stated in three basic phases:

- A probability density function $\pi(\upsilon)$, called the prior distribution is selected, which expresses our belief in the parameter υ before observing and receiving any new data.
- A probability distribution is selected as $p(x|\upsilon)$, which represents the likelihood function between the observations and the model parameter.
- The initial belief will be updated and the posterior distribution as $p(\upsilon|D_n)$ will be obtained after making a series of observations and receiving data as $D_n=\{X_1,\ldots,X_n\}$.

$$\mathbb{P}(\theta|D_n) = \frac{p(X_1,\ldots,X_n|\theta)\pi(\theta)}{p(X_1,\ldots,X_n)} = \frac{\mathcal{L}_n(\theta)\pi(\theta)}{c_n} \quad (1)$$

where L_n is the likelihood function and c_n is a normalizing constant value, often known as evidence or marginal likelihood. Calculating this value explicitly, which entails solving a multiple integral, is one of the most difficult aspects of this procedure, necessitating the use of numerical and simulation approaches. Therefore, in this study, the MCMC approach for chain generation and the Gibbs sampling algorithm, a well-known method for high-dimensional problems as a specific example of the Metropolis-Hastings algorithm technique in which the model parameters will be altered in a series of stochastic stages, are used to numerically solve the Bayesian model [3]. There are two main advantages using Gibbs sampling:

- It is unnecessary to design a proposal distribution;
- Proposals are always accepted [6].

3.2 Bayesian model comparison

The Bayesian Model Comparison (BMC) method can also be used to validate the likelihood function. If D_n is the observation in different Bayesian models with different likelihood functions, based on the Bayes rule, the posterior distribution of each model, $\mathcal{M}_j$, according to the observation, can be stated as follows:

$$\mathbb{P}(\mathcal{M}_j|D_n) = \frac{\mathcal{L}\ (\mathcal{M}_j)\pi(\mathcal{M}_j)}{p(D_n)} \quad (2)$$

where $\pi(\mathcal{M}_j)$ and $\mathcal{L}(\mathcal{M}_j)$ are prior distribution and the likelihood function of each models, respectively. The Bayes factor, BF, between two models is defined as follows if the prior distributions are the same:

$$BF = \frac{p(D_n|\mathcal{M}_j)}{p(D_n|\mathcal{M}_k)} = \frac{\int_{D_n} p(\theta_j|\mathcal{M}_j)p(D_n|\theta_j,\mathcal{M}_j)d\theta_j}{\int_{D_n} p(\theta_k|\mathcal{M}_k)p(D_n|\theta_k,\mathcal{M}_k)d\theta_k} \quad (3)$$

where υ are model parameters and $\mathcal{M}$ is added as the new parameter representing of the rival models. Kas and Raftery interpretated the Bayesian factor based on the common logarithm of Bayes factors according to Table. 1 [16].

Table. 1: Interpretation of Bayes factor results [16]

LOG BF	Strength of Evidence
<-2	Decisively Model 2
-1 to -2	Strongly Model 2
-0.5 to -1	Substantially Model 2
-0.5 to 0.5	Not worth more than a bare mention
0.5 to 1	Substantially Model 1
1 to 2	Strongly Model 1
>2	Decisively Model 1

Although the mathematical definition of the BMC method is simple, its implementation using the MCMC method with the Gibbs sampling is not without flaws. As previously stated, the Gibbs sampling method only selects one variable at a time. Given that the model type is also entered as a variable $\mathcal{M}$ in the model, other parameters are likely to be set on the previous model in the step that is to be taken to a new step on $\mathcal{M}$, and the modification of $\mathcal{M}$ in most cases will not be accepted. In MCMC, this dramatically reduces the effective chain length, and consequently lowers the model's validity. However, numerous strategies have been offered in this area, such as the pseudo-prior distribution approach which has been implemented in this research [17].

4. Case study description

A 20-story reinforced concrete special moment resisting frame designed by Haselton et al. [18] has been selected for 2D models. Details for structural modeling have been provided in Fig. 1. The Opensees software is used to prepare the numerical model of the buildings [19]. The fundamental period of the structure is 2.36 sec. The numbers above the elements are the dimensions (width * height in inches), and the numbers below them represent the percentage of reinforcement.

Based on 68 earthquake events, represented in Table. 2, 994 GMs were chosen from the NGA-WEST2 GM database [20]. A three-dimensional space comprising magnitude, distance, and duration was used to choose GMs that are uniformly distributed in this space. Fig. 2 shows the distribution of the selected GMs. based on the above

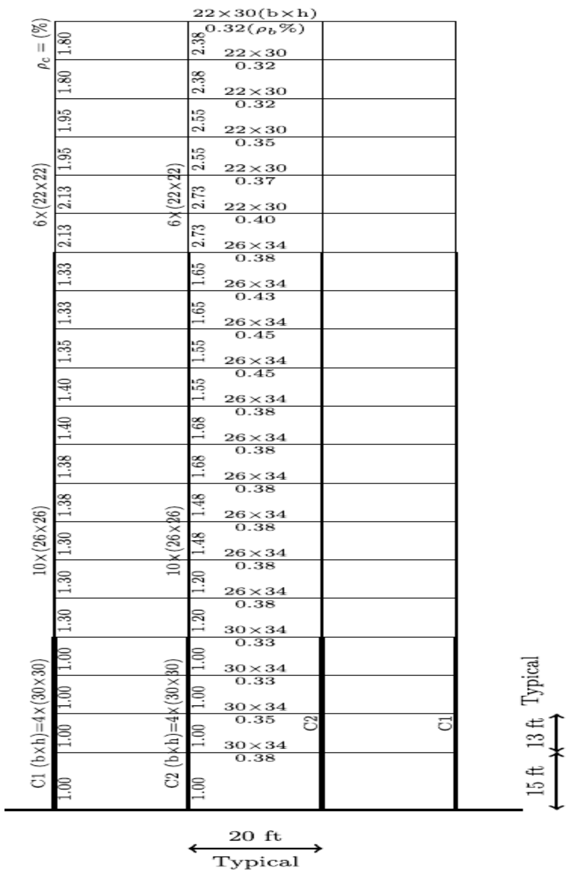


Fig. 1: Sample building

parameters. Also, in accordance with TBI task 12, the reaction of structures at DBE hazard level is evaluated using the site-specific hazard spectrum [21] as represented in Fig. 3.

While various demand parameters can be considered, the most common ones are Maximum Story Drift (MSD) and Peak Floor Acceleration (PFA), which correspond to structural and non-structural loss assessment in the new generation of performance-based design approaches, ATC-58 [1], and are investigated in this study.

5. Identification of predictors

Intensity Measures (IMs) are used to quantify records in the proposed Bayesian model. Due to the fact that the Bayesian model is to be developed only based on linear results, it is not possible to evaluate the accuracy, sufficiency and efficiency of the IMs and consequently, the desired number of IMs are selected. Bayesian approach, unlike traditional frequentist methods, is not very sensitive to the correlation of predictors, although increasing the number of predictors (even uncorrelated) or strong correlation only leads to a decrease in the Credible Interval (CI) of the posterior distributions. Therefore, in order to reduce the dimensions of the model, the Principal Component Analysis (PCA) method is used. Predictors will also be standardized to reduce the correlation of model parameters, termed unknown parameters.

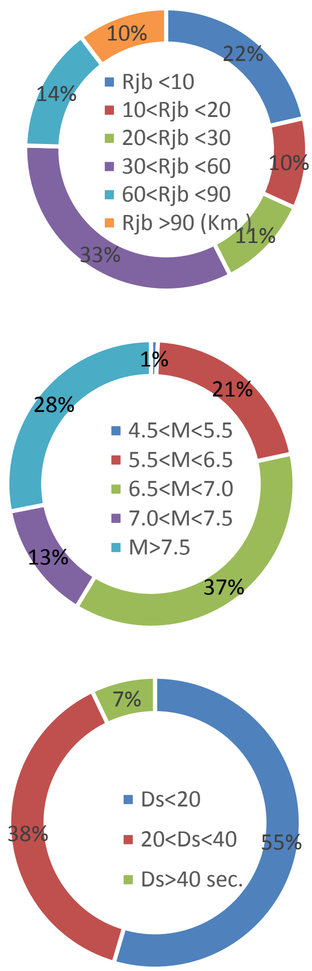


Fig. 2: Distribution of selected ground motions based on distance (R_{jb}), magnitude (M) and duration (D_s)

Table. 2: Earthquake events and number of selected GMs from each event

No.	Earthquake Name	Year	Magnitude	Mechanism	No. of GMs
1	"Kern County"	1952	7.4	Reverse	6
2	"Parkfield"	1966	6.2	strike slip	2
3	"San Fernando"	1971	6.6	Reverse	12
4	"Sitka_ Alaska"	1972	7.7	strike slip	2
5	"Friuli_ Italy-02"	1976	5.9	Reverse	2
6	"Tabas_ Iran"	1978	7.4	Reverse	2
7	"Coyote Lake"	1979	5.7	strike slip	4
8	"Imperial Valley-06"	1979	6.5	strike slip	4
9	"Montenegro_ Yugoslavia"	1979	7.1	Reverse	10
10	"Irpinia_ Italy-01"	1980	6.9	Normal	8
11	"Irpinia_ Italy-02"	1980	6.2	Normal	2
12	"Livermore-01"	1980	5.8	strike slip	6
13	"Mammoth Lakes-01"	1980	6.1	Normal Oblique	4
14	"Mammoth Lakes-02"	1980	5.7	strike slip	2
15	"Mammoth Lakes-03"	1980	5.9	strike slip	2
16	"Mammoth Lakes-04"	1980	5.7	strike slip	4
17	"Mammoth Lakes-06"	1980	5.9	strike slip	2
18	"Victoria_ Mexico"	1980	6.3	strike slip	2
19	"Corinth_ Greece"	1981	6.6	Normal Oblique	2
20	"Coalinga-01"	1983	6.4	Reverse	28
21	"Coalinga-05"	1983	5.8	Reverse	4
22	"Morgan Hill"	1984	6.2	strike slip	10
23	"Pelekanada_ Greece"	1984	5.0	Normal	2
24	"Nahanni_ Canada"	1985	6.8	Reverse	4
25	"Chalfant Valley-02"	1986	6.2	strike slip	2
26	"Kalamata_ Greece-01"	1986	6.2	Normal	2
27	"N. Palm Springs"	1986	6.1	Reverse Oblique	4
28	"San Salvador"	1986	5.8	strike slip	4
29	"Taiwan SMART1(45)"	1986	7.3	Reverse	2
30	"Baja California"	1987	5.5	strike slip	2
31	"New Zealand-02"	1987	6.6	Normal	2
32	"Superstition Hills-02"	1987	6.5	strike slip	2
33	"Whittier Narrows-01"	1987	6.0	Reverse Oblique	2
34	"Loma Prieta"	1989	6.9	Reverse Oblique	44
35	"Upland"	1990	5.6	strike slip	2
36	"Big Bear-01"	1992	6.5	strike slip	10
37	"Cape Mendocino"	1992	7.0	Reverse	10
38	"Joshua Tree_ CA "	1992	6.1	strike slip	2
39	"Landers"	1992	7.3	strike slip	26
40	"Northridge-01"	1994	6.7	Reverse	46
41	"Kobe_ Japan"	1995	6.9	strike slip	4
42	"Kozani_ Greece-01"	1995	6.4	Normal	2
43	"Umbria Marche (aftershock 1) Italy"	1997	5.5	Normal	2
44	"Umbria Marche (aftershock 2) Italy"	1997	5.6	Normal	4
45	"Umbria Marche (aftershock 3) Italy"	1997	5.3	Normal	2
46	"Umbria Marche_ Italy"	1997	6.0	Normal	6
47	"Chi-Chi_ Taiwan"	1999	7.6	Reverse Oblique	268
48	"Chi-Chi_ Taiwan-03"	1999	6.2	Reverse	34
49	"Chi-Chi_ Taiwan-04"	1999	6.2	strike slip	14
50	"Chi-Chi_ Taiwan-05"	1999	6.2	Reverse	4
51	"Chi-Chi_ Taiwan-06"	1999	6.3	Reverse	2
52	"Duzce_ Turkey"	1999	7.1	strike slip	12
53	"Hector Mine"	1999	7.1	strike slip	44
54	"Kocaeli_ Turkey"	1999	7.5	strike slip	12
55	"Tottori_ Japan"	2000	6.6	strike slip	42
56	"Denali_ Alaska"	2002	7.9	strike slip	6
57	"Bam_ Iran"	2003	6.6	strike slip	4
58	"San Simeon_ CA"	2003	6.5	Reverse	12
59	"Niigata_ Japan"	2004	6.6	Reverse	30
60	"Parkfield-02_ CA"	2004	6.0	strike slip	24
61	"Chuetsu-oki_ Japan"	2007	6.8	Reverse	58
62	"14383980"	2008	5.4	Reverse Oblique	2
63	"Iwate_ Japan"	2008	6.9	Reverse	74
64	"L'Aquila (aftershock 1) Italy"	2009	5.6	Normal Oblique	2
65	"L'Aquila_ Italy"	2009	6.3	Normal	10
66	"Darfield_ New Zealand"	2010	7.0	strike slip	6
67	"El Mayor-Cucapah_ Mexico"	2010	7.2	strike slip	20
68	"Christchurch_ New Zealand"	2011	6.2	Reverse Oblique	6

5.1 Intensity measures

Description and formulation of selected IMs is provided in Table. 3.

5.2 Principal component analysis

PCA is a statistical technique which involves converting a set of data of potentially correlated variables into a set of data of linearly uncorrelated variables via orthogonal transformation. PCA is also a tool for condensing multivariate data into smaller dimensions while maintaining the majority of the information [22]. In fact, it is a multivariate methodology for analyzing a given data in which observations are represented by a set of interrelated dependent variables [23]. What this means is that, we begin with a high number of variables, such as 20, and by the conclusion of the process, we have a lower number of variables that also express a significant amount of the information in the original sample.

Additional suggestions have been provided regarding the sample size (N) to the number of components proportion (p) by some researchers. Some of them offered an N/p ratio of 3:1 to 6:1 or a ratio of at least 10:1 ratio [24, 25].

Factor Analysis is in fact a technique used to reduce a large number of variables in a statistical model. In this way, the linear combination of the variables reconstructs each of the main factors and they can be representative of model changes instead of primary variables. In fact, dimensionality reduction emphasizes on performing computations with the fewest possible dimensions so that their attributes are preserved [26].

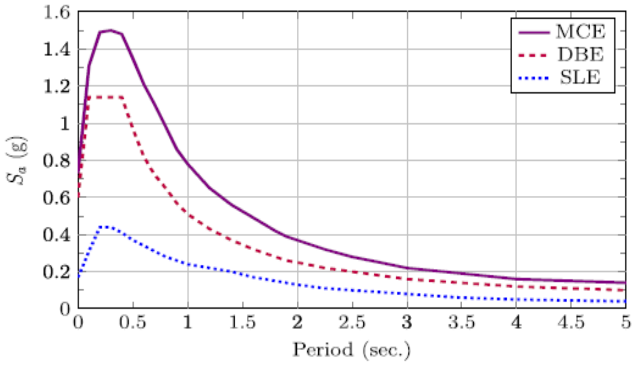


Fig. 3: Site specific hazard levels

With 21 input variables, PCA initially extracted 21 factors or "components". Each component has a quality score called eigen value. Only components with large eigen values are likely to represent a real main factor. Finally, 21 components are aggregated into four principal components following the PCA analysis, each of which is a combination of other components prior to the study. Table. 4 represents the component matrix.

Table. 3: Selected IMs

IM No.	Description	Formulation	Reference
IM.01	Spectral acceleration at the fundamental period	$S_a(T_1)$	-
IM.02	Spectral acceleration at the 2nd period	$S_a(T_2)$	-
IM.03	Spectral acceleration at the 3rd period	$S_a(T_3)$	-
IM.04	DSI	$\int_{2}^{5} S_d(T, 5\%)dt$	[27]
IM.05	SI	$\int_{0.1}^{2.5} S_v(T, 5\%)dt$	[28]
IM.06	ASI	$\int_{0.1}^{0.5} S_a(T, 5\%)dt$	[29]
IM.07	Np	$\frac{(\prod_{i=1}^{n} S_a(T_1))^{\frac{1}{n}}}{S_a(T_1)}$	[30]
IM.08	Peak ground acceleration	PGA	-
IM.09	Peak ground velocity	PGV	-
IM.10	Peak ground displacement	PGD	-
IM.11	Peak ground acceleration to velocity ratio	PGA / PGV	-
IM.12	Peak ground acceleration to displacement ratio	PGA / PGD	-
IM.13	Peak ground velocity to displacement ratio	PGV / PGD	-
IM.14	Arias Intensity	$\frac{\pi}{2g}\int a_t^2 dt$	[31]
IM.15	2nd to 1st spectral acceleration ratio	$S_a(T_2)$ / $S_a(T_1)$	-
IM.16	$S_{1\ldots n}$	$\prod_{i=1}^{n} S_a(T_i)^{MMPR_i}$	[32]
IM.17	IM(1E+2E)	$\sqrt{PF1^2 S_d(T_1,\zeta_1)^2 + PF2^2 S_d(T_1,\zeta_1)^2}$	[33]
IM.18	Duration	D_s (5-95)	-
IM.19	Magnitude	M	[34]
IM.20	Distance	R_{jb}	-
IM.21	Shear wave velocity over the upper 30 meters	V_s (30)	-

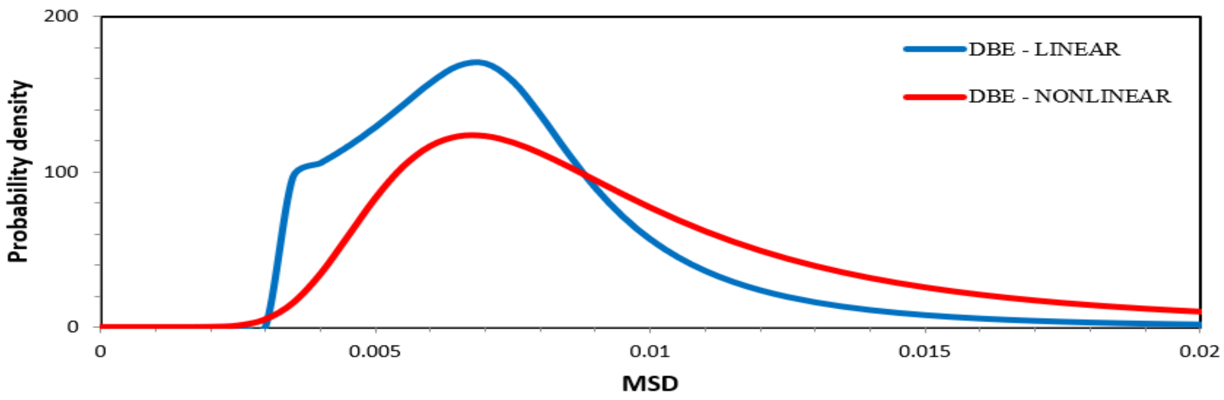


Fig. 4: Comparison of MSD results between LRSA and NTHA (994 analyses)

.6. Likelihood function

6.1 Evaluate linear and nonlinear results

Here, considering that 994 LRSAs and the same number of NTHAs have been performed, there is a significant amount of data from which it is possible to achieve the most appropriate and accurate probability distribution in terms of results.

The best distribution of results is chosen and represented in Fig. 4 and Figure. 5, based on the results of goodness of fit tests. The Kolmogorov-Smirnov test is examined to select the best distribution. The largest distance between the experimental cumulative distribution function and the theory is defined as the Kolmogorov-Smirnov statistic, which is based on the following equation:

$$D_{n,n^{\tau}} = \sup_{x} \left| F_{1,n}(x) - F_{1,n^{\tau}}(x) \right| \quad (4)$$

where $F_{1,n}$ is the experimental cumulative distribution and $F_{1,n^{\tau}}$ is the theoretical cumulative distribution; the sup represents the supremum function [35].

In Table. 5, the best fitted distribution for LRSA and NTHA results are shown. As can be observed, the Dagum, Burr, and Pearson 6 probability distributions have a large frequency after the dominant Wakeby probability distribution, which has a clear advantage over other distributions. Next on the list is the well-known log-normal distribution and its counterpart the log-logistic distribution. A brief explanation of the functions mentioned is provided below:

- Wakeby: a five-parameter distribution that belongs to the advanced distribution category. It is a popular tool for predicting flood flows. Due to the presence of three shape and two location parameters, there is a lot of flexibility in adapting to the data. Its drawbacks include the parameters' dependence on one another and the lack of an explicit relationship with the PDF. The inverse cumulative density function can be defined as:

$$x(F) = \xi + \frac{\alpha}{\beta}\left(1 - (1 - F)^{\beta}\right) - \frac{\gamma}{\delta}\left(1 - \frac{1}{(1 - F)^{\delta}}\right) \quad (5)$$

Where α and φ are location parameters, and β ‹γ & δ are shape parameters. The Wakeby distribution, even if not selected as the best distribution, has always been considered as the first few options and its selection has never been rejected using goodness of fitness tests.

Table. 4: Component matrix of factors

	F1	F2	F3	F4
IM.01	0.97			
IM.02			0.86	
IM.03			0.65	
IM.04		0.83		
IM.05	0.78		-0.51	
IM.06	0.89			
IM.07		-0.77		
IM.08	0.93			
IM.09	0.83			
IM.10		0.82		
IM.11	-0.63	0.42		
IM.12	-0.52	0.67		
IM.13	-0.41	0.77		
IM.14	0.88			
IM.15	-0.49		0.71	
IM.16	0.98			
IM.17	0.62			-0.49
IM.18	0.34	-0.48		0.56
IM.19	0.35	-0.69		
IM.20		-0.36		0.73
IM.21				-0.59

- Burr: a four-parameter distribution and also known as Burr Type XII, is a generalization of the log-logistic distribution family. k & α are two shape parameters, β is the scale parameter and γ is the position parameter [36]. The PDF is defined as follows:

$$P(x) = \frac{\alpha k}{\beta}\left(\frac{x - \gamma}{\beta}\right)^{\alpha - 1}\left(1 + \left(\frac{x - \gamma}{\beta}\right)^{\alpha}\right)^{1 - k} \quad (6)$$

- Dagum: a four-parameter distribution which is developed by changing the burr distribution and is widely used in economics today [37]. Its PDF is defined as follows:

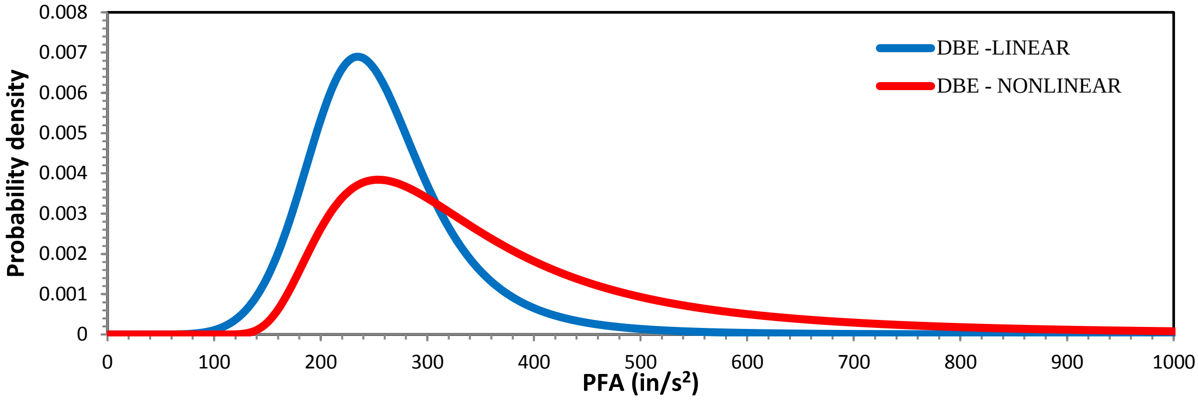


Figure. 5: Comparison of PFA results between LRSA and NTHA (994 analyses)

$$P(x)=\frac{\alpha k(\frac{x-\gamma}{\beta})^{\alpha k-1}}{\beta\left(1+(\frac{x-\gamma}{\beta})^{\alpha}\right)^{1+k}} \quad (7)$$

- Pearson 6: a four-parameter distribution of the Pearson family of distributions. Because of its similarity to the beta distribution, it is sometimes called the second type of beta distribution. α_1 & α_2 are the shape parameters, β is the scale parameter and γ is the location parameter [38]. The PDF is written as follows:

$$P(x)=\frac{1}{\beta B_{(\alpha_1,\alpha_2)}}(\frac{x-\gamma}{\beta})^{\alpha_1-1}\frac{1}{\left(1+\frac{x-\gamma}{\beta}\right)^{\alpha_1+\alpha_2}} \quad (8)$$

- Log-normal: a three-parameter distribution from a series of log distributions. In this distribution, the logarithms in the natural data base follow the normal distribution. This distribution, along with the normal distribution, is one of the most widely used distributions of natural events. Some natural variables, which are themselves the product of other natural variables, follow the log-normal distribution. Usually, for natural variables that have a minimum value but not a maximum value, by leaning towards infinity, the probability of their occurrence decreases and can be modeled with this distribution. A is the shape parameter, β is the scale parameter and γ is the location parameter. Unlike the normal distribution, the kurtosis in the log-normal distribution is not constant. The PDF is as follows:

$$P(x)=\frac{1}{(x-\gamma)\sigma\sqrt{2\pi}}\frac{1}{e^{\frac{1}{2}\left(\frac{ln(x-\gamma)-\mu}{\sigma}\right)^2}} \quad (9)$$

- Log-logistic: a three-parameter distribution from a series of log distributions. In this distribution, logarithms follow the logistic distribution in the natural data base. Some natural variables, which are themselves the product of other natural variables, follow the log-normal distribution. However, if the upstream variables are correlated and the results have less kurtosis than the log-normal distribution, the log-logistic is used. This distribution is a special case of burr distribution. The distribution has a larger tail than the gamma distribution and is widely used in modeling economic and insurance variables. Unlike the logistics distribution, the kurtosis is not constant. The PDF is as follows:

$$P(x)=\frac{\alpha}{\beta}(\frac{x-\gamma}{\beta})^{\alpha-1}\frac{1}{\left(1+(\frac{x-\gamma}{\beta})^{\alpha}\right)^2} \quad (10)$$

Table. 5: Comparison of Kolmogorov-Smirnov goodness of fit test between LRSA and NTHA results (994 analyses)

Rank	MSD Responses				PFA Responses			
	NTHA Results		LRSA Results		NTHA Results		LRSA Results	
	Distributions*	(statistics**)	Distributions	(statistics)	Distributions	(statistics)	Distributions	(statistics)
1	Dagum	(0.0149)	Wakeby	(0.0210)	Frechet	(0.0148)	Log-logistic	(0.0133)
2	Log-logistic	(0.0159)	Burr	(0.0219)	Log-logistic	(0.0149)	Burr	(0.0156)
3	Wakeby	(0.0164)	Log-normal	(0.0231)	Pearson 5	(0.0155)	Wakeby	(0.0199)
4	Burr	(0.0176)	Pearson 5	(0.0243)	Pearson 6	(0.0156)	Pearson 6	(0.0292)
5	Frechet	(0.0191)	Pearson 6	0.0252	Wakeby	(0.0163)	Pearson 5	(0.0293)
6	Pearson 5	(0.0245)	Log-logistic	(0.0281)	Burr	(0.0295)	Log-gamma	(0.0346)
7	Log-Pearson 3	(0.0247)	Dagum	(0.0286)	Log-normal	(0.0302)	Log-normal	(0.0350)
8	Pearson 6	(0.0260)	Inv. Gaussian	(0.0297)	Log-Pearson 3	(0.0340)	Inv. Gaussian	(0.0448)
9	Log-normal	(0.0369)	Fatigue life	(0.0310)	Inv. Gaussian	(0.0487)	Fatigue life	(0.0455)
10	Inv. Gaussian	(0.0469)	Log-Pearson 3	(0.0339)	Fatigue life	(0.0636)	Erlang	(0.0525)

* On the basis of the Kolmogorov-Smirnov statistics, the probability distributions are sorted ascending.

** The numbers in parentheses represents the Kolmogorov-Smirnov statistic for the probability distribution.

6.2 Bayesian model comparison

Although the Wakeby distribution has a clear advantage over other distributions having 994 analysis results available, the correlation of the distribution parameters, and more particularly the implicit PDF formulation, makes it nearly impossible to use this distribution to solve Bayesian problems numerically. On the other hand, using only 11 NTHAs is insufficient to improve and enhance parameters in most distributions. As a rule of thumb, the number of observations should be more than $2n$, where n represents the number of distribution parameters. Therefore, using Burr, Dagum and Pearson 6 are not recommended unless the increase in the number of NTHAs is accepted.

In the regression model, the relationship between independent or predictive variables and dependent or predictable variables through unknown parameters based on log-logistic or log-normal distribution is defined as follows:

$$\mathcal{M}_1\text{: Ln } Y_i \sim Logistic\left(\beta_0 + \sum_j \boldsymbol{\beta}_j \boldsymbol{X}_{j,i}\, , z_\sigma\right) \tag{11}$$

$$\mathcal{M}_2\text{: } Y_i \sim LogNormal\left(\beta_0 + \sum_j \boldsymbol{\beta}_j \boldsymbol{X}_{j,i}\, , z_\sigma\right) \tag{12}$$

Y is the predicted variable (here demand parameters) and $\boldsymbol{X}$ is the predictor variables (here PCA factors). All variables have been standardized, which simply means rescaling the data relative to its mean and standard deviation. $\boldsymbol{\beta}$ is the unknown parameter vector in the standard space of variables. β_0 represents the intercept.

By adding the parameter m, as an indicator of the selected model, according to Bayesian rule, the subsequent distribution of unknown parameters can be written as follows:

$$\mathrm{p}(\boldsymbol{\beta}, \mathcal{M}|D_n) = \frac{\mathcal{L}(D_n|\boldsymbol{\beta}, \mathcal{M}, \sigma)\pi(\boldsymbol{\beta}, \mathcal{M}, \sigma)}{\int_{D_n} \mathcal{L}(D_n|\boldsymbol{\beta}, \mathcal{M}, \sigma)\pi(\boldsymbol{\beta}, \mathcal{M}, \sigma)\mathrm{d}_{\boldsymbol{\beta}}\mathrm{d}_{\mathcal{M}}\mathrm{d}_\sigma} \tag{13}$$

As can be seen in Table. 6, although there is not absolute superiority between the two rival models, log-logistic distribution is more appropriate for the likelihood function. The relative superiority of log-logistics over log-normal was also evident from Fig. 4 and Figure. 5 because the log-logistics has a heavier tail that is evident in the NTHA.

Note that full NTHA results are not available in practice, so BMC should be based on LRSA results. It should be noted, the use of eleven NTHA data is not sufficient for executing BMC and sometimes leads to completely erroneous results. Therefore, it is valid only to estimate the mean value of the response. It can be concluded that the use of LRSA evidence is suitable for implementing BMC of Bayesian probabilistic modelling. A flowchart depicting the steps involved in developing a Bayesian model is shown in Fig. 6. The LRSA results are used to develop an informative prior in the first step, the NTHA results are used for Bayesian updating of the model in the second step, and the updated model predicts the responses in the third step.

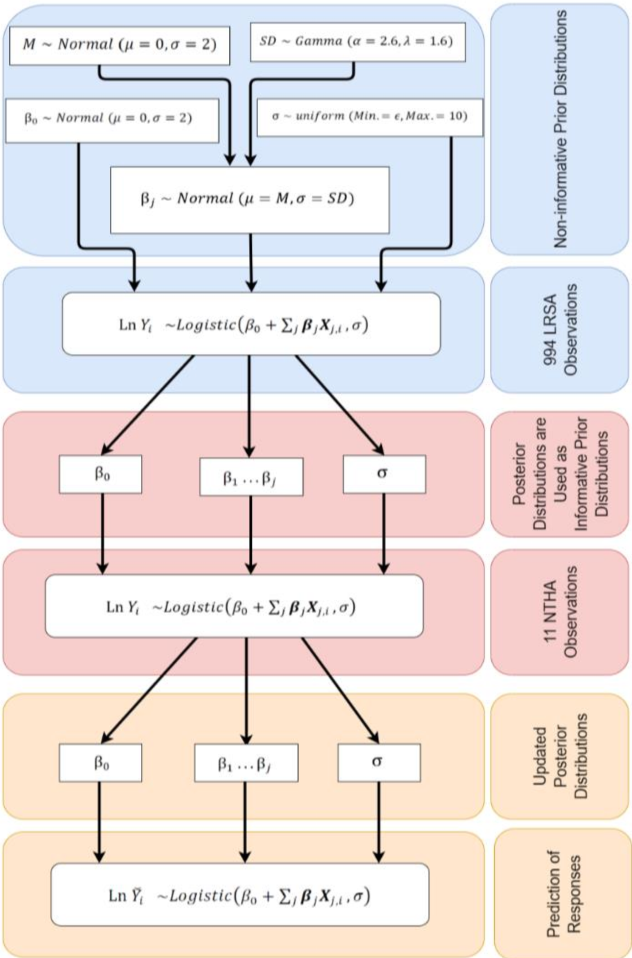


Fig. 6: The proposed Bayesian model flowchart

Table. 6: BMC results using LRSA and NTHA results (994 analyses)

	MSD				PFA			
	NTHA (BF*)		LRSA (BF)		NTHA (BF)		LRSA (BF)	
SLE	Both models are equal	(0.79)	Both models are equal	(0.80)	Strongly Log-logistic	(18.1)	strongly Log-logistic	(22.2)
DBE	Substantially Log-logistic	(4.45)	Both models are equal	(0.57)	Substantially Log-logistic	(4.28)	Substantially Log-logistic	(5.77)
MCE	Strongly Log-logistic	(25.6)	Both models are equal	(0.62)	Substantially Log-logistic	(5.15)	Substantially Log-logistic	(4.82)

** The numbers in parentheses represents the Bayes factor.

7. Bayesian model diagnostics

In Fig. 7 and Fig. 8, posterior distributions for 95% High Density Interval (HDI) of MSD predictors are shown. As demonstrated, the ratio of 95% HDI of priors to posteriors are very high, which indicates the high predictability power of the selected predictors.

The more challenging problem in the MCMC method is to determine how many steps are needed to converge to the stationary distribution within an acceptable error.

The values in the Markov chain must be representative of the posterior distributions. They should not be influenced heavily by the random starting value of the chains or be orphaned and end up in strange parts of the parameter space. Furthermore, the chains should be large enough and formed swiftly to ensure that estimations are accurate and reliable. A visual assessment of the chain trajectory was performed in order to detect unrepresentativeness (See Fig. 9(a) to Fig. 14(a)).

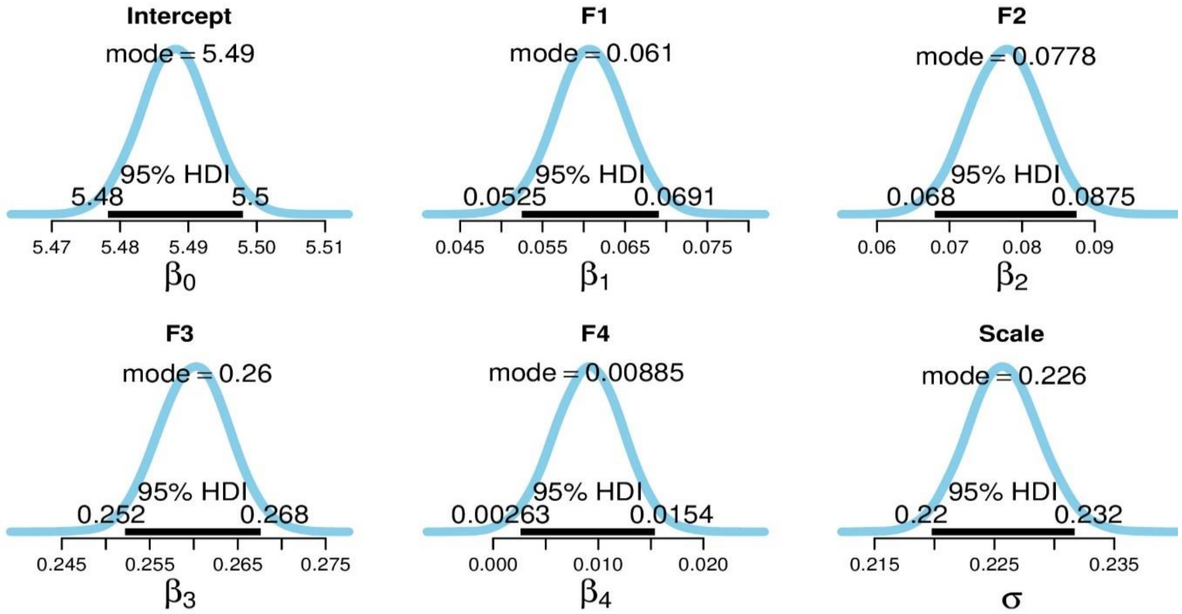


Fig. 7: Posterior distributions of predictors (mode and 95% HDI of predictors have been specified (for PFA).

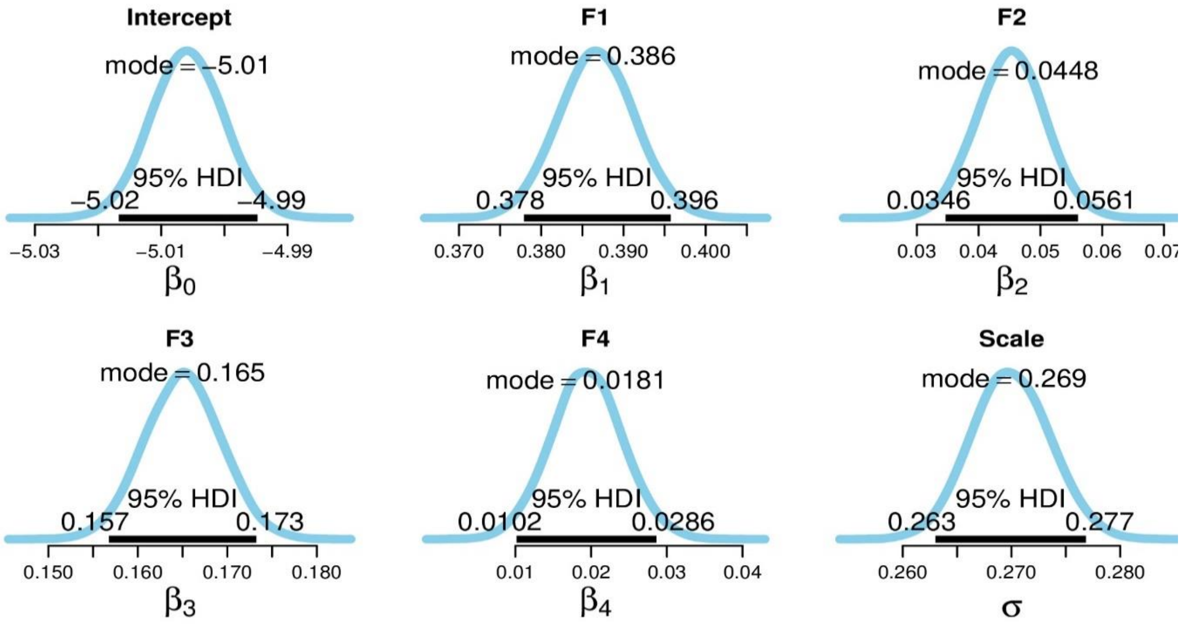


Fig. 8: Posterior distributions of predictors (mode and 95% HDI of predictors have been specified (for MSD).

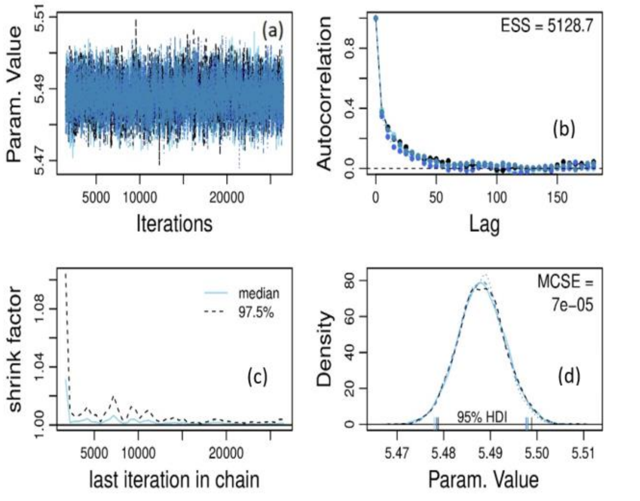


Fig. 9: Diagnostic testing for unknown parameter $\boldsymbol{\beta_0}$ (for MSD).
(a): The trace plot shows that throughout all iterations, all three chains are close to the mean value, with no sudden changes, orphaned chains, or reliance on the initial stating point.
(b): All three chains have low autocorrelations to previous values, and after about 50 lag, there is no significant dependency on previous values. Effective Size Sampling (ESS) has a reasonable value.
(c): The shrink factor is close to one, indicating that there are no discernible differences between the values obtained from different chains.
(d): The PDFs generated in each of the three chains are identical. The Monte Carlo Standard Error (MSCE) is very close to zero.

In addition to trace plots, which demonstrate the dispersion of parameter values in all iterations, generating and inspecting PDF for parameters, in which the probability density for each value in each chain is specified separately, might provide useful information on the representativeness of the chains. The chains should overlap if they are all representative of the posterior distributions (See Fig. 9(d) to Fig. 14(d)).

The Gelman-Rubin statistics, also known as the "shrink factor", is a popular method for testing the MCMC convergence by having a look at Markov multiple chains, simultaneously. This method assesses convergence by comparing variances calculated in two different ways [39]. It is to be done for each model parameter, and it re-estimates the target distribution for the model parameter using the last n iterations of the Gibbs sampler among m multiple parallel chains. Between-chains variance, BC, and within-chains variance, WC, are calculated as follows:

$$BC = \frac{N}{M-1}\sum_{m=1}^{M}\left(\hat{\theta}_m - \hat{\theta}\right)^2 \tag{14}$$

$$WC = \frac{1}{M}\sum_{m=1}^{M}\hat{\sigma}_m{}^2 \tag{15}$$

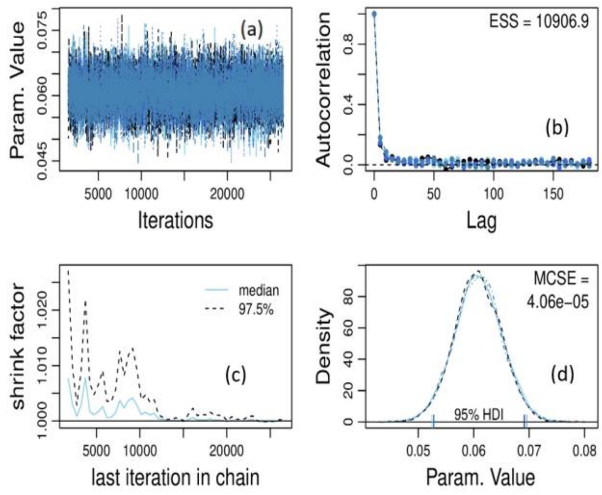


Fig. 10: Diagnostic testing for unknown parameter $\boldsymbol{\beta_1}$(MSD model)

where $\hat{\theta}_m$ is the mean estimated value for model parameters within each chain and $\hat{\theta}$ is the mean estimated value for model parameters among all chains. $\hat{\sigma}_m$ is the standard deviation of the model parameter in each chain. The shrink factor, $\hat{R}$, is calculated as follows:

$$\widehat{Var}_{(\theta)} = \left(1-\frac{1}{n}\right)WC + \frac{1}{n}BC \tag{16}$$

$$\hat{R} = \sqrt{\frac{\widehat{Var}_{(\theta)}}{WC}} \tag{17}$$

Because the chain starting points are over-dispersed in relation to the target density, BC is initially much greater than WC. If no chains become orphaned or stuck in a particular region of parameter space, the BC aims at WC and the shrink factor tends to 1, indicating that all chains have stabilized into a representative sampling (See Fig. 9(c) to Fig. 14(c)).

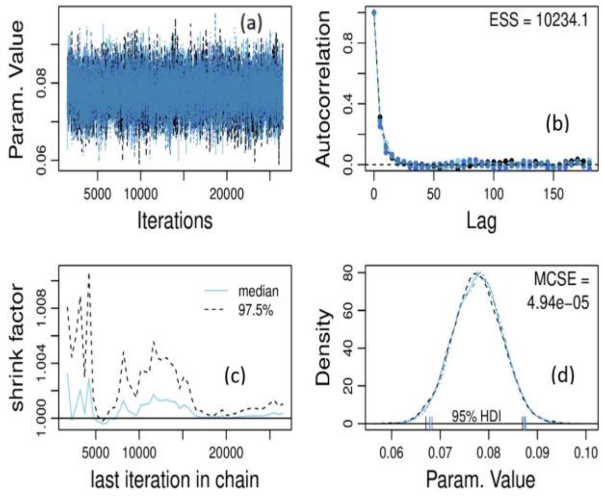


Fig. 11: Diagnostic testing for unknown parameter $\boldsymbol{\beta_2}$(MSD model)

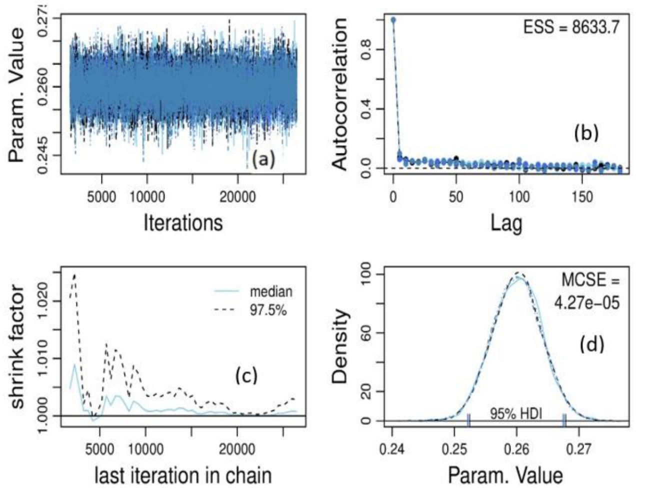


Fig. 12: Diagnostic testing for unknown parameter $\boldsymbol{\beta_3}$(MSD model)

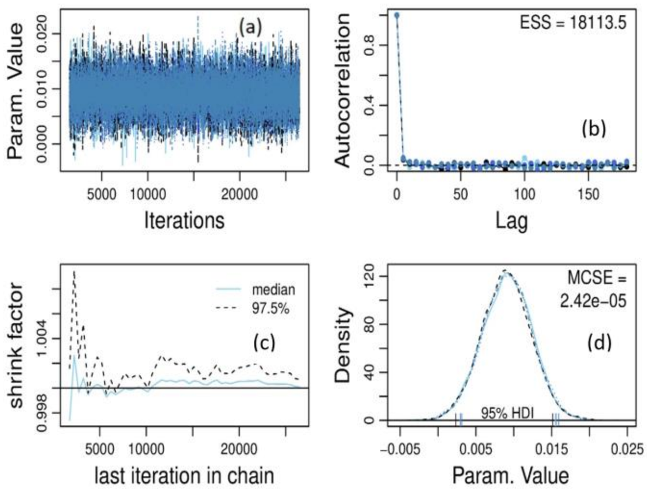


Fig. 13: Diagnostic testing for unknown parameter $\boldsymbol{\beta_4}$(MSD model)

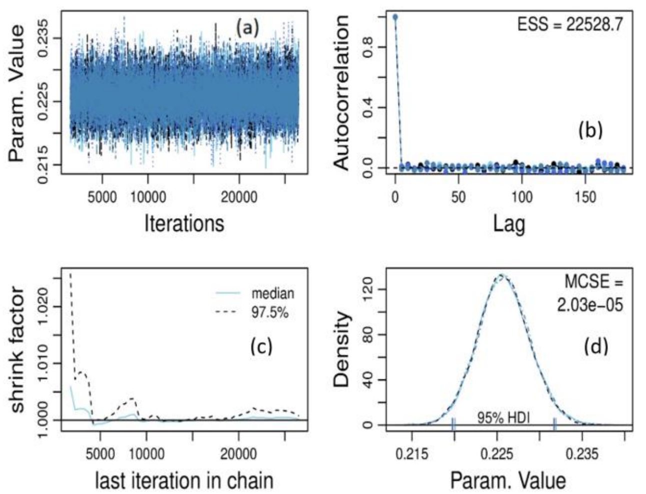


Fig. 14: Diagnostic testing for unknown parameter $\boldsymbol{\sigma}$ (MSD model)

The second core objective is to have a large enough sample for stable and reliable numerical estimations of the distribution with some assurance that the chains are truly representative samples of the posterior distribution. The larger the sample, the more consistent and accurate the estimations will be (on average). Here, for this purpose, the ESS means the effective sample size.

Apart from the ESS, whose higher values are more favorable, the Monte Carlo Standard Error (MSCE), with more favorable lower values, is an important tool for justifying the efficiency and accuracy of chain length. MSCE represents the estimated standard deviation of the sample mean in the chain on the scale of parameter value [40]. In addition, the PDF generated by each chain must be identical See Fig. 9(c) to Fig. 14(c)).

Three chains have been used to reduce the sensitivity to the starting point and also to compare the results of the chains for diagnostic testing. 500 first steps of chains have been deleted for burn-in stage. 20000 steps have been saved during the MCMC process while 5 thin steps, which are performed but not saved, is the process between saved steps. As shown in Fig. 9 to Fig. 14, ESS is very high and, consequently, autocorrelation is near zero in all chains for all predictors. The shrink factor is absolutely near one and there are no significant differences between different chain results which indicate high predictability power of the model. In Fig. 15 and Fig. 16, the correlation matrix of predictors is shown. As is clear, there is no detectable correlation between predictors.

8. Verification of results

Fig. 17 and Fig. 18 demonstrate the results of applying the proposed model to six suites consisting of 11 GMs for PFA and MSD, respectively. Suites were chosen at random from a database of 994 ground motions (see Table. 2) using the setseed (2019) command in the R programming language [41]. The graphs are PDF curves that are used in loss assessment and show the probability of each level of seismic demand. The reciprocal of the units of horizontal axis is the units of probability density and the area below the PDF curve is equal to one.

The black dashed line represents the PDF of 994 LRSA results. This line is used to form an initial belief or informative prior distribution about unknown parameters in the Bayesian model. The blue dashed line represents the PDF of 11 NTHA results for each suite. This line is used for Bayesian updating of the model. This line is utilized in practice for structural assessment with no post-processing of the results. While the mean of responses in each suite is similar, the PDF is highly sensitive to ground motion selection, as shown. The blue solid line represents the results after applying the Bayesian model based on the LRSA results as informative prior distributions and Bayesian updating the model with a limited number of NTHAs. The thick red line indicates the actual response of the structure or NTHAs based on the 994 GMs available in this study.

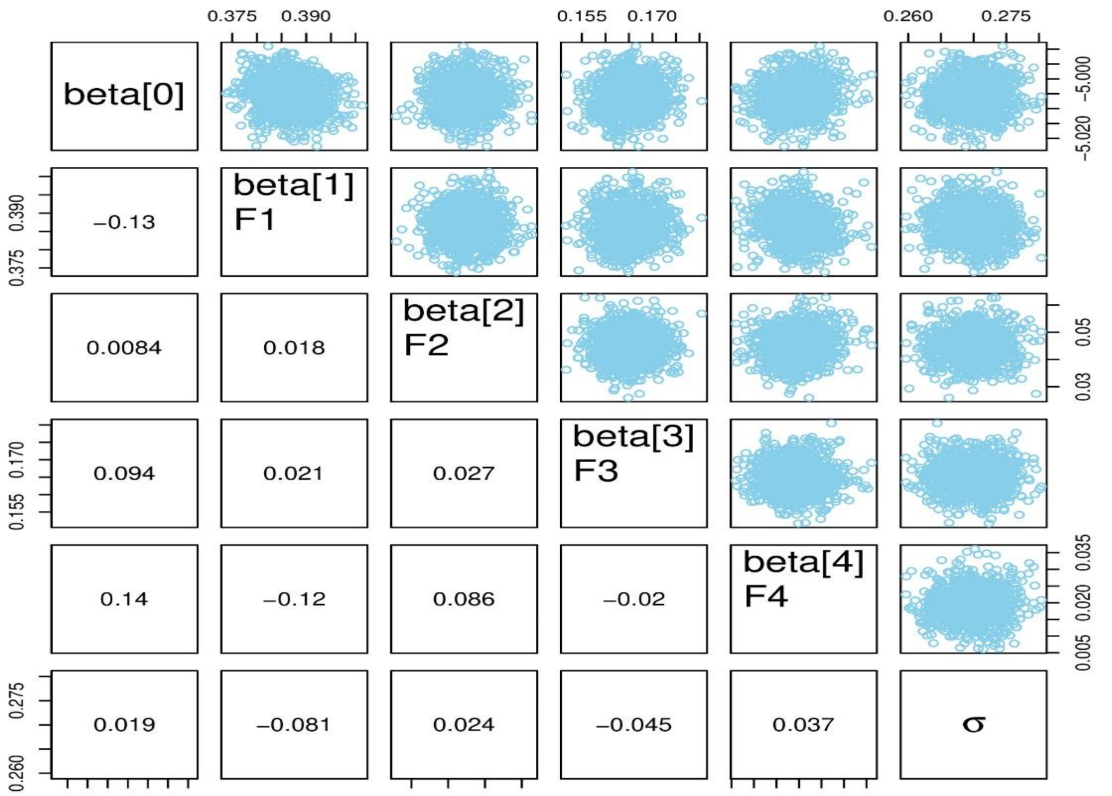


Fig. 15: Correlation matrix of predictors (for MSD).

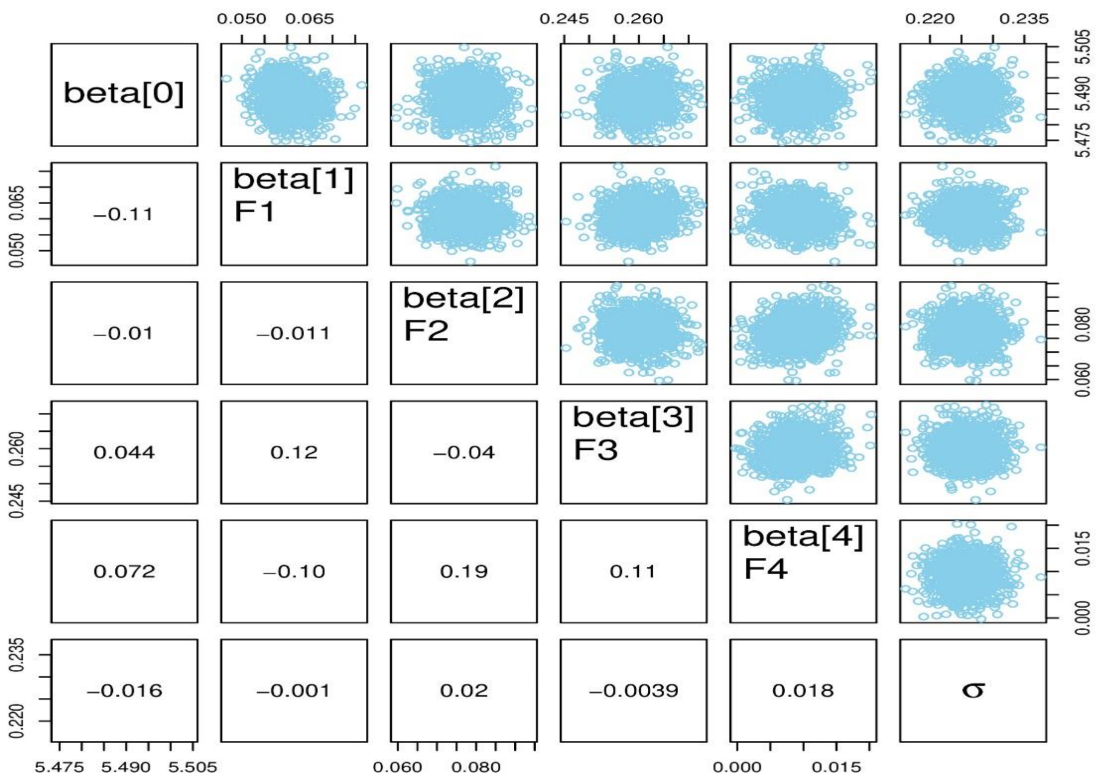


Fig. 16: Correlation matrix of predictors (for PFA).

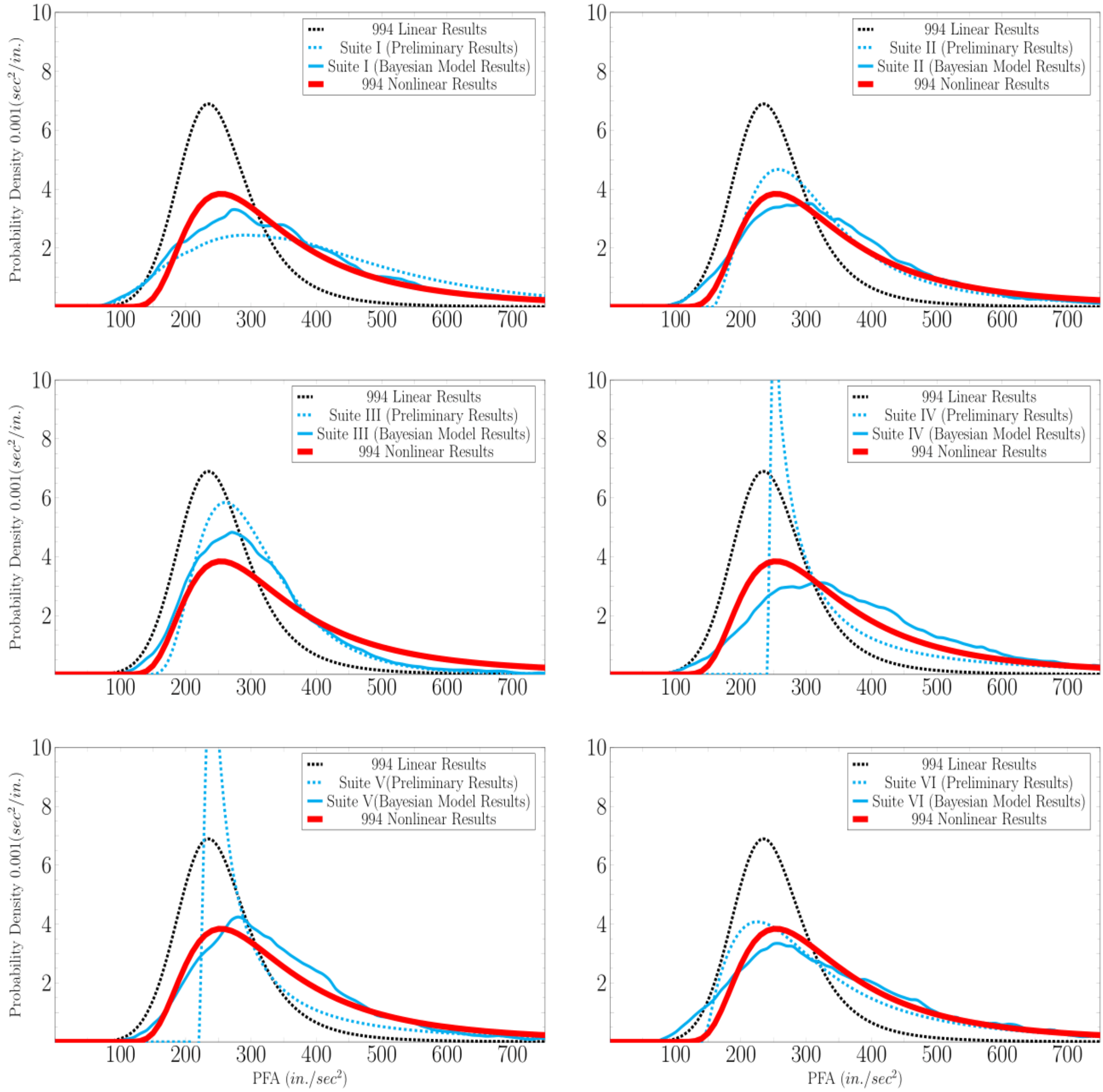


Fig. 17: Comparison of PFA results before and after of implementation of Bayesian model (DBE hazard level)

The Bayesian model has greatly improved the outcomes in all cases, as can be observed.

The method's processing time is independent of the building and site's parameters, and developing and training the model takes roughly 4 hours in all circumstances. As a result, it is extremely computationally efficient, and it is particularly well-suited to complicated and tall structures.

A comparison between non-informative prior distributions for model parameters and the posterior distributions is shown.

Despite the fact that the vertical axis is drawn logarithmically, the posterior distribution diagrams are still spike-shaped, showing the model's great predictability, as shown in Fig. 19.

9. Conclusion

A simplified model was utilized in this research since the objective was to investigate the ability to exploit linear results to develop an accurate and reliable Bayesian model. For other techniques, including Bayesian model averaging, inclusion of predictors' nonlinear interaction, and sequential Bayesian updating, which can enhance the model's quality and robustness, refer to other researches published by the authors of this article [13]. Clear results were obtained from this research which are briefly summarized below:

- To develop an accurate PDF of seismic demand characteristics, hundreds or even thousands of NTHA are necessary, which is nearly impossible to achieve in practice. This is especially true in the new generation of performance-based design methodologies that are primarily concerned with.

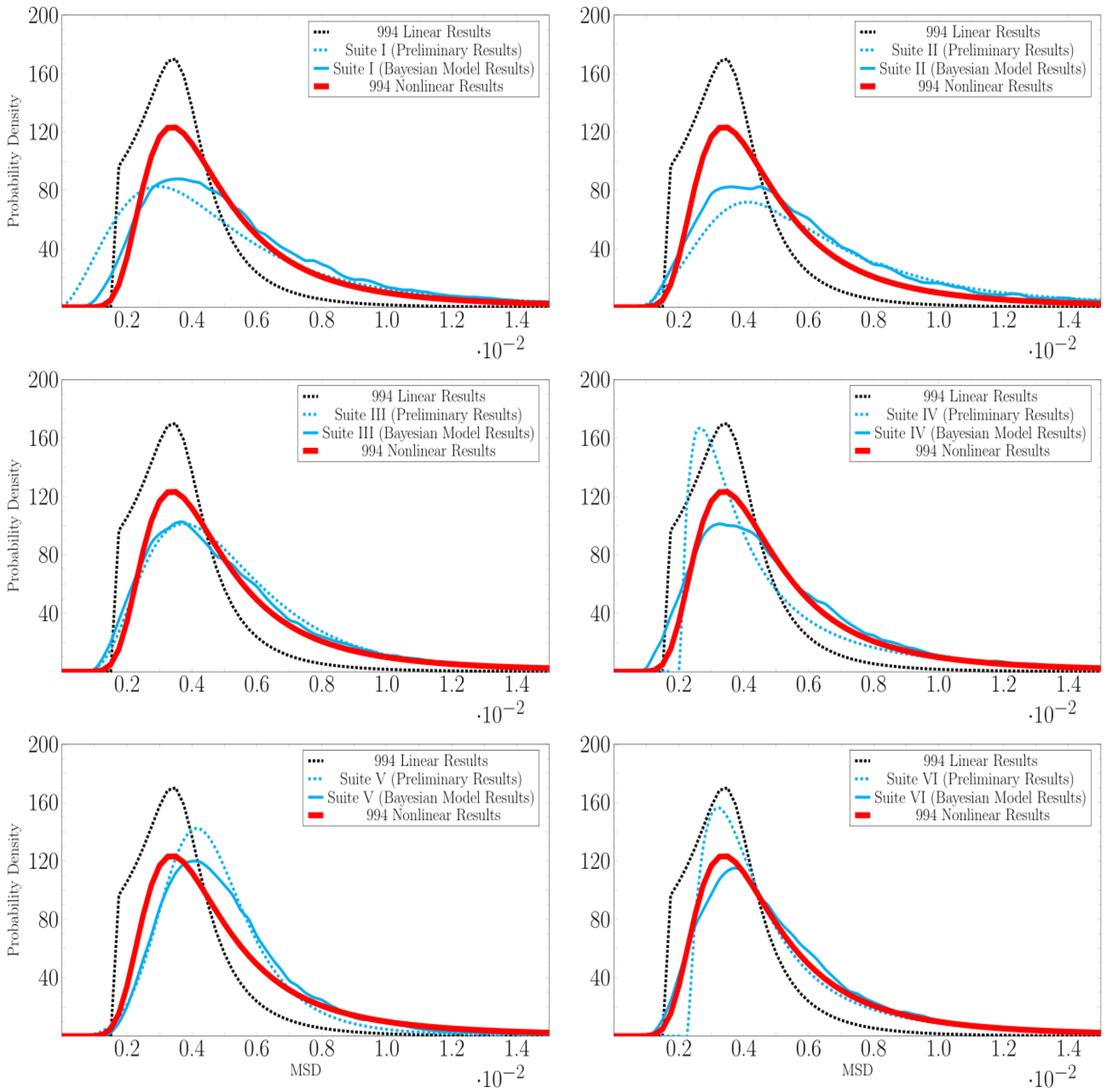


Fig. 18: Comparison of MSD results before and after of implementation of Bayesian model (DBE hazard level)

- loss estimation. Implementation of the Bayesian method, which was based on forming an informative prior distribution with LRSA results and updating the model with a limited number of NTHAs, led to impressive results
- According to the analysis of a significant number of LRSA & NTHA in this research (about 1000) and despite the significant difference between LRSA & NTHA results, the predictability of LRSA results for building a generalized regression Bayesian model is very high. This can be assessed by examining the final results as well as the very high ratio of 95% HDI of priors to posteriors.
- Using both LRSA & NTHA results nearly led to the selection of the same likelihood functions. It seems using the log-logistic likelihood function is more appropriate for selected demand parameters.
- It is possible to select the likelihood function without presumption based on the Bayesian model comparison and only LRSA observations. It was demonstrated that the limited number of NTHA results is not sufficient for decision-making for BMC.
- Implementation of different diagnostic tests indicates the stability and reliability of the developed Bayesian model. No disruption was observed in any of the chains in any of the unknown parameters.

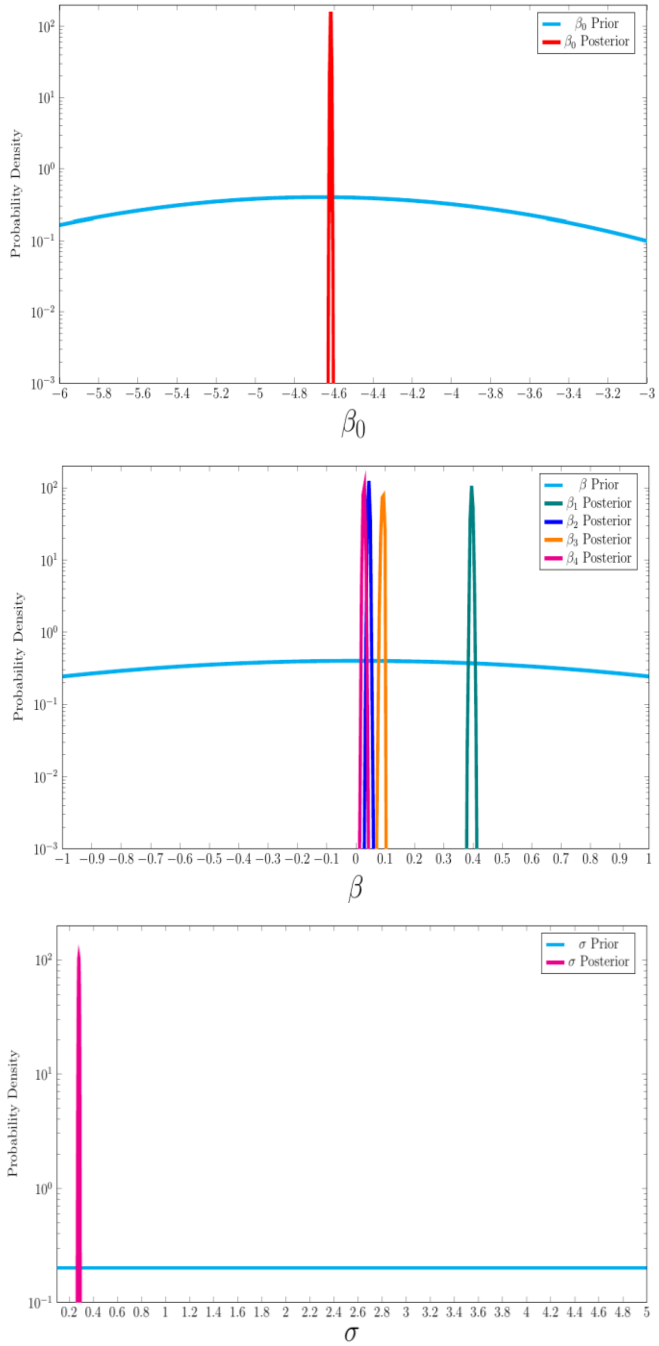


Fig. 19: Comparison of prior and posterior distributions

References

[1] FEMA P58-1, Seismic Performance Assessment of Buildings: Volume1—Methodology, Washington, D.C.,USA, 2018.

[2] F. Guimaraes, «Research: anyone can do it,» PediaPress, Mainz, 2011.

[3] S. Geman et D. Geman, «Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images,» Journal of Applied Statistics, vol. 20, p. 25–62, 1993.

[4] A. Agrawal et K. Gopal, Biomonitoring of water and waste water, Springer Science & Business Media, 2013.

[5] S. E. Fienberg et others, «When did Bayesian inference become" Bayesian"?,» Bayesian analysis, vol. 1, p. 1–40, 2006.

[6] I. Yildirim, Bayesian inference: Gibbs sampling., Technical Note, University of Rochester, 2012.

[7] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari et D. B. Rubin, Bayesian data analysis, CRC press, 2013.

[8] R. McElreath, Statistical rethinking: A Bayesian course with examples in R and Stan, Chapman and Hall/CRC, 2018.

[9] Esmaili, O., Grant Ludwig, L., & Zareian, F. (2016). Improved performance-based seismic assessment of buildings by utilizing Bayesian statistics. Earthquake Engineering & Structural Dynamics, 45(4), 581-597.

[10] S. Kwag, J. Oh, J.-M. Lee et J.-S. Ryu, «Bayesian-based seismic margin assessment approach: Application to research reactor,» Earthquakes Struct., vol. 12, p. 653–663, 2017.

[11] K. Erazo, B. Moaveni et S. Nagarajaiah, «Bayesian seismic strong-motion response and damage estimation with application to a full-scale seven story shear wall structure,» Engineering Structures, vol. 186, p. 146–160, 2019.

[12] H. Gholami, B. Asgarian et S. Asil Gharebaghi, «Practical Approach for Reliability-Based Inspection Planning of Jacket Platforms Using Bayesian Networks,» ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, vol. 6, p. 04020029, 2020.

[13] Sh. Taheri, Prediction of engineering demand parameters using bayesian inference approach, Phd thesis, Department of Civil Engineering, K. N. Toosi University of Technology., Tehran, 2021.

[14] A. Gelman et B. Carpenter, «Bayesian analysis of tests with unknown specificity and sensitivity,» Journal of the Royal Statistical Society: Series C (Applied Statistics), vol. 69, n° %15, pp. 1269--1283, 2020.

[15] American Society of Civil Engineers. (2017, June). Minimum design loads and associated criteria for buildings and other structures. American Society of Civil Engineers.

[16] R. E. Kass et A. E. Raftery, «Bayes factors,» Journal of the American Statistical Association, vol. 90, p. 773–795, 1995.

[17] B. p. Carlin et s. Chib, «Bayesian model choice via Markov chain Monte Carlo methods,» Journal of the Royal Statistical Society: Series B (Methodological), p. 473–484, 1995.

[18] C. B. Haselton, A. B. Liel, G. G. Deierlein, B. S. Dean et J. H. Chou, «Seismic collapse safety of reinforced concrete buildings. I: Assessment of ductile moment frames,» Journal of Structural Engineering, vol. 137, p. 481–491, 6 2011.

[19] F. McKenna, M. H. Scott et G. L. Fenves, «Nonlinear finite-element analysis software architecture using object composition,» Journal of Computing in Civil Engineering, vol. 24, p. 95–107, 6 2010.

[20] T. D. Ancheta, R. B. Darragh, J. P. Stewart, E. Seyhan, W. J. Silva, B. S.-J. Chiou, K. E. Wooddell, R. W.

Graves, A. R. Kottke, D. M. Boore et others, «NGA-West2 database,» Earthquake Spectra, vol. 30, p. 989–1005, 2014.

[21] J. Moehle, Y. Bozorgnia, N. Jayaram, P. Jones, M. Rahnama, N. Shome, Z. Tuna, J. Wallace, T. Yang et F. Zareian, Case Studies of the Seismic Performance ofTall Buildings Designed by Alternative MeansTask 12 Report for the Tall Buildings Initiative, Pacific Earthquake Engineering Research Center, 2011.

[22] S. Karamizadeh, S. M. Abdullah, A. A. Manaf, M. Zamani et A. Hooman, «An overview of principal component analysis,» Journal of Signal and Information Processing, vol. 4, p. 173, 2013.

[23] H. Abdi et L. J. Williams, «Principal component analysis,» Wiley interdisciplinary reviews: computational statistics, vol. 2, p. 433–459, 2010.

[24] K. Y. Hogarty, C. V. Hines, J. D. Kromrey, J. M. Ferron et K. R. Mumford, «The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination,» Educational and psychological measurement, vol. 65, p. 202–226, 2005.

[25] R. B. Cattell, «Fixing the number of factors: The most practicable psychometric procedures,» chez The Scientific Use of Factor Analysis in Behavioral and Life Sciences, Springer, 1978, p. 72–91.

[26] K. K. Vasan et B. Surendiran, «Dimensionality reduction using principal component analysis for network intrusion detection,» Perspectives in Science, vol. 8, p. 510–512, 2016.

[27] B. A. Bradley, «A Generalized Conditional Intensity Measure Approach and Holistic Ground motion Selection,» Earthquake Engineering & Structural Dynamics, vol. 39, p. 1321–1342, 2010.

[28] G. W. Housner, «Spectrum intensities of strong-motion earthquakes,» Bulletin of the Seismological Society of America, vol. 53, p. 403–417, 6 1952.

[29] J. L. Von Thun, «Earthquake Ground Motions for Design and Analysis of Dams,» Earthquake Engineering & Soil Dynamics II-Recent Advances in Ground-Motion Evaluation, 1988.

[30] E. Bojórquez, R. Chávez, A. Reyes-Salazar, S. E. Ruiz et J. Bojórquez, «A new ground motion intensity measure IB,» Soil Dynamics and Earthquake Engineering, vol. 99, p. 97–107, 2017.

[31] A. Arias, «Measure of Earthquake Intensity.,» 1970.

[32] N. Su, X. Lu, Y. Zhou et T. Y. Yang, «Estimating the peak structural response of high-rise structures using spectral value-based intensity measures,» The Structural Design of Tall and Special Buildings, vol. 26, p. e1356, 2017.

[33] N. Luco et C. A. Cornell, «Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions,» Earthquake Spectra, vol. 23, p. 357–392, 2007.

[34] A. Azarbakht, M. Mousavi, M. Nourizadeh et M. Shahri, «Dependence of correlations between spectral accelerations at multiple periods on magnitude and distance.,» Earthquake Engineering and Structural Dynamics, vol. 43, n° %18, pp. 1193-1204, 2014.

[35] I. M. Chakravarty, J. D. Roy et R. G. Laha, «Handbook of methods of applied statistics,» 1967.

[36] I. W. Burr, «Cumulative frequency functions,» The Annals of mathematical statistics, vol. 13, p. 215–232, 1942.

[37] C. Dagum, «A new model of personal income distribution: specification and estimation,» chez Modeling income distributions and Lorenz curves, Springer, 2008, p. 3–25.

[38] K. Pearson, «XI. Mathematical contributions to the theory of evolution.—X. Supplement to a memoir on skew variation,» Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, vol. 197, p. 443–459, 1901.

[39] Cowles, M. K., & Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. Journal of the American Statistical Association, 91(434), 883-904.

[40] Flegal, J. M., Haran, M., & Jones, G. L. (2008). Markov chain Monte Carlo: Can we trust the third significant figure?. Statistical Science, 250-260.

[41] R. Team, "R: A language and environment for statistical computing.," R Foundation for Statistical Computing, Vienna, Austria, 2020.