Hierarchical Bayesian Inversion for Quantification of Mixed Aleatory and Epistemic Uncertainties in Model Parameters

AUTHOR(S):

Kitahara, Masaru; Kitahara, Takeshi; Beer, Michael

CITATION:

Kitahara, Masaru ...[et al]. Hierarchical Bayesian Inversion for Quantification of Mixed Aleatory and Epistemic Uncertainties in Model Parameters. Proceedings of the 20th working conference of the IFIP WG 7.5 on Reliability and Optimization of Structural Systems 2022: 1-9: 3.

ISSUE DATE:

2022-09

URL:

https://doi.org/10.14989/ifipwg75_2022_3

Hierarchical Bayesian Inversion for Quantification of Mixed Aleatory and Epistemic Uncertainties in Model Parameters

Masaru Kitahara
Leigniz University Hannover

Takeshi Kitahara
Kanto Gakuin Univesity

Michael Beer
Leigniz University Hasitynnover, University of Liverpool, Tongji University

ABSTRACT: Uncertainties in the model parameters need to be properly characterized for the reliable and economic performance assesment of structures using a numerical model. Since not all parameters are trivial to measure directly, inverse uncertainty quantification (UQ) techniques, which infer the non-determinism in the model parameters by the measurments of the structural responses, are often necessary. Among such techniques, the class of Bayesian methods has been widely accepted as a coherent probabilistic approach to handle uncertainties in the inverse UQ. However, the main drawback of the conventional Bayesian methods is that they cannot quantify the inherent variability in the model parameters which causes the random failure of the structure. To fill this gap, the hierarchical Bayesian methods have gained increasing attention, in which a proability distribution is assigned to the model parameters to characterize their variability while its hyperparameters are treated as epistemic uncertainty and updated through Bayesian scheme. The first author and his co-workers have recently developed the hierarchical Bayesian approach using the staircase density function (SDF). This approach can consider the lack-of-knowledge on the distribution formats as epistemic uncertainty and infer the true-but-unknown distribution by updating the hyperparameters of SDF. This paper amis to illustrate its fundamental ideas and demonstrate its applicability to the estimation of a broad range of distributions through simple numerical test examples.

1. INTRODUCTION

Nowadays, the performance assesment of existing structures using numerical models is widely studied to determine the necessity and priority of their repair and reinforcement. For the purpose of model vaidation and verification, it is important to calibrate the model parameters so that the model response coincides as close as possble to the actual structural behavior. This procedure is referred to as model updating.

In deterministic model updating (Mottershead et al. 2011), parameter values are estimated to minimize the error between the model response and the observed system behavior. In this case, the observed features are treated as deterministic values, and thus uncertainties in the observed data are not taken into account. In contrast, the Bayesian inference, a type of probabilistic model updating (Mares et al. 2006) that has been widely studied in recent years, estimates parameters as a posterior distribution accounts for the likelihood of multiple observations. It should be noted that the posterior distribution obtained by the Bayesian inference represents the degree of plausibility about which parameter ranges are more probable than others based on observations; thus, it does not represent the probability distribution that the parameters themselves physically and spatially follow.

On the other hand, uncertainties in the model parameters are often not inevitable due to e.g., the manufacturing tolerance, environmental conditions, and aging conditions. In such cases, the objective of model updating is to estimate the probability distribution that the parameters follow to reproduce the variability of the observed data. The hierarchical Bayesian inference (Jia et al. 2022), which treats the hyperparameters of the probability distribution as the parameters to be updated, is effective for this purpose.. However, since the distribution family that the parameters belongs is rarely known a priori, it is common to assume a Gaussian distribution as the unique target distribution for convenience (Jia et al. 2022).

In contrast to that, the first author and his co-wokers have proposed a hierarchical Bayesian inference that does not assume a specific distribution shape by using staircase density functions (SDFs) (Crespo et al. 2018), which can approximate a wide range of distributions discretely (Kitahara et al. 2022). Although the application of SDFs to hierarchical Bayesian inference has been discussed in Kitahara et al. 2022, its applicability to various typs of distributions has not been thoroughly investigated.

In this study, we first examine the usefulness of the proposed method in comparison with the conventional Bayesian method using a simple numerical example with linear, monotonically increasing nonlinear, and convex nonlinear relationships between model parameters and model responses. We then verify the robustness of the proposed method to various distribution shapes, such as asymmetric, flat, sharped, and multimodal distributions, as well as normal distributions.

2. OVERVIEW OF HIERARCHICAL BAYESIAN INFERENCE

2.1. *Hierarchical Bayesian Method*

In general, Bayesian inference updates the prior distribution $P(\boldsymbol{x})$ of the parameter $\boldsymbol{x}$ to the posterior distribution $P(\boldsymbol{x}|\boldsymbol{D})$ using the observed data $\boldsymbol{D}$, which is done based on the Bayes' theorem shown in Equation (1).

$$P(\boldsymbol{x}|\boldsymbol{\mathcal{D}}) \propto \mathcal{L}(\boldsymbol{\mathcal{D}}|\boldsymbol{x})P(\boldsymbol{x}) \tag{1}$$

where the likelihood function $L(\boldsymbol{D}|\boldsymbol{x})$ quantifies the degree of agreement between the observed data $\boldsymbol{D}$ and the model response $\boldsymbol{M}(\boldsymbol{x})$ and is often given by Equation (2), assuming a normal distribution for the modeling error.

$$\mathcal{L}(\boldsymbol{\mathcal{D}}|\boldsymbol{x}) = \prod_{k=1}^{n} \mathcal{N}\left(\boldsymbol{\mathcal{D}}^{(k)} - \mathcal{M}(\boldsymbol{x}), \sigma_\varepsilon^2\right) \tag{2}$$

where $N(Exp, Var)$ represents the normal distribution with the expected value Exp and variance Var, $\boldsymbol{\sigma}_\varepsilon^2$ is the error variance that is often chosen as the variance of the observed data $\boldsymbol{D}$, and n is the number of observed data. In the above formulation, the model response $\boldsymbol{M}(\boldsymbol{x})$ for $\boldsymbol{x}$ is a definite value, and the parameter uncertainty is not considered.

In contrast, the hierarchical Bayesian inference considered in this study assumes a probability distribution $f_x(\boldsymbol{x},\boldsymbol{\theta})$ that characterizes the parameter uncertainty and replaces Equation (1) to the following equation by taking the hyperparameter $\boldsymbol{\theta}$ as the parameters to be updated.

$$P(\boldsymbol{\theta}|\boldsymbol{\mathcal{D}}) \propto \mathcal{L}(\boldsymbol{\mathcal{D}}|\boldsymbol{\theta})f_{\boldsymbol{x}}(\boldsymbol{x},\boldsymbol{\theta})P(\boldsymbol{\theta}) \tag{3}$$

Since the model response $\boldsymbol{M}(\boldsymbol{x},\boldsymbol{\theta})$ is obtained from the probability distribution of $\boldsymbol{x}$ determined for an instance of $\boldsymbol{\theta}$, $\boldsymbol{M}(\boldsymbol{x},\boldsymbol{\theta})$ also follows a probability distribution. Therefore, the derivation of the likelihood function does not require the assumption of modeling error and can be expressed as the conditional probability of the observed data $\boldsymbol{D}$ for an instance $\boldsymbol{\theta}$:

$$\mathcal{L}(\boldsymbol{\mathcal{D}}|\boldsymbol{\theta}) = \prod_{k=1}^{n} P\left(\boldsymbol{\mathcal{D}}^{(k)}|\boldsymbol{x},\boldsymbol{\theta}\right) \tag{4}$$

In general, the posterior distributions in Equations (1) and (3) cannot be obtained analytically, and sampling from the posterior distribution by the Monte Carlo method is commonly used. For example, Kuroda and Nishio (2016) and Matsuoka et al. (2020) used Markov chain Monte Carlo

(MCMC) (Tierney 1994) and Hayashi et al. 2018 used a particle filter (Kitagawa 1996). In this study, we employ the transitional Markov chain Monte Carlo (TMCMC) (Ching and Chen 2007). This method pushes samples from the prior distribution to the posterior distribution step by step by repeatedly sampling from the intermediate distribution, which has a nested relationship with the prior and posterior distributions. In its last step, samples are generated using MCMC where samples that follow the posterior distribution are used as the seeds; thus, the burn-in in MCMC does not necessary to be considered. Compared to MCMC, which samples directly from the posterior distribution, TMCMC is known to be suitable for high-dimensional problems and sampling from complex-shaped distributions.

2.2. *Approximate Bayesian computation based on Bhattacharyya distance*

In the hierarchical Bayesian approach, the likelihood evaluation for each instance of $\boldsymbol{\theta}$ requires the estimation of the probability density function (PDF) of the model response $\boldsymbol{M(x,\theta)}$ using the Monte Carlo method. This hinders its application to the problem where the time consuming model, e.g., finite element model, is employed in the likelihood function. Hence, in this study, we approximate Bayesian computation (ABC) is utilized to reduce the computational burden for the likelihood evaluation (Kitahara et al. 2021) In ABC, instead of the exact likelihood function in Equation (4), an approximate likelihood function based on an arbitrary statistic that quantifies the degree of agreement between the observed data and the model response is used. In this study, the following Equation based on the most common Gaussian kernel function is used as the approximate likelihood function.

$$\bar{\mathcal{L}}(\boldsymbol{\mathcal{D}}|\boldsymbol{x}) \propto \exp\left\{-\frac{S\big(\boldsymbol{\mathcal{D}},\mathcal{M}(\boldsymbol{x})\big)^2}{\varepsilon^2}\right\} \tag{5}$$

where $S(\boldsymbol{D}, \boldsymbol{M(x)})$ is a statistic that expresses the degree of agreement between the observed data $\boldsymbol{D}$ and the model response $\boldsymbol{M(x)}$, and $\boldsymbol{\varepsilon}$ is the centralization coefficient of the posterior distribution. Its optimal value depends on the problem, but generally, a value between 0.01 and 0.1 is considered to be appropriate (Patelli 2017). In this study, it is set to be $\boldsymbol{\varepsilon} = 0.01$. In addition, the Bhattacharyya distance (Bhattacharyya 1946), which quantifies the degree of overlap between two probability distributions, $\boldsymbol{D}$ and $\boldsymbol{M(x)}$, is used as the statistic S. Its theoretical definition is based on the PDFs of two different sample sets. However, both $\boldsymbol{D}$ and $\boldsymbol{M(x)}$ are obtained as discrete distributions; thus, the probability mass function (PMF) is used instead to approximate the Bhattacharyya distance as the following equation:

$$d_B\big(\boldsymbol{\mathcal{D}},\mathcal{M}(\boldsymbol{x})\big) = \sum_{j=1}^{n_{bin}} \sqrt{l_{\boldsymbol{\mathcal{D}}}^{j}\, l_{\mathcal{M}(\boldsymbol{x})}^{j}} \tag{6}$$

where n_{bin} is the number of bins commonly defined to obtain the PMFs of $\boldsymbol{D}$ and $\boldsymbol{M(x)}$, and $l^{j}_{\boldsymbol{D}}$ and $l^{j}_{\boldsymbol{M(x)}}$ are the PMF values of $\boldsymbol{D}$ and $\boldsymbol{M(x)}$ in the jth bin. Since the approximate likelihood function using the Bhattacharyya distance is based on discrete PMFs, it can significantly reduce the computational cost compared to the exact likelihood function.

2.3. *Staircase Density Functions (SDFs)*

As shown in Equation (3), the posterior distribution $P(\boldsymbol{\theta}|\boldsymbol{D})$ depends on the underlying probability distribution $f_x(\boldsymbol{x},\boldsymbol{\theta})$ of $\boldsymbol{x}$. Thus its estimation accuracy depends on the choice of the distribution family for $f_x(\boldsymbol{x},\boldsymbol{\theta})$. However, its optimal choice is generally unknown a priori, which hinders the practical application of the hierarchical Bayesian approach.

Alternatively, we consider to approximate $f_x(\boldsymbol{x},\boldsymbol{\theta})$ by means of SDF, which is a probability distribution defined for a recently proposed class of random variable, called staircase random variable (SRV) (Crespo et al. 2018). Consider a univariate case, SRV has the bounded support domain $[\underline{x}, \overline{x}]$ and four hyperparameters $\boldsymbol{\theta} = [\mu, \sigma^2, \tilde{m}_3, \tilde{m}_4]$, consisting of the mean, variance, skewness, and kurtosis. It is noted that the skewness and kurtosis are defined as $\tilde{m}_3 = m_3/\sigma^3$ and $\tilde{m}_4 = m_4/\sigma^4$ using the third and fourth order central moments $\boldsymbol{m}_3$ and $\boldsymbol{m}_4$, respectively. The hyperparameters $\boldsymbol{\theta}$ should satisfy the constraint conditions given as a series of inequalities presented in Table 1. By dividing the support domain $[\underline{x}, \overline{x}]$ of x equally into n_b bins of length $\kappa = (\overline{x} - \underline{x})/n_b$, the SDF is given by

$$f_x(\boldsymbol{x}) = \begin{cases} l^j & \forall x \in (x^j, x^{j+1}], \forall j = 1,2,\cdots,n_b \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where l^j is the PDF value at the jth bin and $x^j = \underline{x} + (j-1)\kappa$ is the leftmost of the jth bin. For all bins, Equation (7) satisfies $l^j \geq 0$ as well as $\kappa \sum_{j=1}^{n_b} l^j = 1$. The PDF value $\boldsymbol{l} = [l^1, l^2, \cdots, l^{n_b}]$ can be obtained by solving an optimization problem based on the moment matching constraint expressed as follows.

$$\hat{\boldsymbol{l}} = \underset{\boldsymbol{l} \geq 0}{\operatorname{argmin}} \left\{ J(\boldsymbol{l}) : \sum_{j=1}^{n_b} \int_{x^j}^{x^{j+1}} x l^j dx = \mu, \right. \\ \left. \sum_{j=1}^{n_b} \int_{x^j}^{x^{j+1}} (x-\mu)^r l^j dx = m_r, r = 2,3,4 \right\} \tag{8}$$

where $m_2 = \sigma^2$ and $J(\boldsymbol{l})$ is an arbitrary cost function. In this study, the following equation is used to maximize the entropy.

$$J(\boldsymbol{l}) = \kappa \log \boldsymbol{l}^{\mathrm{T}} \boldsymbol{l} \tag{9}$$

Table 1 Hyperparameter constraints

	constraint inequality $g_i \leq 0$
μ	$g_1 = \underline{x} - \mu$, $g_2 = \mu - \overline{x}$
σ^2	$g_3 = -\sigma^2$, $g_4 = \sigma^2 - (\mu - \underline{x})(\overline{x} - \mu)$
$\tilde{m}_3$	$g_5 = \sigma^4 - \sigma^2(\mu - \underline{x})^2 - \sigma^3\tilde{m}_3(\overline{x} - \mu)$, $g_6 = \sigma^3\tilde{m}_3(\overline{x} - \mu) - \sigma^2(\overline{x} - \mu)^2 + \sigma^2$, $g_7 = 4\sigma^4 + \sigma^6\tilde{m}_3^2 - \sigma^4(\overline{x} - \underline{x})^2$, $g_8 = 6\sqrt{3}\sigma^3\tilde{m}_3 - (\overline{x} - \underline{x})^3$, $g_9 = -6\sqrt{3}\sigma^3\tilde{m}_3 - (\overline{x} - \underline{x})^3$
$\tilde{m}_4$	$g_{10} = -\tilde{m}_4$, $g_{11} = 12\sigma^4\tilde{m}_4 - (\overline{x} - \underline{x})^4$, $g_{13} = \sigma^6\tilde{m}_3^2 + \sigma^6 - \sigma^6\tilde{m}_4$, $g_{12} = \{\sigma^4\tilde{m}_4 - \sigma^2(\mu - \underline{x})(\overline{x} - \mu) - (\underline{x} + \overline{x} - \mu)\}\{(\mu - \underline{x})(\overline{x} - \mu) - \sigma^2\} + \{\sigma^3\tilde{m}_3 - \sigma^2(\underline{x} + \overline{x} - \mu)\}^2$

Figure 1 shows the PDFs obtained from the optimization problem in Equation (8) for three hyperparameter sets $\boldsymbol{\theta}^{(1)}$=[1.0,0.04,0,3.0], $\boldsymbol{\theta}^{(2)}$=[1.0,0.33,0,1.8], and $\boldsymbol{\theta}^{(3)}$=[1.0,0.42,0.42,1.37] that satisfy the constraints in Table 1. The support domain is fixed as [0,2] for all the cases. $\boldsymbol{\theta}^{(1)}$ corresponds to a Gaussian distribution with the mean 1.0 and the standard deviation 0.2. The obtained SDF shows good agreement with this distribution given by the dashed line in the figure, indicating that the approximation accuracy of SDF is sufficient. It should be noted that the SDF is defined on the bounded support set [0,2], while the actual Gaussian distribution has infinete support set. Similarly, $\boldsymbol{\theta}^{(2)}$ corresponds to a uniform distribution with the lower bound 0 and the upper bound 2.0 and $\boldsymbol{\theta}^{(3)}$ corresponds to a bimodal distribution. As can be observed, even if the mean is fixed as 1.0, SDF can approximate various distributions by changing the remaining hyperparameters as appropriate. Therefore, by inferring the hyperparameters of SDF through the hierarchical Bayesian approach, it is expected to estimate arbitrary probability distributions that are plausible to reproduce observations without limiting constraints on the distribution families.

Based on the moment constraints in Table 1, the support domain of the hyperparameters conditional to the support domain of $[\underline{x}, \overline{x}]$ of x can be analytically obtained as shown in Table 2. It should noted that the support domain is not defined for the skewness and kurtosis but for their unnormalized parameters, i.e., the third- and fourth-order central moments. In the proposed hierarchical Bayesian approach, the prior distribution of the hyperparameters $\boldsymbol{\theta}$ is given as a uniform distribution over the support domain.

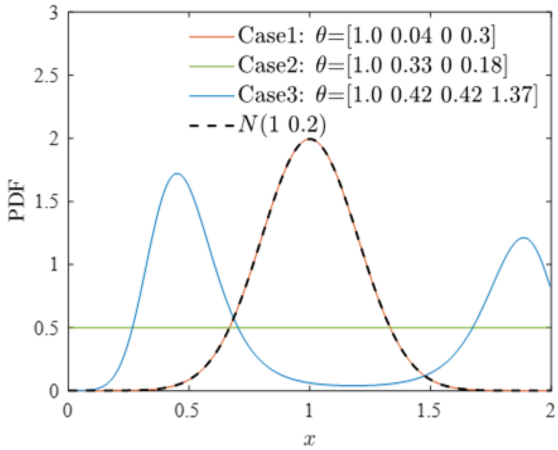

Figure 1 Examples of SDF in [0, 2].

Table 2 Suppor domain for $\boldsymbol{\theta}$

	Interval
μ	$[\underline{x}, \overline{x}]$
σ^2	$\left[0, \frac{(\overline{x}-\underline{x})^2}{4}\right]$
m_3	$\left[-\frac{(\overline{x}-\underline{x})^3}{6}, \frac{(\overline{x}-\underline{x})^3}{6}\right]$
m_4	$\left[0, \frac{(\overline{x}-\underline{x})^4}{12}\right]$

Table 3 Numerical model functions

Linear	$y = 0.2x + 1$
Monotonically increasing nonlinear	$y = \exp(0.2x)$
Convex nonlinear	$y = 0.1(x+1)^2 + 0.5$

3. NUMERICAL EXAMPLES

3.1. *Problem descriptions*

A numerical model *y=f(x)*, which can be described by one parameter for both input and output, is assumed as a preliminary study, and three input-output relationships, i.e., linear, monotonically increasing nonlinear, and convex nonlinear, are considered, as shown in Table 3. Figure 2 shows these three input-output relationships in the range of [-5,5].

Consider the case where the parameter *x* follows a standard normal distribution. From this distribution, 1000 samples of x are generated, and the model responses with the relationships in Table 3 are also computed. Histograms of the model outputs are given in Figure 3. In the linear case, the histogram shows good agreement with the Gaussian distribution with the mean 1.0 and the standard deviation 0.2 given by the dashed line in the figure. In contrast, as the nonlinearity is increased, the histograms show deviations from the Gaussian distribution. Herein, these 1000 samples of the model outputs are employed as the observed data.

3.2. *Comparison of the conventional and hierarchical Bayes approaches*

With the aforementioned problem descriptions, both the conventional and hierarchical Bayesian approaches are employed. In the conventional Bayesian approach, the prior distribution of *x* is assumed to be a uniform distribution within [-5,5], and posterior distribution is estimated by TMCMC. The obtained posterior distributions are summarized in Figure 4 in the order of linear, monotonically increasing nonlinear, and convex nonlinear cases. The dashed lines in the figures

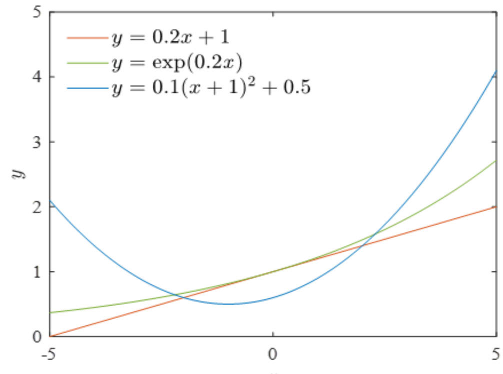

Figure 2 Model functions.

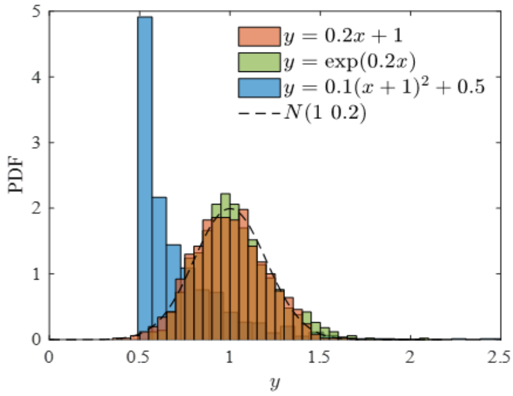

Figure 3 Model output histograms.

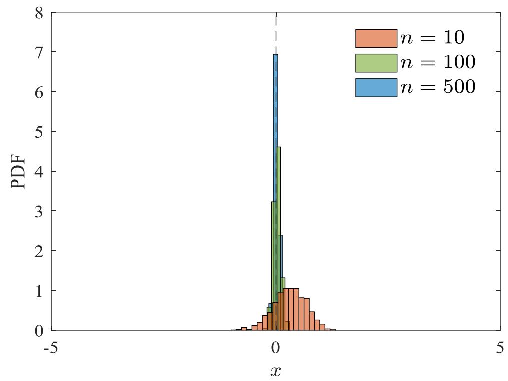

Figure 5 Influence of sample size.

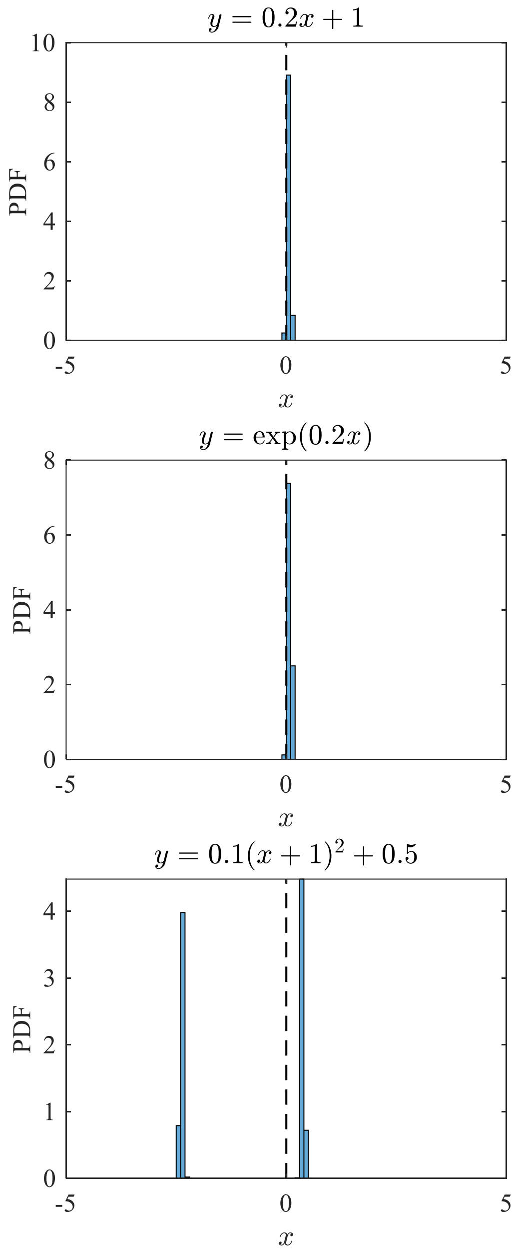

Figure 4 Posterior PDFs by conventional Bayes.

corresponds to the mean (i.e., x=0) of the target standard normal distribution. In the linear and monotonically increasing nonlinear cases, the posterior distribution is peaked around x=0, indicating that the maximum likelihood value of x can be estimated with a very high degree of accuracy due to a sufficient number of observations (1000 samples). In contrast, in the convex nonlinear case, the posterior distribution shows multimodality with peaks around x = -2.4 and $\boldsymbol{x}$ = 0.4. This is due to the non-unique solution for the input-output relationship shown in Figure 2. In addition, a peak x = 0.4 is still bit far from the true mean x=0 compared to the former two cases. This is because the modeling error is assumed to follow a Gaussian distribution while it actually does not the case due to the strong nonlinearlity as shown in Figure 3.

In addition, the number of observations n is changed to 10, 100, amd 500 to investigate its effect on the resultant posterior distribution. The posterior distributions obtained for the linear case are shown in Figure 5. As can be seen, the posterior distribution for n=100 shows a relatively large variance and a deviation is also found between its peak value and the maximum likelihood value of x (i.e., x=0). In contrast, for the cases of n=100 and n=500, the posterior variance is reduced as the number of observations increases and the posterior distribution is sharply distributed around the maximum likelihood value of x. This is because the likelihood function in equation (2) is given by the sum of products of the data, and the contribution of the likelihood function to the posterior distribution becomes more dominant as the number of data increases. These results also confirm that, in the conventional Bayesian approach, the posterior distribution shows the degree of plausibility according to the observations and does not denotes the probability distribution that the parameters follow to reproduce the observed data.

In the proposed hierarchical Bayesian approach, on the other hand, the support domain of x is set to be [-5,5]. Then, the prior distribution of the hyperparameters $\boldsymbol{\theta}$ of the SDF is assumed as a uniform distribution within their support domain that is derived from Table 2. The posterior distribution of $\boldsymbol{\theta}$ is estimated using TMCMC. The numbers of bins for the Bhattacharya distance evaluation and the SDF estimation are set to be n_{bin}=10 and n_b=50, respectively. The obtained posterior distributions are shown in Figure 6 in the order of the linear, monotonically increasing nonlinear, and convex nonlinear cases. The dashed lines in the figure correspond to the mean, variance, skewness, and kurtosis of the standard normal distribution, respectively.

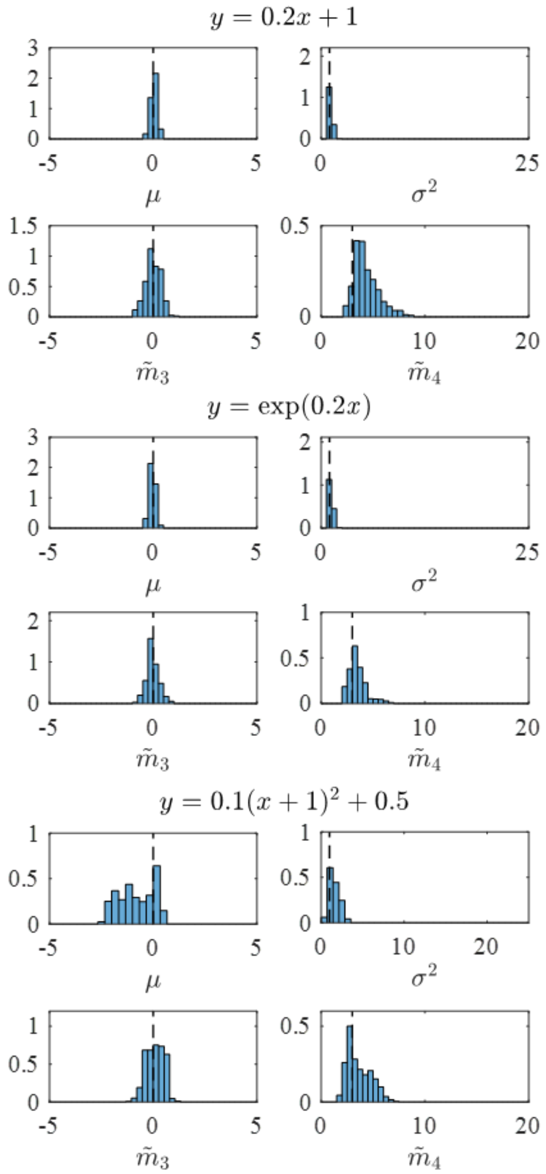

Figure 6 Posterior PDFs by hierarchical Bayes.

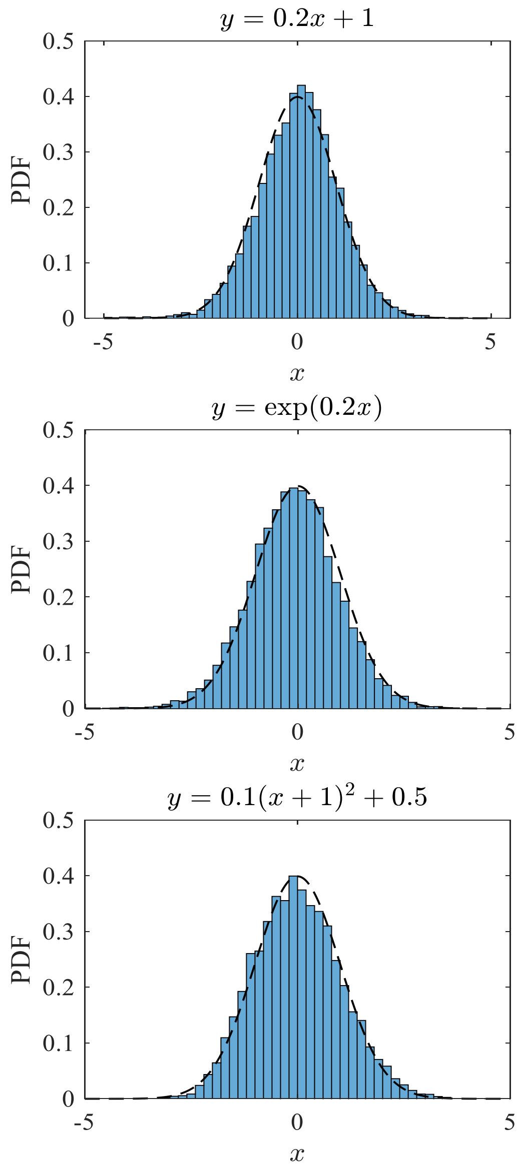

Figure 7 Infuluence of sample size to the hierarchcal Bayes results.

For the linear and monotonically increasing nonlinear cases, the posterior distributions are sharply distributed around the corresponding hyperparameter values of the standard normal distribution. Compared to the mean and variance, the posterior distributions of skewness and kurtosis have larger support sets, indicating that these parameters are relatively insensitive to the variability of the model response. On the other hand, in the convex nonlinear case, the posterior distributions of mean and skewness are flat compared to the above two cases, indicating that the estimation accuracy is relatively low.

Figure 7 illustrates histograms of the SDFs obtained by assining the most probable values of the posterior distributions. For all cases, the SDF shows good agreement with the PDF of the standard normal distribution shown by the dashed line in the figure. This demonstrates that the proposed hierarchical Bayesian approach can accurately estimate the probability distribution of model parameters from observations, even in the case with strong nonlinearity.

3.3. *Robustness to various Distribution Shapes*

Finally, to verify the applicability of the proposed method to various types of distributions, the target probability distribution of x is changed from the standard normal distribution using SDFs. The support set is fixed as [-5,5], and five sets of the hyperparameters $\boldsymbol{\theta}^{(1)}$=[-1.2,1.0,0.5,3.0], $\boldsymbol{\theta}^{(2)}$=[1.2,1.0,-0.5,3.0], $\boldsymbol{\theta}^{(3)}$=[0,2.2,0,2.25], $\boldsymbol{\theta}^{(4)}$=[0,0.5,0,4.0], and $\boldsymbol{\theta}^{(5)}$=[1.0,1.5,0.8,2.0] are then considered. By decreasing the mean and increasing the skewness from the standard normal distribution, $\boldsymbol{\theta}^{(1)}$ corresponds to the SDF biased to the left, and oppsite for $\boldsymbol{\theta}^{(2)}$. In addition, by increasing the variance and decreasing the kurtosis from the standard normal distribution, we can see that $\boldsymbol{\theta}^{(3)}$ provides a flat distribution, and conversely, $\boldsymbol{\theta}^{(4)}$ provides a sharp distribution. Furthermore, $\boldsymbol{\theta}^{(5)}$ corresponds to the SDF with bimodality.

Figure 8 shows histograms of the SDFs obtained by the proposed approach. In all cases, the SDF is in good agreement with the target probability distribution given by the dashed line. This indicates that the proposed approach can infer a wide range of parameter distributions from the indirect measurements. The proposed approach does not require any limiting constraints on the parameter distributions; thus, it has larger applicability compared to the most of the hierarchical Bayesian approach where it is assumed that the model parameters follow Gaussian distributions.

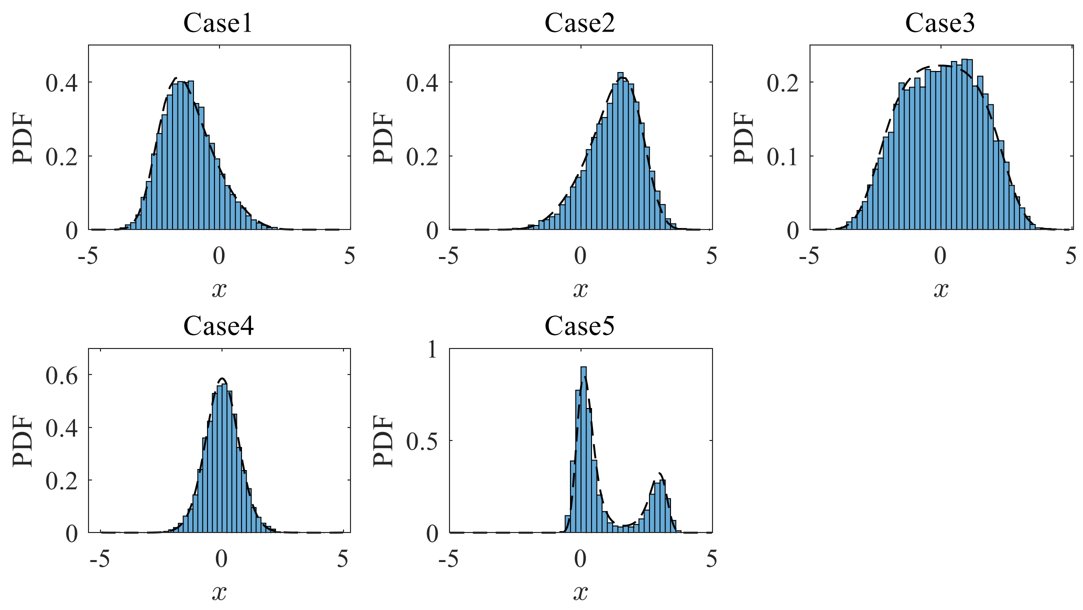

Figure 8 Results of the proposed method for various parameter distributions.

4. CONCLUSIONS

The findings of this study are as follows.

1) The proposed hierarchical Bayesian inversion method can accurately estimate the probability distribution of model parameters even for the case where the input-output relationship of the numerical model shows nonlinearity.
2) By updating the hyperparameters of the SDF, the proposed method could accurately estimate various types of distributions as appropriate.

REFERENCES

Beck, J. L. and Yuen, K. V. 2004. Model Selection Using Response Measurements: Bayesian Probabilistic Approach, *J. Eng. Mech*., 130(2), 192-203.

Betz, W., Papaioannou, I. and Straub, D. 2016. Transitional Markov Chain Monte Carlo: Observations and Improvements, *J. Eng. Mech.*, 142(5), 04016016.

Bhattacharyya, A. 1946. On a measure of divergence between two multinominal populations, *Indian J. Stat*., 7(4), 401-406.

Ching, J. and Chen, Y. C. 2007. Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class Selection, and Model Averaging, *J. Eng. Mech*., 133(7), 816-832,.

Crespo, L. G., Kenny, S. P., Giesy, D. P. and Stanford, B. K. 2018. Random variables with moment-matching staircase density functions, *Appl. Math. Model*., 64, 196-213.

Hayashi, G., Kim, C. W., Mimasu, T., Goi, Y., Yoshida, I. and Sugiura, K. 2018. FE modeling of a steel plate twin girder bridge using particle filter, *Journal of Japan Society of Civil Engineering*, Ser. A2, 74(2). I_705- I_714 (in Japanese).

Jia, X., Sedehi, O., Papadimitriou, C., Katafygiotis, L. S. and Moaveni, B. 2022. Hierarchical Bayesian modeling framework for model updating and robust predictions in structural dynamics using modal features, *Mech. Syst. Signal Process*., 170, 108784.

Kitagawa, G. 1996. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, *J. Comp. Graph. Stat*, 5, 1-25.

Kitahara, M., Broggi, M. and Beer, M. 2021. Residual seismic performance estimation of seismic-isolated bridges based on model updating using approximate Bayesian computation, *Journal of Japan Society of Civil Engineering*, Ser. A1, 77(4), I_61- I_70 (in Japanese).

Kitahara, M., Bi., S., Broggi, M. and Beer, M. 2022. Nonparametric Bayesian stochastic model updating with hybrid uncertainties, *Mech. Syst. Signal Process*., 163, 108195.

Kuroda, R. and Nishio, M. 2016. Reliability assessment of an existing steel girder bridge using posterior distributions of model parameters, *Journal of Japan Society of Civil Engineering*, Ser. A1, 72(3), 380-392 (in Japanese).

Mares, C., Mottershead, J. E. and Friswell, M. I.2006. Stochastic model updating: Part 1-theory and simulated example, *Mech. Syst. Signal Process*., 20(7), 1674-1695.

Matsuoka, K., Shinozuka, S. and Kaito, K. 2020. Structural model update considering uncertainty and reliability assessment of strengthening effect: Application to additional supporting strengthening of a high-speed railway bridge, *Journal of Japan Society of Civil Engineering*, Ser. A1, 76(3), 560-579 (in Japanese).

Mottershead, J. E., Link, M. and Friswell, M. I. 2011. The sensitivity method in finite element model updating: A tutorial, *Mech. Syst. Signal Process*., 25(7), 2275-2296.

Patelli, E., Govers, Y., Broggi, M., Gomes, H. M., Link, K. and Mottershead, J. E. 2017. Sensitivity or Bayesian model updating: comparison of techniques using the DLR AIRMOD test data, *Arch. Appl. Mech*., 87, 905-925.

Tierney, L. 1994. Markov Chains for Exploring Posterior Distributions, *Annal. Stat.*, 22(4)4, 1701-1761.