Distribution-free stochastic model updating with staircase density functions

M. Kitahara The University of Tokyo, Tokyo, Japan

T. Kitahara Kanto Gakuin University, Yokohama, Japan

S. Bi University of Strathclyde, Glasgow, UK

M. Broggi & M. Beer Leibniz University Hannover, Hannover, Germany

ABSTRACT: In stochastic model updating, hybrid uncertainties are typically characterized by the distributional p-box. It assigns a certain probability distribution to model parameters and assumes its hyper-parameters as interval values. Thus, regardless of the updating method employed, the distribution family needs to be known a priori to parameterize the distribution. Meanwhile, a novel class of the random variable, called staircase random variable, can discretely approximate a wide range of distributions by solving moment-matching optimization problem. The first author and his co-workers have recently developed a distribution-free stochastic updating framework, in which model parameters are considered as staircase random variables and their hyper-parameters are inferred in a Bayesian fashion. This framework can explore an optimal distribution from a broad range of potential distributions according to the available data. This study aims to further demonstrate the capability of this framework through a simple numerical example with a parameter following various types of distributions.

1 INTRODUCTION

Model updating has been widely accepted as a fascinating technique to mitigate the discrepancy between model outputs and measurements (Mottershead and Friswell, 1993). The conventional deterministic model updating aims to calibrate model parameters to find their optimal values from a single set of the measurement data. It has been successfully employed in a wide range of practical applications, including the correction of complex finite element models. However, this approach considers the model and measurement data as deterministic, ignoring uncertainties in both modeling and measuring processes.

In contrast, the stochastic model updating aims to calibrate not parameters themselves but their probability distributions, so that corresponding model outputs reproduce the uncertainty characteristics of the multiple sets of the measurement data (Mares et al., 2006). This can be achieve by finding the optimal values of the distribution hyper-parameters that minimize a stochastic distance between model outputs and measurement datasets. Bi et al. (2019) employed the Bhattacharyya distance and developed a Bayesian updating framework that utilizes a distance-based approximate likelihood. The capability of this framework has been demonstrated upon complex applications, e.g., the first edition of NASA UQ problem (Crespo et al., 2014).

More recently, the latest edition of NASA UQ problem (Crespo and Kenny, 2021) has posed a challenge in the stochastic model updating to calibrate the parameter distributions without prior information about their distribution families. Motivated by this, the first author and his co-workers have developed a distribution-free stochastic updating framework, where the parameter distributions

DOI: 10.1201/9781003323020-81

are approximated by staircase density functions (SDFs) (Kitahara et al., 2022a; Kitahara et al., 2022b). SDF is a discrete probability density function (PDF) defined for the staircase random variable (SRV) which has a bounded support set and prescribed values for the first four moments (Crespo et al., 2018). It has no analytical solution but enables to discretely approximate a broad range of distributions by solving moment-matching optimization problem. This distribution-free stochastic updating framework has been demonstrated that it enables to calibrate the parameter distributions with no constraining hypothesis on the distribution formats. However, its capability to estimate various types of distributions has not been thoroughly investigated. Hence, this study aims at further demonstrating the feasibility of this framework through a simple numerical example with a parameter following various types of distributions, including heavy-tailed and multi-modal distributions.

2 OVERVIEW OF DISTRIBUTION-FREE STOCHASTIC UPDATING FRAMEWORK

2.1 Bhattacharyya distance

In the stochastic model updating, the stochastic discrepancy between model outputs and measurement datasets needs to be quantified and minimized. Let $\mathbf{Y}_D = \{\mathbf{y}^{(i)}; i = 1, \dots, N_D\}$ be N_D sets of the measurement data $\mathbf{y} \in \mathbb{R}^m$. Let also $\mathbf{Y}_S = \{\mathcal{M}(\mathbf{x}^{(i)}); i = 1, \dots, N_S\}$ be the corresponding N_S model outputs parameterized through a model parameter vector $\mathbf{x} \in \mathbb{R}^n$. In this study, the stochastic discrepancy between \mathbf{Y}_S and \mathbf{Y}_D is quantified as the Bhattacharyya distance, and its theoretical definition is given as:

$$d_B(\mathbf{Y}_S, \mathbf{Y}_D) = -\log\left[\int_{\mathcal{Y}} \sqrt{f_{\mathbf{Y}_S}(\mathbf{y}) f_{\mathbf{Y}_D}(\mathbf{y})} \mathrm{d}\mathbf{y}\right]$$
(1)

where $f_{(\cdot)}(\mathbf{y})$ represents the PDF of \mathbf{y} ; \mathbf{y} denotes the support domain \mathbf{y} which comprises the *m*-dimensional space. Equation (1) indicates that the Bhattacharyya distance measures the degree of overlap between two different distributions. However, the direct evaluation of Equation (1) is generally impractical because the PDF of \mathbf{Y}_D often cannot be precisely estimated due to the very limited number of available measurement datasets. Hence, Bi et al. (2019) proposed the so-called binning algorithm to discretely evaluate the Bhattacharyya distance as:

$$d_B(\mathbf{Y}_S, \mathbf{Y}_D) = -\log\left\{\sum_{i=1}^{N_{bin}} \sqrt{P_{\mathbf{Y}_S, i} P_{\mathbf{Y}_D, i}}\right\}$$
(2)

where N_{bin} indicates the total number of bins; $P_{(\cdot),i}$ represents the probability mass function of y at the *i*th bin. N_{bin} is set as $N_{bin} = 10^n$ in this study.

2.2 Staircase density functions

Let the model parameters x be independent random variables having the support set $[\underline{x}, \overline{x}]$ and prescribed values for the hyper-parameters $\theta_x = [\mu, m_2, m_3, m_4]$ that consists of the mean vector $\mu \in \mathbb{R}^n$, variance vector $m_2 \in \mathbb{R}^n$, third-order central moment vector $m_3 \in \mathbb{R}^n$, and fourth-order central moment vector $m_4 \in \mathbb{R}^n$. Note that, in practice, the third- and fourthorder central moments are normalized by the variance as the skewness \widetilde{m}_3 and kurtosis \widetilde{m}_4 , respectively, in the updating procedure. Any such variables must satisfy the feasibility conditions $g(\theta_x) \leq 0$ given in Crespo et al. (2018). The realizations of θ_x that satisfies these conditions constitute the feasible domain $\Theta = \{\theta_x : g(\theta_x) \leq 0\}$.

Let also the support set $[\underline{x}, \overline{x}]$ partitioned into n_b sub-intervals with the equal length of $\kappa = (\overline{x} - \underline{x})/n_b$, x can be then considered as SRVs the PDF of which is expressed as:

$$f_X(\mathbf{x}) = \begin{cases} l^j \ \forall \mathbf{x} \in (\mathbf{x}^j, \mathbf{x}^{j+1}], \ \forall j = 1, 2, \cdots, n_b \\ 0 \ \text{otherwise} \end{cases}$$
(3)

where $l^{j}(=\prod_{i=1}^{n} l_{i}^{j})$ is the PDF value at the *j*th bin; $\mathbf{x}^{j} = \mathbf{x} + (j-1)\mathbf{\kappa}$. $n_{b} = 25$ is utilized in this study, The marginal staircase densities are obtained by solving the optimization problem:

$$\underset{l_i}{\operatorname{argmin}} \{ J(l_i) : \mathbf{A}(\boldsymbol{\theta}_{\boldsymbol{x}_i}, n_b) \boldsymbol{l} = \boldsymbol{b}(\boldsymbol{\theta}_{\boldsymbol{x}_i}), \boldsymbol{\theta}_{\boldsymbol{x}} \in \Theta \}$$
(4)

where J denotes the cost function; AI = b are moment matching constraints. This optimization problem is convex when the cost function is a convex function. In this study, the cost function is defined as follows based on the principle of maximum entropy:

$$J(l_i) = \kappa_i \log l_i^T l_i \tag{5}$$

2.3 Approximate Bayesian computation

In the proposed stochastic model updating framework, approximate Bayesian computation (ABC) (Beaumont, 2019) is employed. ABC is based on the well-known Bayes' theorem:

$$P(\boldsymbol{\theta}_{x}|\mathbf{Y}_{D}) = \frac{\tilde{\mathcal{L}}(\mathbf{Y}_{D}|\boldsymbol{\theta}_{x})P(\boldsymbol{\theta}_{x})}{P(\mathbf{Y}_{D})}$$
(6)

where $P(\theta_x)$ denotes the prior distribution of the hyper-parameters θ_x that reflects one's initial beliefs on θ_x ; $P(\theta_x | \mathbf{Y}_D)$ is the posterior distribution of θ_x that represents the posterior state of knowledge on θ_x ; $\tilde{\mathcal{L}}(\mathbf{Y}_D | \theta_x)$; is the so-called approximate likelihood function that serves as the connection between the measurement datasets \mathbf{Y}_D and θ_x ; $P(\mathbf{Y}_D)$ means the evidence ensuring that the integral of the posterior distribution equal to one.

Given the support set $[\underline{x}, \overline{x}]$, the support set of θ_x can be obtained based on the feasibility conditions as:

$$\Omega = [\underline{\theta}_{x}, \overline{\theta}_{x}] = \begin{bmatrix} \underline{\mu}, \overline{\mu} \\ \underline{m}_{2}, \overline{m}_{2} \\ \underline{m}_{3}, \overline{m}_{3} \\ \underline{m}_{4}, \overline{m}_{4} \end{bmatrix} = \begin{bmatrix} \underline{x}, x \\ \mathbf{0}, \frac{(\overline{x} - \underline{x})^{2}}{4} \\ -\frac{(\overline{x} - \underline{x})^{3}}{6\sqrt{3}}, \frac{(\overline{x} - \underline{x})^{3}}{6\sqrt{3}} \\ -\frac{(\overline{x} - \underline{x})^{4}}{6\sqrt{3}}, \frac{(\overline{x} - \underline{x})^{4}}{6\sqrt{3}} \end{bmatrix}$$
(7)

In this study, the hyper-parameters θ_x are assumed to be independent each other and the prior distribution is expressed as:

$$P(\theta_x) = U_{4n}(\Omega) I_{\Theta}(\theta_x) \tag{8}$$

where $U_{4n}(\Omega)$ denotes the PDF of 4n independent multivariate uniform distribution on Ω ; I_{Θ} denotes the indicator function that equals to one if $\theta_x \in \Theta$ and otherwise equals to zero.

The approximate likelihood function is defined using an arbitrary kernel. The principle behind it is that it should return a high value when the stochastic discrepancy between the model outputs and measurement datasets is small and, conversely it penalizes θ_x that leads to a large stochastic discrepancy. In this study, the Gaussian kernel is utilized and the stochastic discrepancy is measured by the Bhattacharyya distance. Thus, the approximate likelihood function is defined as:

$$\tilde{\mathcal{L}}(\mathbf{Y}_D|\boldsymbol{\theta}_x) = \frac{1}{\varepsilon\sqrt{2\pi}} \exp\left\{-\frac{d_B(\mathbf{Y}_S, \mathbf{Y}_D)^2}{2\varepsilon^2}\right\}$$
(9)

where ε indicates the scaling parameter which controls the centralization degree of the posterior distribution. A smaller ε provides a more peaked posterior distribution, which is more likely to converge to the true values, but needs more computation burden for convergence. In this study, it is set as $\varepsilon = 0.02$.

The posterior distribution in Equation (6) generally has no analytical solution and thus needs to be estimated using advanced sampling techniques. In this study, the transitional Markov chain Monte Carlo (TMCMC) sampler (Ching and Chen, 2007) is employed. TMCMC is a sequential

approach sampling from a series of intermediate distributions which will progressively converge to the true posterior distribution. The *i*th intermediate distribution is expressed as:

$$P_j(\boldsymbol{\theta}_{\boldsymbol{x}}) \propto \mathcal{L}(\mathbf{Y}_D | \boldsymbol{\theta}_{\boldsymbol{x}})^{\beta_j} P(\boldsymbol{\theta}_{\boldsymbol{x}})$$
(10)

where β_i indicates the so-called reduction coefficient. Its value starts from $\beta_i = 0$ in the initial step and gradually increases until $\beta_m = 1$ in the final step. β_j is adaptively computed from the samples of the previous step. Markov chains then propagate new samples starting from the ones in the previous step with higher likelihood values. The reader can refer to Ching and Chen (2007) for details of the TMCMC sampler.

NUMERICAL EXAMPLES 3

3.1 Problem descriptions

The proposed stochastic model updating framework is demonstrated upon a simple three degree-of-freedom (DOF) spring-mass system shown in Figure 1. The stiffness coefficients k_1 , k_2 , and k_3 are supposed to be uncertain with the uncertainty characteristics summarized in Table 1. k_1 and k_2 follow Gaussian distributions, whose hyper-parameters, i.e., means and standard deviations, are not fully determined but fall within given intervals as listed in the third column of Table 1. On the contrary, the distribution family of k_3 is unknown before model updating and only the support set is given as [5.0, 7.0]. Hence, it is assumed to be characterized as a SRV and its hyper-parameters are fall within the intervals computed as Equation (7). As a consequence, in total eight hyper-parameters are treated as interval-valued parameters and updated through the proposed procedure. Besides these uncertain parameters, the remaining parameters (i.e., stiffness coefficients k_4 to k_6 and masses m_1 to m_3) are set to be constants with determined values: $k_i = 5.0$ N/m (i = 4, 5, 6), $m_1 = 0.7$ kg, $m_2 = 0.5$ kg, and $m_1 = 0.3$ kg.



Figure 1. 3-DOF spring-mass system.

Table 1. Uncertainty characteristics of the model parameters.

Parameter	Uncertainty characteristics		
	Distribution family	Support set/ hyper-parameters	Target values of hyper-parameters
k_1	Gaussian	$\mu_1 \in [3.0, 7.0], \sigma_1 \in [0.0, 0.5]$	$\mu_1 = 4.0, \ \sigma_1 = 0.3$
k_2	Gaussian	$\mu_2 \in [3.0, 7.0], \sigma_2 \in [0.0, 0.5]$	$\mu_2 = 5.0, \ \sigma_2 = 0.1$
k_3	Unknown	$k_3 \in [5.0, 7.0]$	Given in Table 2
$k_4 - k_6, m_1 - m_3$	Deterministic	_	_

In addition to the prior information on the uncertainty characteristics, target values of the hyperparameters in k_1 and k_2 are shown in the last column of Table 1. On the other hand, to investigate the capability of the proposed approach calibrating a wide range of distributions without the prior knowledge about their distribution families, five different distributions presented in Figure 2 are

considered as the target distribution of k_3 . Properties of these distributions are also provided in Table 2. The first distribution is a (truncated) Gaussian distribution of which hyper-parameters $\theta_1 = [\mu_{11}, m_{21}, \tilde{m}_{31}, \tilde{m}_{41}]$ are given in Table 2. While the distribution is truncated because the support set of k_3 is a closed interval, its mean and variance are determined such that the support set covers more than the 99.99 % confidence interval of the original Gaussian distribution. The remaining distributions are given by SDFs with the hyper-parameters, $\theta_i = [\mu_{1i}, m_{2i}, \tilde{m}_{3i}, \tilde{m}_{4i}], i = 2, \dots 4$ The second distribution is a (left) skewed one having a positive skewness. The third distribution is a flat one having a larger variance and smaller kurtosis compared to the Gaussian distribution. The fourth distribution is a heavy-tailed distribution having a larger kurtosis. Finally, the fifth distribution is a bi-modal distribution.



Figure 2. PDF for the target distributions of k_3 .

Fable 2.	Target	distributions	of <i>k</i> ₃ .
----------	--------	---------------	----------------------------

Distribution format	Hyper-parameters
Truncated Gaussian Skewed Flat Heavy-tailed Bi-modal	$\begin{array}{l} \mu_{11}=6.0,\ m_{21}=0.04,\ \tilde{m}_{31}=0.0,\ \tilde{m}_{41}=3.0\\ \mu_{12}=5.7,\ m_{22}=0.06,\ \tilde{m}_{32}=0.5,\ \tilde{m}_{42}=3.0\\ \mu_{13}=6.0,\ m_{23}=0.14,\ \tilde{m}_{33}=0.0,\ \tilde{m}_{43}=2.25\\ \mu_{14}=6.0,\ m_{24}=0.04,\ \tilde{m}_{34}=0.0,\ \tilde{m}_{44}=4.2\\ \mu_{15}=6.0,\ m_{25}=0.10,\ \tilde{m}_{35}=0.8,\ \tilde{m}_{45}=2.0 \end{array}$

The outputs of the system are the three eigen-frequencies f_1 , f_2 and f_3 . The measurement datasets consisting of these eigen-frequencies are generated through multiple model evaluations with multiple sets of the model parameters sampled from their target distributions. In this study, the number of datasets are set as $N_D = 1000$. Figure 3 depicts the measurement datasets in the plane of the first and third frequencies for the case where the truncated Gaussian distribution is employed as the target distribution of k_3 . The reference range in the figure means the 95 % confidence interval of the sample distribution. Moreover, 1000 samples of the model parameters are generated by assigning a set of randomly selected initial values of the hyper-parameters, and subsequently 1000 initial model outputs are obtained through the model evaluations. These outputs are also presented in Figure 3. As can be seen, the scatters of the initial outputs are clearly apart from the measurement datasets; thus, model updating is necessary to obtain the model outputs as close as the measurement datasets.

3.2 Model updating results

For the case where the target distribution of k_3 is the truncated Gaussian distribution, totally 17 TMCMC iterations are executed to reach convergence. Figure 4 illustrates histograms of



Figure 3. Target relative position of the measurement datasets and initial model outputs.

1000 posterior samples of the eight hyper-parameters, i.e., μ_i and σ_i (i = 1,2) as well as θ_1 . The ranges of the horizontal axes are identical to the intervals of the prior uniform distribution. It can be observed that all the hyper-parameters are significantly updated from the prior distribution. The means of the posterior samples are obtained as the posterior estimates of the hyper-parameters and presented in Table 3. The posterior estimates show good agreement with the target values. It should be noted that a relatively large error in m_{21} can be caused due to its very small target value. The updated distribution of k_3 is then obtained as a SDF with the posterior estimates of the hyper-parameters θ_1 . Figure 5 illustrates the updated distribution as the histogram of samples generated from the SDF. As can be observed from the figure, the updated distribution coincides well with the target truncated Gaussian distribution.



Figure 4. Posterior distribution of the hyper-parameters in histograms.

Figure 6 shows a relative position of the measurement datasets and updated model outputs. The updated outputs are obtained through the model evaluations with 1000 sets of the model parameters sampled from their updated distributions. Compared to the initial model outputs presented in Figure 3, the updated model outputs fit well with the measurement datasets,

Hyper-parameters	Target values	Posterior estimates ^a
μ_1	4.0	4.01/4.00/3.98/4.01/4.04 (1.0) ^b
σ_1	0.3	0.322/0.328/0.329/0.327/0.346 (15.3) ^b
μ_2	5.0	5.01/4.99/5.00/5.00/4.99 (0.2) ^b
σ_2	0.1	0.100/0.098/0.094/0.098/0.088 (12.0) ^b
$\mu_{11}/m_{21}/\tilde{m}_{31}/\tilde{m}_{41}$	6.0/0.04/0.0/3.0	6.02 (0.3)/0.046 (15.0)/0.05/2.94 (2.0)
$\mu_{12}/m_{22}/\tilde{m}_{32}/\tilde{m}_{42}$	5.7/0.06/0.5/3.0	5.69 (0.2)/0.064 (6.7)/0.402 (19.6)/3.36 (12.0)
$\mu_{13}/m_{23}/\tilde{m}_{33}/\tilde{m}_{43}$	6.0/0.14/0.0/2.25	5.99 (0.2)/0.155 (10.7)/-0.006/2.30 (2.2)
$\mu_{14}/m_{24}/\tilde{m}_{34}/\tilde{m}_{44}$	6.0/0.04/0.0/4.2	6.00 (0.0)/0.045 (12.5)/0.174/4.05 (3.6)
$\mu_{15}/m_{25}/\tilde{m}_{35}/\tilde{m}_{45}$	6.0/0.1/0.8/2.0	6.01 (0.2)/0.106 (6.0)/0.740 (7.5)/2.01 (0.5)

^a Percentage errors compared to the target values in parentheses.

^b Posterior estimates for all the five cases in a row and their maximum percentage errors in parentheses.

which demonstrates that the proposed updating procedure enables to calibrate the model so that its outputs represent wholly the uncertainty characteristics of the measurement datasets.



Figure 5. Updated distributions of k_3 .

Similarly, the proposed updating procedure is also performed for the remaining cases with the different target distributions of k_3 . For all the cases, the obtained posterior estimates of the hyper-parameters are summarized in Table 3. All the posterior estimates of the hyper-parameters show good agreement with the target values, including the higher-order moments such as the skewness and kurtosis in the SDFs, and the maximum percentage error compared to the target values is less than 20 %. The updated distributions of k_3 , which assign the posterior estimates of the hyper-parameters are also illustrated in Figure 5 for these cases. It can be seen that the updated distributions coincide well with the target distributions, indicating that the proposed updating procedure can quantify the parameter uncertainty as an appropriate probability distribution including heavy-tailed and multi-modal distributions. Finally, the updated model outputs are obtained through the model evaluations with the updated parameter distributions and compared with the measurement datasets. While relative positions of the updated model outputs and measurement datasets are not further provided for the sake of brevity, it is confirmed that the updated model outputs are properly tuned for all the cases and fit well with the measurement datasets.





4 CONCLUSIONS

In this study, we present a distribution-free stochastic model updating framework to quantify the parameter uncertainty that forms a broad range of probability distributions, including heavy-tailed and multi-modal distributions, without the prior knowledge about their distribution families. The unknown parameter distribution is characterized by SDF, and it is assumed that only its support set is known a priori. Its hyper-parameters, i.e., the first four moments, are then inferred through the ABC procedure aiming at minimizing the Bhattacharyya distance between the model outputs and measurement datasets. The proposed updating framework is demonstrated on a simple 3-DOF spring-mass system, in which five different distributions are assumed as the target distribution of a model parameter. The results demonstrate that the proposed procedure has a potential to calibrate the arbitrarily parameter distribution as appropriate so that the model outputs recreate wholly the uncertainty characteristics of the measurement datasets.

REFERENCES

- Beaumont, M.A. 2019. Approximate Bayesian Computation, Annual Review of Statistics and Its Application 6: 379–403.
- Bi, S. & Broggi, M. & Beer, M. 2019. The role of the Bhattacharyya distance in stochastic model updating. *Mechanical Systems and Signal Processing* 117:437–452.
- Ching, J. & Chen, Y.C. 2007. Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class Selection, and Model Averaging. *Journal of Engineering Mechanics* 133(7): 816–832.
- Crespo, L.G. & Kenny, S.P. & Giesy, D.P. 2014. The NASA Langley multidisciplinary uncertainty quantification challenge. In Proc. 16th AIAA Non-Deterministic Approaches Conf., Maryland.
- Crespo, L.G. & Kenny, S.P. & Giesy, D.P. & Stanford, B.K. 2018. Random variables with moment-matching staircase density functions. *Applied Mathematical Modelling*, 64: 196–213.
- Crespo, L.G. & Kenny, S.P. 2021. The NASA Langley multidisciplinary uncertainty quantification challenge. *Mechanical Systems and Signal Processing* 152: 107405.
- Kitahara, M. & Bi, S. & Broggi, M. & Beer, M. 2022. Nonparametric Bayesian stochastic model updating with hybrid uncertainties. *Mechanical Systems and Signal Processing* 163: 108195.
- Kitahara, M. & Bi., S. & Broggi, M. & Beer, M. 2022. Distribution-free stochastic model updating of dynamic systems with parameter dependencies. *Structural Safety* 163: 108195.
- Mares, C. & Mottershead, J.E. & Friswell, M.I. 2006. Stochastic model updating: Part 1-theory and simulated example, *Mechanical Systems and Signal Processing* 20(7): 1674–1695.
- Mottershead, J.E. & Friswell, M.I. 1993. Model Updating In Structural Dynamics: A Survey. *Journal of Sound and Vibration* 167(2): 347–375.