Advances in Reliability, Safety and Security, Part 1 – Kołowrocki, Dąbrowska (eds.)
 ISBN 978-83-68136-13-5 (printed), ISBN 978-83-68136-00-5 (electronic)

Advances in Reliability, Safety and Security

ESREL 2024
Monograph Book Series

Stochastic Model Updating Using Jenson-Shannon Divergence For Calibration And Validation Under Limited Data

Adolphus Lye[a], Scott Ferson[b], Sicong Xiao[a]

[a]*Singapore Nuclear Research and Safety Institute, National University of Singapore, Singapore*
[b]*Department of Civil and Environmental Engineering, University of Liverpool, Liverpool, United Kingdom*

Abstract

The paper proposes the use of the Jenson-Shannon divergence as the distance function within the distance-based Approximate Bayesian computation framework. Under such framework, the distance function serves to quantify the difference between the observed data and the model predictions and from there, assigns higher statistical importance to model parameter samples which yields high degree of agreement between the model predictions and the observed model and lower statistical models to model parameter samples which achieve otherwise. The objective is to illustrate the feasibility of the Jenson-Shannon divergence towards stochastic model updating for the subsequent model calibration and validation under limited data. To achieve this objective, the paper is divided into two parts: the first part of the paper introduces the mathematical formalism of the Jenson-Shannon divergence along with a review of the adaptive-binning algorithm from the literature; and the second part of the paper presents a case study in the form of the 2014 NASA-LaRC Uncertainty Quantification challenge problem involving a black-box model with uncertain model input parameters to demonstrate the feasibility of the proposed framework in stochastic model updating towards model calibration and validation under limited data.

Keywords: stochastic model updating, approximate Bayesian computation, Jenson-Shannon divergence, probability box, fuzzy set, transitional ensemble Markov chain monte carlo, area metric, validation

1. Introduction

In recent years, the need for model parameter identification under uncertainty has gained significant attention. Such uncertainties may stem from measurement "noise", input model parameter variability, and model incertitude (Lye and Marino, 2023d). Broadly speaking, the underlying uncertainty can be classified into two distinct types (Kennedy and O'Hagan, 2001): 1) Aleatory (Type I uncertainty); and 2) Epistemic (Type II uncertainty). Details to the respective types of uncertainty can be found in the literature by Kennedy and O'Hagan (2001). A well-known approach to quantify such uncertainties is stochastic model updating which has been implemented in numerous areas such as engineering reliability (Faes et al., 2021; Lye et al., 2020, 2023a), structural health monitoring (Lye et al., 2022b, 2023b), and model validation (Ferson et al., 2008). The latter would be of interest in the context of the paper.

A key challenge in engineering is the need for credible computational models which can simulate the actual system of interest. This presents the necessity to develop model calibration and validation capabilities which are robust under limited data. In the context of the paper, calibration refers to the process of "adjusting" the model such the model prediction matches the existing data (i.e., the training data) while validation refers to the process of assessing how well the model prediction agrees with a new data set (i.e., validation data).

To perform stochastic model updating for engineering systems, Bi et al. (2018) proposed the distance-based Approximate Bayesian computation framework involving the Bhattacharyya distance function (Bi et al., 2018). Since then, other distance functions have been implemented in recent literature including the 1-Wasserstein (Gray et al., 2022; Lye et al., 2023a), and the Bray-Curtis (Zhao et al., 2022) distance functions. Details on each of these distance functions can be found in the respective reference. For the work presented in the paper, the

implementation of the Jenson-Shannon divergence as the distance function is proposed as the uncertainty metric to quantify the difference the distribution of the observed data and that of the model prediction.

The research objectives of the paper are: 1) to present a distance-based Approximate Bayesian Computation framework involving the Jenson-Shannon divergence as the distance function for stochastic model updating; 2) to present a review of the mathematical formalism behind the Jenson-Shannon divergence; and 3) to demonstrate the feasibility of the proposed framework through a case study based on the recent 2014 NASA-LaRC Uncertainty Quantification challenge problem involving the task of calibrating and validating a black-box model to be calibrated under limited data. To achieve the above objectives, the paper is outlined as follows: Section 2 reviews the Bayesian model updating framework, the concept of Approximate Bayesian Computation, and the Transitional Ensemble Markov Chain Monte Carlo method; Section 3 reviews the concept of the distance metric and that of the Jenson-Shannon divergence; Section 4 presents the 2014 NASA-LaRC Uncertainty quantification challenge problem along with the results and discussions of implementing the proposed framework towards addressing the challenge; and Section 5 summarises the contents presented in the paper and reiterates the key learning points before drawing the paper to a close.

2. Bayesian model updating

A popular approach towards stochastic model updating would be Bayesian model updating to which the mathematical formalism follows (Beck and Katafygiotis, 1998):

$$P(\boldsymbol{\theta}|\boldsymbol{D},M) = \frac{P(\boldsymbol{\theta}|M)\cdot P(\boldsymbol{D}|\boldsymbol{\theta},M)}{P(\boldsymbol{D}|M)} \tag{1}$$

where $P(\boldsymbol{\theta}|M)$ is the prior distribution reflecting the prior knowledge on the inferred parameter(s) $\boldsymbol{\theta}$ before collecting data $\boldsymbol{D}$, $P(\boldsymbol{D}|\boldsymbol{\theta},M)$ is the likelihood function reflecting the degree of agreement between the observed data $\boldsymbol{D}$ and the prediction from model M given $\boldsymbol{\theta}$, and $P(\boldsymbol{D}|M)$ is the evidence which ensures that the posterior integrates to one. Details on each of the above terms in (1) can be found in Lye (2023c) and McGurk et al. (2024). In general, the inferred parameter(s) can be time-invariant, or time-varying (Lye et al., 2023b). In the context of the paper, the focus is specifically on time-invariant parameters.

Due to $P(\boldsymbol{D}|M)$ being a numerical constant, the term itself can be neglected resulting in the posterior being expressed in its un-normalised form:

$$P(\boldsymbol{\theta}|\boldsymbol{D},M) \propto P(\boldsymbol{\theta}|M)\cdot P(\boldsymbol{D}|\boldsymbol{\theta},M) \tag{2}$$

As a result, direct Monte Carlo sampling technique becomes inapplicable and advanced Monte Carlo sampling techniques need to be implemented such as the Markov Chain Monte Carlo methods, Transitional Markov Chain Monte Carlo, and the Sequential Monte Carlo samplers (Lye et al., 2019). Detailed reviews on these sampling approaches are found in Lye (2023c). For the work presented in the paper, the Transitional Ensemble Markov Chain Monte Carlo (TEMCMC) method (Lye et al., 2022a) will be implemented.

2.1. Transitional Ensemble Markov Chain Monte Carlo

The TEMCMC sampler is a variant of the Transitional Markov Chain Monte Carlo (TMCMC) sampler developed by Ching and Chen (2007) which allows for the generation of samples from complex-shaped posteriors (e.g., very peaked or having multiple peaks) in an iterative manner. This is done using a series of intermediate functions known as transitional distributions P^j which is defined as:

$$P^j \propto P(\boldsymbol{\theta}|M)\cdot P(\boldsymbol{D}|\boldsymbol{\theta},M)^{\beta_j} \tag{3}$$

where $j \geq 0$ is the sampling iteration number, β_j is the tempering parameter such that $0 = \beta_0 < \beta_1 < \cdots < \beta_{m-1} < \beta_m = 1$, and m is the final iteration number. An important aspect of the sampler is the transition step size $\Delta\beta_j = \beta_j - \beta_{j-1}$ which determines how gradual the transition is from the prior to the posterior and, in turn, how well the sample distribution converges to the posterior (Lye et al., 2023d). The optimal value of $\Delta\beta_j$, and therefore the value of β_j, is determined by solving the following optimisation problem (Ching and Chen, 2007):

$$\beta_{j+1} = \text{argmin}_{\beta_{j+1}} \left\{ \left| \frac{\sigma\left(P\left(\boldsymbol{D}|\boldsymbol{\theta}_i^j,M\right)^{\Delta\beta_j}\right)}{\mu\left(P\left(\boldsymbol{D}|\boldsymbol{\theta}_i^j,M\right)^{\Delta\beta_j}\right)} \right| - 1 \right\} \tag{4}$$

where $\sigma(\cdot)$ and $\mu(\cdot)$ are the standard deviation and the mean operators respectively, and $i = 1, \ldots, N$ is the sample index with N being the total sample size.

A key advantage of the TEMCMC sampler over the TMCMC sampler is the improved sample mixing performance by the former owing to the implementation of the Affine-invariant Ensemble sampler (Goodman and Weare, 2010) as the Markov Chain Monte Carlo move kernel and the introduction of an adaptive step-size tuning algorithm to moderate the acceptance rate of the sampler within the optimal bounds of $[0.15, 0.50]$. Information to such study can be found in Lye (2023c) while details on the algorithm and conceptual descriptions to the TEMCMC sampler can be found in Lye et al. (2022a).

2.2. Approximate Bayesian Computation

As seen from (1), an important component of the Bayesian model updating procedure is the likelihood function $P(\boldsymbol{D}|\boldsymbol{\theta}, M)$. Assuming independence between the N_{obs} sets observations, the full analytical likelihood function is defined as follows:

$$P(\boldsymbol{D}|\boldsymbol{\theta}, M) = \prod_{k=1}^{N_{\text{obs}}} P(\boldsymbol{D}_k|\boldsymbol{\theta}, M) \tag{5}$$

However, in instances when the model M is computationally expensive, the evaluation of the full analytical likelihood function in (5) becomes almost impossible to evaluate since it requires a large amount of model evaluations. Such an issue becomes more pronounced with complex computational models involving a large number of input and output features. To overcome such issue, the Approximate Bayesian computation approach is implemented by substituting the full likelihood function in (5) with the distance-based approximate likelihood function proposed by Bi et al. (2018):

$$P(\boldsymbol{D}|\boldsymbol{\theta}, M) = \exp\left[-\frac{d^2}{\varepsilon^2}\right] \tag{6}$$

where d is the distance function, while ε is the width-factor serving as the pre-defined parameter controlling the centralisation degree of the posterior. It is proposed in Bi et al. (2018) that the width-factor should lie within the interval of $[10^{-3}, 10^{-1}]$.

3. Jenson-Shannon divergence

For the work presented in the paper, the Jenson-Shannon divergence is proposed as the distance metric for the distance-based ABC framework. The purpose of such distance metric is to quantify the statistical difference between the distribution of the model prediction and that of the observed data $\boldsymbol{D}$ using information entropy. To provide a clear conceptual understanding on the Jenson-Shannon divergence, Section 3.1 presents the mathematical formalism to compute the distance function, and Section 3.2 presents the mathematical properties of the Jenson-Shannon divergence through a simple illustrative example which also serves as a verification exercise on the distance function algorithm.

3.1. Mathematical formalism

The Jenson-Shannon divergence is based on the Kullback-Leibler divergence d_{KL}. Given two d-dimensional distributions p_1 and p_2, the Kullback-Leibler divergence is defined as (Nielsen, 2019):

$$d_{\text{KL}}(p_1 \,||\, p_2) = \sum_{x_d=1}^{N_{\text{bin}}} \cdots \sum_{x_1=1}^{N_{\text{bin}}} p_1\left(b_{x_1,\ldots,x_d}\right) \cdot \log\left[\frac{p_1\left(b_{x_1,\ldots,x_d}\right)}{p_2\left(b_{x_1,\ldots,x_d}\right)}\right] \tag{7}$$

where N_{bin} is the total number of bins used to approximate the distributions p_1 and p_2. However, the implementation Kullback-Leibler divergence as the distance function would not be optimal for ABC for the following reasons: 1) it does not obey the symmetrical property (i.e., $d_{\text{KL}}(p_1 \,||\, p_2) \neq d_{\text{KL}}(p_2 \,||\, p_1)$); and 2) the function yields infinity in the event that the support of p_1 is not a subset of p_2. As such, the Jenson-Shannon divergence was developed to overcome such drawbacks and is defined as:

$$d_{\text{JS}}(p_1, p_2) = \frac{1}{2} \cdot \left(d_{\text{KL}}(p_1 \,||\, T) + d_{\text{KL}}(p_2 \,||\, T)\right), \qquad \text{for } T = \frac{1}{2} \cdot (p_1 + p_2) \tag{8}$$

In the context of ABC, the interest would be to compute $d_{\text{JS}}(p_M, p_{\boldsymbol{D}})$ where p_M is the distribution of the model prediction while $p_{\boldsymbol{D}}$ is the distribution of the observed data.

An important aspect in the computation of the Jenson-Shannon divergence is the computation of the parameter N_{bin}. To do this, an adaptive-binning algorithm proposed by Zhao et al. (2022) can be implemented to which the procedure follows (Zhao et al., 2022):

1) Compute the parameter Δ^{sim} following:

$$\Delta^{\text{sim}} = \max\left(\max\left|\boldsymbol{D}_{i,m}^{\text{sim}} - \boldsymbol{D}_{j,m}^{\text{sim}}\right|\right) \tag{9}$$

where $i, j = 1, \ldots, N$ and $m = 1, \ldots, d$. Note that N is the total number of model evaluations corresponding to the total sample size from the posterior while $\boldsymbol{D}^{\text{sim}}$ is the simulated data (i.e. model prediction);

2) Compute the Euclidean distance d_E between $\boldsymbol{D}$ and $\boldsymbol{D}^{\text{sim}}$ follwing:

$$d_E = \sqrt{\left(\overline{\boldsymbol{D}^{\text{sim}}} - \overline{\boldsymbol{D}}\right) \cdot \left(\overline{\boldsymbol{D}^{\text{sim}}} - \overline{\boldsymbol{D}}\right)^{\text{T}}} \tag{10}$$

where $\overline{\boldsymbol{D}^{\text{sim}}}$ and $\overline{\boldsymbol{D}}$ are the means of the simulated data and that of the observed data respectively;

3) Compute the bin width parameter w following:

$$w = \frac{\log\left[\Delta^{\text{sim}} + 1\right]}{\max\left(N^{\frac{1}{3}}, N_{\text{obs}}^{\frac{1}{3}}\right)} \times \exp[d_E] \tag{11}$$

4) Finally, compute the number of bins N_{bin} following:

$$N_{\text{bin}} = \left\lceil \frac{\Delta^{\text{sim}}}{w} \right\rceil \tag{12}$$

As an adaptive variable, the number of bins should be bounded such that:

$$2 \leq N_{\text{bin}} \leq \left\lceil \frac{\max(N, N_{\text{obs}})}{10} \right\rceil. \tag{13}$$

3.2. Illustrative example

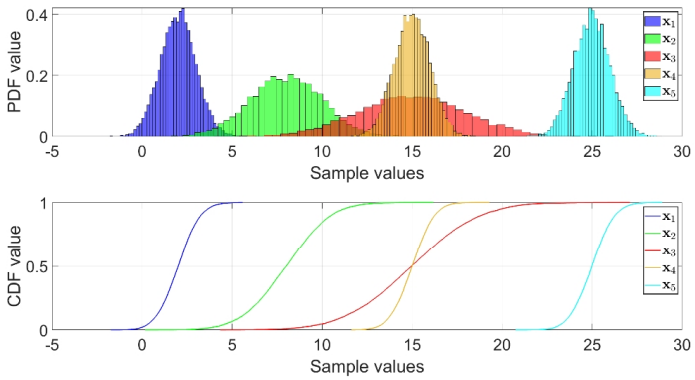

Fig. 1. Illustration to the PDFs and ECDFs of the different Normal distributions presented.

Table 1. Details on the Normal distribution shape parameters each sample set.

Sample	Distribution parameters
x_1	$\mu = 2, \sigma = 1$
x_2	$\mu = 8, \sigma = 2$
x_3	$\mu = 15, \sigma = 3$
x_4	$\mu = 15, \sigma = 1$
x_5	$\mu = 25, \sigma = 1$

For the illustrative example, five distinct sample sets of size $N = 10000$ each are considered and are denoted respectively as: $\{x_1, x_2, x_3, x_4, x_5\}$. Each of these sample sets follow a Normal distribution to which their corresponding shape parameters are defined in Table 1. As an illustration, the histogram representation of the PDFs as well as the Empirical Cumulative Distribution Functions (ECDFs) to the respective sample set are presented in Figure 1. For each pair of sample sets, the Jenson-Shannon divergence will be computed and the numerical results are presented in Table 2.

Table 2. Details on the Normal distribution shape parameters each sample set.

Sample set	x_1	x_2	x_3	x_4	x_5
x_1	0	0.307	0.404	0.693	0.693
x_2	0.307	0	0.314	0.427	0.691
x_3	0.404	0.314	0	0.192	0.347
x_4	0.693	0.427	0.192	0	0.693
x_5	0.693	0.691	0.347	0.693	0

Based on the results in Table 2, the following properties on the Jenson-Shannon divergence are observed: 1) there exists hard bounds on the Jenson-Shannon divergence value as it takes strictly positive values between 0 and 0.693. This is consistent with the theory of the distance function which takes a maximum value of $\log[2]$. In fact, it is seen that the Jenson-Shannon divergence takes the maximum value when there exists no overlap between two given PDFs – e.g., $\{x_1, x_4\}$ and $\{x_1, x_5\}$ sample set pairs; 2) the Jenson-Shannon divergence obeys the identity of indiscernibles which yields 0 when it computes the divergence between a pair of identically-distributed sample sets; 3) the Jenson-Shannon divergence possesses the symmetric property in that $d_{\mathrm{JS}}(x_i, x_j) = d_{\mathrm{JS}}(x_j, x_i)$ for $i \neq j$; and 4) the Jenson-Shannon divergence obeys the triangle inequality – e.g., $d_{\mathrm{JS}}(x_1, x_3) < d_{\mathrm{JS}}(x_1, x_2) + d_{\mathrm{JS}}(x_2, x_3)$. Thus, this makes the Jenson-Shannon divergence a distance metric.

4. Case study: NASA-LaRC Uncertainty Quantification challenge 2014

The challenge revolves about a black-box subsystem which can be described by the following black-box model (Crespo et al., 2014):

$$x_1 = h_1(p_1, p_2, p_3, p_4, p_5) \tag{14}$$

where x_1 is the scalar output and $p_{1:5}$ are the uncertain input parameters whose uncertain characteristics are presented in Table 3.

Table 3. Details on the respective uncertain input parameters.

Parameter	Uncertainty model
p_1	Unimodal Beta: $0.6 \leq \mathrm{E}[p_1] \leq 0.8$; $0.02 \leq \mathrm{V}[p_1] \leq 0.04$
p_2	Interval: $[0, 1]$
p_3	Uniform: $[0, 1]$
p_4, p_5	Normal: $-5 \leq \mathrm{E}[p_j] \leq 5$; $0.0025 \leq \mathrm{V}[p_j] \leq 4$; $\lvert\rho\rvert < 1$ (for $j = 1,2$)

There are two objectives to the case study: 1) to calibrate the uncertainty models of the respective model input parameters $p_{1:5}$ using a given set of 25 observations of x_1 (i.e., training data) corresponding to the "true" uncertainty model of each input model parameter; and 2) to validate the calibrated uncertainty models of each model input parameter against another set of 25 observations of x_1 not used for the calibration (i.e., validation data). A histogram and ECDF representations of the training, validation, and the combined (i.e., training and validation) data are illustrated in Figure 2.

4.1. Bayesian model updating set-up

The calibration of the uncertainty models associated with the model input parameters $p_{1:5}$ would be done through Bayesian model updating and involves reducing the epistemic uncertainty of the following eight parameters: $\boldsymbol{\theta} = \{\mathrm{E}[p_1], \mathrm{V}[p_1], p_2, \mathrm{E}[p_4], \mathrm{V}[p_4], \mathrm{E}[p_5], \mathrm{V}[p_5], \rho\}$. Given that parameter p_1 follows a unimodal Beta distribution which is parameterised by the shape parameters α and β, the shape parameters can be re-expressed in terms of $\mathrm{E}[p_1]$ and $\mathrm{V}[p_1]$ using the method of moments:

$$\alpha = \mathrm{E}[p_1] \cdot \left(\frac{\mathrm{E}[p_1] \cdot (1 - \mathrm{E}[p_1])}{\mathrm{V}[p_1]} - 1\right) \tag{15}$$

$$\beta = (1 - \mathrm{E}[p_1]) \cdot \left(\frac{\mathrm{E}[p_1] \cdot (1 - \mathrm{E}[p_1])}{\mathrm{V}[p_1]} - 1\right) \tag{16}$$

This enables $\mathrm{E}[p_1]$ and $\mathrm{V}[p_1]$ to be updated directly during the model calibration procedure.

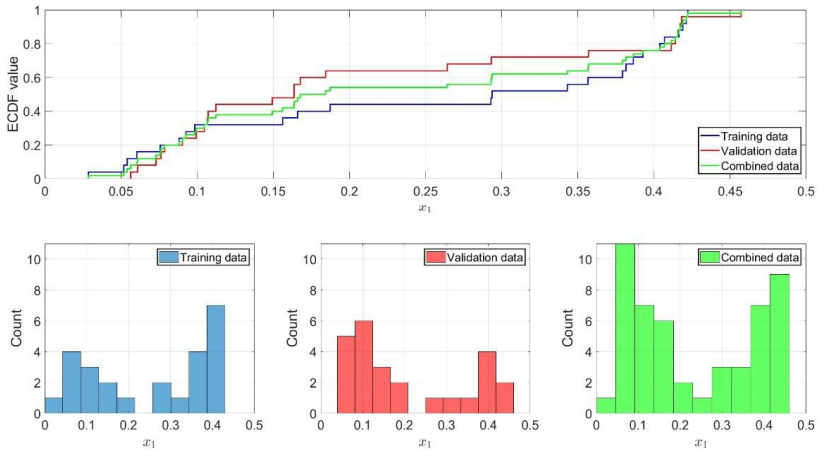

Fig. 2. The ECDF and histogram representations of the training, validation, and the combined data for x_1.

A short-coming of the Bayesian model updating approach is its inapplicability in reducing the epistemic uncertainty if such uncertainty is represented as an interval. To address such short-coming, the epistemic intervals can be re-represented as Uniform distributions. Such approach has also been implemented by Bi et al. (2018) and Lye et al. (2023a). It also needs to be noted that the subsequent posterior distribution obtained can also be interpreted as a fuzzy set where different levels of statistical significance $L_c \in [0,1]$ (i.e., Alpha-cut level) would yield updated intervals of varying width. This is because only the epistemic intervals are updated rather than providing a distribution (i.e., a posterior distribution) over the intervals. Such interpretation has been implemented by Bi et al. (2019) and Lye et al. (2023a).

For the Bayesian model updating set-up, the prior distribution for the epistemic parameters is set as a Uniform distribution with the corresponding bounds defined in Table 3. It is assumed here that the inferred epistemic parameters are independent from one another. The likelihood function is defined as per (6) with the width factor set as $\varepsilon = 0.0004$ to ensure "sufficient" convergence on the resulting posterior distribution (i.e., 8 total sampling iterations by the TEMCMC sampler). The total sample size of samples to be obtained by the TEMCMC sampler is set at $N = 1000$ with the total realizations of the stochastic black-box model output $\boldsymbol{D}^{\text{sim}}$ set at $N_{\text{sim}} = 30$.

4.2. Results and discussions

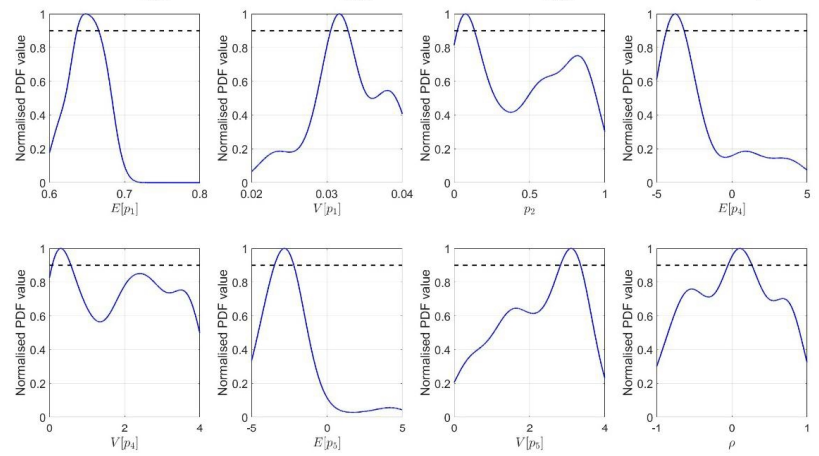

Fig. 3. The normalised PDF for each inferred parameter. Note that the black dotted line denotes the Alpha-cut level at $L_\alpha = 0.9$.

Table 4. Results of the updated epistemic interval for the respective inferred parameter at Alpha-cut level $L_\alpha = 0.9$.

$\boldsymbol{\theta}$	Updated interval
$\mathrm{E}[p_1]$	$[0.637, 0.666]$
$V[p_1]$	$[0.031, 0.033]$
p_2	$[0.021, 0.135]$
$\mathrm{E}[p_4]$	$[-4.351, -3.173]$
$\mathrm{V}[p_4]$	$[0.076, 0.573]$
$\mathrm{E}[p_5]$	$[-3.480, -2.196]$
$\mathrm{V}[p_5]$	$[2.821, 3.355]$
ρ	$[-0.053, 0.268]$

The resulting posterior distribution for each component of $\boldsymbol{\theta}$ is converted into a PDF via the Kernel density estimation approach with a Gaussian kernel. This is done from their respective histograms described with $N_{\mathrm{bin}} = 10$ bins for there to be sufficient resolution in the shape profile of the distribution. The resulting PDF obtained for each component of $\boldsymbol{\theta}$ is then normalised such that the peak value is one to obtain the equivalent fuzzy set and this is illustrated in Figure 3. From the figure, it is observed that the resulting fuzzy sets have multiple peaks and this can be attributed to the following two reasons (Bi et al., 2018): 1) the back-box model h_1 takes in multiple model inputs and returns a single scalar output. This gives rise to parameter unidentifiability since there could exist multiple solutions to the input to give the same model output; and 2) the training data used for the model calibration has a bi-model distribution profile as seen in Figure 2. Such distribution profile is also observed for the validation data and the combined data which suggests that such bi-modal distribution profile is an inherent property of the stochastic model output of h_1.

From the fuzzy sets in Figure 2, the resulting updated interval for each component of $\boldsymbol{\theta}$ is obtained at an alpha-cut level of $L_\alpha = 0.9$. This is to provide a non-conservative estimate on each inferred parameter and to ensure that the resulting bounds are obtained only from one of the peaks in the case whereby the normalized PDF has multiple peaks. The resulting updated intervals for each of the inferred parameters are presented in Table 4.

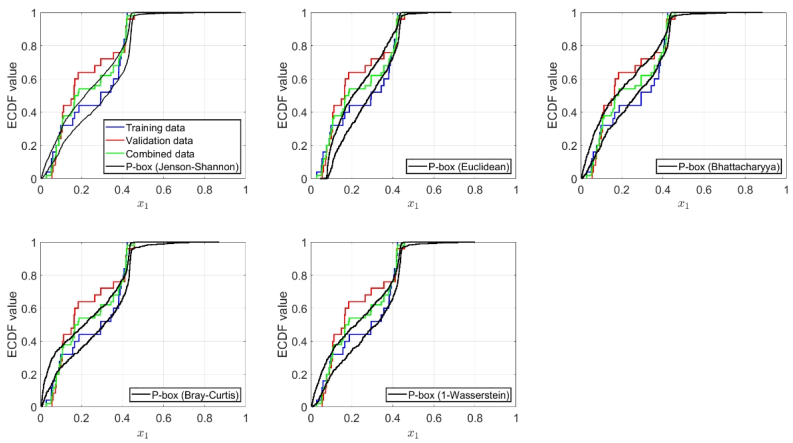

Fig. 4. The resulting p-box obtained from the black-box model h_1 with the calibrated model inputs.

Using the results in Table 4, a probability box (p-box) on the stochastic model output of h_1 is constructed using the double-loop Monte Carlo procedure described as follows (Rocchetta et al., 2018): Based on the updated intervals for the corresponding inferred parameters, an 8-dimensional hyper-rectangle is defined which constitutes the epistemic space. From the resulting hyper-rectangle, a set of $N_e = 1000$ samples are obtained to ensure sufficient exploration of the epistemic space. Each realization of the N_e is subsequently used as inputs to the uncertainty models for $p_{1:5}$ which, in turn, serve as inputs to h_1. From there, a total of $N_a = 1000$ realizations of the stochastic model output from h_1 is obtained to provide sufficient convergence on the distribution describing the aleatory variability on x_1. The above procedure results in N_e distinct ECDFs constructed with N_a stochastic model outputs from which a p-box is obtained and illustrated in Figure 4 (Ferson et al., 2003; Bi et al., 2019).

Based on Figure 4, it is observed that the ECDF of the training and the combined data are enclosed to a significant extent within the p-box while the ECDF associated with the validation data lies mostly outside the p-box. There are two possible reasons to account for such observations: 1) the training data used to calibrate the uncertainty models of $p_{1:5}$ is small resulting in a poor generalization on the distribution of all possible stochastic model outputs from h_1; and 2) the distribution profile of the ECDF is sensitive to the size of the sample data when such size is small (e.g., 25 in this case). This implies that even if the p-box can contain most of the ECDF associated with the training data, it may still fail to contain much of the ECDF associated with the validation data despite both the training and the validation data stemming from the same uncertainty model configurations for $p_{1:5}$ as inputs to h_1. However, when accounting for the fact that the ECDF of the combined data is still largely contained within the resulting p-box, it provides significant evidence that the risk-based decision on the updated epistemic space is sufficiently justified.

Table 5. Results to the corresponding prediction performances by the stochastic model output of h_1 for the respective distance function.

Distance function	Calibration		Validation		Validation (Combined)	
	Mean	Stdev.	Mean	Stdev.	Mean	Stdev.
Jenson-Shannon, d_{JS}	0.021	2.26×10^{-3}	0.056	7.02×10^{-3}	0.030	6.76×10^{-3}
Euclidean, d_E	0.035	3.03×10^{-3}	0.051	6.78×10^{-3}	0.033	4.20×10^{-3}
Bhattacharyya, d_B	0.030	6.35×10^{-3}	0.040	5.33×10^{-3}	0.020	2.15×10^{-3}
Bray-Curtis, d_{BC}	0.026	4.12×10^{-3}	0.059	6.10×10^{-3}	0.032	5.85×10^{-3}
1-Wasserstein, d_{W1}	0.021	2.80×10^{-3}	0.055	5.46×10^{-3}	0.029	5.32×10^{-3}

To quantify the calibration performance against the training data as well as the validation performance against both the validation and the combined data by the $N_a \times N_e$ stochastic model outputs by the respective distance function used in the Bayesian model updating procedure, the discrepancy between each N_e model output ECDFs and the ECDF of the corresponding data type is quantified using the area metric d_A defined as (Ferson et al., 2008):

$$d_A = \int_{-\infty}^{\infty} |F_D(x) - F(x)| \cdot dx \tag{17}$$

where $F_D(\cdot)$ is the ECDF of the data (i.e., training, validation, or combined) while $F(\cdot)$ is the ECDF of the stochastic model output. The smaller the area d_A, the greater the degree of agreement between the two ECDFs. From the procedure, a total of N_e values on d_A is obtained from which the mean and standard deviation are obtained. The resulting statistics relative to the training, validation, and the combined data are presented in Table 5. From the table, it is observed that the mean and standard deviation values on d_A is the smallest for the training data and the largest for the validation data. Such result is consistent with the observation and the conclusion obtained from Figure 4.

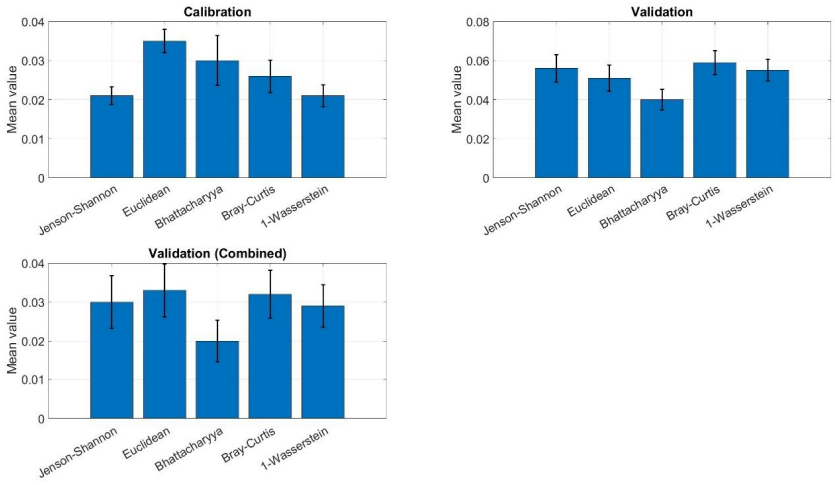

Fig. 5. Bar chart representation of the calibration, validation, and the validation (combined) performance statistics for the respective distance function.

As a further analysis, the model prediction performance by the Jenson-Shannon divergence is compared against that by the following distance functions: 1) Euclidean; 2) Bhattacharyya; 3) Bray-Curtis; and 4) the 1-Wasserstein distance functions to which detailed investigations have been done by Lye et al. (2024). The additional numerical results are presented in Table 5 while the bar chart representation of the performance statistics for the corresponding distance function is presented in Figure 5. As seen from the results, the model calibration performance is the strongest when the Jenson-Shannon divergence is used as the distance function for the Approximate Bayesian computation procedure relative to the other distance functions implemented. While the model validation performance relative to the validation and the combined data using the Jenson-Shannon divergence has a comparable mean value to that of the 1-Wasserstein, it is noteworthy that the standard deviation of the validation performance is significantly higher in the case of the Jenson-Shannon divergence compared to the other four distance functions. This highlights a limitation of the Jenson-Shannon divergence in that it yields a less consistent model validation performance which provides an area for further investigation for future research.

5. Conclusion

In the paper, the Jenson-Shannon divergence is proposed as the distance metric to be implemented within the distance-based Approximate Bayesian computation framework for model calibration and validation under limited data. To provide an understanding of the proposed framework, the paper first reviews the concept of Bayesian model updating and the justifications for Approximate Bayesian computation. From which, the paper introduces the mathematics behind the Jenson-Shannon divergence and illustrated its strengths and properties as a distance metric through a numerical illustrative example.

To demonstrate its feasibility, the proposed distance-based Approximate Bayesian computation framework is implemented on a benchmark problem based on the 2014 NASA-LaRC Uncertainty quantification challenge problem which involves calibrating the uncertainty model to the input variables of a black-box model and subsequently validating the latter under limited data. The results showed that the proposed Approximate Bayesian computation framework was able to achieve an acceptable calibration performance relative to the training data but a relatively poor validation performance relative to the validation data. When compared against the Euclidean, Bhattacharyya, Bray-Curtis, and the 1-Wasserstein distance functions, the Jenson-Shannon divergence yields the strongest calibration performance but the least consistent model validation performance despite having a comparable mean model validation performance to the 1-Wasserstein distance metric.

Further works include the following: 1) making improvements to the adaptive-binning algorithm implemented in the paper as it was initially not developed to account for a small data-set (e.g., 25 in this case). This is a drawback of the current proposed framework which can have significant impact on the model calibration and validation results; 2) to consider the implementation of the Jenson-Shannon divergence as a metric for sensitivity analysis; and 3) to investigate the implementation of the proposed Approximate Bayesian computation framework towards a distribution-free stochastic model updating for risk analysis, model verification and validation.

To allow for a better understanding of the implementation of the Jenson-Shannon divergence as well as the opportunity for the readers to reproduce the results presented in the paper, the MATLAB codes used in the study are made accessible on GitHub via: https://github.com/Adolphus8/Approximate_Bayesian_Computation.git

Acknowledgements

The authors would like to acknowledge the gracious funding from the National Research Foundation Singapore (Funding no.: A-0001360-06-00) and the resource support from the Singapore Nuclear Research and Safety Institute without which the research work presented in the paper would not have been possible.

References

Beck, J. L., Katafygiotis, L. S. 1998. Updating Models and Their Uncertainties I: Bayesian Statistical Framework. Journal of Engineering Mechanics 124, 455-461.

Bi, S., Broggi, M., Beer, M. 2018. The role of the Bhattacharyya distance in stochastic model updating. Mechanical Systems and Signal Processing 117, 437-452.

Bi, S., Broggi, M., Wei, P., Beer, M. 2019. The Bhattacharyya distance: Enriching the P-box in stochastic sensitivity analysis. Mechanical Systems and Signal Processing 129, 265-281.

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, 816-832.

Crespo, L. G., Kenny, S. P., Giesy, D. P. 2014. The NASA Langley Multidisciplinary Uncertainty Quantification Challenge. In Proceedings of the 16th AIAA Non-Deterministic Approaches Conference 1.
Ferson, S., Kreinovick, V., Ginzburg, L., Sentz, F. 2003. Constructing Probability Boxes and Dempster-Shafer Structures. U.S. Department of Energy Office of Scientific and Technical Information.
Ferson, S., Oberkampf, W., Ginzburg, L. 2008. Model validation and predictive capability for the thermal challenge problem. Computer Methods in Applied Mechanics and Engineering 197, 2408–2430.
Goodman, J., Weare, J. 2010. Ensemble samplers with affine invariance. Communications in Applied Mathematics and Computational Science 5, 65-80.
Gray, A., Wimbush, A., de-Angelis, M., Hristov, P. O., Calleja, D., Miralles-Dolz, E., Rocchetta, R. 2022. From inference to design: A comprehensive framework for uncertainty quantification in engineering with limited information. Mechanical Systems and Signal Processing 165, 108210.
Kennedy, M. C., O'Hagan, A. 2001. Bayesian Calibration of Computer Models. Journal of the Royal Statistical Society Series B: Statistical Methodology 63, 425-464.
Lye, A., Cicirello, A., Patelli, E. 2019. A Review of Stochastic Sampling Methods for Bayesian Inference Problems. In Proceedings of the 29th European Safety and Reliability Conference 1, 1866–1873.
Lye, A., Cicirello, A., Patelli, E. 2020. Bayesian Model Updating of Reliability Parameters using Transitional Markov Chain Monte Carlo with Slice Sampling. In Proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference 1, 2734–2741.
Lye, A., Cicirello, A., Patelli, E. 2022a. An efficient and robust sampler for Bayesian inference: Transitional Ensemble Markov Chain Monte Carlo. Mechanical Systems and Signal Processing 167, 108471.
Lye, A., Cicirello, A., Patelli, E. 2022b. On-line Bayesian Inference for Structural Health Monitoring under Model Uncertainty using Sequential Ensemble Monte Carlo. In Proceedings of the 13th International Conference on Structural Safety and Reliability 1.
Lye, A., Gray, A., de-Angelis, M., Ferson, S. 2023a. Robust Probability Bounds Analysis for Failure Analysis under Lack of Data and Model Uncertainty. In Proceedings of the 5th International Conference on Uncertainty Quantification in Computational Science and Engineering 1.
Lye, A., Marino, L., Cicirello, A., Patelli, E. 2023b. Sequential Ensemble Monte Carlo Sampler for On-Line Bayesian Inference of Time-Varying Parameter In Engineering Applications. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part B: Mechanical Engineering 9, 031202.
Lye, A. 2023c. Robust and Efficient Probabilistic Approaches towards Parameter Identification and Model Updating. The University of Liverpool Repository (Doctor of Philosophy Thesis).
Lye, A., Marino, L. 2023d. An investigation into an alternative transition criterion of the Transitional Markov Chain Monte Carlo method for Bayesian model updating. In Proceedings of the 33rd European Safety and Reliability Conference 1, 3556–3563.
Lye, A., Ferson, S., Xiao, S. 2024. Comparison between distance functions for Approximate Bayesian Computation towards Stochastic model updating and Model validation under limited data. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part A: Civil Engineering (Accepted).
McGurk, M., Lye, A., Renson, L., Jie, Y. 2024. Data-driven Bayesian inference for stochastic model identification of nonlinear aeroelastic systems. AIAA Journal (Accepted).
Nielsen, F. 2019. On the Jensen–Shannon Symmetrization of Distances Relying on Abstract Means. Entropy 21, 485.
Rocchetta, R., Broggi, M., Huchet, Q., Patelli, E. 2018. On-line Bayesian Model Updating for Structural Health Monitoring. Mechanical Systems and Signal Processing 103, 174-195.
Zhao, W., Yang, L., Dang, C., Rocchetta, R., Valdebenito, M., Moens, D. 2022. Enriching stochastic model updating metrics: An efficient Bayesian approach using Bray-Curtis distance and an adaptive binning algorithm. Mechanical Systems and Signal Processing 171, 108889.