An investigation towards the Uncertainty Model calibration approaches for NASA-Langley UQ Challenge 2019

Adolphus Lye[1], Alice Cicirello[2], and Edoardo Patelli[3]

[1] Institute for Risk and Uncertainty, University of Liverpool, Peach Street, Liverpool, L7 7BD, UK

[2] Faculty of Civil Engineering and Geoscience, Delft University of Technology, Stevinweg 1, Delft, 2628 CN, Netherlands

[3] Centre for Intelligent Infrastructure, Department of Civil and Environmental Engineering, University of Strathclyde, 75 Montrose Street, Glasgow, G1 1XJ, UK

E-mail: adolphus.lye@liverpool.ac.uk

Abstract. The paper presents a series of analysis based on the recent NASA-Langley Uncertainty Quantification Challenge 2019 aimed towards calibrating an Uncertainty Model through a black-box model. In this research, 4 alternative Uncertainty Models are being proposed to investigate the following factors which contribute to the lowest degree of uncertainty over the aleatory and epistemic uncertain model parameters: 1) the choice of distribution of the aleatory model parameters; 2) the choice of the stochastic distance metric within the likelihood function to model the data variability; and 3) the choice of data type used within the likelihood function for the Bayesian model updating approach. To model the distribution of the aleatory model parameters, 2 distribution functions are considered: Beta vs Staircase Density Function. To quantify the variability of the input data used for model calibration, 2 types of stochastic distances are considered: Wasserstein's distance vs Bhattacharyya's distance. For the input data used for model calibration, 2 data types are considered: Time-domain vs Frequency domain. Based on the results, it was found that the Uncertainty Model incorporating the Beta distribution to model the aleatory model parameters, the Bhattacharyya's distance as the stochastic metric, and the time-based data as the input data, yielded P-box estimates of the aleatory distribution and probabilistic estimates of the epistemic model parameters with the lowest degree of uncertainty.

1. Introduction

1.1. Problem description

The research work presented in the paper is based on the recent NASA-Langley Uncertainty Quantification Challenge 2019. The challenge problem is set in the context of a dynamical black-box subsystem that is modelled by a black-box model $\hat{y} = yfun(\boldsymbol{a}, \boldsymbol{e}, t)$. Here, $\boldsymbol{a}$ is a 5-component vector of aleatory model parameters, $\boldsymbol{e}$ is a 4-component vector of epistemic model parameters, and t is the time parameter. The model parameters $\boldsymbol{a}$ are characterised by a joint distribution f_a defined within the domain of $[0, 2]^5$, while $\boldsymbol{e}$ are defined within a bounded set in the form of $E \subseteq [0, 2]^4$. Thus, an Uncertainty Model (UM) for the subsystem can be obtained from the black-box model and it can be fully described by $\langle f_a, E \rangle$. The challenge

is to calibrate the UM under limited data in the form of a set of 100 response signal output provided for time $t \in [0, 5]$ s. Each signal output is represented as a discrete time history: $y^l(t) = [y^l(0), y^l(dt), \ldots, y^l(5000 \cdot dt)]$; where $l = 1, \ldots, 100$, and $dt = 0.001$ s. As such, the calibration data is denoted as $D_1 = \{y^l(t)\}_{l=1,\ldots,100}$. Further details to the task and challenge are found in [1].

1.2. Research objectives

In [2], 2 different UM set-ups are proposed and calibrated using Bayesian model updating: UM^1_{y0} and UM^2_{y0}. For UM^1_{y0}, a Beta distribution is used to model f_a, D_1 is processed via Fast Fourier Transformation (FFT) into the frequency domain and used for calibration, and the stochastic distance metric used for the likelihood function is the Wasserstein's distance [3]. For UM^2_{y0}, the Staircase Density Function (SDF) is used to model f_a [4], D_1 is used directly in the time-domain for calibration, and the stochastic distance metric used for the likelihood function is the Bhattacharyya's distance [5]. Details to the signal processing approaches into the respective domains are found in [2]. Investigations concluded that UM^2_{y0} yielded more precise estimates to the P-box [6] characterising the uncertainty over f_a as well as the bounds for the individual components of $\boldsymbol{e}$. However, it is not addressed as to what is/are the factor(s) contributing to UM^2_{y0} having less uncertainty compared to UM^1_{y0}. Hence, the paper seeks to analyse whether such observation is due to either of the following: 1) the choice of distribution model for f_a; 2) The choice of stochastic distance metric used in the likelihood function; and/or 3) the choice of the data type used to calibrate the UM.

2. Bayesian Model Updating

Bayesian model updating is a probabilistic model updating approach incorporating Bayes' inference [7]:

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

whereby $\boldsymbol{\theta}$ represents the vector of inferred parameter(s), $\boldsymbol{D}$ represents the vector of observed data, M represents the model that is being considered for updating, $P(\boldsymbol{\theta}|M)$ represents the prior, $P(\boldsymbol{D}|\boldsymbol{\theta}, M)$ represents the likelihood function, $P(\boldsymbol{\theta}|\boldsymbol{D}, M)$ represents the posterior, and $P(\boldsymbol{D}|M)$ is the normalising constant to ensure that the posterior distribution integrates to 1. Details to the above terms can be found in [2].

In general, $P(\boldsymbol{\theta}|\boldsymbol{D}, M)$ is not normalised due to the complexity in computing $P(\boldsymbol{D}|M)$. To sample from the un-normalised $P(\boldsymbol{\theta}|\boldsymbol{D}, M)$, the recently-developed Transitional Ensemble Markov Chain Monte Carlo (TEMCMC) sampler is implemented to which details are found in [8].

2.1. Approximate Bayesian Computation

Due to the availability of a large data-set in D_1, the implementation of an actual Bayesian computation using a full likelihood function can be a computationally-expensive procedure [5]. To reduce such cost, Approximate Bayesian Computation (ABC) is adopted to which details are provided in [9]. This involves the use of an approximate likelihood function defined as [10]:

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

where d is the stochastic distance metric which quantifies the difference between the observed data $\boldsymbol{D}$ and the model output $\hat{y}$, while ε is the width factor of the approximate likelihood function. The choice of values for ε given each d is defined in [2].

Section 1.2 highlights the 2 different stochastic distance metric implemented: 1) the Wasserstein's distance d_W; and 2) the Bhattacaryya's distance d_B. The Wasserstein's distance is mathematically defined as [3]:

$$d_W = \int_{-\infty}^{\infty} |F_{\boldsymbol{D}}(x) - F_{\hat{y}}(x)| \cdot dx \tag{3}$$

whereby $F_{\boldsymbol{D}}(x)$ and $F_{\hat{y}}(x)$ are the respective Empirical Cumulative Distribution Functions (ECDFs) of the data and the stochastic model output of $\hat{y}$, while x is the data variable. From Eq. (3), it is seen that d_W quantifies the enclosed area between both ECDFs. The smaller d_W is, the higher the degree of similarity between the ECDFs of the data and the stochastic prediction by $\hat{y}$ [11]. The Bhattacharyya's distance is mathematically defined as [5]:

$$d_B = -\log\left[\int_{-\infty}^{\infty} \sqrt{p_{\boldsymbol{D}}(x) \cdot p_{\hat{y}}(x)} \cdot dx\right] \tag{4}$$

whereby $p_{\boldsymbol{D}}(x)$ and $p_{\hat{y}}(x)$ are the respective Probability Mass Functions (PMFs) of the data and the stochastic model output of $\hat{y}$.

3. Methodology

3.1. Signal processing methods

To utilise the data D_1 in the time domain, some form of dimension-reduction procedure needs to be implemented given that there is a high number of dimensions on the data (i.e. 5001×100 data-points). This makes it computationally expensive to evaluate the likelihood function $P(\boldsymbol{D}|\boldsymbol{\theta}, M)$. Thus, this section presents the approaches towards reducing the data-size required for calibration in the: 1) Frequency domain; and 2) Time domain.

In the frequency domain, the dimension-reduction procedure is implemented through the following steps [2]:

1) Perform the FFT procedure on the data D_1 to obtain the amplitude term A_q^l and phase angle term ϕ_q^l [2]. The resulting frequency spectra for A_q^l and ϕ_q^l are presented in Figure 1;

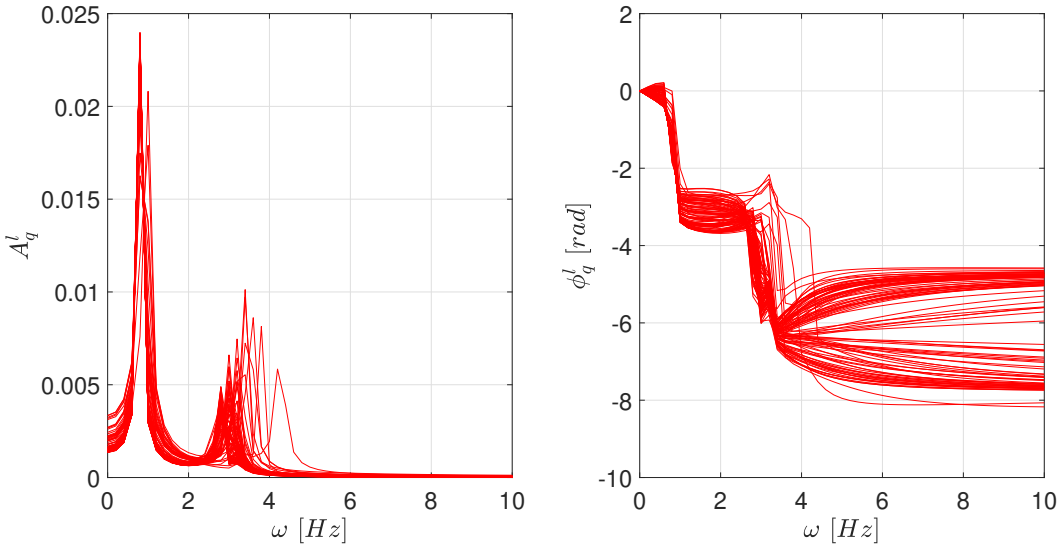


Figure 1: Illustration of the frequency spectra obtained from D_1 via FFT.

2) Neglect the spectra of A_q^l and ϕ_q^l for frequencies $\omega > 5.80\ Hz$ given that the frequency spectra for A_q^l do not show any further perturbations beyond which;
3) Consider only the 30 values of ω between 0 Hz and 5.80 Hz are considered for both A_q^l and ϕ_q^l at which the stochastic distance metric d can be computed between the respective distributions of the frequency spectra and model output at each value of ω.

In the time domain, the dimension-reduction procedure is implemented through the following steps [12]:

1) Divide the data set D_1 into $\left\lceil \frac{N_t}{L_w} \right\rceil$ distinct intervals; where $\lceil \bullet \rceil$ is the ceil operator, $L_w = 50$ is the window length, and $N_t = 5001$ is the length of each signal data;
2) The Root Mean Squared (RMS) values of each interval $\boldsymbol{R} = \left[R_1, \ldots, R_{\left\lceil \frac{N_t}{L_w} \right\rceil} \right]$ is computed from which, the sample set of the RMS values $\boldsymbol{R}_D \in \mathbb{R}^{100 \times \left\lceil \frac{N_t}{L_w} \right\rceil}$ is generated where:

$$\boldsymbol{R}_D = \left[\boldsymbol{R}_D^1, \ldots, \boldsymbol{R}_D^{\left\lceil \frac{N_t}{L_w} \right\rceil} \right], \text{ with } \boldsymbol{R}_D^\nu = [R_{1,\nu}, \ldots, R_{100,\nu}]^T$$

for $\nu = 1, \ldots, \left\lceil \frac{N_t}{L_w} \right\rceil$ while $\boldsymbol{R}_{\hat{y}} \in \mathbb{R}^{N_{sim} \times \left\lceil \frac{N_t}{L_w} \right\rceil}$ where $N_{sim} = 100$ the number of model evaluations by $\hat{y}$ per given set of model inputs $\{\boldsymbol{a}, \boldsymbol{e}\}$;
3) Evaluate the stochastic distance metric d between $\boldsymbol{R}_D^\nu$ and $\boldsymbol{R}_{\hat{y}}^\nu$ for all ν;
4) Obtain the corresponding RMS values of d and use it as the distance metric d_{RMS}.

3.2. Uncertainty model set-up

In this work, 4 new UMs are constructed to be calibrated which are defined respectively as UM_{y0}^I, UM_{y0}^{II}, UM_{y0}^{III}, and UM_{y0}^{IV}. Details to their respective configurations are summarised in Table 1. Here, UM_{y0}^2 serves as the control from which the calibration results will serve as reference to those obtained by the 4 new UMs.

Table 1: Summary of the configurations to the respective UMs.

UM	f_a	Stochastic distance metric	Data type
UM_{y0}^1	Beta	Wasserstein's distance	Frequency domain
UM_{y0}^2 (Control)	SDF	Bhattachryya's distance	Time domain
UM_{y0}^I	SDF	Wasserstein's distance	Frequency domain
UM_{y0}^{II}	Beta	Bhattachryya's distance	Frequency domain
UM_{y0}^{III}	Beta	Wasserstein's distance	Time domain
UM_{y0}^{IV}	Beta	Bhattachryya's distance	Time domain

3.3. Bayesian inference set-up

In calibrating the UM, 2 assumptions are made: 1) the marginal distributions for all a_{i_a} (for $i_a = 1, \ldots, 5$) are independent between components; and 2) the marginals all belong to the same distribution class. For each choice of distribution model for f_a, it introduces additional

parameters to be inferred on top of the 4 epistemic model parameters e_{i_e} (for $i_e = 1, \ldots, 4$). Details to the distribution parameters to be inferred are summarised in Table 2. Hence, this brings a total of 14 parameters to be inferred in the case of f_a being the Beta distribution, and 24 parameters to be inferred in the case of f_a being the SDF.

Due to the lack of information apriori, each inferred parameter is assigned a Uniform prior. The prior bounds for each epistemic model parameter e_{i_e} is defined as $[0, 2]$, while those of the inferred distribution parameters for f_a are listed in Table 2. Given each choice of d, the likelihood function $P(\boldsymbol{D}|\boldsymbol{\theta}, M)$ used is the approximate Gaussian function defined in Eq. (2). From the resulting posterior $P(\boldsymbol{\theta}|\boldsymbol{D}, M)$, due to the significant computational costs in performing the Bayesian model updating via ABC, $N = 500$ samples are obtained via the TEMCMC sampler.

Table 2: Uniform prior bounds of the distribution parameters for the corresponding $f_{\boldsymbol{a}}$.

f_a	Prior distribution	Parameter description
Beta	$\alpha_{i_a} \sim U[0, 100]$	Shape parameter 1
	$\beta_{i_a} \sim U[0, 100]$	Shape parameter 2
SDF	$\mu_{i_a} \sim U[0, 2]$	Distribution mean of a_{i_a}
	$(m_2)_{i_a} \sim U[0, 1]$	2^{nd} central moment of a_{i_a}
	$(m_3)_{i_a} \sim U[-\frac{4}{3\sqrt{3}}, \frac{4}{3\sqrt{3}}]$	3^{rd} central moment of a_{i_a}
	$(m_4)_{i_a} \sim U[0, \frac{4}{3}]$	4^{th} central moment of a_{i_a}

4. Results and Discussions

An alpha-cut at 0.50 level of confidence is performed on the posterior of the distribution parameters while that at 0.025 level of confidence is performed on the posterior of the epistemic model parameters e_1 to e_4.

The resulting inferred bounds for each e_{i_e} along with the volume of the epistemic hyper-rectangle V_E are presented in Table 3. From the table, it can be seen that the interval obtained by UM_{y0}^{IV} for each component of $\boldsymbol{e}$, along with the corresponding V_E, are the smallest. This is followed closely by that of UM_{y0}^{III}. Such result is supported by Figure 2 where it is observed that the histograms corresponding to e_1 to e_4 have generally the least width in the case of UM_{y0}^{IV} and that the UM has yielded estimates on the epistemic model parameters with a significantly higher degree of precision compared to UM_{y0}^{2}.

From the results of the interval of the distribution parameters after taking the alpha-cut, a P-box is constructed for the aleatory model parameters a_1 to a_5 [11, 13]. The resulting P-boxes for each a_{i_a}, given UM_{y0}^{I} to UM_{y0}^{IV}, are presented in Figure 3 and the area enclosed by the P-boxes for each a_{i_a} is computed for each UM and presented in Table 4. From the table, it can be seen that this area is generally the smallest across the components of $\boldsymbol{a}$ for UM_{y0}^{IV}. This is supported by Figure 3 where it can be seen that the P-boxes obtained by UM_{y0}^{IV} are the narrowest compared to the other UMs. This indicates that UM_{y0}^{IV} reflects the lowest degree of uncertainty over the true marginal distributions of f_a.

Finally, to verify the model calibration results for each UM, $N = 500$ samples are generated uniformly and independently from the hyper-rectangle defined by the intervals of the distribution parameters and the epistemic model parameters after accounting for their corresponding alpha-cuts. From there, each realization of samples from this hyper-rectangle will serve as inputs to f_a, and subsequently $y(\boldsymbol{a}, \boldsymbol{e}, t)$, to generate 100 realizations of model output $\hat{y}$ for $t \in [0, 5]$ s.

Table 3: Inferred bounds of each component of $\boldsymbol{e}$ given the respective UM set-up.

UM	e_1	e_2	e_3	e_4	V_E
UM_{y0}^{1}	$[0.2307, 1.4567]$	$[0.3155, 1.4810]$	$[0.0411, 1.4123]$	$[0.0641, 1.9417]$	3.6772
UM_{y0}^{2} (Control)	$[0.4351, 0.7082]$	$[0.5583, 1.0000]$	$[0.0721, 0.5511]$	$[0.6066, 1.6893]$	0.0626
UM_{y0}^{I}	$[0.0149, 1.6714]$	$[0.1245, 1.2491]$	$[0.1441, 1.9873]$	$[0.0083, 1.5558]$	5.3139
UM_{y0}^{II}	$[0.2054, 1.6552]$	$[0.3158, 1.3241]$	$[0.0204, 0.8553]$	$[0.0247, 1.7999]$	1.2855
UM_{y0}^{III}	$[0.4024, 0.6671]$	$[0.6191, 1.2870]$	$[0.0161, 0.6122]$	$[0.1487, 1.9641]$	0.1914
UM_{y0}^{IV}	$[0.4864, 0.6336]$	$[0.7415, 0.9850]$	$[0.1746, 0.4911]$	$[0.1511, 0.5023]$	0.0040

Table 4: Area enclosed by the P-box of each component of $\boldsymbol{a}$ given the respective UM set-up.

UM	a_1	a_2	a_3	a_4	a_5
UM_{y0}^{1}	0.6849	1.2786	0.4638	0.8690	0.7655
UM_{y0}^{2} (Control)	0.3200	0.2946	0.1783	1.2531	0.7159
UM_{y0}^{I}	0.5902	1.0410	0.4886	0.8699	0.5617
UM_{y0}^{II}	1.0101	1.1212	0.7176	0.8109	0.7250
UM_{y0}^{III}	0.5789	0.8208	0.3011	1.0734	0.8568
UM_{y0}^{IV}	0.2217	0.2023	0.1501	0.4416	0.1614

In total, 50000 realizations of $\hat{y}$ are obtained within the domain of t from which the bounds are obtained.

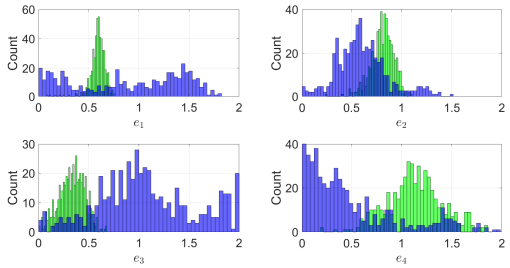
The resulting model output bounds of $\hat{y}$ obtained for each UM are presented in Figure 4 while the area enclosed by these bounds for each UM are computed and presented in Table 5. From the table, it can be seen that UM_{y0}^{IV} yields the smallest area of the region enclosed by the model output bounds of $\hat{y}$. This is supported by Figure 4 where it is observed that the output bounds obtained by UM_{y0}^{IV} is the only one which is narrower than that obtained by UM_{y0}^{2}. This is consistent with expectations given that UM_{y0}^{IV} yields the lowest uncertainty over the P-box estimates for f_a and the intervals of e_1 to e_4. The corresponding uncertainty over the model prediction of $y(t)$ would also be the least as a result.

Table 5: Area enclosed by the model output bounds of $\hat{y}$ given the respective UM set-up.

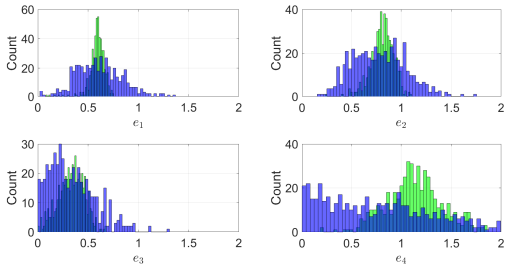
UM	UM_{y0}^{1}	UM_{y0}^{2}	UM_{y0}^{I}	UM_{y0}^{II}	UM_{y0}^{III}	UM_{y0}^{IV}
Area	0.4117	0.4114	0.4452	0.4210	0.3710	0.3518

5. Conclusion

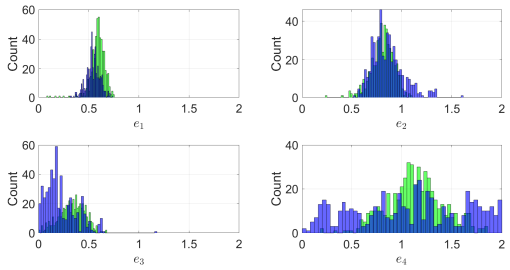
A significant decrease in the uncertainty of the results in $\boldsymbol{e}$ is suggested by UM_{y0}^{III} compared to UM_{y0}^{1} as seen in Table 3 which implies that the type of data used for the calibration plays a



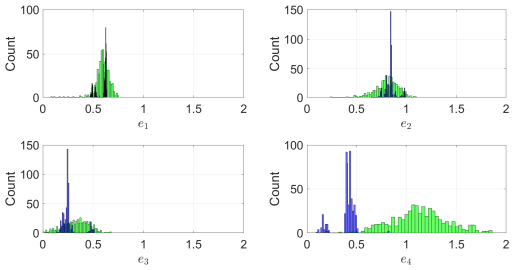
(a) UM^{I}_{y0} (blue) vs UM^{2}_{y0} (green).



(b) UM^{II}_{y0} (blue) vs UM^{2}_{y0} (green).

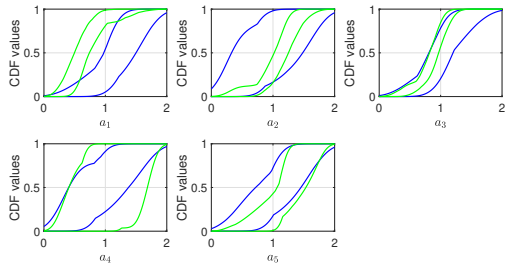


(c) UM^{III}_{y0} (blue) vs UM^{2}_{y0} (green).

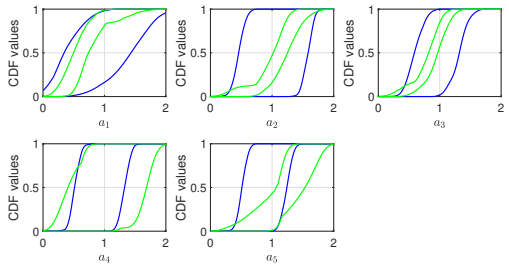


(d) UM^{IV}_{y0} (blue) vs UM^{2}_{y0} (green).

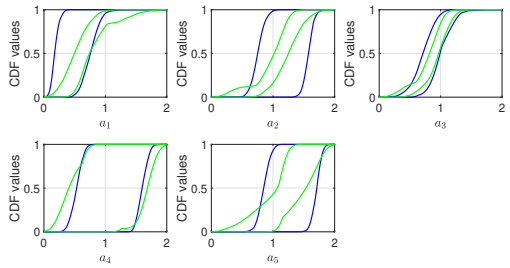
Figure 2: The resulting histograms of e_1 to e_4 obtained by UM^{I}_{y0} to UM^{IV}_{y0} relative to UM^{2}_{y0}.



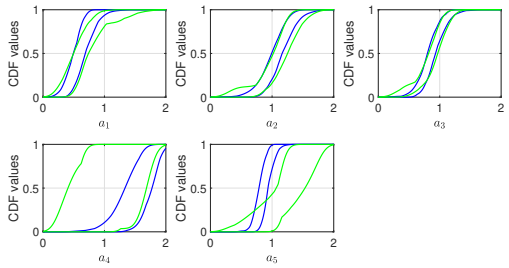
(a) UM^{I}_{y0} (blue) vs UM^{2}_{y0} (green).



(b) UM^{II}_{y0} (blue) vs UM^{2}_{y0} (green).



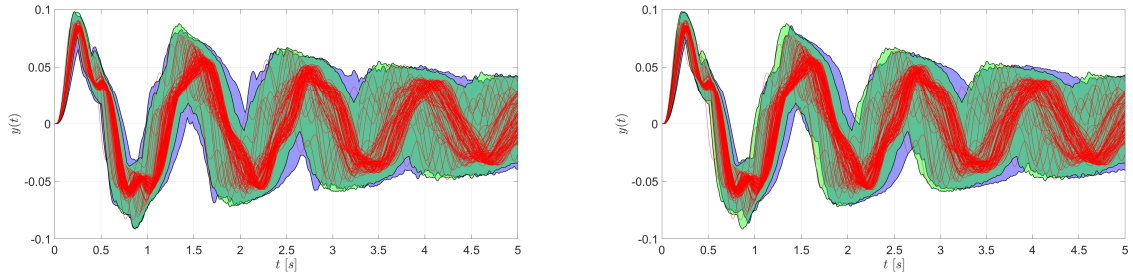
(c) UM^{III}_{y0} (blue) vs UM^{2}_{y0} (green).



(d) UM^{IV}_{y0} (blue) vs UM^{2}_{y0} (green).

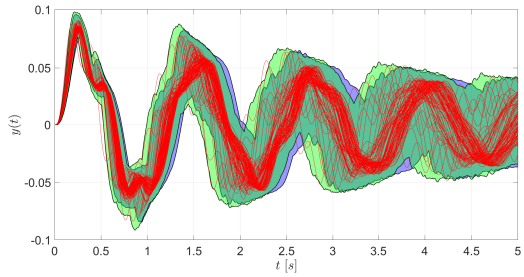
Figure 3: The P-boxes for a_1 to a_5 obtained by UM^{I}_{y0} to UM^{IV}_{y0} relative to UM^{2}_{y0}.

significant role in this aspect. This is expected as there are more data points contained in the time domain data compared to that in the frequency domain. The improvement in the precision of the estimates by UM^{IV}_{y0} from UM^{III}_{y0} is due to the choice in the stochastic distance metric used which shows that the Bhattacharyya's distance is a better metric than the Wasserstein's distance in this context. Such improvement, however, is not as significant compared to that due

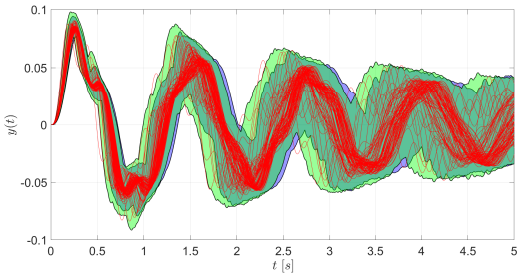


(a) UM^{I}_{y0} (blue) vs UM^{2}_{y0} (green).

(b) UM^{II}_{y0} (blue) vs UM^{2}_{y0} (green).



(c) UM^{III}_{y0} (blue) vs UM^{2}_{y0} (green).



(d) UM^{IV}_{y0} (blue) vs UM^{2}_{y0} (green).

Figure 4: The resulting model output bounds obtained by UM^{I}_{y0} to UM^{IV}_{y0} relative to UM^{2}_{y0}. The red curves are the signal output D_1.

to the use of the data in the time domain instead of the frequency domain. The choice of f_a is of the least relative significance as there is little improvement in the interval estimates for each component of $\boldsymbol{e}$ and the P-box estimates of f_a. In fact, the degree of uncertainty in the estimates yielded by UM^{I}_{y0} is the lowest among the 4 new UMs presented in the paper and the interval estimates on the components of $\boldsymbol{e}$.

Hence, in order of descending significance on the uncertainty of the Uncertainty Model estimates is: 1) Data type; 2) Stochastic distance metric; and 3) the choice of f_a.

For the benefit of the readers, the MATLAB codes to the study presented in the paper are available via: https://github.com/Adolphus8/Transitional_Ensemble_MCMC.git

References

[1] Crespo L G and Kenny S P 2021 *Mech. Syst. Signal Process.* **152** 107405
[2] Lye A, Kitahara M, Broggi M and Patelli E 2022 *Mech. Syst. Signal Process.* **167** 108522
[3] Panaretos V M and Zemel Y 2019 *Annu. Rev. Stat. Appl.* **6** 405–431
[4] Crespo L G, Kenny S P, Giesy D P and Stanford B K 2018 *Appl. Math. Model.* **64** 196–213
[5] Bi S, Broggi M and Beer M 2019 *Mech. Syst. Signal Process.* **117** 437–452
[6] Ferson S, Kreinovick V, Ginzburg L and Sentz F 2002 *Sandia National Laboratories* URL `https://doi.org/10.2172/809606`
[7] Lye A, Cicirello A and Patelli E 2019 *In Proc. of Euro. Saf. and Rel. Conf.* **1** 1866–1873
[8] Lye A, Cicirello A and Patelli E 2022 *Mech. Syst. Signal Process.* **167** 108471
[9] Turner B M and Zandt T V 2012 *J. Math. Psych.* **56** 69–85
[10] Gray A, Wimbush A, Angelis M D, Hristov P O, Calleja D, Miralles-Dolz E and Rocchetta R 2022 *Mech. Syst. Signal Process.* **165** 108210
[11] Ferson S, Oberkampf W and Ginzburg L 2008 *Comput. Methods Appl. Mech. Engrg.* **197** 2408–2430
[12] Kitahara M, Bi S F, Broggi M and Beer M 2021 *ASCE-ASME J. Risk Uncertain. Eng. Syst. A* **7**
[13] Lye A, Cicirello A and Patelli E 2019 *In Proc. of Int. Conf. on Uncertain. Quant. in Comp. Sci. and Eng.* **1**