INSTITUTE FOR RISK AND RELIABILITY

Leibniz Universität Hannover

Uncertainty Quantification Metrics Using the Bhattacharyya Distance

Introduction

Uncertainty quantification (UQ) metrics are significant with the expectation to provide an elaborate measurement of the uncertainty information, which can crucially influence the outcomes of model updating and validation. The Bhattacharyya distance is a stochastic metric computed between two sets of samples and taking into account their probability distributions. It provides a more comprehensive treatment of uncertainties than other more widely employed metrics, e.g. the classical Euclidean distance.

Formulations of the Bhattacharyya Distance

A so-called binning algorithm is proposed to evaluate the probability mass function (PMF) of a discrete distribution:

1) Define the common interval I_i of both $\mathbf{X}_{obs}$ and $\mathbf{X}_{sim}$ according to the i^{th} feature $\mathrm{x}_i, \forall i = 1, \ldots, m$, by finding its maximum and minimum;
2) Within the defined interval, decide the number of bins $n_{bin} \cong \left\lceil \frac{\max(N_{sim}, N_{obs})}{10} \right\rceil$, where $\lceil\blacksquare\rceil$ denotes the upper integer of the investigating values;
3) Count the joint probability mass for each bin. Note that the total number of bins in the joint-PMF space is $N_{bin} = n_{bin}{}^{m}$.

The discrete Bhattacharyya distance between the observational (obs) and simulated (sim) samples is evaluated as

$$d_{\mathrm{B}}(\mathrm{X}_{obs}\mathrm{X}_{sim}) = -\log\left(\sum_{\mathrm{k}=1}^{\mathrm{N}_{bin}} \sqrt{\mathrm{PM}_{obs}^{(\mathrm{k})}\mathrm{PM}_{sim}^{(\mathrm{k})}}\right)$$

where $PM_{\blacksquare}^{(k)}$ is the probability mass of the k^{th} bin.

Approximate Bayesian Computation (ABC) with Stochastic Distance Metrics

In ABC, the Gaussian function is employed to construct an approximate likelihood based on the Bhattacharyya distance metric:

$$P_L(\mathbf{X}_{obs}|\boldsymbol{\theta}) \propto exp\left\{-\frac{d^2}{\varepsilon^2}\right\}$$

where d is the distance metric; ε is the so-called width factor, which is a pre-defined coefficient controlling the centralization degree of the posterior distribution of the parameter.

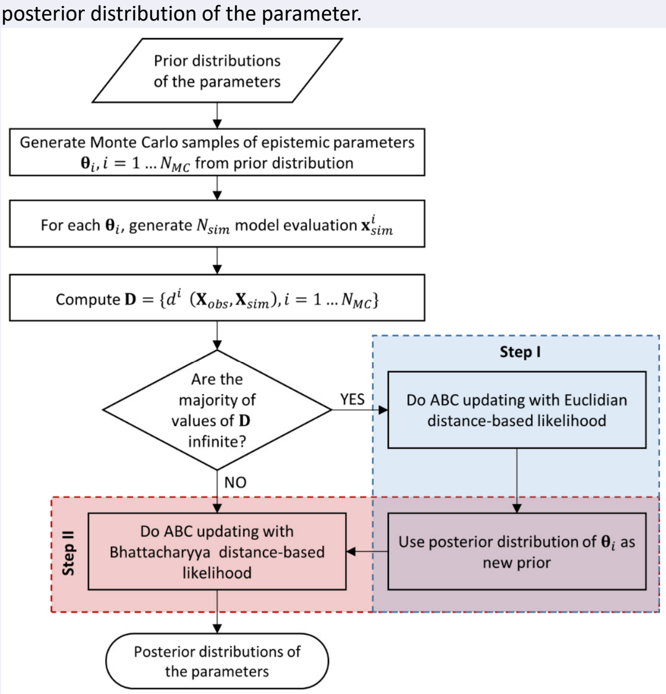
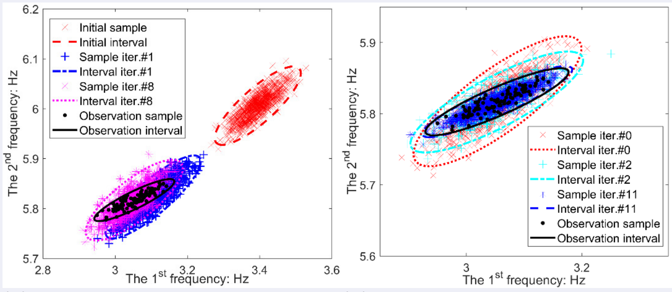


Fig. 1: Schematic of the two-step ABC updating procedure

Principle and illustrative application

A two-step strategy of the proposed ABC updating with both the Euclidian and Bhattacharyya distances as metrics is proposed as illustrated in Fig. 1. This strategy is first demonstrated on a simple spring-mass example. The initial, intermediate, and final output samples are shown in Fig. 2. In the 1st step when the Euclidian distance is taken as updating metric, the initial sample is translated toward the observation sample, while the orientation and dispersion of the simulated samples are not changed. In the 2nd step when the Bhattacharyya distance is employed as updating metric, the simulated samples are rotated toward to the orientation of the observations, and the dispersion is also reduced progressively. The finally updated sample shows a distribution identical to the target sample, implying the Bhattacharyya distance metric has successfully captured the dispersion information of the observations and the model predictions.



(a) with the Euclidian metric (b) with the Bhattacharyya metric

Fig. 2: The updating effect with different updating metrics

The NASA UQ Challenge (NUQC) Problem

The NUQC problem is investigated to demonstrate the capabilities of the Bhattacharyya distance metric in model updating for complex applications. Detailed description and true values of NUQC can be referred in Crespo et al. 2014. Fig. 3 illustrates the posterior distributions of the parameters with the Euclidian and Bhattacharyya distance metrics, respectively. The posterior distributions with the Euclidian metric are still uniform or converging to an opposite direction from the true values, while the posterior distributions with the Bhattacharyya metric are coincident with the true values, demonstrating that the Bhattacharyya metric is capable of calibrating the distributions when the result with the Euclidian metric is inappropriate.

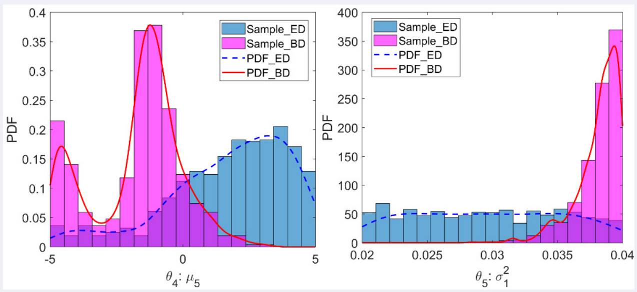


Fig. 3: Posterior distributions with different distance metrics

Reference

Crespo, et al., *16th AIAA Non-Deterministic Approaches Conference*, 2014.

Sifeng Bi, S. Prabhu, S. Cogan, S. Atamturktur, *AIAA Journal*, 2017.

Note

This work is supported by the Alexander von Humboldt Foundation.

Alexander von Humboldt Stiftung/Foundation

Institute for Risk and Reliability

www.irz.uni-hannover.de