Uncertainty Quantification Methods for Model Calibration, Validation, and Risk Analysis

Cosmin Safta* Khachik Sargsyan* Habib N. Najm* Kenny Chowdhary*
Bert Debusschere* Laura P. Swiler† Michael S. Eldred†

In this paper we propose a series of methodologies to address the problems in the NASA Langley Multidisciplinary UQ Challenge. A Bayesian approach is employed to characterize and calibrate the epistemic parameters in problem A, while variance-based global sensitivity analysis is proposed for problem B. For problems C and D we propose nested sampling methods for mixed aleatory-epistemic UQ.

I. Introduction

The NASA Langley Multidisciplinary Uncertainty Quantification Challenge[1] presents several questions, formulated around computer models that describe realistic aeronautical applications. These models are presented in a "black-box" formulation to encourage discipline-independent approaches. The schematic in Fig. 1 illustrates the general structure of the model inputs, outputs, and quantities of interest.

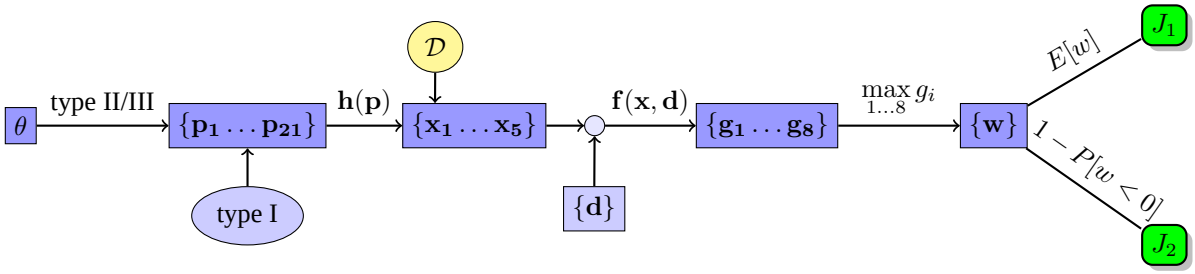

Figure 1: Schematic of the computational workflow with models and parameters

There are 21 uncertain parameters p entering the computational sequence shown in Fig. 1. These parameters are split into three categories: (I) aleatoric parameters with known distributions, (II) epistemic constants that lie within specified bounds, and (III) aleatoric parameters with distributions parameterized by epistemic variables that lie within specified bounds. The parameters of type II and the epistemic variables corresponding to parameters of type III are denoted with $\theta = (\theta_{1,*}, \theta_{2,*}, \dots, \theta_{5,*})$. The dependencies between the intermediate variables x_i, the aleatoric parameters p_j and the epistemics inputs $\theta_{i,*}$ are provided below

$$\begin{aligned}
x_1 &= h_1\left(p_1(\theta_{1,1}, \theta_{1,2}), p_2(\theta_{1,3}), p_3, p_4(\theta_{1,4}, \dots, \theta_{1,8}), p_5(\theta_{1,4}, \dots, \theta_{1,8})\right) \\
x_2 &= h_2\left(p_6(\theta_{2,1}), p_7(\theta_{2,2}, \theta_{2,3}), p_8(\theta_{2,4}, \theta_{2,5}), p_9, p_{10}(\theta_{2,6}, \theta_{2,7})\right) \\
x_3 &= h_3\left(p_{11}, p_{12}(\theta_{3,1}), p_{13}(\theta_{3,2}, \theta_{3,3}), p_{14}(\theta_{3,4}, \theta_{3,5}), p_{15}(\theta_{3,6}, \theta_{3,7})\right) \\
x_4 &= h_4\left(p_{16}(\theta_{4,1}), p_{17}(\theta_{4,2}, \theta_{4,3}), p_{18}(\theta_{4,4}, \theta_{4,5}), p_{19}, p_{20}(\theta_{4,6}, \theta_{4,7})\right) \\
x_5 &= h_5\left(p_{21}(\theta_{5,1}, \theta_{5,2})\right)
\end{aligned} \tag{1}$$

In Eq. (1), parameters $\{p_3, p_9, p_{11}, p_{19}\}$ are of type I, parameters $\{p_2, p_6, p_{12}, p_{16}\}$ are of type II, and the rest of the parameters are of type III. The subsequent computational sequence also depends on a vector of design variables d.

$$g_i = f_i(x_1, \dots, x_5, d), \quad i = 1 \dots 8 \tag{2}$$

*Sandia National Labs, Livermore, CA, 94551, USA
†Sandia National Labs, Albuquerque, NM, 87185, USA

The metrics J_1 and J_2 are based on statistics of w, which is defined as the maximum value across all $g_i = f(p, d)$ for given samples of p and d.

We propose a Bayesian framework[2,3] in Section II to characterize the epistemic variables for the input parameters of submodel h_1. Since the full likelihood evaluations are computationally expensive, we also propose an Approximate Bayesian Computation approach.[4] This approach, while computationally efficient, depends on whether the summary statistics available to construct the approximate likelihood are sufficiently descriptive for the parameters to be inferred.

In Section III we employ variance-based Global Sensitivity Analysis (GSA)[5] to rank both the aleatoric inputs p, and the epistemic variables θ in terms of their contribution to the total variance of intermediate variable x, and the overall Quantities of Interest J_1 and J_2, respectively. Using the GSA results we fix several unimportant parameters p to nominal values since their contribution to the overall variance for the intermediate variables was negligible.

We conclude the model analysis with Section IV where we employ a nested sampling methodology for the combined epistemic-aleatoric inputs to propagate forward the input uncertainties to probability distributions for J_1 and J_2. Conclusions are summarized in Section V.

II. Model Calibration

In this section we address the questions in subproblem A.[1] Specifically we employ a Bayesian framework to compute posterior probabilities for epistemic parameters and answer questions A1 and A3. The relative quality of these posteriors, pertaining to question A4, is assesed later in this section via a CRPS score.

The uncertainties in the model parameters are characterized in a Bayesian framework as

$$p(\theta_1|\mathcal{D}) = L_{\mathcal{D}}(\theta_1)p(\theta_1)/p(\mathcal{D}) \tag{3}$$

Here $p(\theta_1)$ and $p(\theta_1|\mathcal{D})$ are the prior and posterior probability densities of model parameters θ_1, respectively. These represent our knowledge about the values of θ_1 before and after learning from the data $\mathcal{D}$. The likelihood function $L_{\mathcal{D}}(\theta_1) = p(\mathcal{D}|\theta_1)$ is the likelihood of the data $\mathcal{D}$ for a particular instance of model parameters θ_1. The denominator in Eq. (3), $p(\mathcal{D})$, is the "evidence", computed by integrating the numerator over the support of θ_1.

The set of model parameters $\theta_1 = \{\theta_{1,1}, \theta_{1,2}, \ldots, \theta_{1,8}\}$ determines the corresponding uncertain parameters p_1, p_2, p_4, and p_5 as follows:

- $\theta_{1,1}$ and $\theta_{1,2}$ are the expectation and variance for p_1 which is modeled as a Beta distribution, $Beta(\alpha, \beta)$. For given values of $\theta_{1,1}$ and $\theta_{1,2}$, the shape parameters α and β are computed as

$$\alpha = \theta_{1,1}\left(\frac{\theta_{1,1}(1-\theta_{1,1})}{\theta_{1,2}} - 1\right), \quad \beta = (1-\theta_{1,1})\left(\frac{\theta_{1,1}(1-\theta_{1,1})}{\theta_{1,2}} - 1\right) \tag{4}$$

  Corresponding to the characterization of the uncertainty model for p_1 (Category III in the problem description), uniform prior distributions are set for $\theta_{1,1}$ and $\theta_{1,2}$: $\theta_1 \sim U[0.6, 0.8]$ and $\theta_2 \sim U[0.02, 0.04]$.

- $\theta_{1,3}$ is the epistemic input p_2. Its prior distribution is $U[0, 1]$.

- Parameters $\theta_{1,4}$ through $\theta_{1,8}$ characterize a bi-variate normal distribution for p_4 and p_5: $\theta_{1,4}$ and $\theta_{1,5}$ are the expectations of p_4 and p_5, $\theta_{1,6}$ and $\theta_{1,7}$ are their variances, and $\theta_{1,8}$ is their correlation factor. The prior distributions for these parameters are set to

$$\theta_{1,4}, \theta_{1,5} \sim U[-5, 5], \quad \theta_{1,6}, \theta_{1,7} \sim U[0.025, 4], \quad \theta_{1,8} \sim U[-1, 1] \tag{5}$$

Parameter p_3 is presumed aleatoric in the problem setup, with a given PDF, consequently it is not an object of inference.

The data $\mathcal{D}$ consists of two sets of 25 observations of $x_1 = h_1(p_1, p_2, p_3, p_4, p_5)$. These "observations", shown in Fig. 2, are drawn from the "true uncertainty model", i.e. the θ_1 parameters for the distributions of p_1, p_2, p_4, and p_5 were fixed.

Since the data samples are independent, the likelihood $L_{\mathcal{D}}(\theta_1)$ in Eq. (3) is written as

$$L_{\mathcal{D}}(\theta_1) = \prod_{i=1}^{N_d} p(x_{1,i}|\theta_1) \tag{6}$$

where N_d is 25 when only the first set of data is presumed to be known or 50 when both data sets are known. Here $p(x_{1,i}|\theta_1)$ is the probability to observe the value $x_{1,i}$ for a given instance of the epistemic inputs θ_1. Algorithm 1

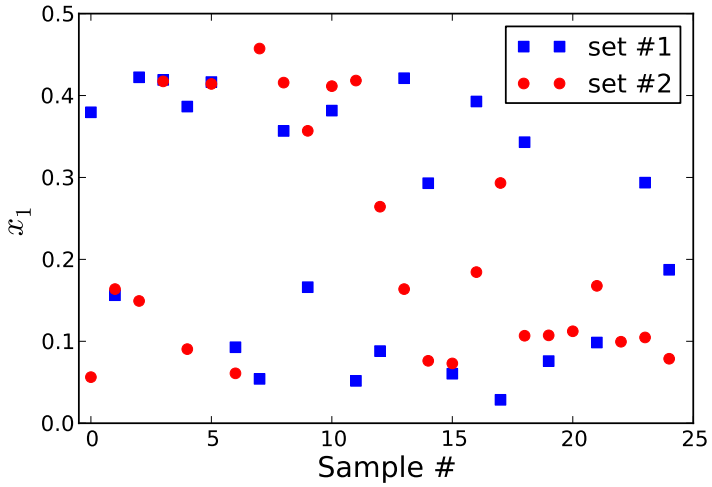

Figure 2: Samples of x_1 provided with the matlab package.

Algorithm 1: Construction of likelihood in eq. 6

Input: $\theta_1 = \{\theta_{1,1}, \ldots, \theta_{1,8}\}$, Algorithm knobs: N_{spl}, N_{bin}
Output: $L_{\mathcal{D}}(\theta_1)$
Generate N_{spl} sets of $(p_1, \ldots, p_5)$:
foreach $k = 1, \ldots, N_{spl}$ **do**
  $p_{1,k} \sim Beta(\alpha(\theta_{1,1}, \theta_{1,2}), \beta(\theta_{1,1}, \theta_{1,2}))$
  $p_{2,k} = \theta_{1,3}$
  $p_{3,k} \sim U[0, 1]$
  $(p_{4,k}, p_{5,k}) \sim N[\mu(\theta_{1,4}, \theta_{1,5}), \Sigma(\theta_{1,6}, \theta_{1,7}, \theta_{1,8})]$
end
$\{x_1|_{\theta_1}\} = \{h_1(\mathbf{p_k}), k = 1, \ldots, N_{spl}\}$
Evaluate $p(x_{1,i})$, $i = 1, \ldots, N_d$ by binning the $\{x_1|_{\theta_1}\}$ ensemble
Evaluate Eq. (6)

describes the steps taken to evaluate $L_{\mathcal{D}}(\theta)$ numerically. This algorithm has two knobs, the number of h_1 model evaluations, N_{spl}, and the number of bins, N_{bin}, used to evaluate the PDFs of x_1 numerically.

Using the likelihood construction provided in Algorithm 1 and the prior information mentioned above, we employ an adaptive-Metropolis Markov Chain Monte Carlo[6] (aMCMC) algorithm to explore the posterior distribution $p(\theta_1|\mathcal{D})$ via sampling. The aMCMC algorithm uses the covariance of the previously visited chain states to adaptively improve proposal distributions, thus exploring the posterior distribution in an efficient manner.

Figure 3 shows marginal PDFs based on 1.5×10^6 samples from the posterior distribution $p(\theta_1|\mathcal{D})$. The samples were generated via the aMCMC algorithm mentioned above. For each sample the likelihood algorithm employed $N_{spl} = 10^5$ evaluations of model h_1 followed by the PDF construction using $N_{bin} = 10^2$ bins. Inspection of the 1D marginal posteriors, in the diagonal frames, reveals a nearly uniform PDF for $\theta_{1,8}$, the correlation of p_4 and p_5. For this parameter, the x_1 data is not sufficiently informative to appreciably update its uniform prior distribution.

The marginal posterior distributions for the other parameters show some modifications compared to their prior distributions. In particular, $\theta_{1,3}$ and $\theta_{1,5}$ exhibit bimodal shapes. Multi-modal posterior distributions generally make it difficult to achieve an efficient Markov chain, hence the large number of samples needed to obtain converged posteriors. Investigation of the joint posterior PDFs in the lower triangle reveals moderate correlations between $\theta_{1,1}$ and $\theta_{1,4}$ and between $\theta_{1,5}$ and $\theta_{1,7}$. These correlations will be evaluated quantitatively in the following section.

Figure 4 shows a comparison of the marginal posterior distribution for select parameters inferred using either set #1 (blue lines) or both sets #1 and #2 (green lines). Evidently, more data leads to sharper PDFs for these parameters. The values for $\theta_{1,6}$ and $\theta_{1,8}$ are not shown since their posterior distributions did not change significantly.

Since the observed data is limited, we adopt a probabilistic test based on the predictive cumulative distribution

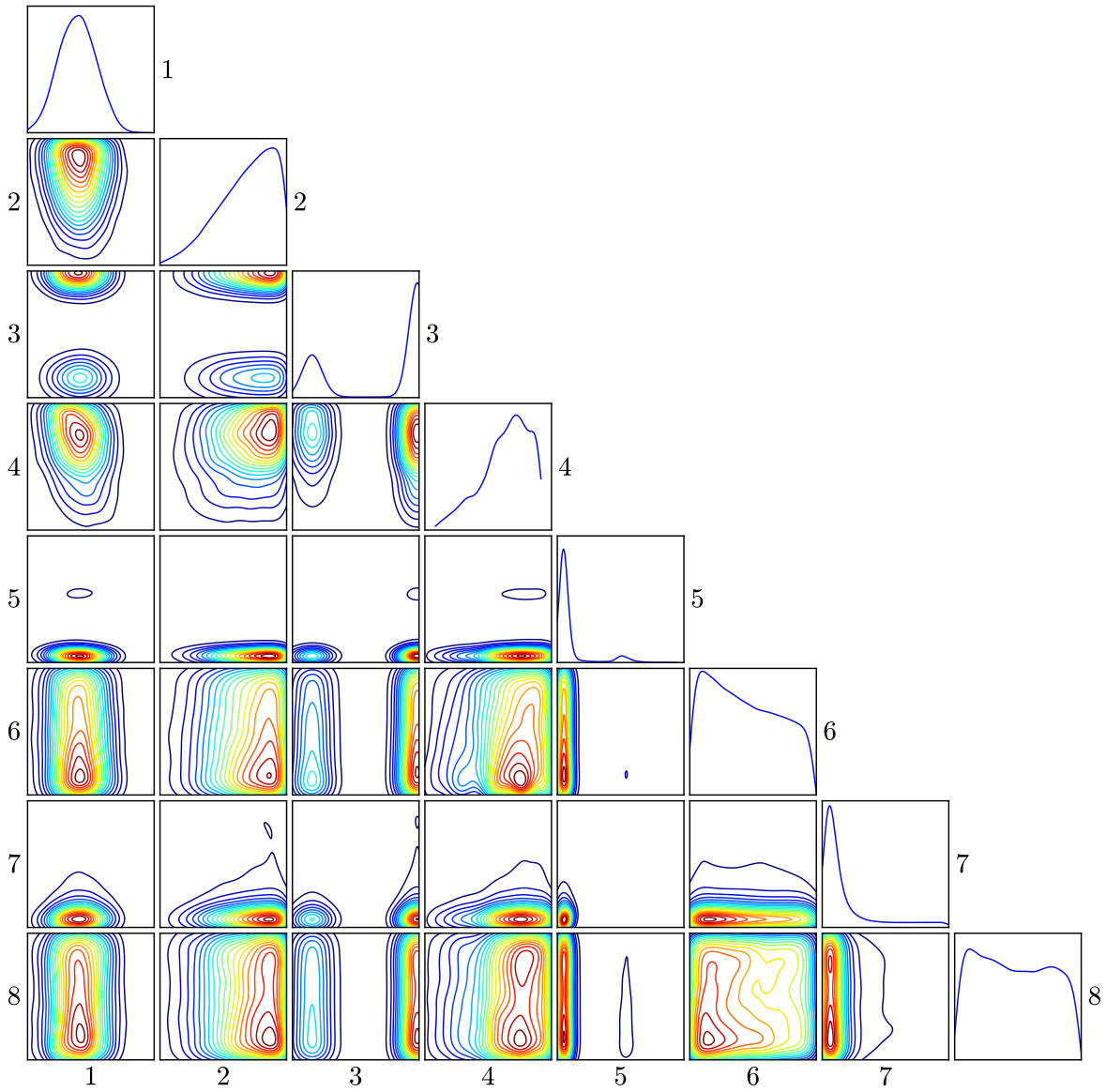

Figure 3: Marginal posterior PDFs for θ_1 calibrated based on all of 50 x_1 samples.

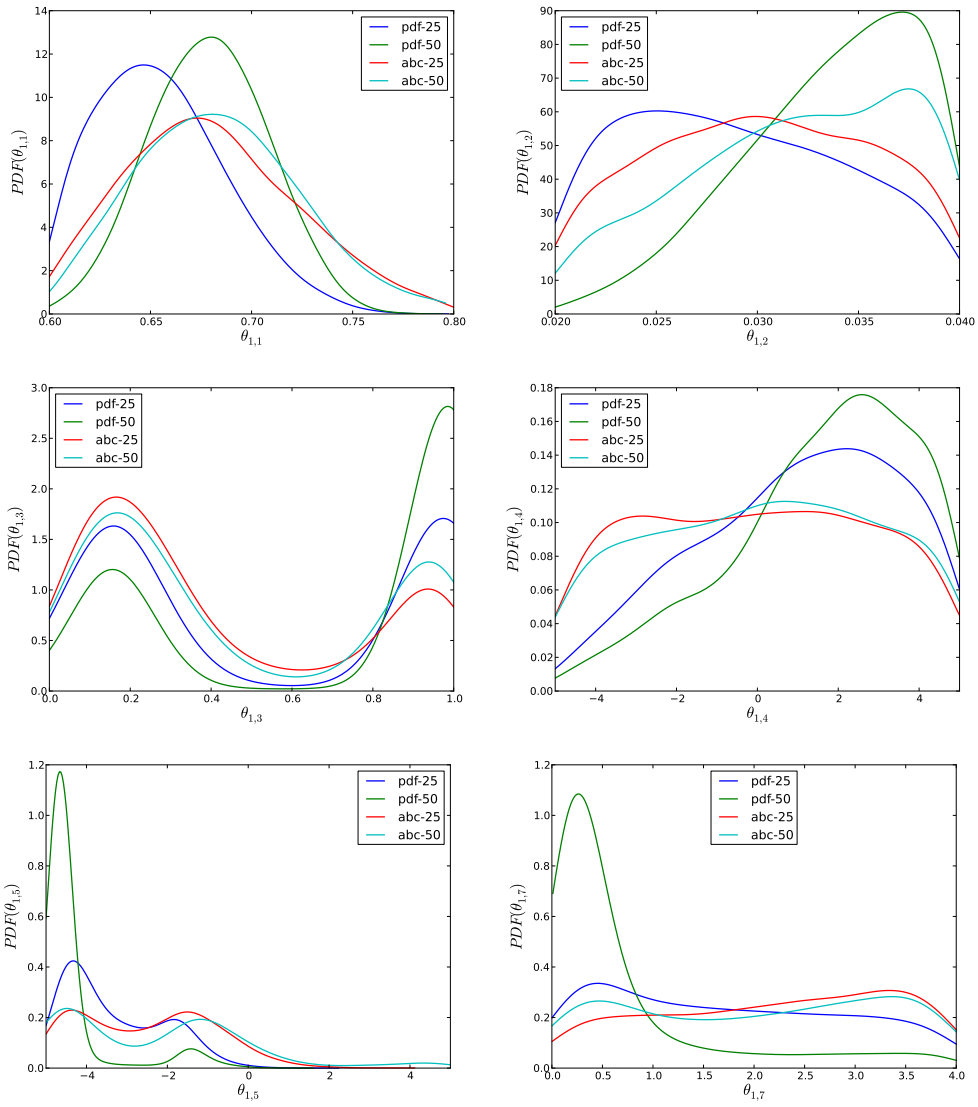

Figure 4: Comparison of marginal posterior PDFs based on 25 and 50 data samples, respectively.

function (CDF) rather than one based on predictive PDF itself. The Continuous Rank Predictive Score (CRPS)[7] measures the difference between the CDF of the provided data and the CDF of the forecast data, i.e. data generated based on the posterior predictive distribution.

$$CRPS = \frac{1}{N_{ds}} \sum_{k=1}^{N_{ds}} \int \left(\mathcal{F}_k(x_1) - \mathcal{F}(x_1)\right)^2 dx_1 \tag{7}$$

Here, $\mathcal{F}(x_1)$ is the CDF based on the data $\mathcal{D}$ shown in Fig. 2. This CDF is piecewise constant, approximated as a sum of Heaviside functions centered at the data locations.[8] Fig. 6 shows the $\mathcal{F}(x_1)$ CDFs based on either 25 or 50 data points, respectively. $\mathcal{F}_k(x_1)$ is the CDF approximated from simulated data constructed as follows. For each sample $\theta_1|_k$ we generate a dataset $\mathcal{D}_k$ of x_1 values through the evaluation of $h_1(p_1 \dots p_5|(\theta_1|_k))$. The $CRPS$ score is an average over a number of datasets, N_{ds}, generated with θ_1 sampled from its posterior distribution. We determined that about $N_{ds} = 10^4$ datasets were necessary to obtain CRPS values with about 10^{-2} relative accuracy. For each dataset we employed 10^3 h_1 model evaluations.

Table 1 displays CRPS values based on posterior distributions obtained with 25 and 50 data samples, respectively. The lower value for the latter case indicates the posterior distribution of model parameters is closer to the "true

uncertainty model" compared to the former case. Also shown in this table are CRPS values for posterior distributions obtained via Approximate Bayesian Computation (ABC) methodologies introduced in the next section. These values are slightly larger than their counterparts mentioned above indicating a slightly lower predictive skill for the ABC-generated distribution.

Table 1: CRPS values computed using Eq. (7).

	full MCMC	ABC-MCMC
set #1	3.1×10^{-3}	3.4×10^{-3}
sets #1,2	2×10^{-3}	2.3×10^{-3}

A. Model Calibration via Approximate Bayesian Computation

The PDF of model x_1 is bimodal, and the two modes exhibit jitter for slight changes in θ_1. Accurate estimation of this PDF requires a significant number of model evaluations, on the order of 10^5, at each MCMC step. Moreover, about $10^5 - 10^6$ MCMC samples are needed to represent the posterior distribution of θ_1 adequately. This leads to a large computational cost for this probabilistic calibration.

Recently, Approximate Bayesian Computation methods have seen significant development.[4] In ABC, the expensive or even intractable likelihood evaluations are replaced by expressions based on select data summaries

$$L_{\mathcal{D}}(\theta) \propto \frac{1}{\epsilon} K\left(\frac{\|s(\mathcal{D}^*(\theta)) - s(\mathcal{D})\|}{\epsilon}\right) \tag{8}$$

Here the data $\mathcal{D}^*(\theta)$ is a proposed data set drawn from the model given the current θ instance, and $s(\mathcal{D})$ is some set of statistics based on the data, e.g. means and/or quantiles. For notational convenience, we will use $s^* = s(\mathcal{D}^*(\theta))$ and $s = s(\mathcal{D})$. The kernel K takes high values when statistics of proposed data sets $\mathcal{D}^*$ are close to the ones based on the original data and conversely, it penalizes samples that are not consistent. The width ϵ plays the role of a "goodness" criterion controlling the radius of the hyper-sphere around the "true" statistic. Several kernel shapes are proposed in the ABC literature. Figure 5 shows the shapes for three kernels described below. Several ABC studies employ a sharp kernel,

$$\frac{1}{\epsilon} K\left(\frac{\|s^* - s\|}{\epsilon}\right) = \begin{cases} \frac{1}{\epsilon} & \text{if } \|s^* - s\| \leq \epsilon \\ 0 & \text{if } \|s^* - s\| > \epsilon. \end{cases} \tag{9}$$

Beaumont et al.[9] introduced the Gaussian and Epanechnikov kernels. The Gaussian kernel is defined as

$$\frac{1}{\epsilon} K\left(\frac{\|s^* - s\|}{\epsilon}\right) = \frac{1}{\epsilon\sqrt{2\pi}} \exp\left(-\frac{\|s^* - s\|^2}{2\epsilon^2}\right), \tag{10}$$

while the Epanechnikov kernel is defined as

$$\frac{1}{\epsilon} K\left(\frac{\|s^* - s\|}{\epsilon}\right) = \begin{cases} \frac{3}{4\epsilon}\left(1 - \frac{\|s^* - s\|^2}{\epsilon^2}\right) & \text{if } \|s^* - s\| \leq \epsilon \\ 0 & \text{if } \|s^* - s\| > \epsilon. \end{cases} \tag{11}$$

In this work we adopt a Gaussian kernel density K given in Eq. (10), with $\|s^* - s\|$ computed using an L_2 norm based on one or more summary statistics

$$\|s^* - s\| = \sqrt{\sum_{i=1}^{N_s} (s_i^* - s_i)^2} \tag{12}$$

The summary statistics employed in this study are quantiles, $s_i^* = x_{1,q_i}$, where x_{1,q_i} are the values associated to select quantile q_i values based on $D^*(\theta)$, and s_i is the corresponding quantile in the data D. Based on the empirical PDFs presented in Fig. 6 we select $q = \{0.05, 0.25, 0.75, 0.95\}$. Table 2 shows the x_{1,q_i} values based on either set #1 (25 samples) or both sets (50 samples).

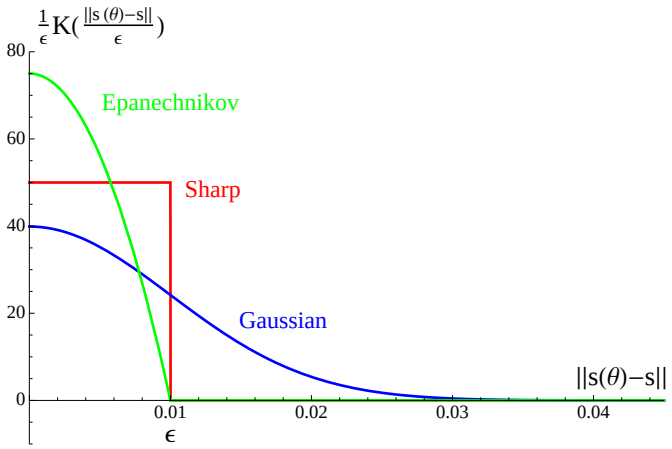

Figure 5: Kernels for Approximate Bayesian Computations.

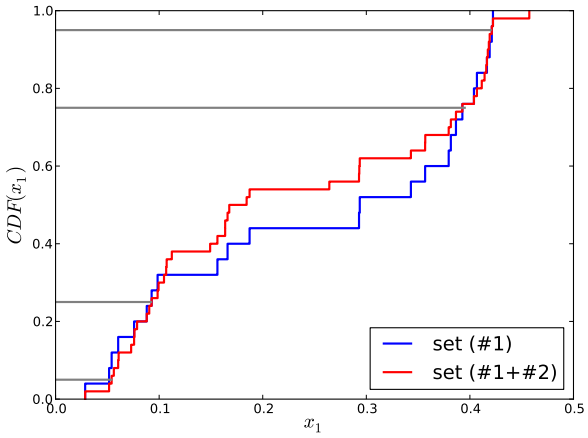

Figure 6: Empirical CDFs corresponding to the x_1 data sets. The blue line is based on the first 25 data points while the red line uses all 50 data points. The horizontal lines correspond to $\{0.05, 0.25, 0.75, 0.95\}$ quantile levels.

We performed several tests using progressively smaller ϵ values from 10^{-4} down to 10^{-6}. We found that $\epsilon = 10^{-5}$ significantly reduces the spread, compared to $\epsilon = 10^{-4}$, in the summary statistics used in the construction of the density kernel in Eq. (10). Beyond this threshold the effect of ϵ becomes negligible.

Figure 7 shows a comparison between summary statistics for data sets obtained via ABC calibration with $\epsilon = 10^{-5}$. The horizontal lines correspond to the values provided in Table 2. The results in this figure indicate that quantile values for $q = 0.05, 0.25, 0.75$ are matched relatively well by the ABC-MCMC results. The $q = 0.95$ seems in general overpredicted which could indicate the fact this quantile is poorly approximated with 25 or 50 data points only.

The marginal probability densities for θ_1 parameters generated by the ABC-MCMC results are compared in Fig. 4 with the corresponding densities based on full likelihood computations. The ABC methodology is about two orders of magnitude more efficient compared to the full Bayesian approach which employs expensive likelihood calculations. The ABC results capture some of the features obtained with full likelihood evaluation, e.g. the bimodal shapes observed for $\theta_{1,3}$ and $\theta_{1,5}$. For other parameters, in particular $\theta_{1,4}$ and $\theta_{1,7}$ the comparison is less favorable. In the following sections, wherever needed, we will be using the results based on the full Bayesian likelihood evaluations presented in Fig. 3.

Table 2: Summary statistics used in the ABC calibration process.

	$x_{1,0.05}$	$x_{1,0.25}$	$x_{1,0.75}$	$x_{1,0.95}$
set #1	0.046	0.0916	0.3955	0.4214
sets #1,2	0.054	0.0927	0.3927	0.4211

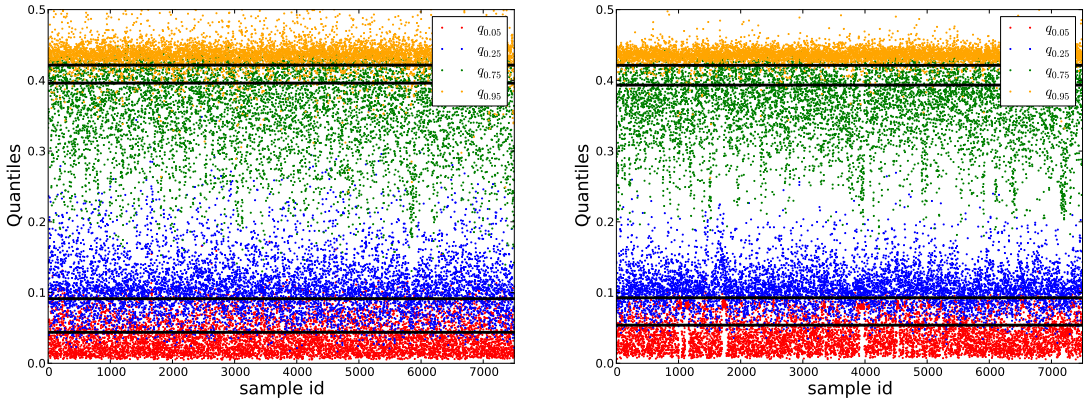

Figure 7: Comparison of summary statistics in Approximate Bayesian Computation inferences. The results in the left frame correspond to summary statistics based on 25 x_1 samples, while the right frame corresponds to statistics based on 50 x_1 samples. The horizontal lines correspond to quantiles shown in Table 2

III. Sensitivity analysis

In this section we employ variance-based Global Sensitivity Analysis (GSA) to understand the effects of the epistemic variables θ and the aleatoric parameters p on the intermediate variables x and the Quantities of Interest (QoI) J_1 and J_2. In the first part of the section the GSA approach is employed to understand the connectivities between p and x and between θ and x, and address question B1. Subsequently we employ a cut-HDMR approach to answer B2 and B3.

In the first step we proceed to determine the effect of parameters p on intermediate variables x. We proceed by first performing a global sensitivity analysis (GSA) study on the components of x to determine their relative dependence on the inputs p. These dependencies are quantified through Sobol indices.[5,10]

$$S_i^k = \frac{\text{Var}[\mathbb{E}(h_k(\mathbf{p})|p_i)]}{\text{Var}[h_h(\mathbf{p})]},\ k = 1 \dots 4 \tag{13}$$

for $i = 1, \dots, N_p$, while the joint sensitivity indices S_{ij} are

$$S_{ij}^k = \frac{\text{Var}[\mathbb{E}(h_k(\mathbf{p})|p_i, p_j)]}{\text{Var}[h_k(\mathbf{p})]} - S_i^k - S_j^k,\ k = 1 \dots 4 \tag{14}$$

for $i, j = 1, \dots, N_p$. We only compute sensitivity coefficients for functions h_1 through h_4 since h_5 depends on only one parameter, $h_5 = p_{21}$. And each of these first four functions depends on 5 p parameters, $N_p = 5$. The variances in the numerators of (13) and (14) are with respect to the parameter p_i or pair of parameters (p_i, p_j), while the expectations are with respect to the rest of the parameters. The sensitivity index S_i^k can be interpreted as the fraction of the variance in $x_k = h_k$ that can be attributed to the i-th input parameter only, while S_{ij}^k is the variance fraction that is due to the joint contribution of i-th and j-th input parameters. Interactions between three or more parameters can be defined in a similar fashion. In realistic models, however, these higher-order interactions are typically negligible.

This decomposition is unique if the input variables are independent. This is true for models h_2 through h_4, but not for h_1, for which the input parameter space is partially correlated as a result of the calibration exercise presented in the previous section. For model h_1 we introduce the total sensitivity index

$$S_i^T = \frac{\mathbb{E}[\text{Var}(h(\mathbf{p})|p_{-i})]}{\text{Var}[h(\mathbf{p})]} \tag{15}$$

which measures the fractional contribution to the total variance due to parameter p_i and its interactions with all other model parameters. This index is useful particularly for dimensionality reduction i.e. for detecting which parameters have negligible overall contribution to the model output.

The Sobol indices (13-15) can be written in integral forms, but these integrals are, in general, not tractable analytically when the input parameter space is high-dimensional. In order to evaluate these indices numerically we employ a Monte-Carlo (MC) approach enhanced by techniques described by Saltelli[11] and Kucherenko *et al.*[12] These techniques employ an efficient re-use of model evaluations to reduce the computational cost of conditional variances.

The MC methodology involves drawing samples of $\mathbf{p}$ from its distribution. For the type II parameters among $p_6 \dots p_{21}$, we employ a uniform distribution $U[0,1]$. For example, using the notation in Fig. 1, $p_6 = \theta_{2,1} \sim U[0,1]$. The distributions for type III parameters among $p_6 \dots p_{21}$ are obtained by marginalization over the corresponding epistemic variables θ. For example, the marginal distribution for p_8 is computed as

$$p(p_8) = \int_{\theta_{2,4}=7.45}^{14.093} \int_{\theta_{2,5}=4.285}^{7.864} p(p_8|\theta_{2,4},\theta_{2,5}) p(\theta_{2,4},\theta_{2,5}) d\theta_{2,4}\theta_{2,5} \tag{16}$$

Here the conditional probability density $p(p_8|\theta_{2,4},\theta_{2,5})$ is a beta distribution with shape parameters $\theta_{2,4}$ and $\theta_{2,5}$, respectively. The joint probability density $p(\theta_{2,4},\theta_{2,5})$ is uniform on the rectangle given by the limits of $\theta_{2,4}$ and $\theta_{2,5}$.

The construction of MC samples for type II parameter p_2 and type III parameters p_1, p_4, and p_5 requires a particular approach since their epistemic variables are now possibly correlated as a result of the calibration against x_1 data. In order to measure their degree of dependence we compute the distance correlation values[13] for $\theta_{1,1} \dots \theta_{1,8}$, based on the MCMC results shown in the previous section. The distance correlation between two random variables X and Y with finite first moments is the non-negative number $\mathcal{R}(X,Y)$ defined as

$$\mathcal{R}(X,Y) = \frac{\vartheta^2(X,Y)}{\sqrt{\vartheta^2(X)\vartheta^2(Y)}} \tag{17}$$

where $\vartheta^2(X,Y)$ is the distance covariance between X and Y and the distance variance $\vartheta^2(X) = \vartheta^2(X,X)$. Szekely et al.[13] provides numerical algorithms to compute $\mathcal{R}(X,Y)$ given samples of random variables X and Y. The results for epistemic variables $\theta_{1,1} \dots \theta_{1,8}$ are shown in Table 3.

	$\theta_{1,2}$	$\theta_{1,3}$	$\theta_{1,4}$	$\theta_{1,5}$	$\theta_{1,6}$	$\theta_{1,7}$	$\theta_{1,8}$
$\theta_{1,1}$	0.06	0.06	**0.17**	0.07	0.05	0.04	0.04
$\theta_{1,2}$	-	0.09	**0.14**	0.09	0.04	0.05	0.06
$\theta_{1,3}$	-	-	0.07	0.03	0.03	0.03	0.04
$\theta_{1,4}$	-	-	-	0.05	0.04	0.06	0.04
$\theta_{1,5}$	-	-	-	-	0.03	**0.39**	0.07
$\theta_{1,6}$	-	-	-	-	-	0.09	0.04
$\theta_{1,7}$	-	-	-	-	-	-	0.04

Table 3: Distance correlation factors for parameters $\theta_{1,1} \dots \theta_{1,8}$, with joint densities shown in Fig. 3.

The distance correlation factor $\mathcal{R}$ quantifies the degree of dependence between epistemic parameters. A value close to zero implies independence, while larger values, up to 1, indicate some degree of dependence between parameters. We highlight in Table 3 the entries greater than 0.1. Based on these results we conclude that the epistemic variables for the type III parameters p_1, p_4, and p_5 are correlated, while the type II $p_2 = \theta_{1,3}$ is independent of the other θ_1 components. With this information, the marginal probability density for p_1, p_4, p_5 is computed as

$$p(p_1,p_4,p_5) = \int p(p_1|\theta_{1,1},\theta_{1,2}) p(p_4,p_5|\theta_{1,4} \dots \theta_{1,8}) p(\{\boldsymbol{\theta_1} \setminus \theta_{1,3}\}) d\{\boldsymbol{\theta_1} \setminus \theta_{1,3}\} \tag{18}$$

where the $p(p_1|\theta_{1,1},\theta_{1,2})$ is a beta distribution with mean $\theta_{1,1}$ and variance $\theta_{1,2}$, while $p(p_4,p_5|\theta_{1,4} \dots \theta_{1,8})$ is a bivariate normal with prescribed means, variances and correlation factor, respectively. Here, the $\{\boldsymbol{\theta_1} \setminus \theta_{1,3}\}$ notation implies all θ_1 parameters except $\theta_{1,3}$. The multi-dimensional integral in Eq. (18) is computed via sampling by re-using the MCMC samples corresponding to the posterior distribution on $\boldsymbol{\theta_1}$. Further, in order to construct conditional samples from the probability density $p(p_1,p_4,p_5)$ required for GSA computations for model h_1, we employ the Rosenblatt

transformation.[14] The distribution for parameter p_2 is computed via Kernel Density Estimate (KDE). This density is also shown in the third diagonal frame of Fig. 3.

The main and and joint sensitivity indices for x_1 through x_4 are shown in Tables 4-7. Coefficients with values less than 0.01% are rounded to zero in these tables. We also included the type I parameters in the sensitivity analysis for completeness. For these parameters, we assumed a U[0,1] distribution.

For model $x_1 = h(p_1 \dots p_5)$, parameters p_2, p_3, and p_4 have a negligible contribution to the variance of x_1. Similarly, parameters $\{p_8, p_9, p_{10}\}$, $\{p_{13}, p_{15}\}$, and $\{p_{19}, p_{20}\}$ have negligible impact on x_2, x_3, and x_4, respectively. Based on these results we fix these parameters to the mode values corresponding to their marginal densities presented above.

	S_i	S_i^T
$\mathbf{p_1}$	**95.80**	**97.20**
p_2	0.20	0.01
p_3	0	0
p_4	3.00	0
$\mathbf{p_5}$	**3.40**	**4.00**

Table 4: Main and total effect sensitivity indices for $h_1(p_1 \dots p_5)$ (%).

	S_i	S_{ij}			
		p_7	p_8	p_9	p_{10}
$\mathbf{p_6}$	**7.57**	**82.50**	0.22	0.08	0.01
$\mathbf{p_7}$	**6.66**	-	0.28	0.14	0.01
p_8	0.27	-	-	0.01	0
p_9	0.11	-	-	-	0
p_{10}	0	-	-	-	-

Table 5: Main and joint Sobol indices for $h_2(p_6 \dots p_{10})$ (%).

	S_i	S_{ij}			
		p_{12}	p_{13}	p_{14}	p_{15}
p_{11}	**1.07**	0.88	0	0.03	0
$\mathbf{p_{12}}$	**92.53**	-	0.07	**2.97**	0.19
p_{13}	0.05	-	-	0	0
$\mathbf{p_{14}}$	**2.22**	-	-	-	0
p_{15}	0.20	-	-	-	-

Table 6: Main and joint Sobol indices for $h_3(p_{11} \dots p_{15})$ (%).

	S_i	S_{ij}			
		p_{17}	p_{18}	p_{19}	p_{20}
$\mathbf{p_{16}}$	**42.76**	**28.02**	**1.63**	0.01	0.17
$\mathbf{p_{17}}$	**11.17**	-	0.40	0.01	0.03
$\mathbf{p_{18}}$	**14.42**	-	-	0	0.05
p_{19}	0.04	-	-	-	0
p_{20}	0.36	-	-	-	-

Table 7: Main and joint Sobol indices for $h_4(p_{16} \dots p_{20})$ (%).

Parameters shown in bold in Tables 4-7 are most relevant to x_1 through x_4 models. Next we proceed to investigate the fractional effect the epistemic variables have on the probability boxes, or p-boxes, of x_1 through x_4. These intermediary variables are functions of aleatoric inputs p, which in turn have their probability distributions parameterized by epistemic variables θ. This results in a distributional p-box,[1] where the x_i is an aleatory uncertainty but the parameters prescribing its mathematical model are epistemic uncertainties. In order to measure these effects we define the quantity of interest to be the "area" of the p-box, computed as follows

$$\int_{\min_i(x_i)}^{\max_i(x_i)} (\max_{\boldsymbol{\theta_i}} F(x_i|\boldsymbol{\theta_i}) - \min_{\boldsymbol{\theta_i}} F(x_i|\boldsymbol{\theta_i}))dx_i \tag{19}$$

Effectively, Eq. (19) defines the area between the leftmost and rightmost envelopes containing all possible CDFs $F(x|\theta)$. Tables 8-11 show the total sensitivity indices for x_1 through x_4. For x_2, x_3, and x_4 sensitivities are with respect to the lower and upper bounds for each epistemic parameter $\theta_{i,j}$, respectively. For model $x_1 = h_1(p_1 \dots p_5)$ the sensitivity coefficients are with respect to the quantiles computed corresponding to the marginal distributions for $\theta_{1,j}$ obtained in the previous section. These results indicate some consistency with the sensitivity analysis results for the underlying functions h_i. For select models, some epistemic bounds are important for the p-box size while the sensitivity coefficient for the corresponding aleatoric variable is negligible, e.g $\theta_{1,3}$ and p_2 for model x_1.

We make use of the sensitivity results presented in this section to identify inputs that have a negligible impact on the model output. If certain parameters p are unimportant for intermediate variables $x = h(p)$, then these parmeters are also negligible for $g = f(x; d)$ and hence they can be assumed to be constant to speedup the computational effort for the subsequent analysis. Specifically, based on results included in Tables 4-7, we set parameters $p_2 - p_4$, $p_8 - p_{10}$, p_{13},

$\mathbf{p_1}$				$\mathbf{p_2}$		p_4, p_5									
$\theta_{\mathbf{1},\mathbf{1}}$		$\theta_{\mathbf{1},\mathbf{2}}$		$\theta_{1,3}$		$\theta_{1,4}$		$\theta_{1,5}$		$\theta_{1,6}$		$\theta_{1,7}$		$\theta_{1,8}$	
0.67	**0.42**	**0.35**	**0.33**	**0.33**	**0.32**	0.03	0.08	**0.47**	**0.56**	0.09	**0.15**	**0.38**	**0.35**	0.03	0.03

Table 8: Total effect sensitivity indices for x_1 with respect to the bounds of each epistemic parameter.

$\mathbf{p_6}$		$\mathbf{p_7}$				p_8				p_{10}			
$\theta_{\mathbf{2},\mathbf{1}}$		$\theta_{\mathbf{2},\mathbf{2}}$		$\theta_{2,3}$		$\theta_{2,4}$		$\theta_{2,5}$		$\theta_{2,6}$		$\theta_{2,7}$	
0.27	**0.52**	**0.23**	0.05	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03

Table 9: Total effect sensitivity indices for x_2 with respect to the bounds of each epistemic parameter.

p_{15} , p_{19}, and p_{20} to nominal values equal to expectations given their marginal distributions. For example, following the example provided in Eq. (16) for p_8,

$$p_8|_{\text{fix}} = \mathbb{E}[p_8] = \int p_8 p(p_8) dp_8 \tag{20}$$

The results presented hereafter in this paper use fixed values for the parameters mentioned above. In the following section we present sensitivity analysis results connecting the epistemic inputs to J_1 and J_2.

A. High-dimensional Model Representation

In order to determine the relative impact of the epistemic variables θ on J_1 and J_2 we need to construct efficient maps between input parameters and model outputs to alleviate the computational cost associated with repeated model evaluations.

To this end we employ High-Dimensional Model Representation (HDMR)[15] techniques. In this approach the model output is expressed in terms of a hierarchy of functions that account for the interaction between model parameters

$$M(\theta) = f_0 + \sum_{i=1}^{n} f_i(\theta_i) + \sum_{1 \le i < j \le n} f_{ij}(\theta_i, \theta_j) + \dots \tag{21}$$

Here M is the model output of interest and $f_0 = \langle M(\theta) \rangle$ is its mean value over the relevant parameter space. The univariate functions f_i represent the independent contribution of each component parameter θ_i, while f_{ij} represents the joint contribution of θ_i and θ_j components to $M(\theta)$. The summation on the *rhs* of Eq. (21) continues with higher-order interaction terms up to the function that describes the joint effect of all input parameters simultaneously.

There are several approaches to represent the component functions f in Eq. (21). In order to avoid expensive evaluations of high-dimensional integrals for the computations of these terms, here we employ the cut-HDMR approximation.[15,16] In this approach, the decomposition presented in Eq. (21) is with respect to an anchor point θ^0 in the

$\mathbf{p_{12}}$		p_{13}				p_{14}				p_{15}			
$\theta_{\mathbf{3},\mathbf{1}}$		$\theta_{3,2}$		$\theta_{3,3}$		$\theta_{3,4}$		$\theta_{3,5}$		$\theta_{3,6}$		$\theta_{3,7}$	
0.95	**0.11**	0	0	0	0	0	0	0	0	0	0	0	0

Table 10: Total effect sensitivity indices for x_3 with respect to the bounds of each epistemic parameter.

$\mathbf{p_{16}}$		p_{17}				p_{18}				p_{20}			
$\theta_{4,1}$		$\theta_{4,2}$		$\theta_{4,3}$		$\theta_{4,4}$		$\theta_{4,5}$		$\theta_{4,6}$		$\theta_{4,7}$	
0.25	**0.83**	0.02	0.01	0.01	0.01	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table 11: Total effect sensitivity indices for x_4 with respect to the bounds of each epistemic parameter.

parameter space, while the component functions f are defined along "cut" lines, planes, etc:

$$\begin{aligned} f_0 &= M(\theta^0) \\ f_i(\theta_i) &= M(\theta_i, \theta^0_{-i}) - f_0 \\ f_{ij}(\theta_i, \theta_j) &= M(\theta_i, \theta_j, \theta^0_{-ij}) - f_i(\theta_i) - f_j(\theta_j) - f_0 \\ &\dots \end{aligned} \tag{22}$$

where

$$\begin{aligned} (\theta_i, \theta^0_{-i}) &= (\theta^0_1, \dots, \theta^0_{i-1}, \theta_i, \theta^0_{i+1}, \dots, \theta^0_n) \\ (\theta_{ij}, \theta^0_{-ij}) &= (\theta^0_1, \dots, \theta^0_{i-1}, \theta_i, \theta^0_{i+1}, \dots, \theta^0_{j-1}, \theta_j, \theta^0_{j+1}, \dots, \theta^0_n) \end{aligned} \tag{23}$$

Eq. (22) outlines a conveninent recipe to evaluate the component functions. Further, we adopt a Polynomial Chaos (PC)[17,18] representation for this functions. Since the θ parameters have bounded support, we focus on Legendre-Uniform PC.[19] Specifically, the univariate functions f_i are constructed as

$$f_i(\theta_i) \approx \sum_{k=0}^{K} c^i_k \Psi_k(\xi_i(\theta_i)) \tag{24}$$

where Ψ_k is the Legendre polynomial of order k and ξ_i is a uniform random variable $\xi_i \sim U[-1, 1]$. For the epistemic parameters corresponding to p_6 through p_{21}, $\xi_i(\theta_i)$ are linear transformations mapping the bounds of the corresponding θ_i to $[-1 : 1]$. For parameters p_1 and p_5 we employ the Rosenblatt transformation to map the correlated probability densities of their epistemic variables to i.i.d. $U[-1, 1]$. The bi-variate functions are constructed as

$$f_{ij}(\theta_i, \theta_j) \approx \sum_{\substack{0 \le k_1, k_2 \\ k_1 + k_2 \le K}} c^{ij}_{k_1, k_2} \Psi_{k_1}(\xi_i(\theta_i)) \Psi_{k_2}(\xi_j(\theta_j)) \tag{25}$$

The PC expansion terms in Eq. (25) are truncated based on a total order isotropic rule, i.e. $k_1 + k_2 \le K$. If necessary, higher-variate functions can be constructed in a similar fashion.

The coefficients of the PC expansions in Eq. (24) are computed by Galerkin projection using the orthogonality of the basis terms.

$$c^i_k = \frac{\langle (M(\theta_i, \theta^0_{-i}) - M_0) \Psi_k \rangle}{\langle \Psi^2_k \rangle} \tag{26}$$

The projection integrals in the equation above are evaluated via Gauss quadrature. The model M, either one of J_1 or J_2, is evaluated at prescribed quadrature points corresponding to the quadrature rule employed. For univariate functions, employing 3-rd order PC expansions, 5 model evaluations are necessary for the quadrature rule corresponding to each function. Numerical tests showed that, at each quadrature point about $N_p = 5 \times 10^4$ aleatoric samples are necessary to reduce the standard deviation for the J_1 statistic to less than 1%, while for J_2, 10^4 samples are sufficient. Due to the computational expense, we only consider the univariate expansion terms in the cut-HDMR expansion. Bi-variate and higher-variate interactions, while potentially important, are henceforth neglected.

Once the PC expansion coefficients are available, the orthogonality of the basis terms facilitates the straightforward computation of the fractional contribution of each input parameter to the total variance of the model:

$$V_i = \sum_{k=1}^{K} (c^i_k)^2 \langle \Psi^2_k \rangle \tag{27}$$

Rank	J_1		J_2	
	p	$V_p/\max(V_p)$	p	$V_p/\max(V_p)$
1	p_{21}	1	p_1	1
2	p_1	0.17	p_{12}	0.26
3	p_5	0.02	p_5	0.14
4	p_7	0.02	p_{14}	2×10^{-3}

Table 12: Ranking of p parameters via a cut-HDMR approach.

The p parameters are then ranked according to the total contributions from their corresponding epistemic parameters θ. These results are shown in Table 12.

We find that for this particular setting, parameters p_{21} and p_1 are most influential for J_1 and J_2, respectively. Parameters p_1 and p_5, corresponding to submodel h_1 are found to be important for both outputs. For J_1, parameters corresponding to submodels h_2 (p_7) and h_5 (p_{21}) are also found to be important, while for J_2 submodel h_3 (p_{12}, p_{14}) is a also a major contributor to its variance. It should be noted however, that these results might be strongly dependent on the choice for the anchor point θ^0 for the *cut-HDMR* approach, and the fact that we explored the first-order effects only. With additional computational resources, a convergence test should be pursued to determine the effect of the reference parameter values and the contributions from higher-order interactions to the variance of the model output.

IV. Uncertainty Propagation

In this section we employ a nested sampling approach to quantify the probability distributions for the Quantities of Interest (QoIs) J_1 and J_2. and answer subproblem C and questions D1 and D2.. The "outer" loop involves samples of the epistemic variables θ, while in the "inner" loop we sample the aleatoric parameters p. For each "outer" sample the computational model is evaluated N_p times to generate the data required to evaluate the statistics for J_1 and J_2.

In the previous section we found that, in order to compute J_1 accurately, N_p should be on the order of $10^4 \ldots 10^5$, while for J_2 the number of required samples is about one order of magnitude less. This range is prohibitive for the nested sampling approach adopted for this section. In order to further explore the accuracy given a certain number of aleatoric samples, we randomly selected 20 epistemic samples θ and proceeded to compute J_1 using a range of N_p samples between 2^5 and 2^{12}. Figure 8 shows convergence results for two of these epistemic samples. The results in the left frame show that for $N_p \geq 10^3$, the standard deviation shown with error bars becomes less than 10% of the mean value of J_1 which is about 0.1 for this case. However, this threshold is not guaranteed. In the right plot, the standard deviation for $N_p = 2^{12}$ is significantly affected by the presence of outliers. Nevertheless, we proceed with $N_p = 10^3$ in the nested sampling approach due to limited computational resources imposed by the fact that the forward model is available in Matlab only. In order to examine the effect of N_p on the extreme values of J_1 and J_2 we will repeat these convergence tests on the epistemic samples corresponding to the extreme values.

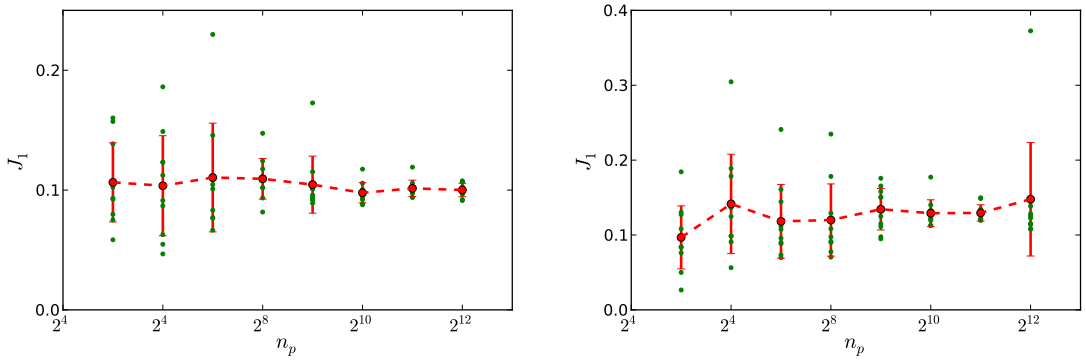

Figure 8: Distribution of J_1 values for select θ samples for a number of 'inner' samples N_p between 2^5 and 2^{12}.

Figure 9 shows probability densities for J_1 and J_2. The normalized histograms employ 100 equally-spaced bins for both J_1 and J_2. The first QoI exhibits a tailed distribution, with most of the probability located around $J_1 = 0.17$.

About 20 samples out of 5000 or 0.4% lead to $J_1 > 20$ and $\max(J_1) \approx 35$. While the smallest value of J_1 is 0.07, the next small values of J_1 are clustered around 0.1. The smallest and largest J_2 values are 0.071 and 0.923, respectively. This QoI exhibits a bimodal distribution with a strong mode peaked at $J_2 \approx 0.2$ and a weak mode at $J_2 \approx 0.6$.

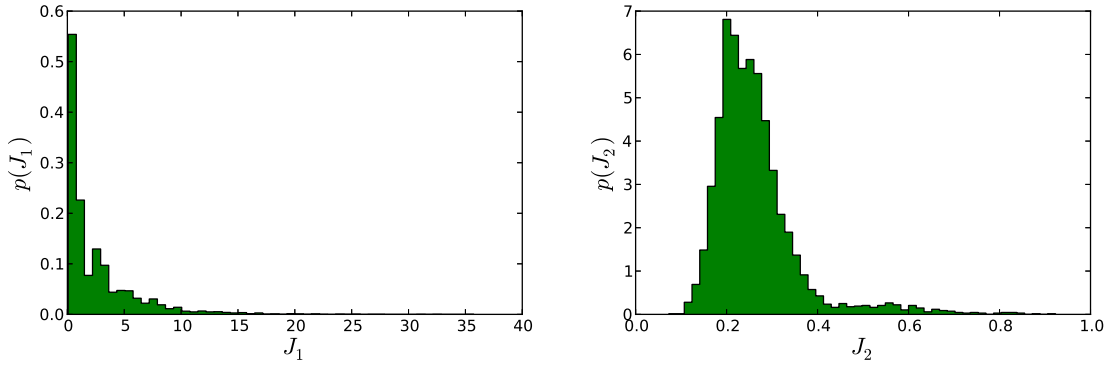

Figure 9: Probability densities for J_1 and J_2 based on a number of outer samples, $N_\theta = 5000$ and a number of 'inner' samples, $N_p = 1000$.

In order to examine the effect of the number of "inner" samples N_p on the extreme J_1 and J_2 values, we select the epistemic values corresponding to these extremes and proceed to re-compute the QoI's using an increasing number of N_p samples. Figure 10 shows normalized histograms for w constructed with $N_p = 10^5$ aleatoric samples. The left frame corresponds to the epistemic sample that lead to the maximum J_1 mentioned above, while the right frame corresponds to the epistemic sample leading to the smallest J_1. For both these epistemic samples, w has a bimodal distribution. The main mode occurs for $w < 1$, while the second mode occurs at $w \approx 10^3$ and is much weaker.

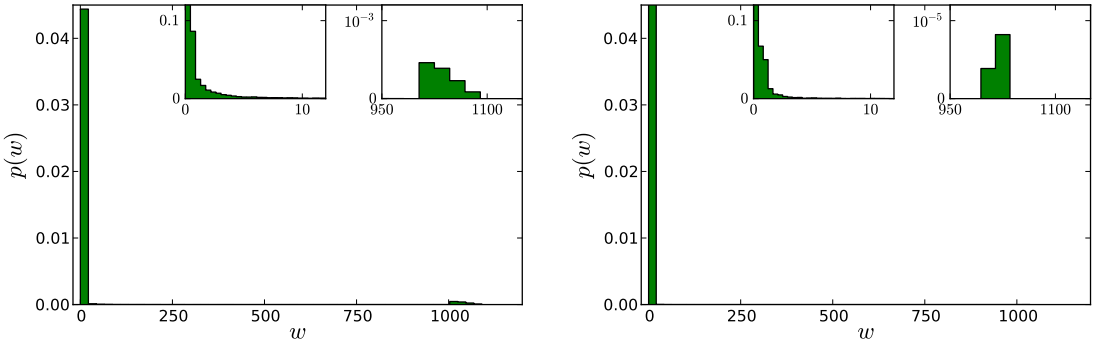

Figure 10: Probability densities for w values used to compute the results for J_1 shown in Table 13. The histograms are based $N_p = 10^5$ samples.

The multimodal distribution for w requires a large number of samples to estimate $J_1 = \mathbb{E}(w)$, since an incorrect balance between the number of samples in the two modes can lead to large errors in J_1. This is compounded by the fact that the second mode is several orders of magnitude weaker than the main mode and thus a larger number of samples is required to have good coverage for both modes.

Table 13 shows values of J_1 and J_2 based on N_p between 10^3 and 10^5. The original values mentioned above are also shown in the second column in the table. For both J_1 extremes, there is a discrepancy between the two sets of results employing $N_p = 10^3$. Additional evaluations lead to converged results around $N_p = 10^4$, for these θ samples. Additional convergence tests, results not shown, using θ samples corresponding to large and small J_1 values from Fig. 9, point to extreme values for this QoI in the vicinity of 30 and 0.1, respectively. The J_2 statistic is much less sensitive to the multimodal behavior of w. For this QoI, the $\min$ / $\max$ values change much less with an increasing number of samples.

Figure 11 shows the densities for J_1 and J_2 based on improved bounds for select epistemic parameters. These densities are estimated using 5×10^3 and 10^3 "outer" and "inner" samples, respectively. For both QoIs the support for their densities shrank compared to the original ranges shown in Fig. 9. In particular J_1 exhibits a multimodal distribu-

	$N_p = 10^3$	$N_p = 10^3$	$N_p = 10^4$	$N_p = 10^5$
$J_1(\theta\|_{\max(J_1)})$	35.7	27.7	27.3	27.4
$J_1(\theta\|_{\min(J_1)})$	0.071	0.12	0.42	0.41
$J_2(\theta\|_{\max(J_2)})$	0.923	0.912	0.910	0.908
$J_2(\theta\|_{\min(J_2)})$	0.071	0.075	0.090	0.089

Table 13: Convergence of J_1 and J_2 values with increasing number of N_p samples. The first $N_p = 10^3$ column shows extreme values extracted from Fig. 9, while subsequent columns shows results from convergence tests using and increasing number of N_p samples.

tion. To verify these results we select several epistemic realizations corresponding to $J_1 = \{0.4, 2.5, 3.5, 4.5, 5.5, 6.5\}$ and resampled the corresponding aleatoric distribution using up to 10^5 samples. For all these tests the "converged" J_1 converged to the mode located between 1 and 2. We suspect, based on these results, that the multimodal distribution for J_1 shown in Fig. 11 is the result of undersampling the aleatoric space. Similar tests for J_2, results not shown, confirm that the results for this QoI are converged and that the new range is approximately $[0.27 \ldots 0.39]$.

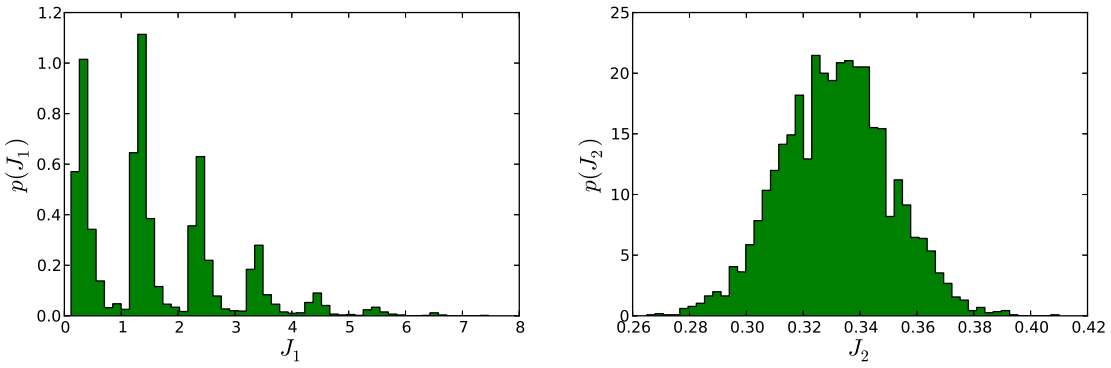

Figure 11: Probability densities for J_1 and J_2 based on a number of outer samples, $N_\theta = 5000$ and a number of 'inner' samples, $N_p = 1000$; results based on a reduced epistemic bounds for select inputs.

Finally, we have preliminary results for subproblem E (the robust design optimization) which we plan to show at the January meeting.

V. Conclusions

In this paper we employ probabilistic methodologies to handle inverse and forward UQ studies for the models posed by the NASA LaRC UQ challenge. Overall, methodologies are challenged by sparse data, the computational expense of some of the submodels and by the multi-modality of model outputs.

For the calibration exercise we propose a Bayesian framework to characterize the posterior distribution for select epistemic parameters. We compare the results based on full likelihood estimation, which is computationally intensive, with results based on Approximate Bayesian Computation (ABC) concepts. In the latter approach the likelihood is based on summary statistics and hence is computationally cheap compared with the former approach where the likelihood is based on the estimation of the probability density for each proposed data set. The ABC results are in qualitative agreement with the full Bayesian results for some of the parameters, while for others they fail to capture the marginal densities. Nevertheless, the ABC algorithm is computationally efficient and a viable alternative when the full likelihood estimation is intractable.

We then proceed with Global Sensitivity Analysis (GSA) for the effects the epistemic and aleatoric inputs have on models h and on the statistics J_1 and J_2. We employ High Dimensional Model Representation (HDMR) decomposition to dissect the fractional contribution of each parameter or combinations thereof. For computationally cheap models the sensitivity coefficients are estimated through exhausitive sampling, while for the models that are expensive we employ a cut-HDMR approach. In this latter approach, we use Polynomial Chaos expansions to represent the parameter contributions to the full model. We find that several epistemic and aleatoric variables have a negligi-

ble contribution to the model output variance and we assign constant values for this parameters in the forward UQ exercise.

In the forward UQ problem, we employ a nested sampling approach to estimate the bounds and the probability density for quantities of interest (QoI) that are based on statistics over several model outputs. One of these statistics, J_1 is based on model outputs that are bimodal, and hence the number of samples required to discover the bounds accurately are prohibitive. We also find that J_2 requires a much smaller number of model samples to obtained information on its bounds and a good estimate on its distribution.

Acknowledgement

Support for this work was provided through Scientific Discovery through Advanced Computing (SciDAC) program funded by U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

[1] "NASA LaRC UQ Challenge 2014," http://http://uqtools.larc.nasa.gov/nda-uq-challenge-problem-2014/.

[2] Sivia, D., *Data Analysis: A Bayesian Tutorial*, Oxford Science, 1996.

[3] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B., *Bayesian Data Analysis*, Chapman & Hall CRC, 2nd ed., 2003.

[4] Sisson, S. A. and Fan, Y., *Handbook of Markov Chain Monte Carlo*, chap. Likelihood-free Markov chain Monte Carlo, Chapman and Hall/CRC Press, 2011.

[5] Sobol, I. M., "Sensitivity Estimates for Nonlinear Mathematical Models," *Math. Modeling and Comput. Exper.*, Vol. 1, 1993, pp. 407–414.

[6] Haario, H., Saksman, E., and Tamminen, J., "An adaptive Metropolis algorithm," *Bernoulli*, Vol. 7, 2001, pp. 223–242.

[7] Gneiting, T. and Raftery, A., "Strictly proper scoring rules, prediction, and estimation," *Journal of the American Statistical Association*, Vol. 102, 2007, pp. 359–378.

[8] Hersbach, H., "Decomposition of the continuous ranked probability score for ensemble prediction systems." *Wea. Forecasting*, Vol. 15, 2000, pp. 559–570.

[9] Beaumont, M. A., Zhang, W., and Balding, D. J., "Approximate Bayesian computation in population genetics," *Genetics*, Vol. 162, No. 4, 2002, pp. 2025–2035.

[10] Campolongo, F., Saltelli, A., Sørensen, T., and Tarantola, S., "Hitchhiker's Guide to Sensitivity Analysis," *Sensitivity Analysis*, edited by A. Saltelli, K. Chan, and E. Scott, Wiley, Chicester, 2000.

[11] Saltelli, A., "Making best use of model evaluations to compute sensitivity indices," *Computer Physics Communications*, Vol. 145, 2002, pp. 280–297.

[12] Kucherenko, S., Tarantola, S., and Annoni, P., "Estimation of global sensitivity indices for models with dependent variables," *Computer Physics Communications*, Vol. 183, 2012, pp. 937–946.

[13] Székely, G., Rizzo, M., and Bakirov, N., "Measuring and Testing Dependence by Correlation of Distances," *Annals of Statistics*, Vol. 35, 2007, pp. 2769–2794.

[14] Rosenblatt, M., "Remarks on a Multivariate Transformation," *Annals of Mathematical Statistics*, Vol. 23, No. 3, 1952, pp. 470 – 472.

[15] Rabitz, H. and Alis, O. F., "General Foundations of High-Dimensional Model Representations," *J. Math. Chem.*, Vol. 25, 1999, pp. 197–233.

[16] Li, G., Rosenthal, C., and Rabitz, H., "High dimensional model representations," *J. Phys. Chem. A*, Vol. 105, 2001, pp. 7765–7777.

[17] Wiener, N., "The use of statistical theory in the study of turbulence," *Fifth International Congress for Applied Mechanics*, Wiley, New York, NY, 1939.

[18] Ghanem, R. and Spanos, P., *Stochastic Finite Elements: A Spectral Approach*, Springer Verlag, New York, 1991.

[19] Xiu, D. and Karniadakis, G., "The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations," *SIAM J. Sci. Comp.*, Vol. 24, No. 2, 2002, pp. 619–644.