Uncertainty Reduction using Bayesian Inference and Sensitivity Analysis: A Sequential Approach to the NASA Langley Uncertainty Quantification Challenge

Shankar Sankararaman*
SGT Inc., NASA Ames Research Center, Moffett Field, CA 94035, USA

This paper presents a computational framework for uncertainty characterization and propagation, and sensitivity analysis under the presence of aleatory and epistemic uncertainty, and develops a rigorous methodology for efficient refinement of epistemic uncertainty by identifying important epistemic variables that significantly affect the overall performance of an engineering system. The proposed methodology is illustrated using the NASA Langley Uncertainty Quantification Challenge (NASA-LUQC) problem that deals with uncertainty analysis of a generic transport model (GTM). First, Bayesian inference is used to infer subsystem-level epistemic quantities using the subsystem-level model and corresponding data. Second, tools of variance-based global sensitivity analysis are used to identify four important epistemic variables (this limitation specified in the NASA-LUQC is reflective of practical engineering situations where not all epistemic variables can be refined due to time/budget constraints) that significantly affect system-level performance. The most significant contribution of this paper is the development of the sequential refinement methodology, where epistemic variables for refinement are not identified all-at-once. Instead, only one variable is first identified, and then, Bayesian inference and global sensitivity calculations are repeated to identify the next important variable. This procedure is continued until all 4 variables are identified and the refinement in the system-level performance is computed. The advantages of the proposed sequential refinement methodology over the all-at-once uncertainty refinement approach are explained, and then applied to the NASA Langley Uncertainty Quantification Challenge problem.

I. Introduction

Research in the area of uncertainty quantification has focused on identifying, representing, and quantifying the various sources of uncertainty that affect the performance of engineering systems, and systematically estimating their effect on the system-level response in order to facilitate risk-informed decision-making. Two types of uncertainty, namely aleatory uncertainty and epistemic uncertainty, have been commonly discussed by several researchers.[1] While aleatory uncertainty refers to the uncertainty arising out of physical variability or true randomness, epistemic uncertainty arises due to lack of knowledge regarding a quantity whose true value is deterministic in nature. Epistemic uncertainty, expressed in the form of interval data, has gained particular attention in the research community during the past ten years. Sandia National Laboratories conducted an epistemic uncertainty workshop[2] that focused on the quantification and propagation of uncertainty in engineering applications where both aleatory uncertainty and epistemic uncertainty (in the form of interval) are present. There is a general consensus that it is necessary to delineate the effect of aleatory uncertainty and epistemic uncertainty on the system-level response,[3]several important questions have come into limelight. While it is straightforward to use tools of probability to represent and quantify aleatory uncertainty, some researchers have suggested the use of alternative techniques[4,5] for the representation of aleatory uncertainty. In general, non-probabilistic techniques are interval analysis-based approaches[5] and are computationally expensive wherein the cost increases exponentially with the number of uncertain variables, and with the increase in non-linearity of the response function that depends on these uncertain variables.

*Research Engineer, Intelligent Systems Technical Division, NASA Ames Research Center, AIAA Member

Some researchers still believe that probabilistic methods are sufficient to represent epistemic uncertainty and different types of probabilistic approaches have been proposed to deal with epistemic uncertainty in the form of interval data. These approaches are either based on the concept of probability boxes[6,7] or by using a family of probability distributions[8,9] to represent epistemic uncertainty. It can be argued that a probabilistic representation for epistemic uncertainty is justified according to the subjective/Bayesian interpretation of probability. Bayesian methods are based on subjective probability and can be used to assign probabilities even to epistemic variables that are not truly random but just unknown.

By definition, epistemic uncertainty is reducible though there may be a significant amount of cost associated with such uncertainty refinement. In general, if there are several epistemic variables associated with a particular engineering application, it may not be possible to refine the uncertainty in all of them, due to budget/time constraints. Therefore, it is necessary to develop a computational methodology for identifying important epistemic variables that have significant effects on system-level performance. An intuitive approach for identifying such important variables is based on variance-based global sensitivity analysis,[10] by computing the sensitivity of system-level performance measures to the various epistemic variables.

However, the use of global sensitivity analysis for the treatment of epistemic uncertainty is not well-established in the literature. The primary goal of this paper is to present a generic framework for uncertainty quantification and sensitivity analysis in the presence of both aleatory and epistemic uncertainty, and to develop a new computational approach for uncertainty refinement, based on the above framework. The most important feature of the proposed approach is that variables for refinement are selected in a sequential manner. The initial results of sensitivity analysis are used to identify only one epistemic variable for refinement; after this epistemic variable is refined, the entire analysis (including uncertainty quantification and variance-based sensitivity analysis) is repeated and the new results are used to identify the second variable for refinement. This approach is continued until all possible refinements have been made or until budget/time constraints are met.

The proposed methodology is illustrated using the NASA Langley Uncertainty Quantification Challenge (referred to NASA-LUQC, in the rest of the paper) problem presented in detail by Crespo et al.[11] The proposed sequential approach for uncertainty refinement is starkly different from existing refinement approaches,[12–16] all of which simultaneously identify all four candidates of refinement; such an approach is referred as all-at-once uncertainty refinement in this paper.

II. Importance of Sequential Uncertainty Refinement

As stated earlier, the all-at-once approach for uncertainty refinement simultaneously identifies all the possible candidates for refinement, based on the results of Bayesian inference (inferring subsystem-level variables) and global sensitivity analysis (identifying important variables that affect system-level performance). On the other hand, the proposed sequential uncertainty refinement approach chooses candidates for refinement one-by-one; the first variable is selected and refined, and then Bayesian inference and global sensitivity calculations are repeated to identify the next candidate for refinement. This procedure is continued until all candidates are selected.

The disadvantage of the all-at-once approach can be easily identified and understood based on the following argument. The various epistemic variables can be ranked using global sensitivity analysis by computing their variance-based contribution to the system-level prediction. Say for example, the highest ranked epistemic variable is chosen for uncertainty refinement and a new refined probability distribution (defined on a domain that is a much smaller subset of the domain before refinement) is available for this variable. Then, this information alters the sensitivity effects of all variables since the sensitivity of system-level performance to any variable depends on the probability distribution of all variables. Therefore, changing information regarding one variable may alter sensitivities completely. For example, the second-highest ranked variable before refinement may not necessarily remain the highest ranked variable after refinement (eliminating the previously highest ranked variable out of the ranking scheme, since it has been refined already) and may have moved farther down in the ranking.

This issue gets further complicated in the NASA-LUQC because of the presence of the subsystem-level model and the corresponding data to infer subsystem-level epistemic quantities. When multiple quantities that need to be simultaneously inferred using Bayesian updating are possible candidates for uncertainty refinement, the importance of sequential refinement increases multi-fold. Consider the case when one quantity is identified for refinement and refined uncertainty information is available; then Bayesian updating can be

performed again using a new prior (based on the newly obtained refined uncertainty estimate) on the refined quantity and new posteriors can be obtained for all other epistemic variables. It is important not to re-use available data, and therefore, Bayesian updating needs to be re-started from the original prior (for all the quantities that have not been refined).

Therefore, when multiple quantities that need to simultaneously inferred using Bayesian updating are possible candidates for uncertainty refinement, it is necessary to adopt a sequential approach for uncertainty refinement. In the NASA-LUQC, this is the case, and that is why a sequential approach for uncertainty refinement is proposed in this paper, and this is the most important contribution of this manuscript.

III. Challenge Problem: Auxiliary Variables and Notation

This section describes the notation that is used to describe and solve the NASA-LUQC. This lies on the concept of auxiliary variables developed by Sankararaman and Mahadevan[17] in order to facilitate global sensitivity analysis in the presence second-order uncertainty.

A. Use of Auxiliary Variables

The NASA-LUQC describes three types of uncertain variables: Type-I (aleatory variable, whose probability distribution is completely defined), Type-II (epistemic variable, whose uncertainty is specified using an interval), and Type-III (aleatory variable, whose probability distribution is defined in terms of distribution parameters that are epistemic variables described using intervals). The auxiliary variable (represented as U, a uniform random variable on the interval [0, 1], for every Type-III variable) is of utmost importance when dealing with Type-III uncertainty. Type-III variables are typically represented using a family of distributions; each member of the family corresponds to the aleatory uncertainty resulting from one particular realization of distribution parameters while the uncertainty in the distribution parameters leads to multiple members of the family. Any quantity that depends on any Type-III variable also follows a family of distributions. The auxiliary variable approach is based on the concept of probability integral transform and is very useful in analyzing such a family of distributions. Further, it can use tools of global sensitivity analysis can quantitatively assess the contributions of aleatory uncertainty and epistemic (distribution parameter) uncertainty.

B. Subsystem Level

In the system-level, there are 21 independent variables denoted as p_i ($i = 1$ to 21), and the five intermediate variables x_j ($j = 1$ to 5) are well defined functions of p_i ($i = 1$ to 21). All of these models have been provided as a part of the NASA-LUQC problem and available as MATLAB files, as explained by Crespo et al.[11] Let $\boldsymbol{x} = \{x_j;\ j = 1 \text{ to } 5\}$, $\boldsymbol{h} = \{h_j;\ j = 1 \text{ to } 5\}$, and $\boldsymbol{p} = \{p_i;\ i = 1 \text{ to } 21\}$.

$$x_1 = h_1(p_1, p_2, p_3, p_4, p_5) \tag{1}$$
$$x_2 = h_2(p_6, p_7, p_8, p_9, p_10) \tag{2}$$
$$x_3 = h_3(p_{11}, p_{12}, p_{13}, p_{14}, p_{15}) \tag{3}$$
$$x_4 = h_4(p_{16}, p_{17}, p_{18}, p_{19}, p_{20}) \tag{4}$$
$$x_5 = p_{21} \tag{5}$$

All random variables are denoted by upper case letters and the realizations of the random variables are denoted by the corresponding lower case letters. Hence, $\boldsymbol{P} = \{P_i;\ i = 1 \text{ to } 21\}$, and $\boldsymbol{X} = \{X_j;\ j = 1 \text{ to } 5\}$ are random variables whose realizations are denoted by $\boldsymbol{p} = \{p_i;\ i = 1 \text{ to } 21\}$ and $\boldsymbol{x} = \{x_j;\ j = 1 \text{ to } 5\}$. Note that functions are always expressed in terms of realizations of random variables.

Amongst the variables p_i ($i = 1$ to 21), some of them are Type-I variables (well-defined probability distributions), some of them are Type-II variables (epistemic, described using an interval), and the rest are Type-III variables (random variables whose distribution parameters are described using an interval each). Following the auxiliary variable framework, aleatory uncertainty is present in all Type-I and Type-III variables. Therefore, an auxiliary variable U_i is assigned to every p_i that is either Type-I or Type-III. An auxiliary variable is not assigned for Type-II variables since they are purely epistemic. There are totally 17

auxiliary variables that include U_1, U_3, U_4, U_5, U_7, U_8, U_9, U_{10}, U_{11}, U_{13}, U_{14}, U_{15}, U_{17}, U_{18}, U_{19}, U_{20}, and U_{21}. All of the variables are uniformly distributed on the interval [0, 1]. Let $\boldsymbol{U}$ denote the vector of these aleatory variables.

There are 31 epistemic quantities that are numbered serially as indicated in Table 1. Let $\boldsymbol{\Theta}$ denote the vector of these variables, and $\boldsymbol{\theta}$ denotes a realization of $\boldsymbol{\Theta}$. When necessary to denote only a subset of variables, the list of variables is also compactly represented as subscript. For example, the variables P_4 and P_5 are jointly represented by $P_{4,5}$, and the variables Θ_4, Θ_5, Θ_6, Θ_7, and Θ_8 are jointly represented by Θ_{4-8}. It is easy to observe from Table 1 that Θ_{4-8} constitute the epistemic components of $P_{4,5}$. Further, realizations of these quantities can also be indicated using the respective lower case letters.

Table 1: List of Epistemic Components of Uncertainty

Symbol	Description	Interval	Symbol	Description	Interval
Θ_1	$E(P_1)$, P_1 is unimodal Beta	$[\frac{3}{5}, \frac{4}{5}]$	Θ_2	$V(P_1)$, P_1 is unimodal Beta	$[\frac{1}{50}, \frac{1}{25}]$
Θ_3	P_2, constant	[0, 1]	Θ_4	$E(P_4)$, P_4 is Gaussian	$[-5, 5]$
Θ_5	$V(P_4)$, P_4 is Gaussian	$[\frac{1}{400}, 4]$	Θ_6	$E(P_5)$, P_5 is Gaussian	$[-5, 5]$
Θ_7	$V(P_5)$, P_5 is Gaussian	$[\frac{1}{400}, 4]$	Θ_8	Correlation between P_4 and P_5	$[-1, 1]$
Θ_9	P_6, constant	[0, 1]	Θ_{10}	a of Beta P_7	[0.982, 3.537]
Θ_{11}	b of Beta P_7	[0.619, 1.080]	Θ_{12}	a of Beta P_8	[7.450, 14.093]
Θ_{13}	b of Beta P_8	[4.285, 7.864]	Θ_{14}	a of Beta P_{10}	[1.520, 4.513]
Θ_{15}	b of Beta P_{10}	[1.536, 4.750]	Θ_{16}	P_{12}, constant	[0, 1]
Θ_{17}	a of Beta P_{13}	[0.412, 0.737]	Θ_{18}	b of Beta P_{13}	[1.000, 2.068]
Θ_{19}	a of Beta P_{14}	[0.931, 2.169]	Θ_{20}	b of Beta P_{14}	[1.000, 2.407]
Θ_{21}	a of Beta P_{15}	[5.435, 7.095]	Θ_{22}	b of Beta P_{15}	[5.287, 6.945]
Θ_{23}	P_{16}, constant	[0, 1]	Θ_{24}	a of Beta P_{17}	[1.060, 1.662]
Θ_{25}	b of Beta P_{17}	[1.000, 1.488]	Θ_{26}	a of Beta P_{18}	[1.000, 4.266]
Θ_{27}	b of Beta P_{18}	[0.553, 1.000]	Θ_{28}	a of Beta P_{20}	[7.530, 13.492]
Θ_{29}	b of Beta P_{20}	[4.711, 8.148]	Θ_{30}	a of Beta P_{21}	[0.421, 1.000]
Θ_{31}	b of Beta P_{21}	[7.772, 29.621]			

Each X_j has at least one input P_i that is either Type-II or Type-III uncertainty. Hence, from the discussion earlier in this Section, it follows that each X_i needs to be represented using a family of distributions. Therefore, auxiliary functions need to be developed for each X_j. The corresponding auxiliary functions are denoted as H_{X_j} ($j = 1$ to 5). Each auxiliary function takes as inputs a set of aleatory variables and a set of epistemic variables.

C. System-level Outputs and Performance Metrics

As specified in Crespo et al.,[11] the system level outputs ($\boldsymbol{g} = \{g_i;\ i = 1 \text{ to } 8\}$) are functions ($\boldsymbol{f} = \{f_i;\ i = 1 \text{ to } 8\}$) of the intermediate variables ($\boldsymbol{x}$) and design variables ($\boldsymbol{d}$), as:

$$\boldsymbol{g} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{d}) \tag{6}$$

Similar to the previous notations, random variables $\boldsymbol{G} = \{G_i;\ i = 1 \text{ to } 8\}$ can also be defined, and their realizations are denoted by $\boldsymbol{g} = \{g_i;\ i = 1 \text{ to } 8\}$.

There are two system-level performance metrics of interest J_1 and J_2, that are defined as:

$$J_1 = E(w(\boldsymbol{p}, \boldsymbol{d}_{baseline})) \tag{7}$$

$$J_2 = 1 - P[w(\boldsymbol{p}, \boldsymbol{d}_{baseline}) < 0] \tag{8}$$

where $w(\boldsymbol{p}, \boldsymbol{d}_{baseline})$ refers to worst case requirement metric that is calculated as:

$$w(\boldsymbol{p}, \boldsymbol{d}_{baseline}) = max(g_i) = max\Big(f_i(\boldsymbol{h}(\boldsymbol{p}), \boldsymbol{d})\Big) \tag{9}$$

The NASA-LUQC provides Eq. 6 – 9 in terms of MATLAB files, and refer to Crespo et al.[11] for detailed description of the various quantities in Eq. 6 – 9

IV. Subproblem A: Uncertainty Characterization

The first subproblem consists of only one response function that is used to calculate x_1. Here, it is important to account for the correlation between p_4 and p_5 systematically; this is straightforward because conditional distributions for correlated Gaussian variables are analytically calculable.

A. Problems A1 and A2

The goal is to update/refine the uncertainty in Θ_i ($i = 1$ to 8) using data (first set of 25 observations denoted by $D_1 = \{x_1^j;\ j = 1 \text{ to } 25\}$) on x_1. This data D_1 is provided as a part of the NASA-LUQC, in the form a MATLAB file, as explained by Crespo et al.[11] Such uncertainty refinement can be easily accomplished through Bayesian updating, using Bayes' theorem as:

$$f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1) = \frac{L(\boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})}{\int L(\boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{10}$$

where $f_{\boldsymbol{\Theta}}()$ is the prior distribution and the likelihood $L(\boldsymbol{\theta})$ defined as:

$$L(\boldsymbol{\theta}) = \prod_{j=1}^{25} f_{X_1}(x_1^j|\boldsymbol{\theta}) \tag{11}$$

In order to compute $f_{X_1}(x_1|\boldsymbol{\theta})$, it is necessary to resort to an uncertainty propagation technique. It can be easily seen that X_1 follows a probability distribution for a given realization of $\boldsymbol{\Theta}$. The corresponding PDF $f_{X_1}(x_1|\boldsymbol{\theta})$ is calculated using Monte Carlo sampling (5000 Latin hypercube samples) in this paper.

The prior distributions are chosen based on the provided interval data and/or using Jeffrey's prior, as and when appropriate. Bayesian updating is performed using the slice sampling technique.[18] After Bayesian updating, the resultant joint probability distribution of Θ_i ($i = 1$ to 8) are referred to as joint mid-posterior distribution (since this calculated only using 25 of available 50 samples of X_1), and the corresponding marginal PDFs $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1)$ are shown in in Fig. 1. It is important not to use the marginal distributions and preserve the information available in the form of joint distribution throughout the solution of the NASA-LUQC.

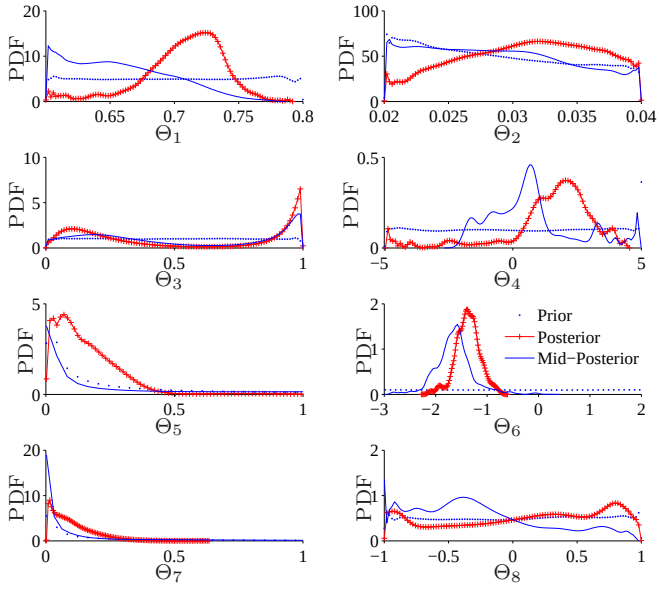

Figure 1: Subproblem A: Uncertainty Characterization

Now it is necessary to validate the estimated $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1)$ using the second set of 25 samples (D_2, as specified in the NASA-LUQC). This data set D_2 is also presented as a MATLAB file, as explained by Crespo et al.[11]

Two routes are proposed for validation: two-sample K-S test[19] and a normalized area metric.[20] It is inferred that the fraction of $\boldsymbol{\Theta}$ values for which the samples in D_2 correspond to the PDF $f_{X_1}(x_1|\boldsymbol{\theta}, D_1)$ is estimated to be equal to 0.6. The alternative normalized area metric is also computed as follows, and shown in Fig. 2.

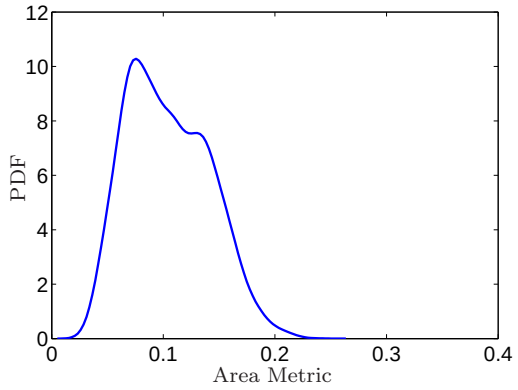

Figure 2: Normalized Area Metric

The uncertainty reflected in Fig. 2 is reflective of the uncertainty in $\boldsymbol{\Theta}$ and its correspondence with the observed data D_2. With either this result, or the result from two-sample KS-test, it is not possible to guarantee the validity of the uncertainty model given by $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1)$, as (1) only 60% of the values of $\boldsymbol{\Theta}$ suggest agreement of model with the data; and (2) the area metric is not very small. However, identifying tolerance levels for acceptance in these cases could be subjective and vary from application to application.

B. Problem A3

The third task requires further refinement the uncertainty model $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1)$ with the data contained in D_2. Instead of updating the mid-posterior, the prior is updated with all 50 data points, and the final posterior PDF $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1, D_2)$ is calculated. The corresponding marginal PDFs are indicated in Fig. 1. One important point to note here is that the entire joint density function denotes the refined uncertainty model; by converting this information to interval format (say, by calculating credible intervals) will result loss of information – information regarding marginal densities and more importantly, the dependence between the variables.

C. Problem A4

In this subproblem, it is required to account for the effect of number of observations on the fidelity of the resulting uncertainty models. This can be accomplished by quantifying the amount of improvement achieved in the solution of A3, compared to the solution of A1.The values of the individual and overall effects of the total aleatory and epistemic components are tabulated in Table 2, for the prior, mid-posterior, posterior distributions.

Table 2: Quantifying Extent of Refinement: Subproblem A4

Uncertainty Model	**Aleatory Uncertainty**		**Epistemic Uncertainty**	
	Individual Effect $S^{X_1}_{I,U_{1,3-5}}$	Overall Effect $S^{X_1}_{I,U_{1,3-5}}$	Individual Effect $S^{X_1}_{I,\Theta_{1-8}}$	Overall Effect $S^{X_1}_{I,\Theta_{1-8}}$
Prior ($f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$)	0.80	0.90	0.10	0.20
Mid-Posterior ($f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1)$)	0.91	0.96	0.04	0.09
Posterior ($f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1, D_2)$)	0.94	0.97	0.03	0.06

If there were no epistemic uncertainty, then the individual and overall contributions of aleatory uncertainty should be equal to unity.[17] To answer the question: "How much better is the model found in A3 as compared to the model found in A1?", it is simply necessary to look into the individual effects of aleatory

uncertainty in the case of these two uncertainty models. The value for model in A1 is already equal to 0.91, and it is necessary to increase the value by 0.09 to achieve perfect refinement whereas the model in A3 increases the value only by 0.03. In terms of percentage, this may be quantified as "the model in A3 has achieved 30% of the maximum refinement possible from the model in A1". (Similarly, the model in A1 has achived a little more than 50% of the maximum refinement possible from the prior.) Obviously, these refinements are measured in terms of the contributions of these quantities to the intermediate response variable X_1. Since these quantities may have significantly different contributions to other response variables, the extent of refinement will also need to be calculated with reference to each of those response variables and interpreted appropriately.

V. Subproblem B: Sensitivity Analysis

This subproblem deals with analyzing the sensitivity of $\boldsymbol{X}$ (to P_i, where $i = 1$ to 5), J_1 (to Θ_i where $i =$ to 31), and J_2 (to Θ_i where $i =$ to 31) in subproblems B1, B2, and B3 respectively.

The intermediate output quantity $\boldsymbol{X}$ follows a probability distribution for a given value of $\boldsymbol{\Theta}$. On the other hand, J_1 and J_2 are point-valued functions of $\boldsymbol{\Theta}$. For the purpose of solving Subproblem B, the posterior density function $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1, D_2)$ is considered for analysis, as mentioned in the NASA-LUQC.

A. Subproblem B1

The first goal in Subproblem B1 is to rank the 4 category II-III parameters (p_1, p_2, p_4, and p_5) according to degree of refinement in the p-box (equivalently, family of distributions, in the approach pursued here) of X_1 which one could hope to obtain by refining their uncertainty models. Using, the auxiliary variable-based framework and global sensitivity analysis, the results for sensitivity analysis of $\boldsymbol{X}$ with respect to $\boldsymbol{P}$ are provided in Tables 3 (for X_1), 4 (for X_2), 5 (for X_3), and 6 (for X_4). Note that the analysis for X_1 indicates the sensitivity of X_1 to P_4 by itself and P_5 by itself (during this computation, the dependency between P_4 and P_5 cannot be considered). This was done in order account for the possible correlation (represented by Θ_8) between P_4 and P_5.

Table 3: Sensitivity of X_1: Subproblem B1

Variable	Individual Effect	Overall Effect	Ranking
P_1	$S^{X_1}_{I,\Theta_{1-2}} = 2.4 \times 10^{-2}$	$S^{X_1}_{O,\Theta_{1-2}} = 5.6 \times 10^{-2}$	I
P_2	$S^{X_1}_{I,\Theta_3} = 8.0 \times 10^{-5}$	$S^{X_1}_{O,\Theta_3} = 2.3 \times 10^{-4}$	IV
P_4	$S^{X_1}_{I,\Theta_{4-5}} = 5.4 \times 10^{-4}$	$S^{X_1}_{O,\Theta_{4-5}} = 1.2 \times 10^{-3}$	III
P_5	$S^{X_1}_{I,\Theta_{6-7}} = 4.0 \times 10^{-3}$	$S^{X_1}_{O,\Theta_{6-7}} = 6.8 \times 10^{-3}$	II

Table 4: Sensitivity of X_2: Subproblem B1

Variable	Individual Effect	Overall Effect	Ranking
P_6	$S^{X_2}_{I,\Theta_9} = 8.2 \times 10^{-2}$	$S^{X_2}_{O,\Theta_9} = 9.0 \times 10^{-1}$	I
P_7	$S^{X_2}_{I,\Theta_{10-11}} = 1.9 \times 10^{-3}$	$S^{X_2}_{O,\Theta_{10-11}} = 1.5 \times 10^{-1}$	II
P_8	$S^{X_2}_{I,\Theta_{12-13}} = 5.8 \times 10^{-4}$	$S^{X_2}_{O,\Theta_{12-13}} = 4.5 \times 10^{-3}$	III
P_{10}	$S^{X_2}_{I,\Theta_{14-15}} = 1.2 \times 10^{-5}$	$S^{X_2}_{O,\Theta_{14-15}} = 1.0 \times 10^{-4}$	IV

B. Subproblems B2 and B3

From hereon, the focus of the NASA-LUQC shifts from focusing on the intermediate variables $\boldsymbol{x}$ to system-level response quantities $\boldsymbol{g}$ and the corresponding performance metrics J_1 and J_2. It was explained earlier in Section III that J_1 and J_2 would have been point valued has there been no epistemic uncertainty. For every

Table 5: Sensitivity of X_3: Subproblem B1

Variable	Individual Effect	Overall Effect	Ranking
P_{12}	$S^{X_3}_{I,\Theta_{16}} = 9.2 \times 10^{-1}$	$S^{X_3}_{O,\Theta_{16}} = 9.6 \times 10^{-1}$	I
P_{13}	$S^{X_3}_{I,\Theta_{17-18}} = 1.7 \times 10^{-5}$	$S^{X_3}_{O,\Theta_{17-18}} = 8.1 \times 10^{-5}$	IV
P_{14}	$S^{X_3}_{I,\Theta_{19-20}} = 1.7 \times 10^{-3}$	$S^{X_3}_{O,\Theta_{19-20}} = 1.1 \times 10^{-2}$	II
P_{15}	$S^{X_3}_{I,\Theta_{21-22}} = 6.3 \times 10^{-5}$	$S^{X_3}_{O,\Theta_{21-22}} = 2.0 \times 10^{-4}$	III

Table 6: Sensitivity of X_4: Subproblem B1

Variable	Individual Effect	Overall Effect	Ranking
P_{16}	$S^{X_4}_{I,\Theta_{23}} = 4.2 \times 10^{-1}$	$S^{X_4}_{O,\Theta_{23}} = 7.3 \times 10^{-1}$	I
P_{17}	$S^{X_4}_{I,\Theta_{24-25}} = 2.8 \times 10^{-3}$	$S^{X_4}_{O,\Theta_{24-25}} = 1.3 \times 10^{-2}$	III
P_{18}	$S^{X_4}_{I,\Theta_{26-27}} = 1.9 \times 10^{-2}$	$S^{X_4}_{O,\Theta_{26-27}} = 5.1 \times 10^{-2}$	II
P_{20}	$S^{X_4}_{I,\Theta_{28-29}} = 4.7 \times 10^{-4}$	$S^{X_4}_{O,\Theta_{28-29}} = 1.1 \times 10^{-3}$	IV

realization of the epistemic quantities ($\boldsymbol{\Theta}$), there exists a unique value of J_1 and J_2. Hence,

$$J_1 = J_1(\boldsymbol{\Theta}) \tag{12}$$

$$J_2 = J_2(\boldsymbol{\Theta}) \tag{13}$$

In order to evaluate the above functions, this paper uses a 17-dimensional Sobol sequence[21] of 500 quasi-random numbers (denoted by $\boldsymbol{U}$). These 17 dimensions correspond to the 17 aleatory components in the different Type-I and Type-III P_i variables (there is no aleatory component in Type-II variable). All the epistemic components in the different Type-II and Type-III variables are contained in $\boldsymbol{\Theta}$ above (there is no epistemic component in Type-I variables). Then for a fixed epistemic realization $\boldsymbol{\Theta}$, the probability distributions of $\boldsymbol{G} = \{G_i; i = 1 \text{ to } 8\}$ are computed by propagating the aforementioned 500 17-dimensional random samples, resulting in 500 8-dimensional samples of $\boldsymbol{g}$.

In order to analyze the sensitivity of J_1 and J_2 to $\boldsymbol{\Theta}$, this paper uses double loop Monte Carlo sampling to estimate quantities like $V(E(J_1|\Theta_1, \Theta_2))$. In improve computational efficiency, Gaussian Process (GP) surrogate models[22] are used to replace the calculations of J_1 and J_2 as functions of $\boldsymbol{\Theta}$.

The results of sensitivity analysis of J_1 and J_2 with respect to all of the epistemic components ($\boldsymbol{\Theta}$) are tabulated in Table 7. Note that the second column contains the vector of epistemic terms within each P_i; this vector is denoted as Θ_{fix} and sensitivity expressions in columns 2 and 3 rely on these fixing factors.

In order to rank the variables according to sensitivity, it is necessary to look at the individual effects in Table 7. In order study the error caused by fixing variables, it is necessary to look at the overall effects in Table 7. Note that, while one variable may have a large contribution to the uncertainty in J_1, it may have negligible contributions to the uncertainty in J_2, and vice-versa.

VI. Subproblem C: Uncertainty Propagation

The goal in this subproblem is to (1) compute the uncertainty in J_1 and J_2; (2) request for refined uncertainty models from Crespo et al.[11] for 4 of the 17 Type-II and Type-III variables; and (3) re-estimate the uncertainty in J_1 and J_2. Though the original NASA-LUQC only specifies to provide ranges on the values of J_1 and J_2 (this is meaningful because the quantities that lead to uncertainty in J_1 and J_2 are themselves specified only using intervals), the approached proposed in this paper leads to probability distributions for J_1 and J_2, that reflects the likelihood of occurrence for every value of J_1 and J_2. Since a sampling-based technique is used to calculate their distributions, the smallest value and the largest value are used to define the range of J_1 and J_2.

In order to identify the four parameters that will result in maximum reduction of uncertainty in J_1 and J_2, it is necessary to examine the individual effects and overall effects of the various quantities in Table 7. The first choice of the parameter is obvious from Table 7, since p_1 has the highest contribution to the

Table 7: Sensitivity Analysis of J_1 and J_2: No Refinement

Type II-III	Epistemic	J_1		J_2	
		Individual	Overall	Individual	Overall
Variable	Quantity (Θ_{fix})	$S^{J_1}_{I,\Theta_{fix}}$	$S^{J_1}_{O,\Theta_{fix}}$	$S^{J_2}_{I,\Theta_{fix}}$	$S^{J_2}_{O,\Theta_{fix}}$
P_1	Θ_1, Θ_2	4.2×10^{-1}	4.5×10^{-1}	3.5×10^{-1}	3.8×10^{-1}
P_2	Θ_3	1.0×10^{-2}	6.0×10^{-2}	9.5×10^{-4}	4.8×10^{-2}
P_4, P_5	Θ_4, Θ_5, Θ_6, Θ_7, Θ_8	5.7×10^{-2}	6.5×10^{-2}	1.7×10^{-2}	2.4×10^{-2}
P_6	Θ_9	4.5×10^{-3}	4.3×10^{-2}	9.3×10^{-3}	4.6×10^{-2}
P_7	Θ_{10}, Θ_{11}	2.5×10^{-3}	3.1×10^{-2}	4.3×10^{-4}	2.7×10^{-2}
P_8	Θ_{12}, Θ_{13}	1.6×10^{-3}	3.1×10^{-2}	3.7×10^{-3}	3.1×10^{-2}
P_{10}	Θ_{14}, Θ_{15}	4.6×10^{-4}	2.9×10^{-2}	5.1×10^{-3}	3.1×10^{-2}
P_{12}	Θ_{16}	1.2×10^{-3}	4.2×10^{-2}	2.9×10^{-1}	3.3×10^{-1}
P_{13}	Θ_{17}, Θ_{18}	2.7×10^{-3}	3.1×10^{-2}	4.0×10^{-3}	3.0×10^{-2}
P_{14}	Θ_{19}, Θ_{20}	3.3×10^{-4}	2.9×10^{-2}	2.9×10^{-3}	2.9×10^{-2}
P_{15}	Θ_{21}, Θ_{22}	3.9×10^{-3}	3.2×10^{-2}	8.8×10^{-3}	3.6×10^{-2}
P_{16}	Θ_{23}	4.4×10^{-4}	3.8×10^{-2}	1.9×10^{-3}	3.5×10^{-2}
P_{17}	Θ_{24}, Θ_{25}	5.2×10^{-3}	3.3×10^{-2}	1.3×10^{-2}	4.0×10^{-2}
P_{18}	Θ_{26}, Θ_{27}	1.2×10^{-3}	3.0×10^{-2}	2.6×10^{-3}	2.9×10^{-2}
P_{20}	Θ_{28}, Θ_{29}	2.1×10^{-3}	3.0×10^{-2}	1.1×10^{-3}	2.7×10^{-2}
P_{21}	Θ_{30}, Θ_{31}	9.9×10^{-2}	1.3×10^{-1}	1.1×10^{-2}	3.9×10^{-2}
P_1, P_2, P_3, P_4, P_5	Θ_1, Θ_2, Θ_3, Θ_4, Θ_5, Θ_6, Θ_7, Θ_8	5.1×10^{-1}	5.1×10^{-1}	3.3×10^{-1}	3.4×10^{-1}

uncertainty in both J_1 and J_2. Other than p_1, p_{12} has a large contribution to J_2 and p_{21} has a significant contribution to J_1. Hence, it is tempting to select these three parameters and a fourth parameter that has a significant contribution to either J_1 or J_2 or both. In fact, it is easy to that the choice of the fourth parameter is neither straightforward nor clear, because several parameters have comparable estimates of individual/overall effects. Further, one important issue is that, refining the uncertainty model of p_1 shrinks the family of distributions that p_1 represents, and this may significantly alter the probability distributions of the epistemic components of p_2, p_4, and p_5 (resulting from Bayesian inference) in order better predict the 50 samples of x_1 that were used for uncertainty characterization in Subproblem A. As a result, such altering of probability distributions causes a change in the sensitivities of not only the epistemic components of p_i ($i = 1$ to 5) but also the epistemic components of p_i ($i = 6$ to 21), because the sensitivity indices depend on the probability distributions for the epistemic components of all p_i ($i = 1$ to 21) as well. In order address these challenges in selecting four parameters, a sequential approach for model refinement is pursued in this paper.

In the first level of refinement, P_1 is refined. Bayesian updating is reperformed, and the sensitivities are computed. Refining the uncertainty model of P_1 would also alter all the distributions of Θ_i ($i = 1$ to 8), and in turn, further alter the sensitivity of J_1 and J_2 to $\mathbf{\Theta}$. An advantage of the sequential refinement procedure is that it provides the analyst the information to make decisions regarding refinement of uncertainty. If the goal is to simply reduce the uncertainty in J_1, then it is possible to make continuous choices to meet that goal. On the other hand, if the goal is to simply reduce the uncertainty in J_2, then it is possible to make continuous choices to meet that goal as well. In the NASA-LUQC, subjective choices need to made in order to reduce the uncertainty in both J_1 and J_2.

In the second level of refinement, both P_5 and P_{12} are considered for refinement. This is because P_5 affects J_1 and P_{12} affects J_2; more importantly, P_5 does not affect J_2 and P_{12} does not affect J_1 significantly. Hence, in the second refinement, refining the uncertainty in P_5 almost fully contributes to the reduction of uncertainty in J_1, while the refining the uncertainty in P_{12} almost fully contributes to the reduction of uncertainty in J_2. After obtaining the refined uncertainty models for P_5 and P_{12}, new priors are constructed for Θ_i ($i = 1$ to 8), and re-updated using Bayes' theorem. Once the posteriors of $\mathbf{\Theta}_8$ are obtained, sensitivity analysis of J_1 and J_2 are repeated.

Finally, for the third and final level of refinement, it is necessary to choose the 4^{th} variable. It is determined that refining P_2 would lead to reducing the uncertainty in J_1, while refining P_{14} would lead to reducing the uncertainty in J_2. Based on the amounts of refinement achieved in J_1 and J_2 so far (numerical

values are indicated in the next section), the fourth variable is chosen to be P_2. Once the refined model is obtained, new priors are constructed, Bayesian updating is reformed, and the uncertainty in J_1 and J_2 are recomputed.

The progressive refinement in J_1 and J_2 are shown in Fig. 3 (for J_1) and Fig. 4 (for J_2), and in Table 8. Note that the proposed approach only computes probability distributions; since a sampling-based approach is pursued in this paper, the smallest and largest values are indicated in Table 8.

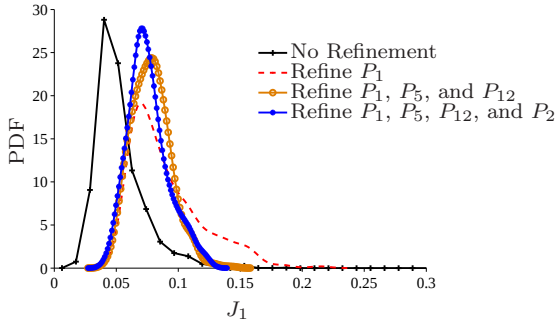

Figure 3: Reduction in Uncertainty of J_1

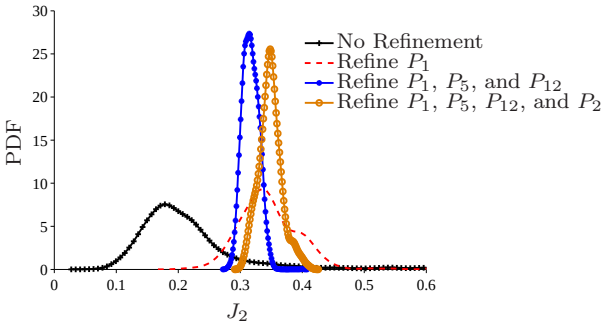

Figure 4: Reduction in Uncertainty of J_2

Table 8: Refinement of Uncertainty in J_1 and J_2

Description	J_1	J_2
No Refinement, Data	[0.02, 1.11]	[0.09, 0.70]
Refine P_1, Data	[0.05, 0.22]	[0.20, 0.60]
Refine P_1, P_5, P_{12}	[0.04, 0.15]	[0.28, 0.41]
Refine P_1, P_5, P_{12}, and P_2	[0.04, 0.15]	[0.30, 0.41]

The observed extents of refinement are also in accordance with the choices made while choosing the variables for refinement. For example, during the first two steps of refinement, both J_1 and J_2 were expected to be refined. However, the last choice P_2 was made knowing well that it may not have an impact on J_2.

VII. Subproblem D: Extreme Case Analysis

This subproblem focuses on identifying epistemic realizations that lead to the maximum and minimum values of J_1 and J_2, and studying failure scenarios. Since a sampling-based approach is pursued in this paper, the epistemic realizations that lead to extreme values of J_1 and J_2 are computed by simply examining the samples. The exact values are discussed as a part of a much more detailed journal manuscript by Sankararaman.[20]

In this subproblem, it is necessary to identify realizations of $\boldsymbol{X}$ that would lead to $J_2 > 0$, i.e., lead to failure. Failure is said to occur when the value of g_i is greater than zero for atleast one "i". In other words, failure is said to occur when any one of the outputs g_i ($i = 1$ to 8) is positive. In order to be able to generate

failure scenarios, it is necessary to find out realizations of $\boldsymbol{X}$ that would lead to at least one of the outputs being positive.

In order to accomplish this goal, consider the posterior PDF $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1, D_2)$ estimated earlier in Section B. Using this PDF, estimate the predictive distribution[9,17] of X_1, as:

$$f_{X_1}(x_1) = \int f_{X_1}(x_1|\boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}|D_1, D_2) d\boldsymbol{\theta} \tag{14}$$

Note that only Θ_{1-8} are used in the above equation to calculate the PDF of X_1, from which samples of X_1 can be generated. For Θ_{9-31}, uniform distributions are assumed on the interval provided by the NASA-LUQC, and the predictive distributions of X_2, X_3, X_4, and X_5 can also be computed. Since 5000 samples were generated using slice sampling in Section B, 5000 samples of each X_i ($i = 1$ to 5) are calculated. Then, these samples are used to compute 5000 corresponding realizations of g_i ($i = 1$ to 8); these results are used to study the behavior of the system.

Out of 5000 samples, 1067 samples resulted in system failure. Out of these 1067 samples, "$g_1 > 0$" in 326 samples, "$g_2 > 0$" in 9 samples, "$g_3 > 0$" in 176 samples, "$g_4 > 0$" in 1067 samples, "$g_5 > 0$" in 149 samples, "$g_6 > 0$" in 352 samples, "$g_7 > 0$" in 14 samples, and "$g_8 > 0$" in 2 samples. This suggests that g_4 is extremely critical while analyzing system reliability (union or intersection of a set of failure events – union, in the NASA-LUQC); even if g_4 is not the worst case metric in some samples and some other g_i ($i \neq 4$) is the worse case metric, g_4 still turns out to be positive whenever failure happens. In other words, a failure realization can be easily generated by simply focusing on generating positive values of g_4, because:

$$P(G_4 > 0|J_2 > 0) \approx 1 \tag{15}$$

The above equation was also verified by increasing the number of samples from 5000 to 10000, and the same behavior was observed. Therefore, it is possible to study the event $G_4 > 0$ in detail. Consider the function:

$$g_4 = f_4(\boldsymbol{x}) \tag{16}$$

The limit state multi-dimensional curve represented by the equation $g_4 = 0$ divides the event space into two: $g_4 > 0$ corresponding to failure and $g_4 < 0$ corresponding to safety. First-order reliability analysis[23] was used to calculate the value of $P(G_4 < 0) = 0.70$, which is reasonably approximate (Monte Carlo led to an estimate of 0.79) given that the first-order reliability method uses a linear approximation of f_4. Further, the Most Probable Point[23] is also calculated as $x_1 = 0.20$, $x_2 = 0.99$, $x_3 = 0.88$, $x_4 = 0.95$, and $x_5 = 0.02$. More importantly, the first-order reliability method also enables the calculation of local sensitivities in the 5-dimensional space, and the sensitivity to X_1, X_2, X_3, X_4, and X_5 are 0.21, 0.01, -0.65, 0, and 0. This calculation directly suggests that X_1 and X_3 are the most important quantities. This verifies the choice of the four refinement models in Section VI, i.e., the four models P_1, P_2, P_5, and P_{12} directly contribute to X_1 and X_3.

By analyzing the failure samples, it is possible to make several interesting observations. With respect to X_1, the likelihood of failure is higher only when the value of x_1 is large (around $x_1 = 0.4$). With respect to X_2, while in general X_2 has a higher likelihood of occurrence at higher values of x_2, failure occurs only at higher values of x_2. With respect to X_3, failure is spread throughout the range of x_3; thus, it is intuitive that, by refining the uncertainty in P_{12} that contributes to the uncertainty in X_3, it is significantly possible to reduce the range of failure probability. In the case of the variable X_4, the likelihood of failure increases at larger values of x_4. Finally, with respect to X_5, failure occurs at particularly lower values of x_5.

VIII. Conclusion

This paper presented a computational framework for uncertainty characterization, uncertainty propagation, and sensitivity analysis in engineering systems, and developed a novel, systematic approach using which important epistemic variables can be selected for refinement in order to reduce the uncertainty in the system-level performance. The proposed refinement approach selects candidates for refinement in a sequential manner, i.e., once the first candidate for refinement is selected, this candidate is refined, and all uncertainty analysis is repeated. Then, the new results are used to select the next candidate for refinement. This procedure is continued until all possible quantities are refined or until budget/time constraints are met. The proposed methodology is significantly different from existing approaches for refinement where all

candidates for refinement are simultaneously chosen at the end of the initial uncertainty analysis (all-at-once refinement). It has been proved[20] that the sequential refinement approach may not only result in a different set of refinement candidates but will also result in better reduction of uncertainty in the system-level performance, in comparison with all-at-once uncertainty refinement approaches. Finally, the proposed methodology using the NASA Langley Uncertainty Quantification Challenge problem. Future work may consider the extension of the proposed framework robust design optimization and other related activities.

Acknowledgments

The work reported herein was in part funded by the NASA System-wide Safety Assurance Technologies (SSAT) project under the Aviation Safety (AvSafe) Program of the Aeronautics Research Mission Directorate (ARMD).

References

[1]Kiureghian, A. D. and Ditlevsen, O., "Aleatory or epistemic? Does it matter?" *Structural Safety*, Vol. 31, No. 2, 2009, pp. 105–112.

[2]Oberkampf, W., Helton, J., Joslyn, C., Wojtkiewicz, S., and Ferson, S., "Challenge problems: Uncertainty in system response given uncertain parameters," *Reliability Engineering & System Safety*, Vol. 85, No. 1, 2004, pp. 11–19.

[3]Ferson, S. and Ginzburg, L., "Different methods are needed to propagate ignorance and variability," *Reliability Engineering & System Safety*, Vol. 54, No. 2, 1996, pp. 133–144.

[4]Shafer, G., *A mathematical theory of evidence*, Vol. 1, Princeton university press, 1976.

[5]Bae, H., Grandhi, R., and Canfield, R., "An approximation approach for uncertainty quantification using evidence theory," *Reliability Engineering & System Safety*, Vol. 86, No. 3, 2004, pp. 215–225.

[6]Ferson, S., Kreinovich, V., Hajagos, J., Oberkampf, W., and Ginzburg, L., "Experimental uncertainty estimation and statistics for data having interval uncertainty," Tech. rep., Sandia National Laboratories, Albuquerque, New Mexico, 2007.

[7]Montgomery, V., *New statistical methods in risk assessment by probability bounds*, Ph.D. thesis, Durham University, Durham, UK., 2009.

[8]Zaman, K., Rangavajhala, S., McDonald, M., and Mahadevan, S., "A probabilistic approach for representation of interval uncertainty," *Reliability Engineering & System Safety*, Vol. 96, No. 1, 2011, pp. 117–130.

[9]Sankararaman, S. and Mahadevan, S., "Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data," *Reliability Engineering & System Safety*, Vol. 96, No. 7, 2011, pp. 814–824.

[10]Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S., *Global sensitivity analysis: the primer*, John Wiley & Sons, 2008.

[11]Crespo, L. G., Kenny, S. P., and Giesy, D. P., "The NASA Langley Multidisciplinary Uncertainty Quantification Challenge," *SciTech, 16th AIAA Non-Deterministic Approaches Conference, Paper No. 2014 - 1347*, 2014.

[12]Chaudhuri, A., Waycaster, G., Matsumura, T., Price, N., and Haftka, R., "Framework for Quantification and Risk Analysis for Layered Uncertainty using Optimization: NASA UQ Challenge," *SciTech, 16th AIAA Non-Deterministic Approaches Conference, Paper No. 2014 - 1347*, 2014.

[13]Liang, C. and Mahadevan, S., "Bayesian Framework for Multidisciplinary Uncertainty Quantification and Optimization," *SciTech, 16th AIAA Non-Deterministic Approaches Conference, Paper No. 2014 - 1347*, 2014.

[14]McFarland, J. M., Bichon, B. J., and Riha, D. S., "A Probabilistic Treatment of Multiple Uncertainty Types: NASA UQ Challenge," *SciTech, 16th AIAA Non-Deterministic Approaches Conference, Paper No. 2014 - 1347*, 2014.

[15]Ghanem, R., Thimmisetty, C., Yadegaran, I., Keshavarzzadeh, V., Masri, S., Red-Horse, J., Moser, R., Oliver, T., Spanos, P., and Aldraihem, O., "A Probabilistic Approach to the NASA Langley Multidisciplinary Uncertainty Quantification Challenge Problem," *SciTech, 16th AIAA Non-Deterministic Approaches Conference, Paper No. 2014 - 1347*, 2014.

[16]Safta, C., Sargsyan, K., Najm, H., Chowdhary, K., Debusschere, B., Swiler, L., and Eldred, M., "Uncertainty Quantification Methods for Model Calibration, Validation, and Risk Analysis," *SciTech, 16th AIAA Non-Deterministic Approaches Conference, Paper No. 2014 - 1347*, 2014.

[17]Sankararaman, S. and Mahadevan, S., "Separating the contributions of variability and parameter uncertainty in probability distributions," *Reliability Engineering & System Safety*, Vol. 112, 2013, pp. 187–199.

[18]Neal, R. M., "Slice sampling," *Annals of statistics*, 2003, pp. 705–741.

[19]Massey Jr, F. J., "The Kolmogorov-Smirnov test for goodness of fit," *Journal of the American statistical Association*, Vol. 46, No. 253, 1951, pp. 68–78.

[20]Sankararaman, S., "Sequential Refinement of Uncertainty Through Bayesian Inference and Global Sensitivity Analysis," *Journal of Aerospace Information Systems*, Vol. 12, No. 1, 2015, pp. 49–72.

[21]Sobol, I. M., "On quasi-monte carlo integrations," *Mathematics and Computers in Simulation*, Vol. 47, No. 2, 1998, pp. 103–112.

[22]Rasmussen, C. E. and Williams, C. K., *Gaussian Process for Machine Learning*, MIT Press, 2006.

[23]Haldar, A. and Mahadevan, S., *Probability, reliability, and statistical methods in engineering design*, John Wiley & Sons, Inc., 2000.