AIAA SciTech Forum
13-17 January 2014, National Harbor, Maryland
16th AIAA Non-Deterministic Approaches Conference

10.2514/6.2014-0124

Uncertainty Quantification and Output Prediction in Multi-level Problems

Submitted for inclusion in the Special Session on Model Validation and Uncertainty Quantification

Chenzhao Li[1], and Sankaran Mahadevan[2]
Vanderbilt University, Dept. of Civil & Environmental Engineering, VU Station B #351831, Nashville, TN 37235

The calibration of model parameters is essential to predict the output of a complicated system, but the lack of data at the system level makes it impossible to conduct this quantification directly. This situation drives analysts to obtain information on model parameters using experimental data at lower levels of complexity which share the same model parameters with the system of interest. To solve this multi-level problem, this paper first conducts model calibration using lower level data and Bayesian inference to obtain the posterior distribution of each model parameter. However, lower level models are not perfect; thus model validation is also needed to evaluate the model that was used in model calibration. In the model validation, this paper extends the model reliability metric by using a stochastic representation of model reliability, and model with multivariate output is also considered. Another contribution of this paper is the consideration of physical relevance through sensitivity analysis, in order to measure the extent to which a lower level test represents the physical characteristics of the actual system of interest so that the calibration results can be extrapolated to the system level. Finally all the information from calibration, validation and relevance analysis is integrated to quantify the uncertainty in the system level prediction.

Nomenclature

$\boldsymbol{\theta}_m$	=	set of model parameters
$\boldsymbol{\theta}$	=	set of parameters to calibrate
θ	=	a single model parameter
$f(\cdot)$	=	probability density function
D_i^C, D_i^V	=	calibration and validation data at level *i*
$\boldsymbol{D}_R$	=	vector of model reliability values
$P(G_i)$	=	probability that the model at level *i* is correct
$P(G_i')$	=	probability that the model at level *i* is incorrect
$P(S_i)$	=	physical relevance between level *i* and system level
z	=	random variable of experimental observation
y	=	corrected model prediction
$\boldsymbol{x}$	=	model input
δ	=	model error
ε_m	=	output measurement error
ε_{su}	=	surrogate model error
F	=	computational model
σ_m	=	standard deviation of output measurement error
σ_{su}	=	standard deviation of surrogate model error
λ	=	tolerance in model reliability metric
K	=	number of output quantities
V_l, V_s	=	sensitivity vectors at lower level and system level

[1] Graduate student, Dept. of Civil & Environmental, Vanderbilt University, AIAA student member.
[2] Professor, Dept. of Civil & Environmental, Vanderbilt University, AIAA Associate Fellow.

I. Introduction

Model parameters of complicated systems are often hard to quantify due to lack of experimental data of the entire system; but generally it is possible to obtain data at lower levels, such as the subsystem or component level. This paper considers such a multi-level problem that has several lower levels and a system level; the lower levels and the system level share a set of model parameters ($\boldsymbol{\theta}_m$) to be calibrated, but the input and output of each level are not connected. Since the objective is to predict the system level output while data are only available at lower levels, a reasonable method is to quantify the uncertainty of model parameters using lower level data, and propagate the results through the computational model at the system level. Questions that need to be considered and solved include: 1) model calibration; 2) model validation; 3) relevance of the lower levels to the system level; 4) integration of all the available information for uncertainty quantification.

Model calibration focuses on the inference of model parameters so that the prediction by the computational model matches the experimental data; and the uncertainty quantification of error terms such as model error and output measurement error is also included in this process. Furthermore, in many cases the computational model is replaced by a surrogate model in order to promote computation efficiency, and this activity also brings surrogate model error. In a multi-level problem, when different levels share the same set of model parameters, the model calibration becomes more complicated. Model calibration can be conducted using the data at a single level, or using the data from different levels simultaneously. Generally, if data are available at *n* different levels, $2^n - 1$ model calibrations are possible to quantify the uncertainty of model parameters[1].

Model validation quantifies the extent to which the model prediction is supported by the experimental data. A model reliability metric has been proposed by Rebba and Mahadevan[2]: the model is treated as valid if the difference between model prediction and observed data is less than the predefined tolerance. Since stochastic error terms exist in both model prediction and data observation, the model reliability metric gives a reliability value of the current model. This paper proposes a stochastic model reliability metric in which the model reliability is treated as a stochastic variable since epistemic uncertainty also exists due to sparse validation data; and the basic concept of the model reliability metric is also extended in this paper to deal with multivariate output.

The purpose of uncertainty integration is to combine all the information from calibration, validation and other sources. A roll-up methodology has been proposed by Sankararaman[3] to integrate the information from model calibration and validation. For example, if the multi-level problem contains data and models at two lower levels, the integrated distribution of a model parameter θ is:

$$\begin{aligned} f(\theta|D_1^c, D_2^c, D_1^v, D_2^v) = {} & P(G_1)P(G_2)f(\theta|D_1^c, D_2^c) + P(G_1')P(G_2)f(\theta|D_2^c) \\ & + P(G_1)P(G_2')f(\theta|D_1^c) + P(G_1')P(G_2')f(\theta) \end{aligned} \tag{1}$$

Eq. (1) is expressed as the sum of four terms: in the first term the posterior distribution $f(\theta|D_1^c, D_2^c)$ using the calibration data of Level 1 and Level 2 has the weight of $P(G_1)P(G_2)$ which means that both of the models are correct; similarly, the second and third term contains the posterior distribution $f(\theta|D_i^c)(i = 1,2)$ using the calibration data at Level i and its weight means that the model at level i is correct but the model of another level is wrong; the last term contains the prior distribution $f(\theta)$ which has the weight of $P(G_1')P(G_2')$ meaning both of the models are wrong. Obviously, the weight of each PDF on the right side of Eq. (1) is purely decided by the model reliability, so an implicit assumption is that we have the same confidence in each of the lower level tests before model validation. However, if the physical structure at one level is more relevant to the system level than the other, it is reasonable to have more confidence in the calibration result at this level. Thus an additional consideration is the physical relevance between the lower levels and the system level, i.e. the extent to which the lower levels represent the physical characteristics of the system level so that the calibration results can be extrapolated to the system level. This paper introduces an index of physical relevance using global sensitivity analysis and includes it in the uncertainty integration.

II. Model Calibration

In model calibration, the Kennedy O'Hagan (KOH) framework[4] explicitly connects computational model prediction, model error terms, and experimental observation z as:

$$z = F(\boldsymbol{\theta}_m; \boldsymbol{x}) + \delta(\boldsymbol{x}) + \varepsilon_m \tag{2}$$

The KOH framework contains the quantification of two sources of error: model error δ and output measurement error ε_m. The model error is generally input dependent, and can be modeled as a Gaussian process. In this paper, the model input is assumed to be fixed so the model error δ is only a deterministic but unknown value. On the other hand, the output measurement error is generally unbiased and has a normal distribution of zero mean, i.e., $\varepsilon_m \sim N(0, \sigma_m)$. As a result, the parameters to be calibrated are model parameter $\boldsymbol{\theta}_m$, model error δ and the standard deviation σ_m of the output measurement error. In a multi-level problem where each lower level may give data on multiple output

quantities, an output quantity at any level corresponds to a model error δ and a standard deviation σ_m of output measurement error, which are calibrated if data of this quantity are used in model calibration.

Model calibration can be conducted by Bayesian inference, given available data $\boldsymbol{D}$ and prior distributions of parameters. The posterior distribution of $\boldsymbol{\theta}$ can be expressed as:

$$f''(\boldsymbol{\theta}) \propto L(\boldsymbol{\theta}) f'(\boldsymbol{\theta}) \tag{3}$$

where $L(\boldsymbol{\theta})$ is the likelihood function of $\boldsymbol{\theta}$ and $f'(\boldsymbol{\theta})$ is the joint prior PDF of $\boldsymbol{\theta}$. In addition, the Markov Chain Monte Carlo (MCMC) algorithm[5] generates samples of the posterior distribution without calculating the normalization constant, and the posterior distributions of the model parameters are constructed using these samples by some techniques such as the kernel density estimation[6].

By assuming all calibration parameters to be independent of each other, the joint prior distribution $f'(\boldsymbol{\theta})$ is simply the product of the marginal prior PDF of each calibration parameter. The prior distribution of model parameters $\boldsymbol{\theta}_m$ is defined based on experience or expert knowledge. For the error terms, since there is no more information on δ, a uniform prior distribution is assumed as $\delta \sim U(a, b)$. Based on the Fisher information criterion, the prior distribution of σ_m can be chosen as $f'(\sigma_m) \propto 1/\sigma_m$, which is invariant under reparameterization and of special interest for use with scale parameter, such as σ_m here[7].

Another key point in implementing Eq. (3) is the establishment of the likelihood function $L(\boldsymbol{\theta})$, which is proportional to $P(\boldsymbol{D}|\boldsymbol{\theta})$ [8], i.e., the probability of observing data set $\boldsymbol{D}$ given the parameter $\boldsymbol{\theta}$. By assuming independence among the observed data points, for the multi-level problem in which data are observed at different lower levels and each lower level may provide data from multivariate output quantities, $L(\boldsymbol{\theta})$ will be the product of the likelihood functions corresponding to each individual data point.

Furthermore, in many cases the computational model $F(\boldsymbol{\theta}_m; \boldsymbol{x})$ will be replaced by a surrogate model in order to promote computation efficiency, and this also brings a surrogate model error ε_{su}. This paper applies Gaussian process regression, which gives a stochastic prediction with mean value y_{GP} and normally distributed surrogate error $\varepsilon_{su} \sim N(0, \sigma_{su}^2)$.

III. Model Validation

A. Model Reliability Metric

Assume that a data point D is observed. Due to the measurement error ($\varepsilon_m \sim N(0, \sigma_m)$), the measurement is actually a random variable. This random variable is denoted by d, with mean value D and standard deviation σ_m, i.e. $d \sim N(D, \sigma_m)$. The model reliability metric defines that the model is correct if the difference between the model prediction y and the measurement observation d is less than a predefined tolerance λ. Let G denote the event that the model is correct, then the model reliability is the probability of event G:

$$P(G|D) = P(-\lambda \le y - d \le \lambda) \tag{4}$$

Both y and σ_m can be stochastic. In this paper, the KOH framework is used in the calibration so the model prediction y refers to the computational model output corrected by the model error, i.e. $y = F(\boldsymbol{\theta}_m; \boldsymbol{x}) + \delta$. Although model input $\boldsymbol{x}$ is known, the model prediction y is still stochastic due to the uncertainty of δ and $\boldsymbol{\theta}_m$. Furthermore, another calibration parameter σ_m is also uncertain. In this case, the model reliability is computed as:

$$P(G|D) = \int P(G|D, \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{5}$$

in which the term $P(G|D, \boldsymbol{\theta})$ is the model reliability conditioned on data point, model parameters, and error terms.

If multiple data points $\boldsymbol{D} = \{D_1, D_2, \dots, D_n\}$ of one output quantity are available in model validation, the model reliability metric defines the model reliability conditioned on $\boldsymbol{D}$ as the probability that all the observations support the model:

$$P(G|\boldsymbol{D}) = \int P(G|\boldsymbol{D}, \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} = \int \left(\prod\nolimits_{i=1}^{n} P(G|D_i, \boldsymbol{\theta}) \right) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{6}$$

However, Eq. (6) will cause model reliability to decline due to the multiplication term $\prod_{i=1}^{n} P(G|D_i, \boldsymbol{\theta})$. This contradicts our intuition, since model reliability should be stable for a particular model and independent of the number of validation data. In addition, if the data from multivariate output quantities are available, the difference between model prediction and the observed data is a vector, therefore the comparison of the prediction and observation has to be extended to from a scalar in Eq. (4) to a vector. To solve these two challenges, the following two sections propose a stochastic model reliability metric and extend the model reliability metric to multivariate output by using the Mahalanobis distance.

B. Stochastic Model Reliability Metric

Recall the calibration of model parameters and error terms. All the parameters to be calibrated ($\boldsymbol{\theta} = \{\boldsymbol{\theta}_m, \delta, \sigma_m\}$) have deterministic but unknown values, so their uncertainty arises only from lack of knowledge and data, which is referred to as epistemic uncertainty. Experimental data are used in model calibration to estimate these parameters, and the epistemic uncertainty declines as the number of calibration data increases. This also happens to model reliability. The reliability of a model is a deterministic but unknown value. Assume a set of data $\boldsymbol{D} = \{D_1, D_2, \dots, D_n\}$ is available for model validation, by applying Eq. (5) each data point gives a value of model reliability, and all these model reliability values can be used to construct a distribution of model reliability, which represents the contribution of epistemic uncertainty. In sum, this approach gives a stochastic representation of model reliability.

In this paper, the model reliability is assumed to have a beta distribution since its value lies on the interval [0, 1]. Once the model reliability values $\boldsymbol{D}_R = \{D_{R1}, D_{R2}, \dots, D_{Rn}\}$ given by each data point have been computed, many methodologies are available in the literature to construct the PDF of model reliability, such as maximum likelihood method, method of moments[9], or Bayesian updating. This paper makes use of the method of moments to construct the PDF of model reliability.

C. Model Reliability Metric with Multivariate Output

Data observations of multivariate output indicate that in an experiment more than one quantity of output can be observed. Under this circumstance, we need to extend the theory of model reliability metric to problems with multivariate output.

Eq. (4) shows the model reliability metric for univariate output, and it can be rewritten as:

$$P(G|D) = P\left(\left(\sqrt{\frac{(z-D)^2}{\sigma_z^2}} < \frac{\delta}{\sigma_z}\right)\right) \tag{7}$$

where σ_z^2 is the variance of z. Eq. (7) indicates that the model reliability is the probability that the Mahalanobis distance between the variable z representing the observation and the data point D is less than $\delta\sigma_z^{-1}$. This new definition for the model reliability metric can be also used for the case of multivariate output. If K output quantities are observed in a validation experiment, we build a model group including K models sharing the same input and model parameters:

$$\boldsymbol{z} = \begin{Bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{Bmatrix} = \begin{Bmatrix} F_1(\boldsymbol{\theta}_m; \boldsymbol{x}) + \delta_1 + N(0,\ \sigma_{m_1}) \\ F_2(\boldsymbol{\theta}_m; \boldsymbol{x}) + \delta_2 + N(0,\ \sigma_{m_2}) \\ \vdots \\ F_K(\boldsymbol{\theta}_m; \boldsymbol{x}) + \delta_K + N(0,\ \sigma_{m_K}) \end{Bmatrix} \tag{8}$$

Assume n experiments are conducted. In the i^{th} experiment, data points for each of the K quantities in the model group are observed so we obtain a data set $\boldsymbol{D_i} = \{D_{i1}, D_{i2}, .., D_{iK}\}$ from this experiment. Consistent with the case of univariate output as shown in Eq. (7), we define the model as correct if the Mahalanobis distance between $\boldsymbol{z}$ and $\boldsymbol{D_i}$ is less than $\sqrt{\boldsymbol{\lambda}^T\boldsymbol{\Sigma_z^{-1}}\boldsymbol{\lambda}}$:

$$P(\boldsymbol{G}|D) = P\left(\sqrt{(\boldsymbol{z}-\boldsymbol{D_i})^T\boldsymbol{\Sigma_z^{-1}}(\boldsymbol{z}-\boldsymbol{D_i})} < \sqrt{\boldsymbol{\lambda}^T\boldsymbol{\Sigma_z^{-1}}\boldsymbol{\lambda}}\right) \tag{9}$$

In Eq. (9), $\boldsymbol{\Sigma_z}$ is the covariance matrix of $\boldsymbol{z}$; $\boldsymbol{\delta}$ is a vector of predefined tolerances for each quantity so that $\boldsymbol{\delta} = \{\delta_1, \dots, \delta_K\}^T$. By using the Mahalanobis distance, Eq. (9) shows two advantages: 1) the correlation between each output quantity is considered; and 2) the original data from different quantities are transformed to a comparable scale.

D. Summary

Section III proposes a stochastic treatment of the model reliability metric and extends the model reliability metric to deal with multivariate outputs. These enhancements of the model reliability metric can be summarized into three steps, in order to compute the model reliability of each lower level in a multi-level problem, where data of multivariate output quantities are observed in every experiment:

1. Implement the model reliability metric of multivariate outputs to obtain the model reliability value given by one experiment;
2. Repeat Step 1 for each experiment to obtain all the model reliability values;
3. Construct the distribution of the model metric.

IV. Uncertainty Integration

A. Physical Relevance

As explained in Section 1, information of uncertainty quantification obtained from the levels with closer physical relevance with the system level deserves higher weights in the prediction of system output. This section proposes a method to quantify the physical relevance using global sensitivity analysis.

For a model $Y = G(\boldsymbol{X})$ where Y is the output and $\boldsymbol{X} = \{X_1, X_2, \dots, X_N\}$ is a vector containing all the model inputs or parameters, the sensitivity analysis apportions the uncertainty of Y into the uncertainty of $\boldsymbol{X}$. Two global sensitivity indices[10] have been developed in the literature based on the variance decomposition theorem: the first-order index and the total effects index. Both indices reflect the contribution of each input in the uncertainty of the output; and the higher the sensitivity index of X^i is, the more important it is. In this paper, the total effects index is used to develop a method to quantify the physical relevance; thus where sensitivity index is mentioned in the following sections, it implies the total effects index.

Without loss of generality, this paper takes a multi-level problem with two lower levels and a system level for the illustration of physical relevance. Assume that the ultimate objective is to predict the output of quantity y_s at the system level, and the same quantity is also measured at lower levels. If the model input at each level is fixed, the three computational models for this quantity at different levels are expressed as $y_{L_1} = F_{L_1}(\boldsymbol{\theta}_m), y_{L_2} = F_{L_2}(\boldsymbol{\theta}_m), y_s = F_s(\boldsymbol{\theta}_m)$ where $\boldsymbol{\theta}_m$ is a vector containing N model parameters. The sensitivity analysis of $\boldsymbol{\theta}_m$ at each lever will give an N-dimensional sensitivity index vector, which contains the physical structural information at each level. So the comparison of these vectors will reflect the relevance between a lower level and the system level. This paper defines the physical relevance as the square of cosine similarity[11] of the sensitivity index vectors, which is the standardized dot product of the sensitivity vectors obtained at each level by the sensitivity analysis for the output quantity to be predicted. Thus if the sensitivity vectors at a lower level and the system level are V_l and V_s, the relevance between them is defined as:

$$R = \left(\frac{V_l \cdot V_s}{|V_l||V_s|}\right)^2 \tag{10}$$

In other words, the relevance is the square of the cosine value of the angle α between two sensitivity index vectors. If the angle between the two sensitivity index vectors is zero, the relevance will be 1, which means the physical structures at the two levels are the same and the lower level gives a 100% confidence before model validation from a viewpoint of uncertainty propagation; if the two vectors are perpendicular, the physical structures are not relevant at all and the confidence is zero.

A further question arises in the application of this physical relevance definition using sensitivity analysis. The global sensitivity analysis considers the entire distribution of the variable, so we need to know the distributions of $\boldsymbol{\theta}_m$, which are unknown. In fact, if we know the actual distributions, there is no need to conduct the activities discussed in the previous and current sections. In order to solve this problem, a simple iterative algorithm to compute the physical relevance is proposed:

1. Set initial values of R between a lower level and the system level;
2. Obtain the integrated distribution of each model parameter using the current physical relevance and the proposed roll-up formula in the next section;
3. Use the integrated distributions from step 2 in the sensitivity analysis and compute a updated physical relevance;
4. Repeat step 2 and step 3 until the physical relevance index converges.

So far, all possible sources of information including the posterior parameter distributions from model calibration, distribution of model reliability from model validation, and physical relevance between the lower and the system level using sensitivity analysis have been obtained. The next task is to construct the integrated distribution of each parameter and predict the system output.

B. Roll-up methodology

The distributions of model parameters need to be constructed before propagating them through the computational model at the system level to predict the system output. Eq. (1) expresses a roll-up methodology to integrate the results from model calibration and model validation; and this paper modifies this methodology to include two additional concepts:

1) Stochastic model reliability: the event that the model at level i is correct is denoted by G_i, whose probability is $P(G_i)$;
2) Physical relevance: defined as the square of the cosine value, thus the corresponding non-relevance is the square of the sine value; since their sum is the unity, we can treat the physical relevance similar to probability or a weight in the roll-up equation. This paper denotes the event that level i is relevant to the system level by

$S_i(i = 1, 2)$, and its probability $P(S_i)$ is the value of physical relevance; and S_i'denotes the event of non-relevance.

Take the multi-level problem with two lower levels as an example. The integrated distribution of a model parameter θ conditioned on the calibration and validation data is:

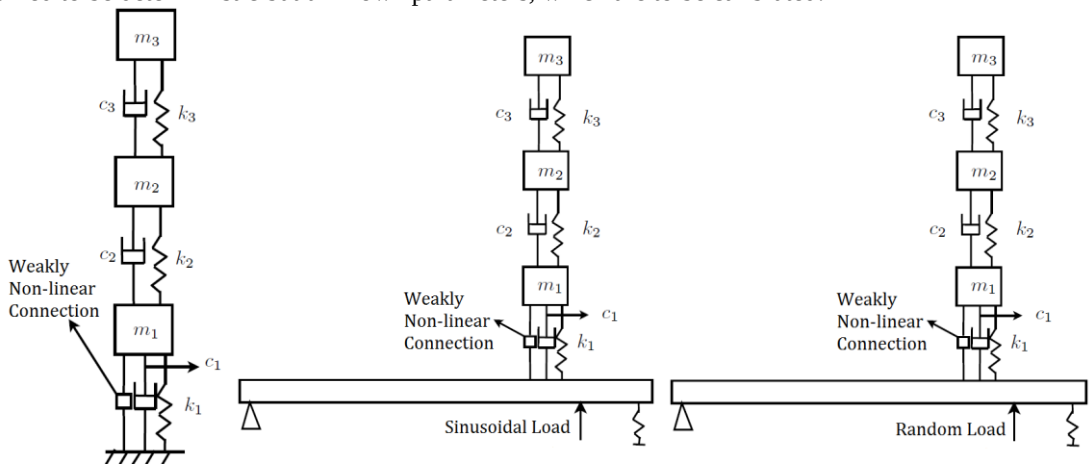
$$f(\theta|D_1^C, D_2^C, D_1^V, D_2^V) = P(G_1G_2S_1S_2)f(\theta|D_1^C, D_2^C) + P\big(G_1S_1 \cap (G_2' \cup S_2')\big)f(\theta|D_1^C) + P\big(G_2S_2 \cap (G_1' \cup S_1')\big)f(\theta|D_2^C) + P((G_1' \cup S_1') \cap (G_2' \cup S_2'))f(\theta) \quad (11)$$

From the view of generating samples, Eq. (11) indicates two criteria: 1) whether a level is relevant to the system level; 2) whether a level has the correct model. A sample of θ is generated from $f(\theta|D_1^C, D_2^C)$ only when both levels satisfy both criteria; a sample of θ is generated from $f(\theta|D_i^C)$ if level i satisfies both criteria but the other level does not; and a sample of θ is generated from the prior $f(\theta)$ if no level achieves both criteria.

After obtaining the integrated distributions of the system model parameters, the final step is to propagate these unconditional distributions through the computational model at the system level to predict the system level output. This can be done simply by Monte Carlo sampling. Due to the uncertainty of model parameters, the predicted quantity of system performance will also be stochastic.

V. Numerical Example

A multi-level structural dynamics problem provided by Sandia National Laboratories[12–14] is used to illustrate the proposed methodology in this paper. As shown in Figure 1, Level 1 contains three mass-spring-damper dynamic components in series, and a sinusoidal force input is applied on m_1. At Level 2, the dynamic system is mounted on a beam supported by a pin at one end and by a spring at the other end; a sinusoidal force input is applied on the beam. The configuration of the system level is the same as Level 2, but the input is a random process loading. All levels share six model parameters: three spring stiffnesses $k_i(i = 1, 2, 3)$ and three damping ratios $\zeta_i(i = 1, 2, 3)$; and they are assumed to be deterministic but unknown parameters, which are to be calibrated.



(a) Level 1 (b) Level 2 (c) System Level

Figure 1. Dynamics challenge problem

Experimental data of six quantities at each lower level are available for model calibration and validation:

1) $A_i(i = 1 \text{ to } 3)$: the maximum acceleration in the *i*-th mass;
2) $D_i(i = 1 \text{ to } 3)$: the energy dissipated by the *i*-th damper in a time interval of 1000 seconds.

The objective in this numerical example is to predict the maximum acceleration at m_3 at the system level, which can be achieved by using the experiment data and the methodologies explained in this paper.

Since as many as six quantities are measured, we can choose any combination of these six quantities in the analysis. As more quantities are employed, uncertainty in the system output prediction will drop, but the computational effort will also increase since each quantity will bring two more error terms to calibrate. For the sake of brevity, only the calibration and validation results using the data of all six quantities are provided below. We also provide a plot showing the reduction in the uncertainty of system output prediction as more quantities are employed.

The calibration results of k_i and ζ_i using the calibration data of the six output quantities are shown in Figure 2. As more data are used in the calibration, the uncertainty of the model parameters will decline, so Figure 2 shows that the posterior distributions using the data of both levels always have less uncertainty than that using data of single level. The difference between the posterior distributions in each sub-figure also indicates that the posterior distribution is a best-fitting results in the sense of representing the data, and we cannot consider it as the true value.

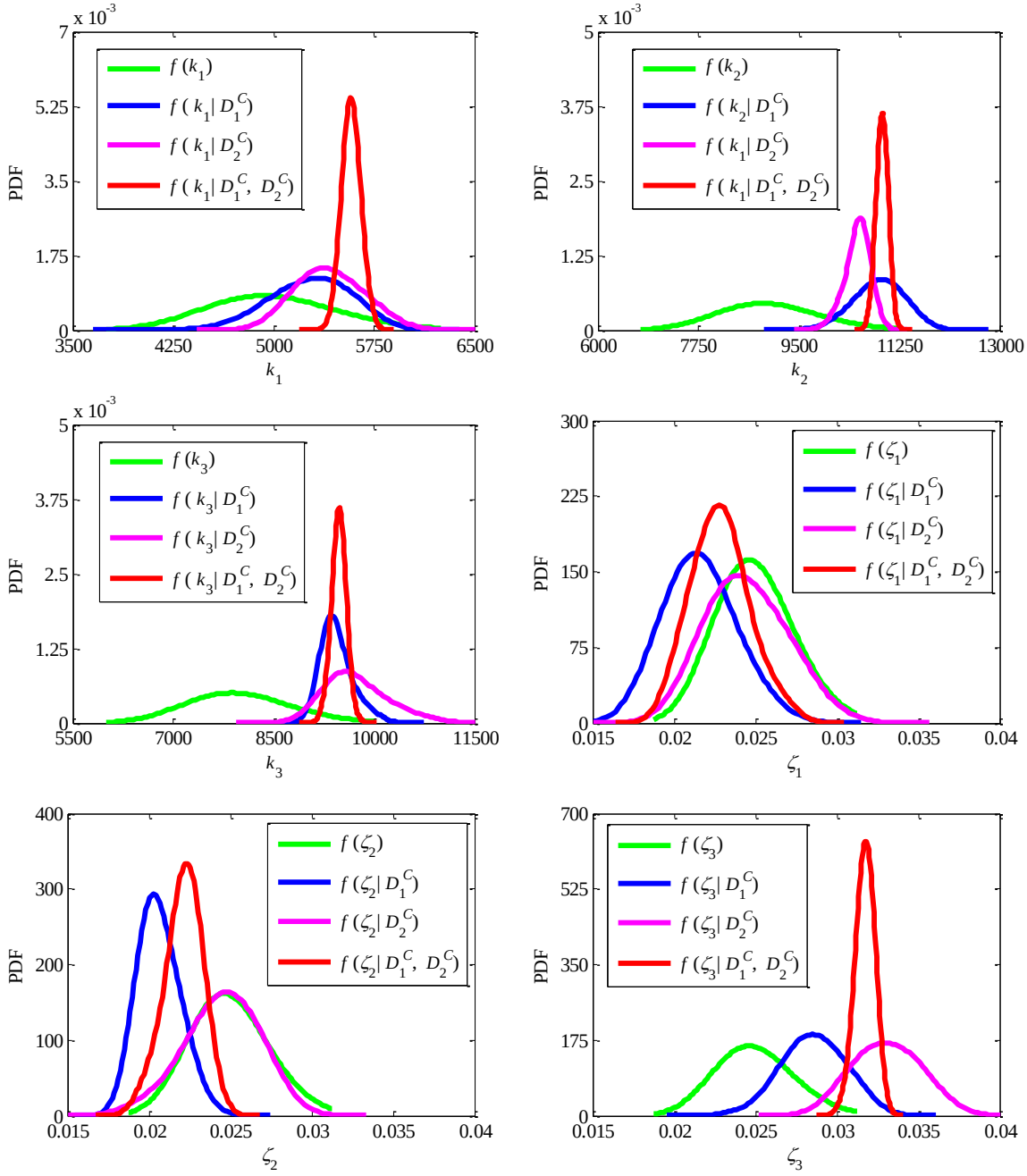


Figure 2. Posterior distributions of model parameters

The next step is model validation using the stochastic model reliability metric with multivariate output. The tolerance for each quantity is 15% of validation data. The model reliability values given by the validation data are listed in Table 1, and they are used to construct the distribution of model reliability at Level 1 and Level 2 shown in Figure 3.

Table 1. Model reliability values

Level 1	0.9702	0.9580	0.9398	0.9828	0.9800
Level 2	0.9616	0.8564	0.9208	0.9796	0.7904

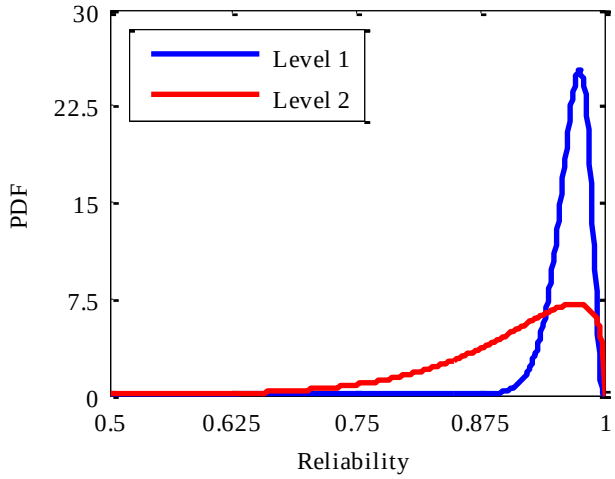


Figure 3. Model reliability

Next the physical relevance between the lower levels and the system level is computed. The initial values of physical relevance are set as 1. For Level 1, the algorithm for computing physical relevance converges after three iterations, and after five iterations for Level 2. The results are: $P(S_1) = 0.5790, P(S_2) = 0.8979$.

Based on all the information from calibration, validation and physical relevance, the integrated distributions of all six model parameters are computed using Eq. (11). And the system output is predicted by propagating the integrated distributions of model parameters through the computational model of the system level. Figure 4 gives not only the prediction using the data of all six quantities but also the prediction by other combinations of quantities whose names are shown in the legend. As expected, the uncertainty of system output prediction declines as more quantities are employed.

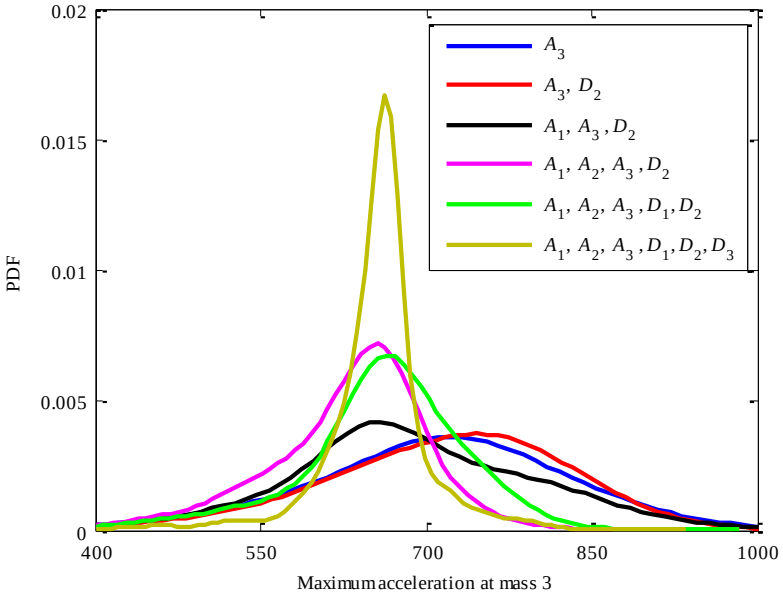


Figure 4. System output prediction

Conclusion

This paper developed a methodology to predict the system level output in a multi-level problem. Since data is unavailable at the system level, the challenge is to obtain and integrate the information on the uncertainty of model parameters from different activities at lower levels of complexity. The main contributions of this paper are the new methodologies for model validation and uncertainty integration, and the definition of physical relevance.

For model validation, the proposed definition of a stochastic model reliability metric solves the problem that model reliability declines as more validation data are considered in the current model reliability metric. This stochastic model reliability metric results in a PDF of model reliability. Then this paper also extends the concept of model reliability metric to deal with data of multivariate output.

The definition of physical relevance in this paper is an important step in uncertainty integration. The physical relevance quantifies whether the lower level represents the structural characteristics of the system level, and determines the weight of each posterior distribution in the uncertainty integration. In the proposed method, the relevance is computed using global sensitivity analysis, and defined as the square of the cosine of the angle between two sensitivity index vectors.

The last contribution of this paper lies in the development of the roll-up formula (Eq. (11)) to integrate the information from three sources: 1) posterior distributions of model parameters by model calibration; 2) stochastic model reliability by model validation; 3) and physical relevance of lower levels to the system level.

In summary, model calibration obtains posterior distributions of each parameter within and across different lower levels; model validation evaluates the model reliability at each lower level separately; the physical relevance analysis reveals the similarity between a lower level and the system level; all the above activities provide information for uncertainty integration which results in the integrated distribution of each model parameter; finally the system level output distribution is predicted by propagating the integrated distributions through the computational model of the system level.

Acknowledgments

The research in this paper is partially supported by funds from Sandia National Laboratories through contract no. BG-7732 (Technical Monitor: Dr. Angel Urbina). This support is gratefully acknowledged. The authors also thank Shankar Sankararaman, Joshua Mullins and You Ling for valuable discussions.

References

[1]Mullins, J., Li, C., Sankararaman, S., and Mahadevan, S., "Probabilistic Integration of Validation and Calibration Results for Prediction Level Uncertainty Quantification: Application to Structural Dynamics," *54th AIAA/ASME/ASCE/AHS/ ASC Structures, Structural Dynamics, and Materials Conference*, 2013.

[2]Rebba, R., and Mahadevan, S., "Computational methods for model reliability assessment," *Reliability Engineering & System Safety*, vol. 93, Aug. 2008, pp. 1197–1207.

[3]Sankararaman, S., "Uncertainty Quantification and Integration in Engineering Systems," Vanderbilt University, 2012.

[4]Kennedy, M., and O'Hagan, A., "Bayesian calibration of computer models," *Journal of the Royal Statistical Society*, vol. 63, 2001, pp. 425–464.

[5]Mcfarland, J. M., "Uncertainty Analysis for Computer Simulations throuth Validation and Calibraion," 2008.

[6]Rosenblatt, M., "Remarks on some nonparametric estimates of a density function," *The Annals of Mathematical Statistics*, 1956, pp. 832–837.

[7]Jeffreys, H., "An Invariant Form for the Prior Probability in Estimation Problems," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 186, Sep. 1946, pp. 453–461.

[8]Edwards, A. W. F., *Likelihood*, Cambridge University Press, 1972.

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

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

[11]Tan, P.-N., Steinbach, M., and Vipin Kumar, *Introduction to Data Mining*, Addison-Wesley, 2006.

[12]McFarland, J., and Mahadevan, S., "Error and variability characterization in structural dynamics modeling," *Computer Methods in Applied Mechanics and Engineering*, vol. 197, May. 2008, pp. 2621–2631.

[13]Red-Horse, J. R., and Paez, T. L., "Sandia National Laboratories Validation Workshop: Structural dynamics application," *Computer Methods in Applied Mechanics and Engineering*, vol. 197, May. 2008, pp. 2578–2584.

[14]Paez, T. L., and Red-Horse, J., "Structural Dynamics Challenge Problem: Summary," *Computer Methods in Applied Mechanics and Engineering*, vol. 197, May. 2008, pp. 2660–2665.