Chapter 9
Uncertainty Quantification and Integration in Multi-level Problems

Chenzhao Li and Sankaran Mahadevan

Abstract Calibration of model parameters is an essential step to in predicting the output of a complicated system, but the lack of experimental data at the system level makes it impossible to conduct this quantification directly. This situation drives the analyst to obtain information on model parameters from test data at lower levels of complexity which share the same model parameters with the system of interest. However, the validity and relevance of lower level models and experiments for the calibration of system model parameters need to be examined. This chapter integrates calibration and validation for models with multivariate outputs, using a stochastic representation of model reliability. The investigation of physical relevance is done through sensitivity analysis, thus measuring 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.

Keywords Calibration • Validation • Uncertainty • Bayesian • Model reliability

9.1 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 of complexity, such as the subsystem or component level. In this multi-level problem, the lower levels and the system share a set of model parameters (θ_m), but the input and output of each level are not connected. Since the objective is to predict the system level output, a reasonable route is to quantify the uncertainty of θ_m using lower level data, and propagate the results to system output prediction. In such situation, several questions need to be considered and solved [1]: (1) model calibration; (2) model validation; (3) integration of information from calibration and validation to quantify the uncertainty in model parameters.

Model calibration focuses on the inference of model parameters so that the prediction by the computational model matches the experimental data [2]. The computational model $y_c = F(\theta_m;x)$ describes the physical configuration at a single level, where y_c is the output of the computational model, and θ_m is a set of unknown model parameters, and $\boldsymbol{x}$ is the model input. The difference between $\boldsymbol{y_c}$ and the true value is defined as model error or model discrepancy, which is generally input dependent and denoted as $\delta(\boldsymbol{x})$. When measuring experimental data, the measurement error ε_m with a zero mean normal distribution $N(0,\sigma^2)$ is also added into the observed value of the output. Kennedy and O'Hagan [2] expressed the relationship between the experimental data z, computational model and error terms as $z = F(\theta_m;x) + \delta(x) + \varepsilon_m$. This chapter applies Bayesian inference within this framework to estimate the posterior distributions of model parameters. Model calibration is more complicated in a multi-level problem since it can be conducted using data at a single level, or using data from multiple levels simultaneously. Figure 9.1 shows a multi-level problem of two lower levels (level 1 and level 2), so three calibrations are possible: (1) calibration using the data and model at level 1 alone; (2) calibration using the data and model at level 2 alone; (3) calibration using the data and models at both level 1 and 2.

Model validation quantifies the degree to which the model supported by the experimental data [3], and it provides information that is necessary to understand the limitation of calibration [3]. Many quantitative metrics for model validation

C. Li • S. Mahadevan (✉)
Department of Civil & Environmental Engineering, Vanderbilt University, VU Station B #351831, Nashville, TN 37235-1831, USA
e-mail: chenzhao.li@vanderbilt.edu; sankaran.mahadevan@vanderbilt.edu

H.S. Atamturktur et al. (eds.), *Model Validation and Uncertainty Quantification, Volume 3: Proceedings of the 32nd IMAC, A Conference and Exposition on Structural Dynamics, 2014*, Conference Proceedings of the Society for Experimental Mechanics Series, DOI 10.1007/978-3-319-04552-8_9,

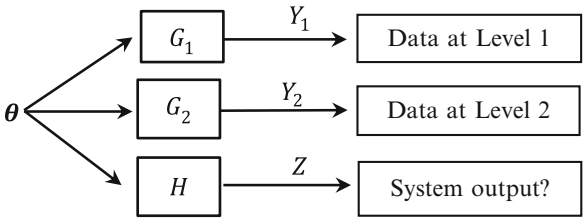

Fig. 9.1 Multi-level problem

have been developed in the literature, such as the classic hypothesis testing [4], the Bayesian hypothesis testing [5], and the area metric [6]. This chapter applies the model reliability metric proposed by Rebba and Mahadevan [7] and further developed by Sankararaman and Mahadevan [8]. The model reliability metric gives the probability that the difference between model prediction and observed data is less than a predefined tolerance. The previous literature on model reliability [7, 8] metric computes a deterministic model reliability value, but this chapter proposes a stochastic model reliability metric such that the resultant distribution of model reliability captures the epistemic uncertainty due to sparse validation data. In addition, validation metric dealing with multivariate output is necessary for a multi-level problem but not well developed in the literature, so this chapter will also extend the basic concept of the model reliability metric to this case.

The objective of uncertainty integration is to aggregate all the available information from calibration and validation activities to quantify the uncertainty of model parameters and output. An integration (roll-up) methodology was proposed in [9] to quantity the uncertainty in model parameters. For example, if the multi-level problem contains two lower levels, the integrated distribution of a system model parameter θ is:

$$f\left(\theta\middle|D_1^{C,V},D_2^{C,V}\right)=$$

$$P\left(G_1\right)P\left(G_2\right)f\left(\theta\middle|D_1^{C},D_2^{C}\right)+P\left(G_1'\right)P\left(G_2\right)f\left(\theta\middle|D_2^{C}\right)+P\left(G_1\right)P\left(G_2'\right)f\left(\theta\middle|D_1^{C}\right)+P\left(G_1'\right)P\left(G_2'\right)f\left(\theta\right) \quad (9.1)$$

Equation (9.1) 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 both levels 1 and 2 has the weight of $P(G_1)P(G_2)$ which is the probability that both of the models are correct; in the second and third terms, $f(\theta|D_i^C)$ is the posterior distribution using the calibration data at level i alone and its weight is the probability that the model at level i is correct but the model at another level is wrong; and the last term contains the prior distribution $f(\theta)$ with the weight $P(G_1^{'})P(G_2^{'})$ which is the probability that both of the models are wrong. It is assumed that the validity of the two models are independent.

Obviously, the weight of each PDF on the right hand side of (9.1) is purely decided by model validation, so an implicit assumption is that we have the same confidence in the calibration results of each lower level before the validation analysis. However, recall that the only reason that we calibrate the model parameters by the data at lower levels is that we have no data available at the system level. If we have enough data at the system level, we would directly calibrate model parameters at the system level. In other words, we hold a 100 % confidence in the calibration results at the system level, thus we can neglect the information at lower levels. Furthermore, an inherent assumption in this 100 % confidence is that the data at the system level reflect its physical characteristics completely since the physical structure used to measure the data and the physical structure to be predicted are the same. Similarly, if the physical structure at one level is more relevant to the system level than the other, it is reasonable to assign more confidence to the calibration result of this level. Therefore an important issue is the physical relevance between the lower levels and the system level, i.e. the extent of belief that the lower levels represent the physical characteristics of system level so that the calibration results can be reliably used in the system level prediction. The quantification of physical relevance needs to be studied and included in the uncertainty integration. This chapter develops a methodology to quantify the physical relevance using global sensitivity analysis. Finally, all the information from model validation, model calibration and physical relevance analysis are integrated using a new roll-up formula to predict the system level output.

9.2 Model Calibration

The first step in model calibration is to identify the parameters to calibrate. To promote computational efficiency, the computational model is generally replaced by a surrogate model such as the Gaussian process regression model [10], therefore the parameters of this surrogate model need to be estimated. In addition, the model error term $\delta(x)$ can also be

modeled as a GP model. Thus the parameters to calibrate in the KOH framework are: (1) system model parameters θ_m; (2) parameters of the GP model that replaces the computational model; (3) parameters of the GP model for $\delta(x)$; and (4) the standard deviation σ of the measurement error. The GP model parameters are also referred as hyper-parameters to distinguish them from θ_m. The presence of so many calibration parameters may be challenging if the calibration data is sparse.

However, McFarland [11] reported that the uncertainty resulting from the hyper-parameters of the GP model is negligible compared to the uncertainty in the model parameters, thus we can first estimate the hyper-parameters of the GP model using the physical model and then fix them as deterministic values in the subsequent calibration. In addition, if the model input is fixed then the input dependent model discrepancy function $\delta(x)$ will become to a single parameter δ to calibrate and no GP model is needed for it. In the numerical example of this chapter, the input is fixed and so the calibration parameters are: (1) model parameters θ_m; (2) model error δ; and (3) the standard deviation σ of ε_m. Here the entire set of calibration parameters is denoted as $\theta = \{\theta_m, \delta, \sigma\}$.

Using Bayesian inference, given available data D and prior distributions of parameters, the posterior distribution of θ is:

$$f''(\theta) \propto L(\theta)\, f'(\theta) \tag{9.2}$$

In (9.2), The likelihood function $L(\theta)$ is proportional to $P(D|\theta)$, i.e. the probability of observing data set D given the parameter θ; while the joint prior PDF $f'(\theta)$ is the product of the prior marginal distribution of each calibration parameter by assuming their independence, and prior marginal distributions are defined by researchers' knowledge. With the construction of $L(\theta)$ and $f'(\theta)$, the Markov Chain Monte Carlo (MCMC) [12] methods can be used to generate samples of $f''(\theta)$, and the posterior distributions of the model parameters can be constructed using kernel density estimation [13].

In a multi-level problem where each lower level may give data on multivariate output quantities, each output quantity at any level has a model error δ and a measurement error standard deviation σ, both of which need to be calibrated. For example, assume that a multi-level problem has two lower levels, and data regarding two output quantities are observed at each level. If calibration is conducted using the data of these two quantities at a single level, two model error terms and two measurement standard deviation terms will appear in model calibration; when calibration uses the data of these two quantities at both levels, four model error terms and four measurement standard deviation terms will appear in model calibration.

9.3 Model Validation

With model parameters calibrated, a conventional method to predict system level output is to simply propagate the posterior distributions of model parameters through the computational model at the system level and compute the system output, but the hidden assumption here is that the calibration results are correct, i.e. the model used for calibration is correct. However, there is no evidence yet to assess the validity of the calibration. In addition, a multi-level problem with n lower levels can give $2^n - 1$ combinations of model calibration and result in corresponding posterior distributions of model parameters, but model calibration cannot answer the question regarding how to select or combine these posterior distributions. Thus model validation becomes necessary to evaluate the validity of the model and construct appropriate integrated distributions of model parameters for system output prediction.

In model validation, it would be useful for a decision maker to have a quantitative probabilistic measure of the validity of the model [7]. To achieve this objective, the model reliability metric defines that the model is correct if the difference between the model prediction y and the measurement observation is less than a predefined tolerance λ. Assume that there is a data point D available. Due to the measurement error ($\varepsilon_m \sim N(0,\sigma)$), the measurement is actually a random variable denoted by d, with mean value $\mu_d = D$ and standard deviation σ, i.e. $d \sim N(\mu_d,\sigma)$. Let G denote the event that the model is correct, then the model reliability is the probability of event G:

$$P\left(G\middle|D\right) = P\left(-\lambda \le y - d \le \lambda\right) \tag{9.3}$$

First, consider the case where the model prediction y and the measurement error standard deviation σ are deterministic. By (9.3), the model reliability is expressed in (9.4), where ε is a dummy variable for integration:

$$P\left(G\middle|D\right) = \int_{-\lambda}^{\lambda} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\varepsilon - (y - D))^2}{2\sigma^2}\right] \mathrm{d}\varepsilon \tag{9.4}$$

In this chapter, 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. $y = G(\theta_m;x) + \delta(x)$. Therefore the model prediction y is stochastic due to the uncertainty in δ and θ_m. Furthermore, σ is also uncertain due to sparse data and need to be calibrated. In this case, the model reliability is computed as:

$$P\left(G\middle|D\right) = \int P\left(G\middle|D,\theta\right) f''(\theta)\,\mathrm{d}\theta \tag{9.5}$$

where $P(G|D,\theta)$ is given by the right hand side (9.4).

The computational model may be replaced by a GP model, which gives a normal distribution with mean value m and variance S at each input. Then in (9.4) y will be replaced by $m+\delta$ and σ^2 will replaced by σ^2+S.

If multiple data points $D=\{D_1,D_2,\cdots,D_n\}$ of one output quantity are available in model validation, a simplistic definition of the model reliability is the probability that all the observations support the model:

$$P\left(G\middle|D\right) = \int P\left(G\middle|D,\theta\right) f''(\theta)\,\mathrm{d}\theta = \int \left(\prod\nolimits_{i=1}^{n} P\left(G\middle|D_i,\theta\right)\right) f''(\theta)\,\mathrm{d}\theta \tag{9.6}$$

However, the problem with (9.6) is that as the number of data points applied in validation increases, the model reliability will decline due to the multiplication term $\prod_{i=1}^{n} P(G|D_i,\theta)$. This contradicts our intuition, since model reliability should be stable for a certain model, and independent of the number of validation data. Recall that all the calibration parameters ($\theta=\{\theta_m,\delta,\sigma\}$) are deterministic but unknown values, so their uncertainty arises only from lack of knowledge and data, which is 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 particular model is a deterministic but unknown value, and epistemic uncertainty is introduced into the model reliability due to sparse validation data. Here we apply (9.4) and (9.5) to each data point and obtain values of model reliability with respect to each data point. Then all these model reliability values can be used to construct a distribution of model reliability, which represents the epistemic uncertainty. In sum, this approach gives a stochastic representation of model reliability, i.e. $P(G)$ is not a single value but represented by a probability distribution due to epistemic uncertainty.

In this chapter, $P(G)$ is assumed to have a beta distribution with parameters q and r since the value of beta distribution lies on the interval [0, 1]. Once the model reliability values $\boldsymbol{D}_R=\{D_{R1},D_{R2},\cdots,D_{Rn}\}$ at each validation data point have been computed, several methods can be used to construct the PDF of model reliability, such as maximum likelihood, method of moments, or Bayesian inference. This chapter uses the Bayesian method to construct the PDF of $P(G)$ along the following steps:

1. Assume prior distributions of beta distribution parameters: $f'(q),f'(r)$.
2. Construct the likelihood function of q and r:

$$L(q,r) \propto P\left(D_R\middle|q,r\right) \propto \prod\nolimits_{i=1}^{n} \beta\left(D_{R_i}\middle|q,r\right) \tag{9.7}$$

where $\beta(\cdot|q,r)$ is the PDF of the beta distribution with parameters q and r.

3. Apply Bayesian inference to obtain the posterior distribution of q and r: $f^{'}(q),f^{''}(r)$.
4. Compute the unconditional distribution of $P(G)$:

$$f(P(G)) = \int \beta\left(P(G)\middle|q,r\right) f''(q) f''(r)\mathrm{d}q\mathrm{d}r \tag{9.8}$$

In a multi-level problem, if each lower level provides data of multivariate output quantities, we obtain a data set $\boldsymbol{D}_i=\{D_{i1},\ldots,D_{ij},\ldots,D_{iK}\}$ in the i^{th} experiment. The difference between the model prediction and observation will be a K-dimensional vector: $\boldsymbol{H}=\{h_{i1},h_{i2},\ldots,h_{iK}\}=\{y_1-d_{i1},\ldots,y_j-d_{ij},\ldots,y_K-d_{iK}\}$ where $d_{ij}\sim N(D_{ij},\sigma_j^2)$. Then the concept of model reliability metric can be easily extended by defining that the model is correct if the norm of $\boldsymbol{H}$ is less than a tolerance λ; and λ can be defined as the norm of the tolerance vector $\{\lambda_1,\lambda_2,\ldots,\lambda_K\}$ where λ_j is the tolerance for the j^{th}quantity($j=1toK$). Here we redefine G as the event that the model group is correct, then the model reliability can be expressed as:

$$P\left(G\middle|D_i\right) = P\left(\left(\sum_{j=1}^{K}\left|h_j\right|^2\right)^{1/2} < \lambda\right) where\ \lambda = \left(\sum_{j=1}^{K}\left|h_j\right|^2\right)^{1/2} \tag{9.9}$$

9.4 Uncertainty Integration

With the model calibration and model validation conducted at each lower level, the next question is how to integrate all the information on uncertainty collected from these activities and obtain the distribution of each model parameter for future system output prediction. This section first examines the concept of physical relevance, which measures the closeness of the lower level to the system level; and then proposes a new roll-up method to integrate all the available information and predict the system output.

9.4.1 Physical Relevance

A roll-up methodology for uncertainty integration was proposed in (9.1). In this roll-up methodology, the weight of each PDF is purely determined by model validation results. In Sect. 9.1, it was also argued that it is reasonable to assign more weight to the level physically "closer" or more similar to the system level than the other. Hence this chapter introduces the quantification of physical relevance, which measures the degree that the experimental configuration at a lower level represents the physical characteristics of the system. At present such determination is only intuitive and qualitative; an objective quantitative measure of relevance is needed in order to be useful in uncertainty quantification.

Consider a model $Y = G(\boldsymbol{X})$ where Y is the output and $\boldsymbol{X} = \{X^1, \ldots, X^N\}$ is a vector containing all the model inputs (and model parameters if they are uncertain). The objective in sensitivity analysis is to measure the contribution of each component of $\boldsymbol{X}$ in the uncertainty of $\boldsymbol{Y}$ [14]. Compared to local sensitivity analysis, global sensitivity analysis considers the entire probability distribution of the input, not just the contribution at a local point. Two global sensitivity indices have been developed in the literature based on the variance decomposition theorem, namely, first-order index and total effects index. For a particular model input X^i, the first-order index is defined as $S_1^i = V(E(Y|X^i))/V(Y)$; and the total effects index is defined as $S_T^i = 1 - V(E(Y|X^{-i}))/V(Y)$ where X^{-i} means all the inputs other than X^i. The first-order index measures the contribution of the variable X^i by itself, so the sum of first-order indices of all variables is always less than or equal to unity. The difference between this sum and unity is the contribution of the interaction between input variables. A higher first-order index value signals an important contribution, but a small value does not mean the variable is negligible because it may contribute significantly by interacting with other variables. In contrast, the total effect index contains not only the contribution of X^i, but also the interaction effect of X^i with other inputs. So in this study, the total effects index is used to develop a method to quantify the physical relevance; thus in the following discussions the term sensitivity index indicates the total effects index. The sum of total effect indices is always greater than or equal to unity.

Without loss of generality, this chapter takes a multi-level problem with two lower levels and a system level for the illustration of physical relevance. Assuming that the objective is to predict the system output y_s (such as the maximum acceleration at mass three in the numerical example in Sect. 9.5), and the same quantity is also measured at lower levels (in the numerical example the maximum acceleration at mass three is also measured at Level 1 and Level 2), then the three computational models for this quantity at different levels are expressed as $y_{L_1} = G_{L_1}(\theta_m, x)$, $y_{L_2} = G_{L_2}(\theta_m, x)$, $y_s = G_s(\theta_m, x)$, where θ_m is model parameters and $\boldsymbol{x}$ is model inputs. Assuming θ_m and x contain N elements in total, the sensitivity analysis at each level will give an N-dimensional sensitivity index vector, which contains the physical structural information at each level. We can consider this sensitivity vector as a characterization of each level, so the comparison of these vectors will reflect the relevance between each level. Here we define the physical relevance as the square of cosine similarity [15] of the sensitivity index vectors, and the cosine similarity is the standardized dot product of two vectors. Thus if the sensitivity vectors at a lower level and the system level are V_l and V_s, the relevance index between them is defined as:

$$R = \left(\frac{V_l \cdot V_s}{||V_l||\ ||V_s||}\right)^2 \tag{9.10}$$

In other words, the relevance is the square of the cosine value of the angle between two sensitivity vectors. If the angle between the two sensitivity index vectors is zero, the relevance is 1, which means the physical structures at the two levels

are the same and the lower level has a 100 % relevance from the viewpoint of uncertainty integration; if the two vectors are perpendicular, the physical structures are not similar at all and the relevance is zero.

In addition, this definition of physical relevance gives a value on the interval [0, 1]; and its opponent, the square of sine value, indicates physical non-relevance; hence the sum of "relevance" and "non-relevance" is the unity. In the roll-up methodology proposed in Sect. 9.4.2, this characteristic gives the possibility to treat the physical relevance similar to a probability and include the physical relevance in the uncertainty integration.

A further question arises in the application of this physical relevance definition using sensitivity analysis. The global sensitivity analysis considers the entire distribution of variable, so we need to know the distributions of θ_m *and* x, but the distribution of θ_m is unknown. In order to solve this problem, a simple iterative algorithm to compute the physical relevance index R 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 Sect. 9.4.2.
3. Use the unconditional distributions from step 2 in the sensitivity analysis and compute the updated physical relevance R.
4. Repeat step 2 and step 3 until the physical relevance R converges.

9.4.2 Roll-up Methodology

The integrated distributions of model parameters need to be constructed before propagating them through the computational model at the system level to obtain the system output. Equation (9.1) expressed a roll-up methodology to integrate the results from model calibration and model validation; and this chapter modifies this methodology to include two additional concepts:

1. Stochastic model reliability: the event that the model prediction at Level i is correct is denoted by G_i, whose probability is $P(G_i)$; and the stochastic reliability defines $P(G_i)$ as a stochastic variable with PDF $f(P(G_i))$ which has been explained in Sect. 9.3.
2. Physical relevance: defined as the square of the cosine value of the angle between the sensitivity vectors. 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 chapter denotes the event that Level i is relevant to the system level by S_i, and its probability $P(S_i)$ is the value of the 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 and a realization of model reliability $P(G_i)(i = 1, 2)$ is:

$$\begin{aligned} & f\left(\theta \middle| D_1^{C,V}, D_2^{C,V}, P\left(G_1\right), P\left(G_2\right)\right) \\ & = P\left(G_1 G_2 S_1 S_2\right) f\left(\theta \middle| D_1^C, D_2^C\right) + P\left(G_1 S_1 \cap \left(G_2' \cup S_2'\right)\right) f\left(\theta \middle| D_1^C\right) \\ & \quad + P\left(G_2 S_2 \cap \left(G_1' \cup S_1'\right)\right) f\left(\theta \middle| D_2^C\right) + P\left(\left(G_1' \cup S_1'\right) \cap \left(G_2' \cup S_2'\right)\right) f\left(\theta\right) \end{aligned} \tag{9.11}$$

From the view of generating samples, (9.11) indicates two criteria: (1) whether a level is relevant to the system level; (2) whether a level has the correct model prediction. A sample of θ is generated from $f(\theta|D_1^C, D_2^C)$ only when both levels satisfy both criteria; and a sample of θ is generated from $f(\theta|D_i^C)$ if level i satisfies both criteria but the other level does not; a sample of θ is generated from the prior $f(\theta)$ if neither level satisfies both criteria.

The so-called integrated distribution of θ, which is conditioned on the calibration and validation data, are computed as:

$$f\left(\theta \middle| D_1^{C,V}, D_2^{C,V}\right) = \iint\!\!\iint f\left(\theta \middle| D_1^{C,V}, D_2^{C,V}, P\left(G_1\right), P\left(G_2\right)\right) f\left(P\left(G_1\right)\right) f\left(P\left(G_2\right)\right) \mathrm{d}P\left(G_1\right) \mathrm{d}P\left(G_2\right) \tag{9.12}$$

Equations (9.11) and (9.12) express the proposed approach to integrate calibration, validation and relevance results. The analytical expression of $f(\theta|D_1^{C,V}, D_2^{C,V})$ is difficult to derive since the results we collect in model calibration and validation are all numerical. A single loop sampling approach is proposed below to construct $f(\theta|D_1^{C,V}, D_2^{C,V})$ numerically:

1. Generate a sample of each of $P(G_1)$ and $P(G_2)$ from their distributions.

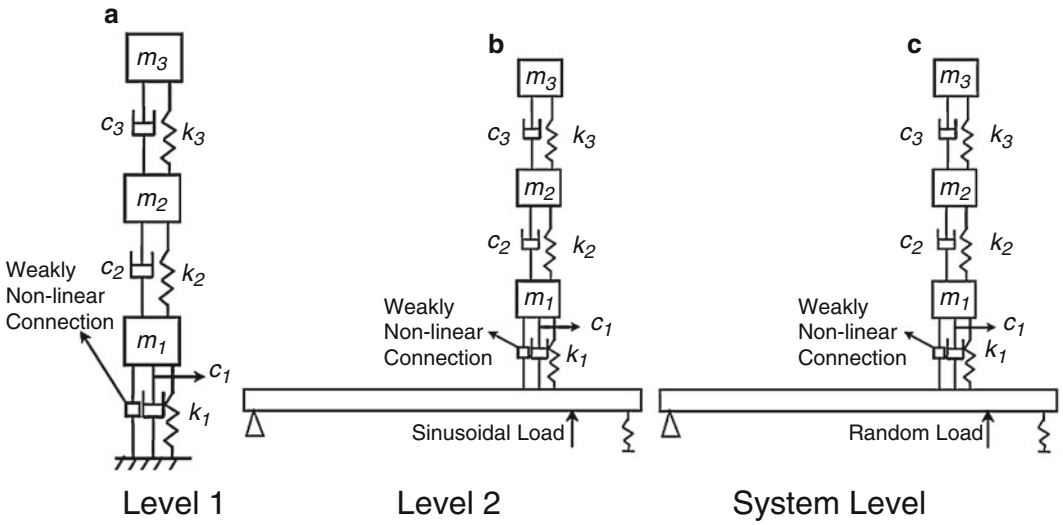

Fig. 9.2 Structural dynamics challenge problem

2. Assuming G_1, G_2, S_1, S_2 are independent of each other, compute the weights $P(G_1G_2S_1S_2)$, $P(G_1S_1 \cap (G_2' \cup S_2'))$, $P(G_2S_2 \cap (G_1' \cup S_1'))$, and $P((G_1' \cup S_1') \cap (G_2' \cup S_2'))$ in (9.11). Divide the interval [0, 1] into four domains and the length of each domain is equal to the value of each weight.
3. Generate a random number from the uniform distribution $U(0,1)$.
4. Generate a sample of θ from $f(\theta|D_1^C, D_2^C)$ if the random number in step 3 is located in the first domain; from $f(\theta|D_1^C)$ if located in the second domain; from $f(\theta|D_2^C)$ if located in the third domain; from $f(\theta)$ if located in the fourth domain.
5. Repeat step 1 to step 4 to obtain multiple samples of θ; then construct the PDF $f(\theta|D_1^{C,V}, D_2^{C,V})$ by any method such as kernel density estimation.

After obtaining the integrated distributions of all model parameters, the final step is to propagate the integrated distributions through the computational model of the system of interest to predict the system level output. This can be done using Monte Carlo sampling. Due to the uncertainty in model parameters, the predicted quantity of interest will also be stochastic, and its distribution can also be constructed by kernel density estimation.

9.5 Numerical Example

A multi-level structural dynamics challenge problem provided by Sandia National Laboratories [16] is used to illustrate the methodology developed in Sects. 9.2, 9.3, and 9.4. As shown in Fig. 9.2, Level 1 contains three mass-spring-damper dynamic components in series, and a sinusoidal force input is applied to m_1. At Level 2, the dynamic system is mounted on a beam supported by a hinge at one end and a spring at the other end; another 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. Here Level 1 and Level 2 are defined as lower levels, and experimental data are only available at lower levels. All levels share six model parameters: three spring stiffnesses $k_i (i = 1, 2, 3)$ and three damping ratios $\zeta_i (i = 1, 2, 3)$. The units of all quantities are unified.

Five experiments for calibration and five experiments for validation are conducted at Level 1 and Level 2, and six quantities at each lower level are measured:

1. $A_i (i = 1, 2, 3)$: the maximum acceleration in the i^{th} mass.
2. $D_i (i = 1, 2, 3)$: the energy dissipated by the i^{th} damper in 1,000 time units.

The parameters to be calibrated in this example are: the spring stiffnesses $k_i (i = 1, 2, 3)$; the damping ratios $\zeta_i (i = 1, 2, 3)$; model error δ and the output measurement error's standard deviation σ_m if the data of the corresponding quantity are used in model calibration. Based on expert judgment, the prior distribution of each k_i and ζ_i is assumed to be lognormal with a coefficient of variation of 10 % and mean values of $\mu_{k_1} = 5000$, $\mu_{k_2} = 9000$, $\mu_{k3} = 8000$, $\mu_{\zeta_i} = 0.025$ $(i = 1, 2, 3)$.

Table 9.1 Model reliability values

Level 1	0.8434	0.8518	0.8526	0.8638	0.8284
Level 2	0.9360	0.9392	0.9438	0.9516	0.8962

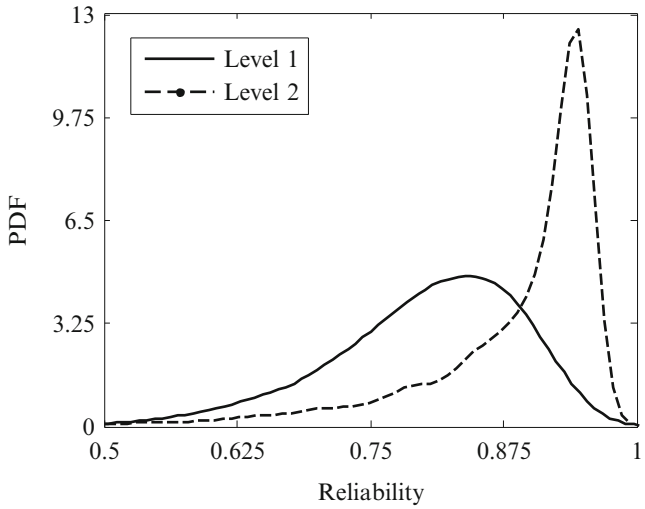

Fig. 9.3 Model reliability

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

Since as many as six quantities are measured, we can choose any combination of these six quantities in the analysis. As more output 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 model reliability and the integrated distribution of model parameters using the data of all six quantities are provided below. A plot showing the reduction in the uncertainty of system output prediction with the increase of information is also provided at the end.

We use the proposed stochastic model reliability metric with multivariate output in the model validation. The tolerance for each quantity is 15 % of the mean value of validation data. The model reliability values given by the validation data are listed in Table 9.1, and they are used to construct the distribution of model reliability at Level 1 and Level 2 shown in Fig. 9.3.

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 in Sect. 9.4.1 converges after three iterations, and after five iterations for Level 2. The results are: $P(S_1) = 0.7001, P(S_2) = 0.8929$.

Based on all the information from calibration, validation and physical relevance, the integrated distribution of all six model parameters are constructed and shown in Fig. 9.4. It shows that the proposed roll-up method is more conservative than the roll-up method in (9.1), since we put one more criterion of physical relevance for sampling generation from the posterior distribution.

The system output is predicted by propagating the integrated distributions of model parameters through the computational model of the system level. Figure 9.5 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 in the system output prediction declines as more quantities are employed.

9.6 Summary

This chapter establishes a methodology to quantify the uncertainty in 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 chapter are the new methodologies of model validation and uncertainty integration, and the definition and computation of physical relevance. For model validation, this chapter proposes a stochastic model reliability metric and extends its application to multivariate output. The physical relevance quantifies whether the lower level represents the structural characteristics of the system level,

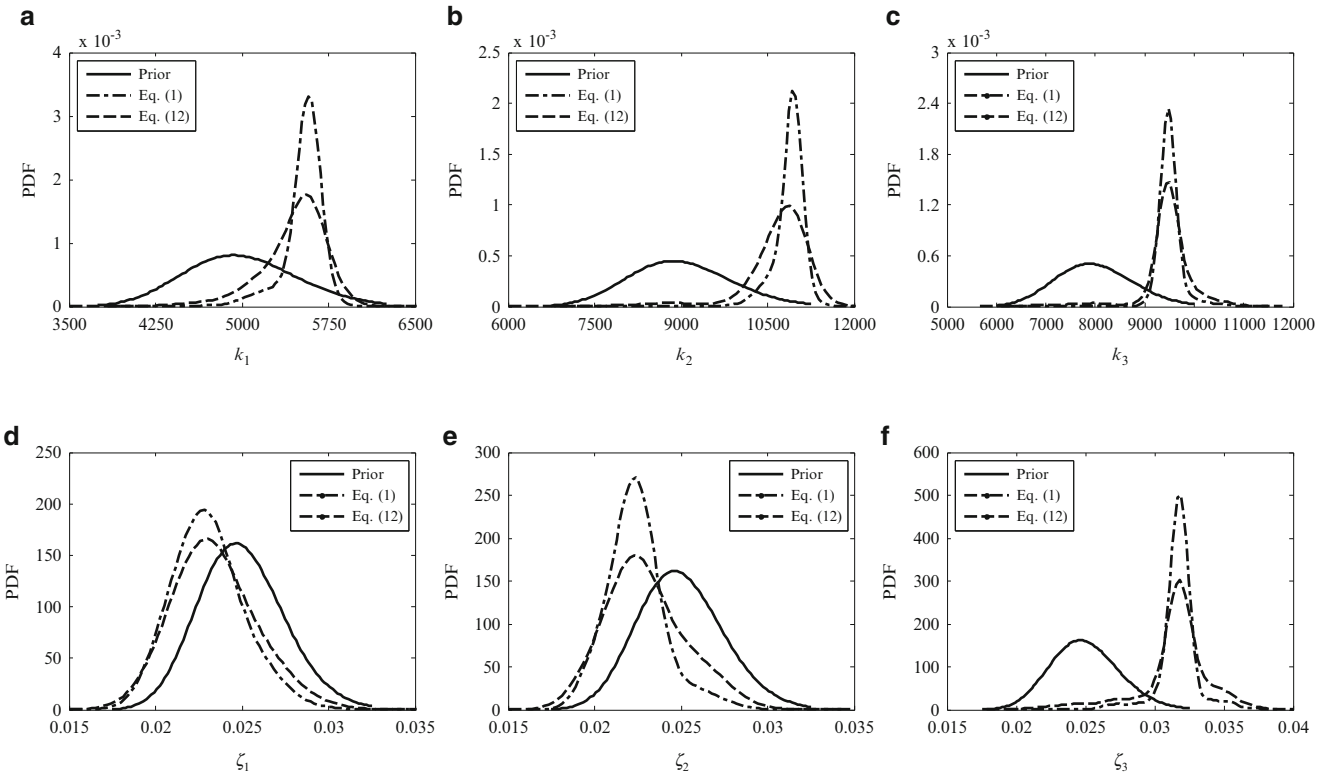

Fig. 9.4 Integrated distributions of model parameters

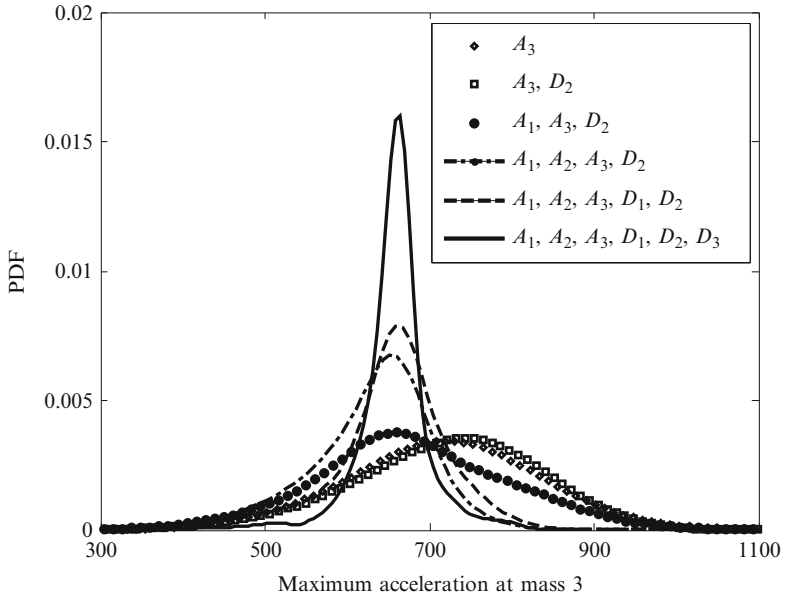

Fig. 9.5 System output prediction

and determines the weight of each posterior distribution in the uncertainty integration. The last contribution of this chapter lies in the development of the roll-up formula 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 magnitude of physical relevance of each level. This methodology can be used for resource allocation decision such how many calibration data or validation data are needed at each level to achieve a desired level of uncertainty in the system level prediction.

Acknowledgements The research in this chapter 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 appreciate valuable discussions with Shankar Sankararaman (NASA Ames), You Ling (GE) and Joshua Mullins (Vanderbilt University).

References

1. Urbina A, Paez TL, Gregory D, Resor B, Hinnerichs T, Gorman CO (2006) Validation of a combined non-linear joint and viscoelastic encapsulating foam model Sandia national laboratories. In: Society for experimental mechanics
2. Kennedy M, O'Hagan A (2001) Bayesian calibration of computer models. J R Stat Soc 63(3):425–464
3. Trucano TG, Swiler LP, Igusa T, Oberkampf WL, Pilch M (2006) Calibration, validation, and sensitivity analysis: what's what. Reliab Eng Syst Saf 91(10–11):1331–1357
4. Schervish M (1996) P values: what they are and what they are not. Am Stat 50(3):203–206
5. Mahadevan S, Rebba R (2005) Validation of reliability computational models using Bayes networks. Reliab Eng Syst Saf 87(2):223–232
6. Ferson S, Oberkampf WL, Ginzburg L (2008) Model validation and predictive capability for the thermal challenge problem. Comput Methods Appl Mech Eng 197:2408–2430
7. Rebba R, Mahadevan S (2008) Computational methods for model reliability assessment. Reliab Eng Syst Saf 93(8):1197–1207
8. Sankararaman S, Mahadevan S (2013) Assessing the reliability of computational models under uncertainty. In: Fifty fourth AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and materials conference, p 1–8
9. Sankararaman S, Mahadevan S (2012) Comprehensive framework for integration of calibration, verification and validation. In: Fifty third AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference, no. April, p 1–12
10. Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning, vol 14, no 2. MIT Press, Cambridge, MA, p 266.
11. Mcfarland JM (2008) Uncertainty analysis for computer simulations through validation and calibration. Vanderbilt University, Nashville, TN
12. Tierney L (1994) Markov chains for exploring posterior distributions. Ann Stat 22(4):1701–1728
13. Rosenblatt M (1956) Remarks on some nonparametric estimates of a density function. Ann Math Stat 27:832–837
14. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S (2008) Global sensitivity analysis: the primer. Wiley-Interscience, New York
15. Singhal A (2001) Modern information retrieval: a brief overview. IEEE Data Eng Bull 24(4):35–43
16. Red-Horse JR, Paez TL (2008) Sandia national laboratories validation workshop: structural dynamics application. Comput Methods Appl Mech Eng 197(29–32):2578–2584