

Uncertainty Quantification of System Model Parameters with Component Level and Sub-System Level Tests

Chenzhao Li and Sankaran Mahadevan
Vanderbilt University

ABSTRACT

An essential step in predicting the output of a complicated system is the calibration of model parameters, but this step cannot be conducted directly if the data at the system-level are not available. In such a situation, a reasonable route is to quantify the system model parameters using tests at lower levels of complexity which share the same model parameters with the system, and propagate the results through the computational model at the system level. For such a multi-level problem, this paper proposes a methodology to quantify the uncertainty in the system-level model parameters by integrating model calibration, model validation and sensitivity analysis at different levels. The proposed approach considers the validity of the models used for parameter estimation at lower levels, as well as the relevance at the lower level to the prediction at the system level. The model validity is evaluated using a model reliability metric, which is extended to handle the uncertainty in model validity and models with multivariate output. The relevance is quantified by comparing Sobol indices at the lower level and system level, thus measuring the extent to which a lower level test represents the characteristics of the actual system so that the calibration results can be reliably used in the system level. Finally a roll-up method is proposed to integrate the results of calibration, validation and relevance analysis and predict the system output.

CITATION: Li, C. and Mahadevan, S., "Uncertainty Quantification of System Model Parameters with Component Level and Sub-System Level Tests," *SAE Int. J. Mater. Manf.* 9(2):2016, doi:10.4271/2016-01-0287.

INTRODUCTION

Parameters of computational models are often calibrated using experimental data. For a complicated system it may be difficult to conduct full-scale experiments; but it may be possible to obtain data at lower levels of complexity (e.g., isolated physics or simpler configurations). Figure 1 shows such a multi-level problem with two lower levels (G_1 , G_2) and a system level (H). The lower levels and the system level constitute a hierarchy, sharing a set of model parameters (θ_m) to be calibrated. In order to predict the system level output when data are only available at lower levels, a reasonable route is to quantify the model parameters using lower level data, and propagate the results through the computational model at the system level. Several issues need to be addressed in realizing such a multi-level problem. First, even if model input and output are measured in the lower level tests, thereby forming pairwise input-output data, the calibration result can still be uncertain due to several sources, including: 1) errors in the lower level computational models; 2) measurement errors in the experiments; and 3) sparse experimental data. Second, the existence of multiple lower levels provides multiple possibilities to conduct model calibration and leads to multiple calibration results. A pertinent question is how to select among or combine these calibration results, and its answer depends on two factors: 1) the quality of the calibration at each level; and 2) the relationship between the lower levels and the system level. The first factor can be quantified by model validation [1], which measures model validity by comparing model prediction with experimental data. In order to quantify the second factor, a lower level may be

defined to be "closer" to the system level if it has similar physical configuration and/or experimental environment as the usage condition of interest at the system level. A quantitative method of physical relevance analysis is proposed in this paper. With these two factors quantified, an appropriate uncertainty integration methodology is needed to aggregate all the information from lower levels to quantify the prediction uncertainty at the system level. Thus, the proposed solution methodology for the above of multi-level problems includes four activities: 1) model calibration at lower levels; 2) model validation at lower levels; 3) physical relevance analysis; and 4) uncertainty integration across all the levels.

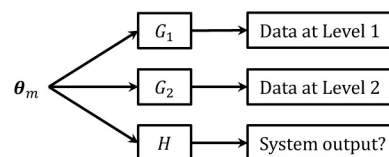


Figure 1. Multi-level problem

Model calibration focuses on the inference of model parameters so that the prediction by the computational model matches the experimental data [2]. This paper uses Bayesian inference within the Kennedy and O'Hagan (KOH) [2] framework to estimate the posterior distributions of model parameters. This framework facilitates the inclusion of different uncertainty sources and available data in model calibration [3]. In a multi-level problem, model calibration can be conducted using the data from a single level or

multiple levels. For the hierarchy in Figure 1 with two lower levels, 3 calibrations are possible: 1) calibration using the data and model from Level 1 alone; 2) calibration using the data and model from Level 2 alone; 3) calibration using the data and models from both Level 1 and Level 2. Generally, if data are available at n different levels, $2^n - 1$ model calibrations are possible to quantify the uncertainty of model parameters [4]. Thus an important problem is how to integrate these competing calibration results. In this paper, the distribution of model parameter propagated through the computational model of the system is referred to as the integrated distribution, and it is represented as a weighted average of all $2^n - 1$ posterior distributions; the weight terms depend on model validity and physical relevance at each level. As a result, the resultant integrated distribution aggregates the information from multiple levels and multiple activities.

Model validation quantifies the degree to which the model is an accurate representation of reality; in practice, model validity is assessed by comparing model prediction against experimental data [5], and validation also provides necessary information to understand the limitation of calibration [5]. Comprehensive reviews on model validation can be found in [1][6]. The result of classical hypothesis testing is an acceptance/rejection decision. Validation metrics resulting in a single quantitative value indicating the degree of model validity have also been developed. In Bayesian hypothesis testing [3] [6], the posterior distribution obtained by model calibration is used as the null hypothesis and an alternative distribution is selected for the alternative hypothesis. Regarding the formulation of the null hypothesis, the Bayesian hypothesis testing can be classified as equality-based (model prediction equal to observation) and interval-based (the difference between model prediction and observation less than a tolerance) [7][8]. The result of Bayesian hypothesis testing is a Bayes factor (which is the likelihood ratio between the null and alternate hypothesis), measuring the support from validation data to the null and alternate hypothesis. The model validation metric used in this paper is the model reliability metric proposed by Rebba and Mahadevan [7] and further developed by Sankararaman and Mahadevan [9], which gives an objective probabilistic measure of model validity. The model reliability metric computes the probability that the difference between model prediction and observed data is less than a predefined tolerance. Since the model reliability metric gives a probability value, it can be conveniently used as a weighting term in subsequent uncertainty integration across multiple levels. Measurement error and variability across different experiment settings cause uncertainty in model reliability. Thus this paper proposes a stochastic model reliability metric in which the model reliability is treated as a stochastic variable. In addition, this paper extends the model reliability metric to consider multivariate output.

In this paper, the relationship between the lower level and the system level is quantified by a proposed physical relevance analysis. The necessity of physical relevance analysis is explained here. The only reason that model parameters are calibrated using lower level data is the unavailability of system level data. In case of enough system level data, we would directly conduct model calibration at the system level. In other words, high confidence is assigned to the calibration using system level data, and the information at lower levels may be neglected if adequate system level data is available. An inherent assumption in this high confidence is that the tests at the system level

completely reflect the physical characteristics of system since the physical structure used to measure the data and the physical structure whose response is to be predicted are the same. Therefore, if the physical structure at one level is more similar to the system level than the other, it is reasonable to assign higher confidence to the calibration result at this level. Thus the physical relevance of the lower level to the system level is an important issue, i.e., the degree to which the experimental configuration at a lower level reflects the physical characteristics of the system so that the calibration results can be reliably used in the system level prediction. The physical relevance measure decides the confidence on the calibration at lower levels, and influences the uncertainty integration. This paper proposes a method to quantify the physical relevance using global sensitivity analysis and the cosine similarity of sensitivity vectors.

The objective of uncertainty integration is to aggregate all the available information from model calibration, model validation and physical relevance analysis. An integration (roll-up) methodology was proposed in [10][11] to quantify the uncertainty in model parameters. For the multi-level problem in Figure 1, the integrated distribution of a model parameter θ was proposed as:

$$f(\theta|D_1^{C,V}, D_2^{C,V}) = P(G_1)P(G_2)f(\theta|D_1^C, D_2^C) + P(G_1')P(G_2)f(\theta|D_1^C) + P(G_1)P(G_2')f(\theta|D_2^C) + P(G_1')P(G_2')f(\theta) \quad (1)$$

In Eq. (1) the integrated distribution $f(\theta|D_1^{C,V}, D_2^{C,V})$ is a weighted average of multiple posterior distributions and contains four terms: in the first term the posterior distribution $f(\theta|D_1^C, D_2^C)$ using the calibration data of both Level 1 and Level 2 has the weight $P(G_1)P(G_2)$ which is the probability that both of the models are valid; similarly, the second and third terms contain the posterior distribution $f(\theta|D_1^C)$ using the calibration data at Level i alone and its weight is the probability that the model at Level i is valid but the model at another level is invalid; 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 invalid. The validities of the models at two lower levels are assumed to be independent. Obviously, the weight of each PDF on the right hand side of Eq. (1) is purely decided by model validation. This paper proposes an extension of Eq. (1) by considering the physical relevance, the necessity of which has been explained earlier.

In summary, the motivation of this paper is to quantify the uncertainty in the system output by solving the challenges in model calibration, model validation, physical relevance analysis, and uncertainty integration. The posterior distributions of model parameters are computed by Bayesian inference. The integration of multiple posterior distributions for each model parameter is assisted by model validation and physical relevance analysis, and realized in a proposed new roll-up method. This paper develops a methodology to compute the physical relevance using global sensitivity analysis and cosine similarity of vectors. In model validation, the model reliability concept is extended to capture the uncertainty in model reliability and to consider multivariate output. Finally, the integrated distributions of model parameters are propagated through the computational model at the system level to predict the system output and quantify its uncertainty.

MODEL CALIBRATION

This section provides a brief summary of Bayesian model calibration, as conducted in this paper. No new developments are reported; however, since calibration is the first step of the proposed methodology, a brief summary is given here for the sake of completeness.

Suppose the physical input-output relationship at a single level is described by a computational model $y_c = F(\theta_m; \mathbf{x})$, where y_c is the computational model output, and θ_m is a set of unknown model parameters, and $\mathbf{x}$ is the model input. Kennedy and O'Hagan (KOH) [2] expressed the relationship between the experimental observation z and the computational model as:

$$z = F(\theta_m; \mathbf{x}) + \delta(\mathbf{x}) + \varepsilon_m \quad (2)$$

where $\delta(\mathbf{x})$ is the model error (input-dependent), and ε_m measurement error which is usually assumed to be Gaussian distribution $N(0, \sigma_m^2)$. The model error $\delta(\mathbf{x})$ can be modeled using different formulations [12], which introduces more parameters. In addition, to reduce the computational effort, the computational model $F(\theta_m; \mathbf{x})$ may also be replaced by a surrogate model; several options such as polynomial response surface [13], polynomial chaos expansion [14], Gaussian process (GP) model [15][16] etc. are available. This paper uses the GP model. The parameters of $\delta(\mathbf{x})$ and of the surrogate model for $F(\theta_m; \mathbf{x})$ are also uncertain and need to be estimated. These parameters are also called hyper-parameters to distinguish them from model parameters θ_m . In sum, all the parameters to calibrate include: 1) model parameters θ_m ; 2) hyper-parameters in the surrogate model for $F(\theta_m; \mathbf{x})$; 3) hyper-parameters θ_δ of the model error $\delta(\mathbf{x})$; and 4) standard deviation σ_m of ε_m . The presence of so many calibration parameters is challenging if calibration data are sparse.

However, if a GP model is chosen as the surrogate for $F(\theta_m; \mathbf{x})$, the uncertainty resulting from the hyper-parameters of the GP model is often negligible compared to the uncertainty in the actual model parameters [17]. Thus we can first estimate the hyper-parameters of the GP model and then fix them as deterministic values in the subsequent calibration of model parameters θ_m . In addition, if the model input is fixed then the input dependent model discrepancy $\delta(\mathbf{x})$ will become a single parameter δ . In the numerical example of this paper, for each lower level calibration test, the model/experimental input is fixed and so the calibration parameters θ includes: 1) model parameters θ_m ; 2) model error δ ; and 3) the standard deviation σ_m of measurement error ε_m .

In a multi-level problem, each lower level may provide data for multivariate output quantities, and each output quantity at any level has a corresponding model error $\delta(\mathbf{x})$ and measurement error standard deviation σ_m to be calibrated. In Figure 1, if calibration data consist of two output quantities at Level 1, model calibration includes two model error terms and two measurement error terms; and if two output quantities at Level 2 are also included for calibration, model calibration includes four model error terms and four measurement error terms.

For the model error $\delta(\mathbf{x})$, we need to select the prior distribution for each hyper-parameter in the above formulation. But if model input $\mathbf{x}$ is fixed, we only need to select a prior distribution for δ . In the numerical example of this paper, since there is no information available on δ , a uniform prior distribution is assumed as $\delta \sim U(a, b)$ where a and b are the lower and upper bounds of the uniform distribution. The prior distribution of σ_m is chosen as the non-informative Jeffrey's prior [18] $f'(\sigma_m) \propto 1/\sigma_m$, which is invariant under reparameterization. In addition, the prior distributions for θ_m are constructed based on from expert opinions.

With prior distributions for $\theta = \{\theta_m, \theta_\delta, \sigma_m\}$ defined and experimental data at lower levels obtained, the Bayesian inference expresses the posterior distribution of θ as:

$$f''(\theta) = \frac{L(\theta)f'(\theta)}{\int L(\theta)f'(\theta) d\theta} \quad (3)$$

where $L(\theta)$ is the likelihood function of θ and $f'(\theta)$ is the joint prior PDF of θ . The samples of $f''(\theta)$ are often generated numerically by the Markov Chain Monte Carlo (MCMC) method [19].

MODEL VALIDATION

With model parameters calibrated, the common approach to predict system level output is to propagate the posterior distribution of model parameters through the computational model at the system level and compute the system output. However, there is no evidence yet to assess the validity of the calibration. In addition, a multi-level problem with n lower levels can provide $2^n - 1$ combinations of model calibration data (individual levels and their combinations) and the corresponding posterior distributions of model parameters, but model calibration cannot answer the question regarding how to integrate these posterior distributions. Thus model validation is necessary to account for the validity of the model calibration before using the calibrated model parameters for system output prediction. This paper applies the model reliability metric [7][9] for model validation, the result of which is the probability of model being valid and can be conveniently used in Eq. (1) and the proposed new roll-up method for uncertainty integration.

Model Reliability Metric

The uncertainty integration method in Eq. (1) expresses the integrated distribution of model parameter θ as the weighted average of multiple posterior distributions. To make the integrated distribution in Eq. (1) a valid PDF, the sum of the weight terms computed in model validation should be unity. A quantitative, probabilistic measure of model validity is needed to achieve this objective. The model reliability metric directly facilitates this calculation: the model is defined to be valid if the difference between the model prediction y and the measurement is less than a predefined tolerance λ . Due to the measurement error ($\varepsilon_m \sim N(0, \sigma_m^2)$), the measurement is actually a random variable. For a single observed value D , this random variable is denoted by d with mean value D and standard deviation σ_m , i.e. $d \sim N(D, \sigma_m^2)$. Let G denote the event that the model is valid, then the model reliability is the probability of event G .