



Selection of model discrepancy priors in Bayesian calibration



You Ling, Joshua Mullins, Sankaran Mahadevan*

Department of Civil and Environmental Engineering, Vanderbilt University, TN 37235, United States

ARTICLE INFO

Article history:

Received 13 December 2013

Received in revised form 6 June 2014

Accepted 2 August 2014

Available online 7 August 2014

Keywords:

Bayesian calibration

Model uncertainty

Identifiability

Validation

ABSTRACT

In the Kennedy and O'Hagan framework for Bayesian calibration of physics models, selection of an appropriate prior form for the model discrepancy function is a challenging issue due to the lack of physics knowledge regarding model inadequacy. Aiming to address the uncertainty arising from the selection of a particular prior, this paper first conducts a study on possible formulations of the model discrepancy function. A first-order Taylor series expansion-based method is developed to investigate the potential redundancy caused by adding a discrepancy function to the original physics model. Further, we propose a three-step (calibration, validation, and combination) approach in order to inform the decision on the construction of model discrepancy priors. In the validation step, a reliability-based metric is used to evaluate posterior model predictions in the validation domain. The validation metric serves as a quantitative measure of how well the discrepancy formulation captures the missing physics in the model. In the combination step, the posterior distributions of model parameters and discrepancy corresponding to different priors are combined into a single distribution based on the probabilistic weights derived from the validation step. The combined distribution acknowledges the uncertainty in the prior formulation of model discrepancy function.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Model calibration is the process of adjusting unknown model parameters in order to improve the agreement between model output and observed data [1], and it is a widely used technique in the analysis of engineering systems [2]. Kennedy and O'Hagan [3] developed a Bayesian calibration framework (referred to as the KOH framework) to include various sources of uncertainty. One of the main features of this framework is the use of a model discrepancy function to explicitly account for the uncertainty due to model inadequacy. It has been shown that by including an appropriate model discrepancy function in calibration, bias and overfitting in the estimation of physical parameters can be mitigated or avoided [4]. Based on the KOH framework, efficient and rigorous calibration methods have been developed for realistic problems [5–8], and practical applications can be found in many areas of science and engineering, such as environmental management [9], heat transfer [10], astronomy [11], hydrology [12], geochemistry [13], fatigue [14], and aerothermal modeling [15].

Despite the extensive development of calibration methods and applications, the formulation of the model discrepancy function remains a challenging issue. As shown by Brynjarsdottir and O'Hagan [4], the calibration results may not be satisfying if the prior assumption of discrepancy does not capture the effect of missing physics in the model. Moreover,

* Corresponding author. Tel.: +1 615 322 3040; fax: +1 615 322 3365. Postal address: Vanderbilt University, VU Station B 356077, Nashville, TN 37235-6077, United States.

E-mail address: sankaran.mahadevan@vanderbilt.edu (S. Mahadevan).

the addition of a discrepancy term to the physics model may sometimes lead to parameter non-identifiability [7,8,12]. Several different prior assumptions of model discrepancy have been used in previous studies, including constant bias [9], physics-based deterministic function [15], Gaussian random variable [6,12–14,16,17], uncorrelated random vector [11], random walk [9], and Gaussian random process [5,7,10]. However, a rigorous comparison between these different prior formulations of model discrepancy has not been conducted, and a general guideline for choosing the appropriate formulation is not currently available.

In this paper, we will investigate Bayesian calibration with five different prior formulations of model discrepancy function: (1) constant, (2) Gaussian random variable with fixed mean and variance, (3) Gaussian random variable with input-dependent mean and variance, (4) Gaussian random process with stationary covariance function, and (5) Gaussian random process with non-stationary covariance function. We consider these five formulations only for the sake of illustration. Other formulations of model discrepancy function are also possible [9,15]. A three-step method is proposed in order to assess these prior formulations of discrepancy, and also obtain a probability distribution of model parameters that accounts for the uncertainty due to the lack of knowledge of model inadequacy. First, Bayesian calibration is performed using each of the five options of the model discrepancy function, and five posterior distributions of model parameters and model discrepancy are obtained. In the second step, we use a quantitative model validation metric, namely a reliability-based metric [18], to assess the posterior model predictions corresponding to the distributions of model parameters and discrepancy obtained in the first step. In the third step, if multiple options are of comparable validity, the corresponding posterior distributions of model parameters and discrepancy are combined into a single distribution using the weights derived from the validation comparison.

The other focus of this paper is model identifiability. A model is non-identifiable if there are infinite “best” (depending on the criterion chosen) estimates for the model parameters. In the Bayesian model calibration framework, the typical sign of non-identifiability is that the posterior PDFs of some of the model parameters are close to the prior PDFs, which indicates that the marginal likelihoods of these parameters are nearly flat and that there is an infinite number of maximum likelihood estimates of the model parameters. Such calibration results may not be desirable for the analyst, since there is little reduction in uncertainty in the model parameters based on the observed data.

In general, model non-identifiability can be classified into two types, namely structural non-identifiability and practical non-identifiability [19]. The first type of non-identifiability, structural non-identifiability is due to the redundant parameterization of the model structure. Even if the model is structurally identifiable, a second type of non-identifiability, practical non-identifiability, can still arise due to the insufficient amount and quality of observation data. The quality of data is related to the bias and noise in the data due to the imprecision of measurement techniques. Successful detection of structural non-identifiability may help reduce model redundancy, i.e., modify the form of the model or fix the values of some model parameters. Also, by detecting the existence of practical non-identifiability, the analyst may be able to overcome this issue by collecting additional data, developing better design of experiments, or improving data quality [20].

It is usually straightforward to detect structural non-identifiability if the analytical expression of a model is available; however, in many problems, the analytical expression of the model is not readily available. One example is a dynamic model without an explicit steady state solution. Another possible case occurs when we add a discrepancy function to the numerical solutions of some governing equations, which in fact forms a new model without any closed-form analytical expression [7]. Various analytical and numerical methods have been developed to detect the structural non-identifiability of dynamic models [21–23], whereas the second possible case does not appear to have been studied in the literature. This section addresses this case.

Since the second type of non-identifiability, practical non-identifiability, is related to both model structure and observation data, it is necessary to inspect the likelihood function in order to determine whether some parameters of a model are practically non-identifiable. In fact, rigorous definitions of model non-identifiability can be constructed based on the analytical properties of likelihood functions [24–26]. In addition to the theoretical analysis of likelihood functions, Raue et al. [19,27] developed a numerical approach based on the concept of “profile likelihood” [28], which has been shown to be effective in detecting practical non-identifiability. When the analytical expression of the likelihood function is available, or its numerical evaluation is trivial, it may be preferable to apply the profile likelihood-based method and determine the practical non-identifiability directly. But this method becomes cumbersome when the construction of the likelihood function is computationally expensive, since repetitive evaluations of the likelihood function are required to compute the profile likelihood.

In this paper, we propose a first-order Taylor series expansion-based method, which can detect structural non-identifiability for models without analytical expressions, and can detect practical non-identifiability due to insufficient amount of data. This method does not involve computing the likelihood function or the posterior probability density of model parameters, and thus is simpler to implement and less computationally demanding. The proposed method can also be viewed as complementary to the well-established global and local sensitivity analysis methods [29,30]. For example, $y = (\theta_1 + \theta_2)x$ is a model with the output y highly sensitive to the two parameters θ_1 and θ_2 . Sensitivity analysis can tell us that both θ_1 and θ_2 have significant main effect as well as strong interaction. By applying the proposed first-order Taylor series expansion-based approach, we can further detect that these two parameters together are non-identifiable.

In Section 2, we present the basic framework of Bayesian model calibration. In Section 3, the five prior formulations of model discrepancy are discussed, and the corresponding likelihood functions are derived. Section 4 focuses on the issue of model parameter non-identifiability caused by adding a model discrepancy function to the original physics model.

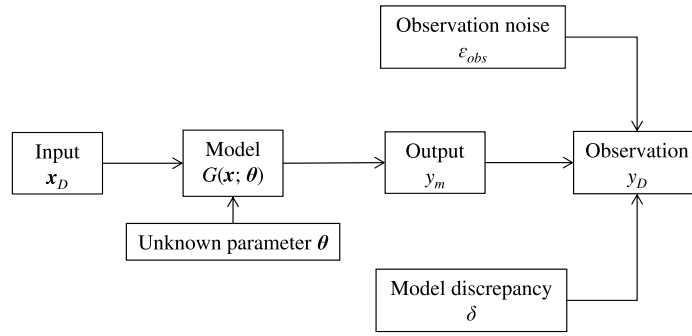


Fig. 1. Relationship between a computer model and corresponding experimental observation.

The first-order Taylor series expansion-based method is developed in order to help choose a model discrepancy function without causing parameter non-identifiability. Section 5 illustrates the three-step method to address uncertainty in model discrepancy formulation. The proposed approach is applied to two numerical examples: (1) calibration of an analytical beam model with synthetic data, and (2) calibration of a dielectric charging model for a microelectromechanical system (MEMS) device [31] with actual experimental data.

2. Bayesian model calibration

Consider a computer model, $y_m = G(\mathbf{x}; \boldsymbol{\theta})$, with input \mathbf{x} , parameter $\boldsymbol{\theta}$, and output y_m . This model is constructed to predict a physical quantity of interest y , which is observable through experiments. The model input \mathbf{x} is a set of quantities that modelers consider unnecessary to calibrate, since these quantities can be either measured in experiments or computed from other physics models. Therefore, the model input \mathbf{x} is set to be \mathbf{x}_D in the process of model calibration, where \mathbf{x}_D is a set of measured/computed values of \mathbf{x} . Note that in the presence of measurement error in experiments, or other sources of uncertainty in the models used to compute \mathbf{x} , some elements of \mathbf{x}_D may be treated as random variables with known probability distributions. In contrast to \mathbf{x} , the model parameter set $\boldsymbol{\theta}$ is considered unknown due to the lack of direct measurement or physical knowledge, and the objective of model calibration is to estimate these parameters based on available input–output observations.

In the KOH framework, the relationship between experimental observation y_D , true value of the quantity y , and model output y_m is described as (Fig. 1)

$$y_D = y + \varepsilon_{obs} \quad (1)$$

$$y = y_m + \delta = G(\mathbf{x}_D, \boldsymbol{\theta}) + \delta(\mathbf{x}_D) \quad (2)$$

where ε_{obs} represents measurement uncertainty and is often treated as a zero-mean Gaussian random variable with variance σ_{obs}^2 . Uncertainty due to model inadequacy is represented by a model discrepancy term δ , which could be a function of model input \mathbf{x}_D . Since σ_{obs} and δ are usually unknown, they may also need to be calibrated. Note that y_D is treated as a random variable in Eq. (1), and the samples of y_D are the actual observation data of y . Assuming m samples of y_D (denoted as $\mathbf{D} = [D_1, D_2, \dots, D_m]$) are collected for a single input setting \mathbf{x}_D , we can calibrate the unknown parameters $\boldsymbol{\theta}$, σ_{obs} , and δ using Bayes' theorem as

$$\pi(\boldsymbol{\theta}, \sigma_{obs}, \delta | \mathbf{D}) = \frac{\mathcal{L}(\boldsymbol{\theta}, \sigma_{obs}, \delta) \pi(\boldsymbol{\theta}) \pi(\sigma_{obs}) \pi(\delta)}{\int \mathcal{L}(\boldsymbol{\theta}, \sigma_{obs}, \delta) \pi(\boldsymbol{\theta}) \pi(\sigma_{obs}) \pi(\delta) d\boldsymbol{\theta} d\sigma_{obs} d\delta}$$

$$\mathcal{L}(\boldsymbol{\theta}, \sigma_{obs}, \delta) \propto \prod_{i=1}^m \pi(y_D = D_i | \mathbf{x}_D, \boldsymbol{\theta}, \sigma_{obs}, \delta) \quad (3)$$

where $\pi(\boldsymbol{\theta})$, $\pi(\sigma_{obs})$, $\pi(\delta)$ are the prior PDFs of $\boldsymbol{\theta}$, σ_{obs} , and δ respectively, representing prior knowledge of these parameters before calibration; $\pi(\boldsymbol{\theta}, \sigma_{obs}, \delta | \mathbf{D})$ is the joint posterior (or calibrated) PDF of $\boldsymbol{\theta}$, σ_{obs} , and δ ; the joint likelihood function of $\boldsymbol{\theta}$, σ_{obs} , and δ , which is denoted as $\mathcal{L}(\boldsymbol{\theta}, \sigma_{obs}, \delta)$, is proportional to the conditional probability of observing the data \mathbf{D} given these parameters. Note that $\pi(\cdot)$ denotes probability density function in this paper.

In the case where experimental observations are taken for multiple input settings, paired data $\{\mathbf{x}_{Di}, y_{Di}\}_{i=1}^n$ may be available. Let $\mathbf{X}_D = [\mathbf{x}_{D1}, \mathbf{x}_{D2}, \dots, \mathbf{x}_{Dn}]$, and the corresponding output observation $\mathbf{y}_D = [y_{D1}, y_{D2}, \dots, y_{Dn}]$. Note that \mathbf{y}_D becomes a random vector. Assuming that m realizations for \mathbf{y}_D (denoted as $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_m]$) are collected in the experiments, the likelihood function in Eq. (3) becomes

$$\mathcal{L}(\boldsymbol{\theta}, \sigma_{obs}, \delta) \propto \prod_{i=1}^m \pi(\mathbf{y}_D = \mathbf{D}_i | \mathbf{X}_D, \boldsymbol{\theta}, \sigma_{obs}, \delta) \quad (4)$$

where $\pi(\mathbf{y}_D = \mathbf{D}_i | \mathbf{X}_D, \boldsymbol{\theta}, \sigma_{obs}, \delta)$ is the joint PDF of the random vector \mathbf{y}_D evaluated at \mathbf{D}_i , and this probability distribution is conditioned on $\mathbf{X}_D, \boldsymbol{\theta}, \sigma_{obs}$, and δ .

In the original KOH framework, both the model (or simulator) $G(\mathbf{x}; \boldsymbol{\theta})$ and the model discrepancy function δ are approximated using Gaussian process models, and these two Gaussian process models are calibrated simultaneously. In this paper, we consider the cases when the physics model is inexpensive to evaluate, or a surrogate model has been built and validated to replace the expensive simulator. Therefore, calibrating the emulator hyperparameters is not considered in this paper.

3. Formulation of model discrepancy function

Section 2 reviewed the basic framework of Bayesian model calibration with a generalized discrepancy term. To establish the likelihood function in Eq. (4), the model discrepancy function δ needs to be formulated prior to calibration. This section explores several options to parameterize model discrepancy, and the corresponding likelihood functions are derived. These formulations can be used to represent the modeler's prior knowledge of model inadequacy, and the coefficients of the parameterized model discrepancy function can be included in Bayesian calibration in order to obtain *a posteriori* estimation of model discrepancy.

3.1. Model discrepancy as a constant bias

The simplest formulation is to treat model discrepancy as a constant, i.e., the bias between model prediction and the real process is assumed to be independent of the input \mathbf{x} . Using the assumption that measurement error $\varepsilon_{obs} \sim \mathcal{N}(0, \sigma_{obs}^2)$ and the linear relationship specified in Eqs. (1) and (2), experimental observation $(\mathbf{y}_D | \mathbf{X}_D, \boldsymbol{\theta}, \sigma_{obs}, \delta)$ is normally distributed:

$$(\mathbf{y}_D | \mathbf{X}_D, \boldsymbol{\theta}, \sigma_{obs}, \delta) \sim \mathcal{N}(G(\mathbf{X}_D; \boldsymbol{\theta}) + \delta, \sigma_{obs}^2 \mathbf{I}) \quad (5)$$

where \mathbf{I} represents identity matrix, i.e., the elements of \mathbf{y}_D are conditionally independent of each other. The likelihood function in Eq. (4) can then be obtained by evaluating the PDF of the above Gaussian distribution with $\mathbf{y}_D = \mathbf{D}_i$.

3.2. Model discrepancy as i.i.d. Gaussian random variables with fixed mean and variance

Bias between model and experiment is rarely an input-independent constant. Instead, it may be different at different input settings. If this variation is purely random, model discrepancy can be treated as independent and identically distributed (i.i.d.) Gaussian random variables over the input domain, with fixed mean μ_δ and variance σ_δ . Therefore, for any input \mathbf{x} , we have

$$\delta(\mathbf{x}) \sim \mathcal{N}(\mu_\delta, \sigma_\delta^2) \quad (6)$$

This type of model discrepancy not only adds a mean correction (μ_δ) to the model prediction but also increases the prediction variance by σ_δ^2 . The conditional probability distribution of \mathbf{y}_D then becomes

$$(\mathbf{y}_D | \mathbf{X}_D, \boldsymbol{\theta}, \sigma_{obs}, \mu_\delta, \sigma_\delta) \sim \mathcal{N}(G(\mathbf{X}_D, \boldsymbol{\theta}) + \mu_\delta, (\sigma_{obs}^2 + \sigma_\delta^2) \mathbf{I}) \quad (7)$$

3.3. Discrepancy as Gaussian random variable with input-dependent mean and variance

If there is some evidence/knowledge suggesting the existence of input dependence in the model discrepancy function, the assumption of i.i.d. random variables may no longer be valid. The input dependence can be explicitly accounted for by assuming the mean and variance of δ to be functions of \mathbf{x} , i.e.,

$$\delta \sim \mathcal{N}(\mu_\delta(\mathbf{x}; \boldsymbol{\phi}), \sigma_\delta^2(\mathbf{x}; \boldsymbol{\varphi})) \quad (8)$$

where $\boldsymbol{\phi}$ is the set of coefficients in the mean function $\mu_\delta(\cdot)$, and $\boldsymbol{\varphi}$ is the set of coefficients in the variance function $\sigma_\delta^2(\cdot)$. The choice of the form of the mean and variance functions can be rather subjective and is typically problem-specific. One such choice will be demonstrated in the numerical example.

In this case, the conditional probability distribution of \mathbf{y}_D becomes

$$(\mathbf{y}_D | \mathbf{X}_D, \boldsymbol{\theta}, \sigma_{obs}, \boldsymbol{\phi}, \boldsymbol{\varphi}) \sim \mathcal{N}(G(\mathbf{X}_D, \boldsymbol{\theta}) + \mu_\delta(\mathbf{X}_D; \boldsymbol{\phi}), [\sigma_{obs}^2 + \sigma_\delta^2(\mathbf{X}_D; \boldsymbol{\varphi})] \mathbf{I}) \quad (9)$$

3.4. Discrepancy as a Gaussian process with stationary covariance function

The aforementioned formulations of model discrepancy in Sections 3.2 and 3.3 assume that model discrepancies at different input points are statistically independent of each other. However, if a model makes a poor prediction at one input point, it is not uncommon to find that it also fails at a nearby input point, which suggests the existence of statistical correlation between model discrepancies at these two input points. Instead of a set of independent random variables, it may be desirable to treat δ as a Gaussian process:

$$\delta \sim \mathcal{GP}(m(\mathbf{x}; \boldsymbol{\phi}), k(\mathbf{x}, \mathbf{x}'; \boldsymbol{\varphi})) \quad (10)$$

where $m(\cdot)$ is the mean function of this Gaussian process δ , and ϕ is the set of coefficients of $m(\cdot)$; $k(\cdot)$ is the covariance function of δ , and φ is the set of coefficients of $k(\cdot)$. Covariance functions with different properties (e.g., order of continuity, stationary/non-stationary, isotropic/anisotropic) have been developed. The most widely used one within the machine learning literature is the squared exponential function, which is infinitely differentiable, stationary, and isotropic [32]. An example squared exponential function can be written as

$$k(\mathbf{x}, \mathbf{x}'; \varphi) = \lambda \exp\left(-\sum_{i=1}^q \frac{(x_i - x'_i)^2}{2l_i^2}\right) \quad (11)$$

where $\varphi = [\lambda, l_1, l_2, \dots, l_q]$, and q is the dimension of the inputs $\mathbf{x} = [x_1, x_2, \dots, x_q]$; λ is the variance of this Gaussian process; l_i is the length-scale parameter corresponding to the input variable x_i . Higher values of l_i indicate higher statistical correlation along the input dimension x_i .

As the discrepancy function is assumed to be a Gaussian process, the elements of \mathbf{y}_D become correlated, and they follow a multivariate Gaussian distribution as

$$(\mathbf{y}_D | \mathbf{X}_D, \theta, \sigma_{obs}, \varphi) \sim \mathcal{N}(G(\mathbf{X}_D; \theta) + m(\mathbf{X}_D; \phi), \Sigma + \sigma_{obs}^2 \mathbf{I}) \quad (12)$$

where

$$m(\mathbf{X}_D; \phi) = \begin{bmatrix} m(\mathbf{x}_{D1}; \phi) \\ \vdots \\ m(\mathbf{x}_{Dn}; \phi) \end{bmatrix}, \quad \Sigma = \begin{bmatrix} k(\mathbf{x}_{D1}, \mathbf{x}_{D1}; \varphi) & \dots & k(\mathbf{x}_{D1}, \mathbf{x}_{Dn}; \varphi) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_{Dn}, \mathbf{x}_{D1}; \varphi) & \dots & k(\mathbf{x}_{Dn}, \mathbf{x}_{Dn}; \varphi) \end{bmatrix} \quad (13)$$

In order to compute the likelihood function $\mathcal{L}(\theta, \sigma_{obs}, \delta)$, the PDF of the multivariate Gaussian distribution in Eq. (12) needs to be evaluated, which requires computing the determinant and inverse of the covariance matrix $\Sigma + \sigma_{obs}^2 \mathbf{I}$. As the number of data points increases, this covariance matrix may become ill-conditioned and lead to significant numerical errors in the computation of the likelihood function. In this paper, we use a Cholesky decomposition-based strategy to compute the determinant and inverse of the covariance matrix, which helps mitigate the numerical difficulty in likelihood evaluation [33].

3.5. Discrepancy as a Gaussian process with non-stationary covariance function

The Gaussian process in Section 3.4 is stationary since the squared exponential covariance function satisfies $k(\mathbf{x}, \mathbf{x}'; \varphi) = k(\mathbf{x} + \Delta\mathbf{x}, \mathbf{x}' + \Delta\mathbf{x}; \varphi)$, i.e., the marginal variances of model discrepancy at different input points are the same. If we consider variance as an indicator of the degree of uncertainty, it may be desirable to allow the variance of model discrepancy to vary with model input, since the uncertainty regarding model discrepancy at an input setting depends on the amount of data points available and the smoothness of the underlying discrepancy function. In order to account for the non-stationary trend of variance in the Gaussian process discrepancy formulation, we can add an input-dependent term to a stationary covariance function. For example, if time is one of the input variables and one has reasons to believe that the variance of discrepancy decreases with time, the following covariance function may be used

$$k(\mathbf{x}, \mathbf{x}'; \varphi) = \begin{cases} k_1(\mathbf{x}, \mathbf{x}'; \varphi_1), & t \neq t' \\ k_1(\mathbf{x}, \mathbf{x}'; \varphi_1) + k_2(t; \mathbf{w}), & t = t' \end{cases} \quad (14)$$

where $k_1(\cdot)$ is the squared exponential function in Eq. (11); the second term $k_2(t; \mathbf{w}) = w_1 \exp(-w_2 t)$ is an exponential decreasing function with respect to time when $w_1 > 0$ and $w_2 > 0$, which essentially reduces the values of the diagonal entries of the covariance matrix Σ in Eq. (13).

Note that other types of non-stationary covariance functions, such as linear/polynomial dot product kernels and neural network kernels [32], can also be used depending on the features of a specific problem.

Substituting the new covariance function into Eq. (13), we can obtain the corresponding covariance matrix, which can be further used to compute the likelihood function as described in Section 3.4.

4. Identifiability of model parameters

The focus of this section is to develop a fast method for the detection of model non-identifiability without actually performing full Bayesian model calibration (which may be computationally expensive).

Suppose the physics model to be calibrated is $y_m = G(\mathbf{x}; \theta)$, and a Gaussian process discrepancy function $\delta \sim \mathcal{GP}(m(\mathbf{x}; \phi), k(\mathbf{x}, \mathbf{x}'; \varphi))$ is added to the model. Thus, a new model is formed as $G_{new}(\mathbf{x}; \psi) = G(\mathbf{x}; \theta) + m(\mathbf{x}; \phi)$, where $\psi = [\theta, \phi]$ includes the parameters of the physics model (θ) and the parameters of the mean of the discrepancy function (ϕ). Note that the covariance function $k(\mathbf{x}, \mathbf{x}'; \varphi)$ is not included in the new model, since the proposed method is not applicable to statistical models. In the case that the measurement noise is a zero mean random variable, G_{new} is the expectation of observation y_D according to Eqs. (1) and (2). Further assume that the analytical form of G_{new} is not avail-

able. In such cases, the first-order Taylor series expansion of this model (as shown in Eq. (15)) can be used as an efficient approximation when the model is believed to be not highly nonlinear:

$$E[y_D] = G_{new}(\mathbf{x}_D; \boldsymbol{\psi}) \approx G_{new}(\mathbf{x}_D; \hat{\boldsymbol{\psi}}) + \sum_{i=1}^p \left. \frac{\partial G_{new}}{\partial \psi_i} \right|_{\boldsymbol{\psi}=\hat{\boldsymbol{\psi}}} \psi_i \quad (15)$$

where $\hat{\boldsymbol{\psi}}$ is the point of expansion, and p is the number of model parameters. In this paper, we use the mean of the prior of $\boldsymbol{\psi}$ as $\hat{\boldsymbol{\psi}}$ since it is readily available. A better option for $\hat{\boldsymbol{\psi}}$ may be the maximum *a posteriori* (MAP) estimation of the parameters if it is available. To find the MAP estimation for the parameters of a general model, global optimization algorithms such as simulated annealing [34] and particle swarm optimization [35] can be used, which may require hundreds of model evaluations to reach convergence.

Suppose there are n data points available, i.e., experimentally observed values $\mathbf{D} = [D_1, D_2, \dots, D_n]$ corresponding to different input settings $\mathbf{X}_D = [\mathbf{x}_{D1}, \mathbf{x}_{D2}, \dots, \mathbf{x}_{Dn}]$. Without considering measurement uncertainty and the variance of the model discrepancy function, we can construct a linear system as

$$\mathbf{A}\boldsymbol{\psi}^T = \mathbf{D} - G_{new}(\mathbf{x}_D; \hat{\boldsymbol{\psi}}), \quad \mathbf{A} = \begin{bmatrix} \left. \frac{\partial G_{new}}{\partial \psi_1} \right|_{\mathbf{x}_{D1}, \hat{\boldsymbol{\psi}}} & \cdots & \left. \frac{\partial G_{new}}{\partial \psi_p} \right|_{\mathbf{x}_{D1}, \hat{\boldsymbol{\psi}}} \\ \vdots & \ddots & \vdots \\ \left. \frac{\partial G_{new}}{\partial \psi_1} \right|_{\mathbf{x}_{Dn}, \hat{\boldsymbol{\psi}}} & \cdots & \left. \frac{\partial G_{new}}{\partial \psi_p} \right|_{\mathbf{x}_{Dn}, \hat{\boldsymbol{\psi}}} \end{bmatrix} \quad (16)$$

The linear system in Eq. (16) can be underdetermined or determined, depending on the rank of the matrix \mathbf{A} (denoted as r_A). If $r_A < p$, the system is underdetermined and there will be an infinite number of $\boldsymbol{\psi}$ values satisfying Eq. (16); if $r_A = p$, the system is determined and there will be a unique vector $\boldsymbol{\psi}$ satisfying Eq. (16). The latter case suggests that the model is practically identifiable given the available data points (assuming the quality of the data does not cause non-identifiability). The former case suggests that the model is non-identifiable either due to the model structure or insufficient data. If the inequality $r_A < p$ continues to hold as we increase the number of observation data, then it can be inferred that the model is structurally non-identifiable.

In order to help the analyst reduce model redundancy once a model is detected as structurally non-identifiable, it may be of interest to know which set of parameters can/cannot be identified. Using the formulation of the linear system in Eq. (16), we can retrieve this information by checking the linear dependence between the column vectors of the matrix \mathbf{A} , since the i -th column of \mathbf{A} corresponds to the parameter ψ_i . For example, if the i -th column vector \mathbf{a}_i and the j -th column vector \mathbf{a}_j are linearly dependent, it is apparent that the corresponding parameters ψ_i and ψ_j are non-identifiable using the linear model in Eq. (15). We can also find one set of identifiable parameters using the simple algorithm below:

Algorithm 1 Find one set of identifiable parameters.

Input: The first-order derivative matrix \mathbf{A}

Output: The index set of identifiable parameters \mathbf{I}

```

 $\mathbf{A}_{temp} = \mathbf{A}$ 
 $\mathbf{I} = \text{empty set}$ 
for  $i = 1$  to  $p - 1$  do
     $r_1 = \text{the rank of } \mathbf{A}_{temp}$ 
    Remove the first column from  $\mathbf{A}_{temp}$ 
     $r_2 = \text{the rank of } \mathbf{A}_{temp}$  (with the first column removed)
    if  $r_1 > r_2$  then
        Add the value of  $i$  to the set  $\mathbf{I}$  as an element
    end if
    if (the number of elements in  $\mathbf{I}$ ) =  $r_A$  then
        Break
    else
        if  $i = p - 1$  then
            Add the value of  $i + 1$  to the set  $\mathbf{I}$  as an element
        end if
    end if
end for
return  $\mathbf{I}$ 

```

Algorithm 1 takes one column from the matrix \mathbf{A}_{temp} ($= \mathbf{A}$ at the beginning) at a time, and saves the column index only if this column is linearly independent of all the remaining columns in \mathbf{A}_{temp} . By repeating this procedure until the number of saved column indices is equal to the rank of \mathbf{A} (r_A), we obtain r_A number of linearly independent columns from the original matrix \mathbf{A} , which represents a set of identifiable parameters. Note that there may be multiple sets of identifiable parameters. To identify all the sets, we can permute the columns of \mathbf{A} , and apply the algorithm to each of the $p!$ permuted matrices. In practice, usually only one set of identifiable parameters is needed, and the analyst can reformulate the model with this set of identifiable parameters, by either removing the other parameters from the model formulation, fixing the values of the other parameters, or combining the parameters between which there are strong evidences of linear correlation. It should also be noted that in the presence of significant model uncertainty, the analyst may choose to keep

the parameter set with non-identifiability, since the tradeoffs between the physics model and the model discrepancy term introduce additional uncertainty in model parameters and model form.

In order to illustrate the proposed method, consider a simple example:

$$y_m = \theta_1 x + \theta_2 x^2 \quad (17)$$

$$\delta = \phi_1 + \phi_2 x \quad (18)$$

It is not difficult to see from Eq. (17) that the two parameters of model (θ_1 and θ_2) are identifiable given no less than two pairs of input–output data. However, if we add the model discrepancy δ in Eq. (18) to the model in Eq. (17), θ_1 and ϕ_2 will become unidentifiable no matter how many data points are available.

Let $\psi = [\theta_1, \theta_2, \phi_1, \phi_2]$. Suppose measurements are available at five input points: $x_D^1 = -1.0$, $x_D^2 = 2.0$, $x_D^3 = 3.5$, $x_D^4 = 4.0$, and $x_D^5 = 6.0$ (superscripts refer to data point number). We can calculate the derivatives $\partial(y_m + \delta)/\partial\psi_i$ numerically (e.g., forward difference or central difference) at these input points for given values of the parameters, and thus obtain the matrix \mathbf{A} . For example, for $[\theta_1, \theta_2, \phi_1, \phi_2] = [2.0, 0.5, 4.0, 1.5]$, the matrix \mathbf{A} for the combination of the physics model and δ is

$$\mathbf{A} = \begin{bmatrix} -1.0 & 1.0 & 1.0 & -1.0 \\ 2.0 & 4.0 & 1.0 & 2.0 \\ 3.5 & 12.25 & 1.0 & 3.5 \\ 4.0 & 16.0 & 1.0 & 4.0 \\ 6.0 & 36.0 & 1.0 & 6.0 \end{bmatrix}$$

The rank of \mathbf{A} is equal to 3, indicating that the parameters become unidentifiable after the combination of y_m and δ . We can also use the program in Algorithm 1 to infer that the parameters will become identifiable once the term $\phi_1 x$ is removed from the model discrepancy function.

It should be noted that the proposed method for model identifiability detection has several limitations: (1) it uses a linear approximation of the model, and hence may fail to detect non-identifiability if the model is highly nonlinear; (2) it can only detect local non-identifiability as the Taylor series expansion is constructed based on the derivatives at a single point. It is possible that some parameters may be non-identifiable in certain regions of the parameter space but become identifiable in other regions. In such cases, we suggest to examine the rank of \mathbf{A} for multiple choices of ψ ; (3) it is applicable only when the output of a model is deterministic for given values of the model inputs and parameters, i.e., it is not applicable to statistical models; and (4) it does not cover practical non-identifiability due to the *quality* of data.

5. Assessment and combination of calibration results using a three-step approach

If the discrepancy term captures all the missing physics in a model, the posterior distribution of model parameters is expected to converge to the true distribution given sufficient data. However, precise knowledge of model discrepancy is rarely achievable in practice. In this section, we propose a three-step heuristic approach based on a quantitative model validation metric and Bayesian model averaging, as explained below.

- (1) The available experimental data are partitioned into two sets, one of which will be used for calibration (denoted as $\{\mathbf{X}_D^C, \mathbf{y}_D^C\}$) whereas the other one will be used for validation (denoted as $\{\mathbf{X}_D^V, \mathbf{y}_D^V\}$).
- (2) We perform Bayesian calibration on model parameters with the various formulations of model discrepancy function using data set D_C as discussed in Section 3. By imposing the i -th prior formulation of the discrepancy function (denoted as δ_i) and adding it to the model G , we obtain the corrected model $M_i = G + \delta_i$. The corresponding joint posterior distribution of model parameters and discrepancy is denoted as $\pi(\theta, \delta | D_C, M_i)$. Model predictions based on $\pi(\theta, \delta | D_C, M_i)$ are then validated using a reliability-based metric, which calculates the probability of model predictions being within a specified distance from the validation data.
- (3) The probability of model prediction satisfying a specified tolerance can be used either to select the best model discrepancy formulation, or can be further used to obtain an “average” posterior distribution of model parameters based on Bayesian model averaging.

5.1. Reliability-based model validation metric

The purpose of validation activity in this section is to assess the quality of predictions resulting from calibration with different prior formulations of the model discrepancy function. The reliability-based metric was proposed by Rebba and Mahadevan [18] as a measure of model predictive capability. It is defined as the probability of the absolute difference between model prediction and observed data being less than a specified tolerance ϵ , i.e.,

$$r(\mathbf{x}) = \Pr(|y_D - y_m| < \epsilon) \quad (19)$$

where $r(\mathbf{x})$ is the reliability metric for a given input point \mathbf{x} within the validation domain, i.e., $\mathbf{x} \in \mathbf{X}_D^V$; y_D is the observation corresponding to \mathbf{x} and thus $y_D \in \mathbf{y}_D^V$; y_m is the model prediction at \mathbf{x} . The tolerance limit ϵ represents the level of acceptable prediction error to the analyst, i.e., larger values of ϵ correspond to lower accuracy requirement. Subject domain matter

knowledge is needed to choose the appropriate ϵ for further decision-making. Some discussion on comparing model validation metric to accuracy requirement can be found in [36]. The computation of $r(\mathbf{x})$ requires the probability distributions of y_D and y_m , which are discussed below.

We first discuss how to obtain the prediction y_m and the corresponding probability distribution. As illustrated in Sections 2 and 3, the computer model $G(\mathbf{x}; \theta)$ and the discrepancy function $\delta(\mathbf{x})$ can be calibrated using the observed values of \mathbf{y}_D^C (denoted as \mathbf{D}_C), and then we obtain the posterior probability distribution $\pi(\theta, \delta | \mathbf{D}_C)$ of model parameters and δ . The prediction y_m in Eq. (19) is based on the extrapolation of the calibrated computer model and discrepancy function into the validation domain, i.e.,

$$y_m | \mathbf{x}, \theta, \delta = G(\mathbf{x}; \theta) + \delta(\mathbf{x}), \quad \mathbf{x} \in \mathbf{X}_D^V$$

$$\pi(y_m | \mathbf{x}) = \int \pi(y_m | \mathbf{x}, \theta, \delta) \pi(\theta, \delta | \mathbf{D}_C) d\theta d\delta \quad (20)$$

The integration in Eq. (20) can be numerically evaluated using Monte Carlo integration method [37].

Eq. (1) can be used to derive the probability distribution of y_D . If the true value y is treated as a constant for the given input \mathbf{x} , and the measurement error is treated as a Gaussian random variable with zero mean and fixed variance, y_D will also be a Gaussian random variable. The mean and variance of y_D can be estimated using repetitive measurements at the same input \mathbf{x} . If only one measurement is taken at \mathbf{x} , we may assume that $y_D \sim \mathcal{N}(D, \sigma_{obs}^2)$, where D is the measured value of y_D , and σ_{obs}^2 is the variance of the measurement error estimated in the calibration step, i.e., the measurement error is assumed to be the same in the calibration and validation domain.

Based on the posterior distribution of the model prediction y_m and the measurement y_D , the reliability-based metric can be rewritten as

$$r(\mathbf{x}) = \int_{|y_D - y_m| < \epsilon} \pi(y_D | \mathbf{x}) \pi(y_m | \mathbf{x}) dy_D dy_m \quad (21)$$

Here the integral can also be evaluated using Monte Carlo integration.

Eq. (19) defines the reliability metric as a function of the input \mathbf{x} , and we can compute the values of r at each point within the validation domain, e.g., $r(\mathbf{x}_{D1}), r(\mathbf{x}_{D2}), \dots$. These individual values of r can inform decisions on point-wise model selection [38]. However, the focus of this section is to assess the formulations of model discrepancy in Bayesian calibration based on validation results. Thus, it is desirable to establish a single measure of model predictive capability over the entire validation domain, which can be achieved by taking a weighted average of all the reliability values of the model

$$\mu_r = \frac{\sum_i w_i r(\mathbf{x}_{Di})}{\sum_i w_i} \quad (22)$$

where the weight w_i is proportional to the probability density of \mathbf{x} , e.g., $w_i \propto 1$ if the input points are uniformly distributed. The averaged reliability μ_r represents the expected probability of prediction error being within tolerance over the validation domain, and thus can be considered as an overall validation metric. Note that the weighted average of model reliability in Eq. (22) does not impose a huge penalty on models with bad prediction at one point. In a complicated physics problem where model uncertainty may be significant, it is likely that none of the available models can achieve nontrivial reliability at every validation point. In such cases, the proposed model validation metric may be more appropriate since it evaluates the global performance of the model and does not become zero because of poor model predictions at a few points. In a more comprehensive model validation practice, in addition to the overall validation metric, the analyst may also need to evaluate the performance of the model at different regions of the validation domain using the original reliability metric, which has explicit dependence on the input \mathbf{x} .

By substituting the results of Bayesian calibration into Eq. (20), the validation metric μ_r can be computed with respect to each of the various model discrepancy formulations illustrated in Section 3. If the μ_r corresponding to a particular formulation δ_i is significantly higher than the others, it is suggested that the calibration using δ_i leads to better prediction in the validation domain, and thus δ_i may be the best approximation to the actual model discrepancy. However, different formulations can sometimes lead to similar values of μ_r , in which case one may use a weighted averaging approach as discussed below.

5.2. Combining the posterior probability distributions of model parameters and discrepancy

When predictions based on different formulations of model discrepancy function lead to similar validation results, it is not clear which formulation of δ and the corresponding calibration result one should select for further prediction. In such cases, we suggest combining the various posterior distributions of model parameters and discrepancy based on a weighted averaging approach [39]. This approach accounts for the uncertainty induced by the lack of knowledge regarding the model discrepancy formulation, and can be viewed as an extension to the Bayesian framework proposed by Sankararaman and Mahadevan [40], which demonstrated the use of model validation results to combine the prior and posterior probability distributions of model parameters in multi-level systems [41,42].

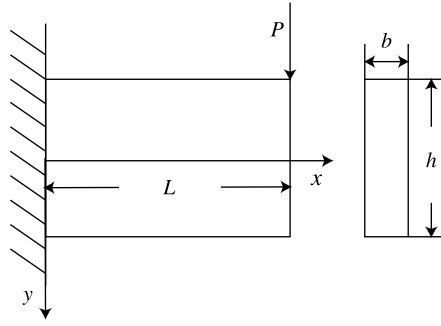


Fig. 2. A thick cantilever subjected to point load at the free end.

Since the overall reliability metric μ_r described in the previous section is a probabilistic metric which quantifies the predictive capability of M_i , we can obtain the probabilistic weight w_{M_i} of each formulation by normalizing μ_r^i as

$$w_{M_i} = \frac{\mu_r^i}{\sum_{j=1}^N \mu_r^j} \quad (23)$$

Each M_i corresponds to a posterior probability distribution of model parameters and discrepancy $\pi(\theta, \delta | \mathbf{D}_C, M_i)$. Based on the probabilistic weights estimated in Eq. (23), the density functions of these posterior distributions can be combined to obtain an averaged density function $\pi(\theta, \delta | \mathbf{D}_C)$ as

$$\pi(\theta, \delta | \mathbf{D}_C) = \sum_{i=1}^N \pi(\theta, \delta | \mathbf{D}_C, M_i) w_{M_i} \quad (24)$$

Note that the probability density function $\pi(\theta, \delta | \mathbf{D}_C)$ is not conditioned on any specific model discrepancy formulation, and is expected to be wider than the conditional density functions $\{\pi(\theta, \delta | \mathbf{D}_C, M_i)\}_{i=1}^N$ since the uncertainty due to model discrepancy formulation is included.

6. Numerical examples

We investigate the various options of model discrepancy formulation using two numerical examples. In Section 6.1, calibration of Young's modulus using an analytical model of a cantilever beam with synthetic data is considered. In Section 6.2, we calibrate a more complicated physics model (MEMS device dielectric charging) with actual experimental data. In both examples, the Metropolis–Hastings Markov chain Monte Carlo (MCMC) method [43] is used to generate samples from the posterior distribution.

6.1. Calibration of Young's modulus using an Euler–Bernoulli beam model

In this section, we will examine the effect of model discrepancy formulation on the calibration of Young's modulus using an Euler–Bernoulli beam model. A thick microcantilever beam subjected to a point load P is considered (Fig. 2, $L = h = 400 \mu\text{m}$, $b = 1 \mu\text{m}$). By solving the Euler–Bernoulli equation, the static vertical deflection u_y along the x -axis ($y = 0$) can be obtained as [44]

$$u_y = \frac{P(3L - x)x^2}{6EI} \quad (25)$$

The solution in Eq. (25) does not account for shear deformation. A more accurate solution of u_y can be obtained from the Timoshenko beam theory as [45,46]

$$u_y = \frac{P}{6EI} \left[(4 + 5\nu) \frac{h^2 x}{4} + (3L - x)x^2 \right] \quad (26)$$

In this example, we assume for the sake of illustration that the solution in Eq. (26) represents the true value of u_y . 20 experimental data points are synthetically generated by adding a measurement noise term $\varepsilon_{obs} \sim \mathcal{N}(0, 0.1)$ to Eq. (26) with Young's modulus $E = 200 \text{ GPa}$ and the applied load $P = 2.5 \mu\text{N}$. Note that this value of Young's modulus and the “true” solution in Eq. (26) are assumed to be unknown in the calibration exercise. The Euler–Bernoulli beam solution in Eq. (25) is assumed to be the only available physics model, and the synthetic data are used to calibrate E using Eq. (25). The discrepancy function between the physics model in Eq. (25) and the “true” solution in Eq. (26) is a linear function of x as

$$\delta_{true} = \frac{P}{6EI} (4 + 5\nu) \frac{h^2 x}{4} \quad (27)$$

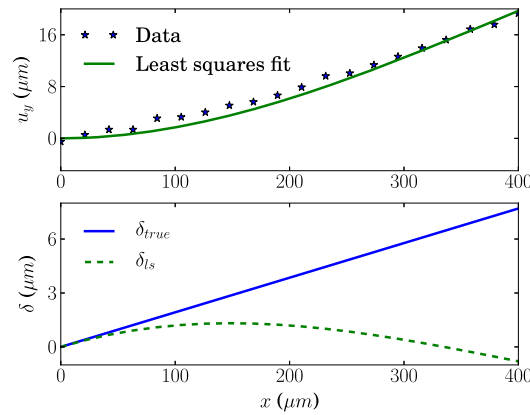


Fig. 3. Actual model discrepancy versus least squares model discrepancy.

Assuming δ_{true} in Eq. (27) is unknown, a common practice to acquire prior knowledge of model discrepancy is by examining the difference between calibration data and the least-squares model prediction (i.e., prediction using Eq. (25) with E estimated using the least-squares method). The upper subplot of Fig. 3 shows the calibration data and least-squares model prediction, and the lower subplot shows the actual model discrepancy δ_{true} and the least-squares model discrepancy δ_{ls} . We can observe that both δ_{true} and δ_{ls} have functional dependence on the input x . However, δ_{ls} appears to be a nonlinear function of x , whereas the “true” model discrepancy δ_{true} is a linear function of x , which suggests that we may not be able to infer the correct form of the model discrepancy function from δ_{ls} .

In order to study the effect of model discrepancy assumptions on the calibration results, we consider the various options discussed in Section 3: (1) an unknown constant δ_1 ; (2) i.i.d. Gaussian random variable with input-independent mean and variance, i.e., $\delta_2 \sim \mathcal{N}(\mu_{\delta_2}, \sigma_{\delta_2}^2)$; (3) Gaussian random variable with input-dependent mean and variance, i.e., $\delta_3 \sim \mathcal{N}(\mu_{\delta_3}(x), \sigma_{\delta_3}(x))$; (4) a Gaussian process with squared exponential covariance function, i.e., $\delta_4 \sim \mathcal{N}(m(x), k(x, x'))$. Gaussian process with non-stationary covariance function as discussed in Section 3.5 is not considered here, since the “true” model discrepancy is just a simple function and there is no need to overparameterize δ . The dependence of model discrepancy on the input x which we observed in the plot of δ_{ls} can be accounted for by using δ_3 or δ_4 , since both options include input-dependent mean and variance function. In this example, $\mu_{\delta_3}(x)$, $\sigma_{\delta_3}(x)$, and $m(x)$ are assumed to be linear functions of x . Note that if these functions are assumed to be third degree polynomials of x , the model will become unidentifiable. It can be easily verified using the first-order Taylor series expansion-based method illustrated in Section 4. For the combination of u_y in Eq. (25) and a third degree polynomial $\phi_1 + \phi_2 x + \phi_3 x^2 + \phi_4 x^3$, the corresponding matrix \mathbf{A} is computed, and the rank $r_A = 4$, which is less than the total number of parameters ($= 5$). By applying Algorithm 1 to randomly permuted \mathbf{A} , it is found that E , ϕ_3 , and ϕ_4 can not be identified at the same time, i.e., one of them will need to be removed from the formulation to ensure the identifiability of the model.

The unknown coefficients in the various model discrepancy formulations are calibrated along with Young's modulus E , and independent uniform prior distributions are used (the prior range of E is [50 GPa, 300 GPa]). The marginal posterior PDFs of Young's modulus corresponding to the various options of model discrepancy are shown in Fig. 4. It can be observed that calibration without model discrepancy results in the most biased estimation of E , which is expected since the physics model is imperfect. The posteriors of E obtained from calibration with δ_1 or δ_2 are also away from the true value, since these two options assume that model discrepancy is input-independent whereas the actual δ is input-dependent. (Note that the use of δ_2 leads to a slightly wider posterior of E compared to δ_1 , since treating model discrepancy as i.i.d. random variables (δ_2) overestimates the uncertainty due to model error.) Calibration with δ_3 or δ_4 gives similar posteriors of E , and both posterior PDFs cover the true value. We can also observe that these two posterior PDFs are significantly wider than the others, which is mainly due to the uncertainty regarding the coefficients of $\mu_{\delta_3}(x)$ and $m(x)$. If stronger priors are enforced on these coefficients (i.e., the ranges of their uniform prior distributions are set to be smaller), the posterior variance of E will decrease correspondingly.

Due to the simplicity of the beam models, the calibrated model with model discrepancy corrections match well with the calibration data. In order to assess the various options of model discrepancy formulations, we examine the predictive capability of the calibrated model outside the domain of calibration ($P = 2.5 \mu\text{N}$). 20 validation data points are generated with the applied load $P = 3.5 \mu\text{N}$, and the three-step approach proposed in Section 5 is implemented. A graphical comparison between the validation data and the calibrated model predictions ($M_i = G + \delta_i$, $i = 0, 1, 2, 3, 4$) is shown in Fig. 5(a)–(e), where the solid lines represent mean prediction and the dashed lines are 90% probability bounds.

It is interesting to observe from Fig. 5 that the predictions M_3 and M_4 are outperformed by the other three in the validation domain, although M_3 and M_4 include better estimations of E and more accurate assumptions on model discrepancy. The reason why M_3 and M_4 fail to match the validation data is that δ_3 and δ_4 only depend on x , whereas the actual δ is a function of both x and P . Thus, when we extrapolate the model from the calibration domain ($P = 2.5 \mu\text{N}$) to the validation domain ($P = 3.5 \mu\text{N}$), δ_3 and δ_4 are no longer good representations of model discrepancy. Since the Euler–Bernoulli model

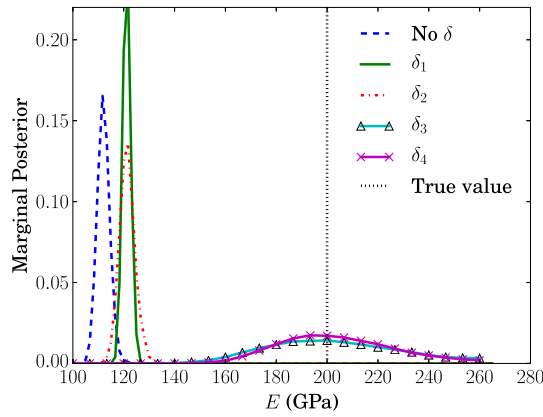


Fig. 4. Marginal posterior PDFs of Young's modulus ("True" value = 200 GPa).

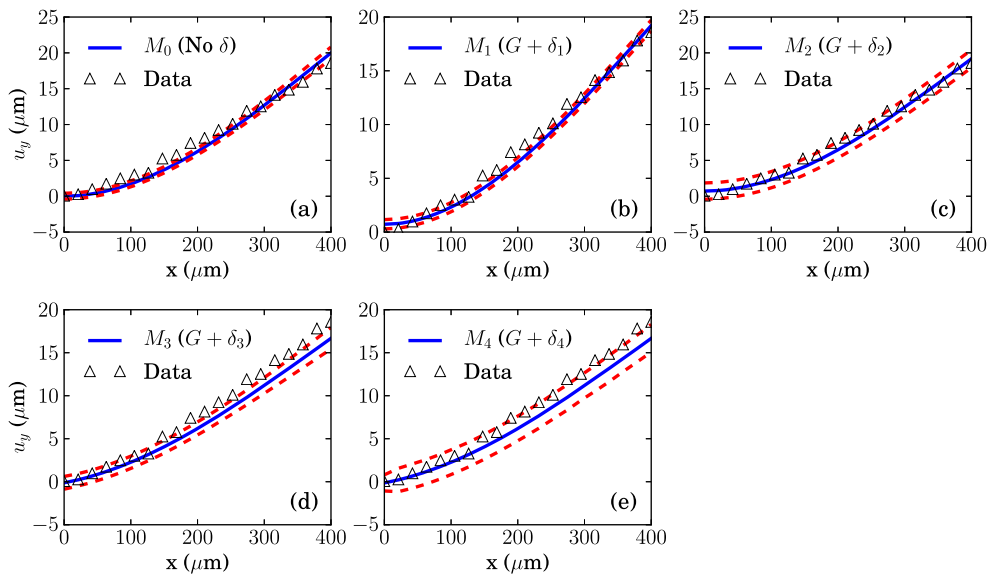


Fig. 5. Comparison between the predictions of calibrated beam model and validation data.

Table 1

Overall reliability of model predictions (calibration data collected with applied load $P = 2.5 \mu\text{N}$).

	M_0 (No δ)	M_1	M_2	M_3	M_4
μ_r^i	0.50	0.66	0.55	0.36	0.34

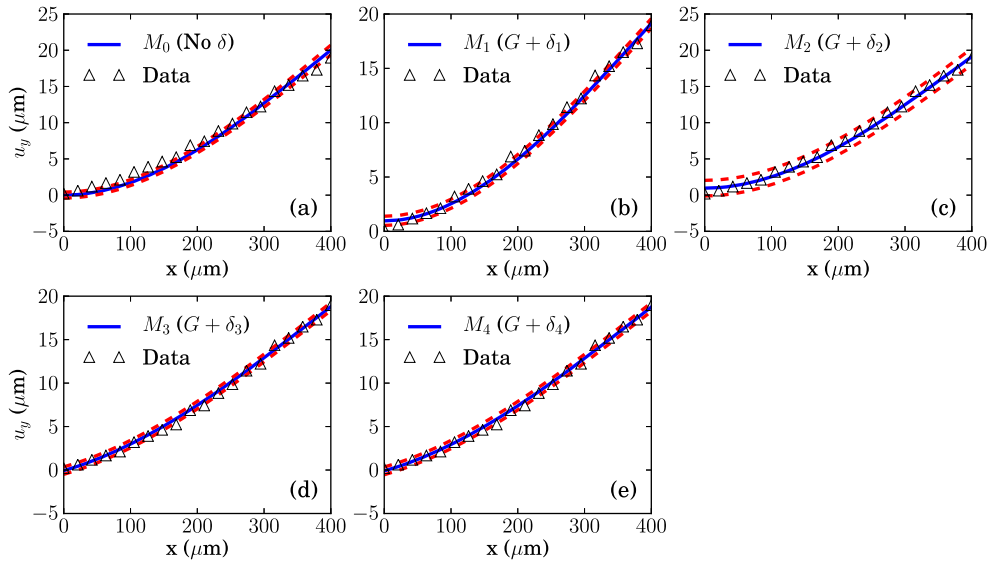
even with the true value of E has significant errors as shown in Fig. 3, inaccurate estimation of the model error will not be able to correct the model prediction. In contrast, calibration without δ or with a simple formulation of δ may produce a model with relatively small errors (e.g., δ_{ls} in Fig. 3), and the prediction is more sensitive to the physics model than to the model discrepancy function δ . Consequently, we may see more consistent performance from the models calibrated with simple δ or without δ given that the behavior of the physical quantity of interest does not change dramatically outside the calibration domain (e.g., from linear to highly nonlinear).

Table 1 lists the values of the overall validation metric μ_r computed with the tolerance level $\epsilon = 0.75 \mu\text{m}$ for the model predictions plotted in Fig. 5. The model calibrated with δ_1 (constant bias) is found to have the highest value of reliability. If obtaining the best possible prediction is the main objective, the constant bias assumption is preferred for model discrepancy in this example.

Combining the various posterior distributions of E and δ via Eq. (24) accounts for the uncertainty in model discrepancy formulation. However, the outcome of this combination will be a multi-modal distribution as suggested in Fig. 4, which has little physics interpretation. In addition, if high prediction accuracy is desired, none of the validation results showed in Table 1 would be satisfying. In such cases, we need to look for possible improvement of the model discrepancy formulation

Table 2Overall reliability of model predictions (calibration data collected with applied load $P = 2.5 \mu\text{N}$, and $P = 3.0 \mu\text{N}$).

	M_0 (No δ)	M_1	M_2	M_3	M_4
μ_r^i	0.50	0.75	0.60	0.95	0.93

**Fig. 6.** Comparison between the predictions of calibrated beam model (with improved model discrepancy formulation) and validation data.

in order to reduce the model form uncertainty. In this example, the direction of improvement is apparent, since the validation results suggest that the model discrepancy also depends on the applied load P . We can modify δ_3 and δ_4 so that their mean and variance/covariance functions are also load-dependent. Note that data corresponding to multiple load values are needed in order to identify the load-dependent model discrepancy. For the purpose of illustration, another 20 data points are generated with $P = 3.0 \mu\text{N}$ in addition to the data set generated with $P = 2.5 \mu\text{N}$. Then, we perform calibration using all the 40 data points and the five model discrepancy formulations (δ_3 and δ_4 are modified). The predictions from the calibrated models are again validated using the data set generated with $P = 3.5 \mu\text{N}$. As shown in Table 2 and Fig. 6, the predictions based on δ_3 and δ_4 have improved notably and are significantly better than the other three options. Since these two formulations lead to similar estimations of the physical parameter E and also similar predictions, selection between them will not make much difference in this example.

6.2. Calibration of dielectric charging parameters using a compact model

Dielectric charging has been identified as an important failure mechanism of radio frequency MEMS switches, causing the switches to either remain stuck or fail to actuate [47]. In this section, we will focus on the calibration of a compact dielectric charging model developed in [31]. The model has three input variables (voltage V , temperature T , and time t), seven unknown parameters (trap density N_T , barrier height Φ_B , capture cross section σ , Frenkel–Poole (FP) attempt frequency γ , high frequency dielectric constant ϵ_{INF} , effective mass m^* , and trap activation energy E_A), and a single output variable (transient current density J). Experiments were conducted on a 200-nm silicon nitride (Si_3N_4) dielectric with $2 \text{ mm} \times 2 \text{ mm}$ area for 12 different combinations of V and T , and the transient current density was measured at about 190 discrete time points between 0 and 100 s. These experiments were repeated four times, and thus a data set with size $n = 12 \times 4 \times 190 = 9120$ is available. Assuming that the measurement error depends only on V and T , we can estimate the variance of measurement error for a particular combination of V and T as

$$\sigma_{obs}^2 = \frac{1}{n_1(n_2 - 1)} \sum_{j=1}^{n_1} \sum_{i=1}^{n_2} \left[J_{D,t_j}^i - \frac{1}{n_2} \sum_{i=1}^{n_2} (J_{D,t_j}^i) \right]^2 \quad (28)$$

where n_1 is the number of time points, and n_2 is the number of repeated experiments. Note that if Bayesian inference is used to estimate σ_{obs} , with different formulations of model discrepancy we may obtain different posterior distributions of σ_{obs} . In such cases, physics domain knowledge on the measurement system as well as subjective judgment will be needed to combine the posterior distributions of σ_{obs} into a single distribution so that σ_{obs} can be independent of model discrepancy formulation.

Table 3

Prior ranges of dielectric charging model parameters.

$\log(N_T)$	Φ_B	$\log(\sigma)$	$\log(\gamma)$	ε_{INF}	m^*	E_A
[55.3, 61.5]	[0.8, 1.4]	[−55.3, −50.6]	[16.1, 25.3]	[0.1, 15.0]	[0.01, 1.0]	[0.01, 1.0]

We again consider the various options of model discrepancy discussed in Section 3, and the same notations are used as in Section 6.1 (i.e., $\delta_1, \delta_2, \dots$). Note that we also include δ_5 as a Gaussian process with non-stationary covariance function. The mean function and variance function of δ_3 are chosen to be linear functions of the model input combined with exponential functions of time as

$$\begin{aligned}\mu_{\delta_3} &= \phi_1 V + \phi_2 T + \phi_3 t + \phi_4 \exp(\phi_5 t) \\ \sigma_{\delta_3}^2 &= \varphi_0 + \varphi_1 V + \varphi_2 T + \varphi_3 t + \varphi_4 \exp(\varphi_5 t)\end{aligned}\quad (29)$$

The same form of mean function is used in δ_4 and δ_5 . We still use a squared exponential function as the covariance function of δ_4 , and Eq. (14) is used as the covariance function of δ_5 . Based on the first-order Taylor series expansion-based method developed in Section 4, a check of identifiability is performed while selecting the form of the mean function μ_{δ_3} , since the addition of the discrepancy function to the original model may cause non-identifiability. In this example, there are 12 unknown parameters in $G + \mu_{\delta_3}$, i.e., $p = 12$, and the corresponding matrix \mathbf{A} is full rank, which suggests that the combination of the dielectric model and this mean function is identifiable. In fact, the reason that there is no constant term in the mean function is because the constant term was found to be unidentifiable.

Note that the model needs to be evaluated at all the calibration experiment input points in order to compute the likelihood function for a given parameter set. Due to limited computational resources, we use only a subset of the available experimental data (40 time points for each combination of V and T ; total number of data points = $12 * 40 = 480$). As discussed in Section 5, we further partition the selected data set into two sets (each with 240 data points), one for calibration and the other for validation. The calibration and validation data points are uniformly distributed in time domain, i.e., for each combination of V and T , the points corresponding to time t_1, t_3, \dots, t_{39} belong to the calibration data and the remaining 20 points belong to the validation data set. Therefore, there is no significant extrapolation from the calibration domain in this example.

Uniform priors are used for all the calibration parameters, and the ranges of the 7 physical parameters are listed in Table 3. 500,000 MCMC samples are collected and the fitted marginal posterior PDFs of the dielectric charging model parameters are shown in Fig. 7. Note that calibration using δ_1 cannot be implemented in this example because the likelihood function remains zero for any set of parameters, which suggests that the difference between the model and data is input-dependent and the assumption of constant model error is not valid. It can be observed that the various choices of model discrepancy give significantly different posteriors for some parameters. However, it is not clear which option is better until we perform the validation step.

The aforementioned validation data set (240 data points) is used to assess the predictive capability of the calibrated models. The reliability metrics are shown in Table 4. For the purpose of illustration, we also show a graphical comparison between model predictions and data in Fig. 8(a)–(e). Due to the limit of space, only the plots corresponding to a single combination of V and T ($V = 20$ V and $T = 360$ K) are shown.

The graphical comparison and the validation metric suggest that calibration without δ or with δ_3 (input-dependent Gaussian random variables) gives the best posterior prediction of current density. However, it can be observed from Fig. 8 that the 95% probability bounds (dashed line) of M_0 and M_3 are both small, whereas the probability bounds of M_2 , M_4 , and M_5 give better coverage of the validation data. Thus, it is preferable to conduct the third step of the proposed method in Section 5.2, which combines the various posteriors of model parameters and model discrepancy based on the validation results. The model prediction based on the combined distribution of model parameters and δ is shown in Fig. 8(f), of which the probability bound also gives a good coverage of the data.

6.3. Discussion

In this section, we demonstrated via two examples how the various choices of model discrepancy function (δ) can affect the calibration result, and how the three-step approach proposed in Section 5 can help the selection of the appropriate discrepancy formulation. In the first example, we showed that calibration with inaccurate assumption on δ can lead to biased estimation of physical parameters. However, posterior predictions based on simpler and less accurate assumptions of δ can still match data well outside the calibration domain. This is because the prediction becomes more sensitive to δ if complicated formulations (i.e., δ_3 and δ_4) are used, and thus the capability of δ to extrapolate from the calibration domain to prediction domain becomes crucial for achieving accurate predictions. Significant improvement in prediction accuracy was observed after we modified δ_3 and δ_4 by introducing the dependence on applied load P . In contrast, prediction based on a simpler formulation of δ tends to rely more on the physics model, which may have better extrapolation capability than a purely statistical term δ . In the second example, a complicated physics model with little information regarding the model error is calibrated with experimental data. We applied the first-order Taylor series expansion-based method in order to avoid parameter non-identifiability due to the addition of a model discrepancy function. We also showed that the various options

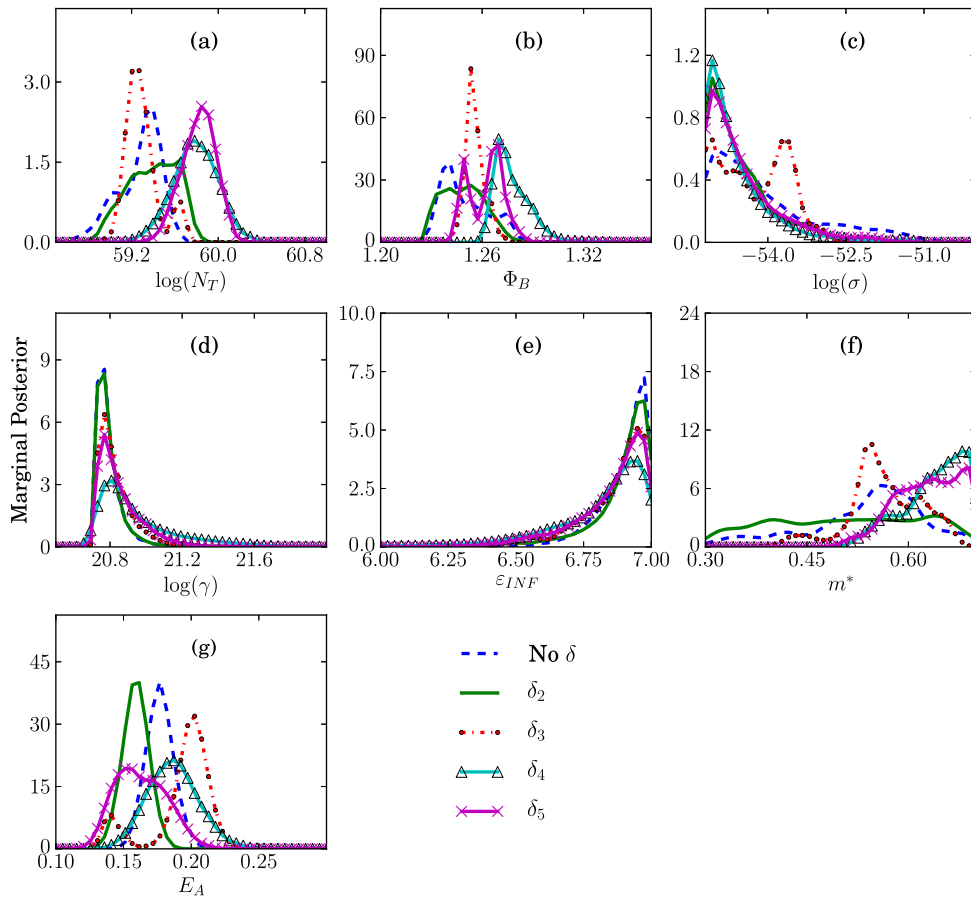


Fig. 7. Marginal PDFs of dielectric charging model parameters.

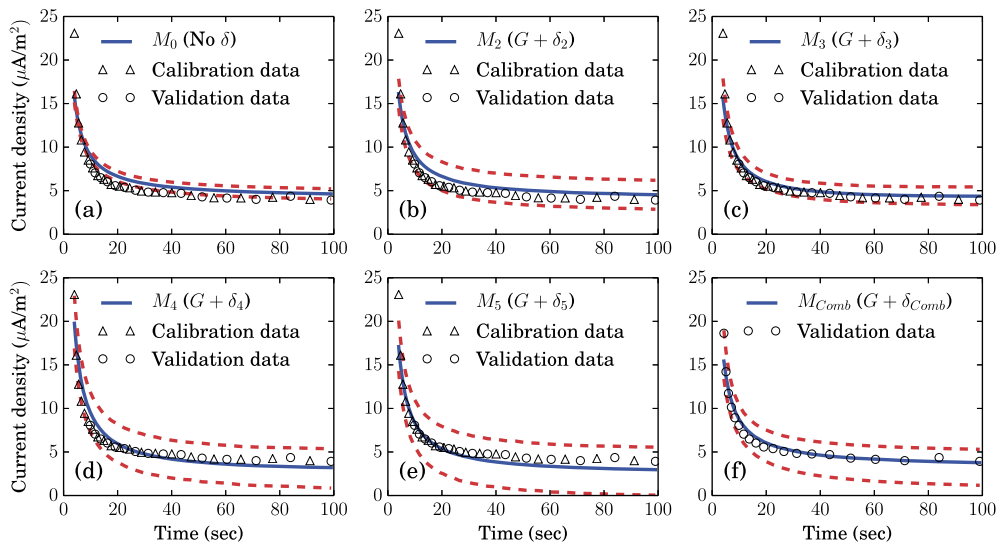


Fig. 8. Comparison between the predictions of calibrated dielectric charging model and data.

of δ can lead to different calibration results, and the proposed three-step approach is implemented to combine the posterior distributions of model parameters and model discrepancy, which leads to accurate prediction while also accounting for the uncertainty in the form of model error.

Table 4

Overall reliability of model predictions.

	M_0 (No δ)	M_1	M_2	M_3	M_4	M_5
μ_r^i	0.86	–	0.58	0.82	0.37	0.36

7. Conclusion

This paper investigated various formulations of the model discrepancy function for use in Bayesian calibration. A first-order Taylor series expansion-based method has been developed to determine the identifiability of model parameters, especially when one tries to add a model discrepancy function to a model that does not have a closed-form analytical expression. The two numerical examples show that different choices of model discrepancy formulation can result in significantly different calibration results, and that the estimation of physical parameters may be biased if the formulation is over-simplified. It is also observed that the model calibrated using simpler formulations of δ may outperform the one calibrated using more complicated formulations outside the calibration domain. In order to facilitate the selection of model discrepancy formulation and also accommodate the uncertainty due to this selection, we developed a three-step approach using a reliability-based model validation metric and Bayesian model averaging. This approach combines the posterior probability distributions of model parameters and δ resulting from the various options of δ into a single distribution, which is useful especially when the various options of δ have similar performance. Future research efforts may include: (1) examining more options of model discrepancy, and investigating the extrapolation capability of model discrepancy to higher level of system hierarchy, (2) extending the first-order Taylor series expansion-based method to investigate the identifiability of models containing hyperparameters (e.g., when the model discrepancy is approximated using a Gaussian process, the coefficients of the covariance function need to be taken into account).

Acknowledgements

This paper is based upon research supported by the U.S. Department of Energy [National Nuclear Security Administration] under Award Number DE-FC52-08NA28617 to Purdue University (Principal Investigator: Prof. Jayathi Murthy, Deputy Director: Prof. Alejandro Strachan), and subaward to Vanderbilt University. The support is gratefully acknowledged. The authors also thank the following personnel at the U. S. DOE (NNSA) PSAAP Center for Prediction of Reliability, Integrity and Survivability of Microsystems (PRISM) at Purdue University for providing the dielectric charging compact model, experimental data, and valuable discussions: Prof. Muhammad Alam, Sambit Palit, Prof. Alejandro Strachan, Dr. Lin Sun, and Ravi Vedula. The numerical examples in this paper were conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University, Nashville, TN.

References

- [1] K. Campbell, Statistical calibration of computer simulations, *Reliab. Eng. Syst. Saf.* 91 (10) (2006) 1358–1363, <http://dx.doi.org/10.1016/j.res.2005.11.032>.
- [2] J.M. McFarland, *Uncertainty analysis for computer simulations through validation and calibration*, Ph.D. thesis, Vanderbilt University, 2008.
- [3] M.C.M. Kennedy, A. O'Hagan, Bayesian calibration of computer models, *J. R. Stat. Soc., Ser. B, Stat. Methodol.* 63 (3) (2001) 425–464, <http://dx.doi.org/10.1111/1467-9868.00294>.
- [4] J. Brynjarsdottir, A. O'Hagan, Learning about physical parameters: the importance of model discrepancy, *Inverse Probl.* to appear in Special issue: Bayesian methods in inverse problems.
- [5] D. Higdon, J. Gattiker, B. Williams, M. Rightley, Computer model calibration using high-dimensional output, *J. Am. Stat. Assoc.* 103 (482) (2008) 570–583, <http://dx.doi.org/10.1198/016214507000000888>.
- [6] P. Koutsourelakis, A multi-resolution, non-parametric, Bayesian framework for identification of spatially-varying model parameters, *J. Comput. Phys.* 228 (17) (2009) 6184–6211, <http://dx.doi.org/10.1016/j.jcp.2009.05.016>.
- [7] P.D. Arendt, D.W. Apley, W. Chen, Quantification of model uncertainty: calibration, model discrepancy, and identifiability, *J. Mech. Des.* 134 (10) (2012) 100908, <http://dx.doi.org/10.1115/1.4007390>.
- [8] Y. Ling, S. Mahadevan, Challenging issues in Bayesian calibration of multi-physics models, in: *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Boston, MA, 2013.
- [9] G. Arhonditsis, D. Papantou, W. Zhang, G. Perhar, E. Massos, M. Shi, Bayesian calibration of mechanistic aquatic biogeochemical models and benefits for environmental management, *J. Mar. Syst.* 73 (1–2) (2008) 8–30, <http://dx.doi.org/10.1016/j.jmarsys.2007.07.004>.
- [10] J. McFarland, S. Mahadevan, Multivariate significance testing and model calibration under uncertainty, *Comput. Methods Appl. Mech. Eng.* 197 (29–32) (2008) 2467–2479, <http://dx.doi.org/10.1016/j.cma.2007.05.030>.
- [11] R.G. Bower, I. Vernon, M. Goldstein, A.J. Benson, C.G. Lacey, C.M. Baugh, S. Cole, C.S. Frenk, The parameter space of galaxy formation, *Mon. Not. R. Astron. Soc.* 407 (4) (2010) 2017–2045, <http://dx.doi.org/10.1111/j.1365-2966.2010.16991.x>.
- [12] B. Renard, D. Kavetski, G. Kuczera, M. Thyer, S.W. Franks, Understanding predictive uncertainty in hydrologic modeling: the challenge of identifying input and structural errors, *Water Resour. Res.* 46 (5) (2010) W05521, <http://dx.doi.org/10.1029/2009WR008328>.
- [13] S. Sarkar, D. Kosson, S. Mahadevan, J. Meeussen, H.V. der Soot, J. Arnold, K. Brown, Bayesian calibration of thermodynamic parameters for geochemical speciation modeling of cementitious materials, *Cem. Concr. Res.* 42 (7) (2012) 889–902, <http://dx.doi.org/10.1016/j.cemconres.2012.02.004>.
- [14] S. Sankararaman, Y. Ling, C. Shantz, S. Mahadevan, Inference of equivalent initial flaw size under multiple sources of uncertainty, *Int. J. Fatigue* 33 (2) (2011) 75–89, <http://dx.doi.org/10.1016/j.ijfatigue.2010.06.008>.
- [15] E.C. DeCarlo, S. Mahadevan, B.P. Smarslok, Bayesian calibration of aerothermal models for hypersonic air vehicles, in: *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Boston, MA, 2013.

- [16] I. Park, H.K. Amarchinta, R.V. Grandhi, A Bayesian approach for quantification of model uncertainty, *Reliab. Eng. Syst. Saf.* 95 (7) (2010) 777–785, <http://dx.doi.org/10.1016/j.res.2010.02.015>.
- [17] M. Riley, R. Grandhi, Quantification of model-form and predictive uncertainty for multi-physics simulation, *Comput. Struct.* 89 (11–12) (2011) 1206–1213, <http://dx.doi.org/10.1016/j.compstruc.2010.10.004>.
- [18] R. Rebba, S. Mahadevan, Computational methods for model reliability assessment, *Reliab. Eng. Syst. Saf.* 93 (8) (2008) 1197–1207, <http://dx.doi.org/10.1016/j.res.2007.08.001>.
- [19] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, J. Timmer, Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, *Bioinformatics* 25 (15) (2009) 1923–1929, <http://dx.doi.org/10.1093/bioinformatics/btp358>.
- [20] P.D. Arendt, D.W. Apley, W. Chen, D. Lamb, D. Gorsich, Improving identifiability in model calibration using multiple responses, *J. Mech. Des.* 134 (10) (2012) 100909, <http://dx.doi.org/10.1115/1.4007573>.
- [21] M. Grewal, K. Glover, Identifiability of linear and nonlinear dynamical systems, *IEEE Trans. Autom. Control* 21 (6) (1976) 833–837.
- [22] E. Walter, L. Pronzato, On the identifiability and distinguishability of nonlinear parametric models, *Math. Comput. Simul.* 42 (2) (1996) 125–134.
- [23] J.-f. Zhang, Q.-h. Yan, T. Zhai, Numerical approach to identifiability test of parametric models in nonlinear mechanical systems, *J. Dyn. Syst. Meas. Control* 133 (5) (2011) 051002, <http://dx.doi.org/10.1115/1.4004062>.
- [24] M. Gu, J. Lu, A note on identifiability of the regression parameter and validity of the partial likelihood approach in general relative risk regression, *Biometrika* 81 (4) (1994) 802–806.
- [25] C. Paulino, C. de Bragança Pereira, On identifiability of parametric statistical models, *Stat. Methods Appl.* 3 (1) (1994) 125–151, <http://dx.doi.org/10.1007/BF02589044>.
- [26] M. Little, W. Heidenreich, G. Li, Parameter identifiability and redundancy: theoretical considerations, *PLoS ONE* 5 (1) (2010) e8915, <http://dx.doi.org/10.1371/journal.pone.0008915>.
- [27] A. Raue, C. Kreutz, T. Maiwald, U. Klingmüller, J. Timmer, Addressing parameter identifiability by model-based experimentation, *IET Syst. Biol.* 5 (2) (2011) 120–130, <http://dx.doi.org/10.1049/iet-syb.2010.0061>.
- [28] S. Murphy, A. Van der Vaart, On profile likelihood, *J. Am. Stat. Assoc.* 95 (450) (2000) 449–465.
- [29] J.E. Oakley, A. O'Hagan, Probabilistic sensitivity analysis of complex models: a Bayesian approach, *J. R. Stat. Soc., Ser. B, Stat. Methodol.* 66 (3) (2004) 751–769, <http://dx.doi.org/10.1111/j.1467-9868.2004.05304.x>.
- [30] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, *Global Sensitivity Analysis: The Primer*, vol. 76, John Wiley & Sons, Inc., 2008.
- [31] S. Palit, M. Alam, Theory of charging and charge transport in “intermediate” thickness dielectrics and its implications for characterization and reliability, *J. Appl. Phys.* 111 (5) (2012) 054112, <http://dx.doi.org/10.1063/1.3691962>.
- [32] C.E. Rasmussen, C.K.I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006.
- [33] L.J. Haarhoff, S. Kok, D.N. Wilke, Numerical strategies to reduce the effect of ill-conditioned correlation matrices and underflow errors in Kriging, *J. Mech. Des.* 135 (4) (2013) 044502, <http://dx.doi.org/10.1115/1.4023631>.
- [34] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, *Science* 220 (4598) (1983) 671–680, <http://dx.doi.org/10.1126/science.220.4598.671>.
- [35] J. Kennedy, R. Eberhart, Particle swarm optimization, *Proceedings of ICNN'95 – International Conference on Neural Networks*, vol. 4, IEEE, Perth, WA, 1995, pp. 1942–1948.
- [36] W. Oberkampf, M. Barone, Measures of agreement between computation and experiment: validation metrics, *J. Comput. Phys.* 217 (1) (2006) 5–36, <http://dx.doi.org/10.1016/j.jcp.2006.03.037>.
- [37] D. Zwillinger, *Handbook of Integration*, 1st ed., Jones and Bartlett Publishers, Boston, MA, 1995.
- [38] V.K. Hombal, S. Mahadevan, Model selection among physics-based models, *J. Mech. Des.* 135 (2) (2013) 021003, <http://dx.doi.org/10.1115/1.4023155>.
- [39] J. Hoeting, D. Madigan, A. Raftery, C. Volinsky, Bayesian model averaging: a tutorial, *Stat. Sci.* 14 (4) (1999) 382–401.
- [40] S. Sankararaman, S. Mahadevan, Comprehensive framework for integration of calibration, verification and validation, in: 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Honolulu, HI, 2012.
- [41] S. Sankararaman, Uncertainty quantification and integration in engineering systems, Ph.D. thesis, Vanderbilt University, 2012.
- [42] J. Mullins, C. Li, S. Sankararaman, S. Mahadevan, A. Urbina, Probabilistic integration of validation and calibration results for prediction level uncertainty quantification: application to structural dynamics, in: 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Boston, MA, 2013.
- [43] W. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* 57 (1) (1970) 97–109, <http://dx.doi.org/10.1093/biomet/57.1.97>.
- [44] J.M. Gere, B.J. Goodno, *Mechanics of Materials*, 7th ed., Cengage Learning, 2009.
- [45] S. Timoshenko, J. Goodier, *Theory of Elasticity*, 3rd ed., McGraw-Hill, New York, 1970.
- [46] C.E. Augarde, A.J. Deeks, The use of Timoshenko's exact solution for a cantilever beam in adaptive analysis, *Finite Elem. Anal. Des.* 44 (9–10) (2008) 595–601, <http://dx.doi.org/10.1016/j.finel.2008.01.010>.
- [47] A. Jain, S. Palit, M.A. Alam, A physics-based predictive modeling framework for dielectric charging and creep in RF MEMS capacitive switches and varactors, *J. Microelectromech. Syst.* 21 (2) (2012) 420–430, <http://dx.doi.org/10.1109/JMEMS.2011.2174418>.