Comput Mech
DOI 10.1007/s00466-017-1448-6

ORIGINAL PAPER

Uncertainty aggregation and reduction in structure–material performance prediction

Zhen Hu[1] · **Sankaran Mahadevan**[1] · **Dan Ao**[1]

Received: 8 March 2017 / Accepted: 8 July 2017

Abstract An uncertainty aggregation and reduction framework is presented for structure–material performance prediction. Different types of uncertainty sources, structural analysis model, and material performance prediction model are connected through a Bayesian network for systematic uncertainty aggregation analysis. To reduce the uncertainty in the computational structure–material performance prediction model, Bayesian updating using experimental observation data is investigated based on the Bayesian network. It is observed that the Bayesian updating results will have large error if the model cannot accurately represent the actual physics, and that this error will be propagated to the predicted performance distribution. To address this issue, this paper proposes a novel uncertainty reduction method by integrating Bayesian calibration with model validation adaptively. The observation domain of the quantity of interest is first discretized into multiple segments. An adaptive algorithm is then developed to perform model validation and Bayesian updating over these observation segments sequentially. Only information from observation segments where the model prediction is highly reliable is used for Bayesian updating; this is found to increase the effectiveness and efficiency of uncertainty reduction. A composite rotorcraft hub component fatigue life prediction model, which combines a finite element structural analysis model and a material damage model, is used to demonstrate the proposed method.

Keywords Uncertainty aggregation · Uncertainty reduction · Bayesian calibration · Model validation · Life prediction

1 Introduction

Structure–material models [1] typically combine a structural analysis model that predicts the structural response under different load conditions, with a material damage model that estimates the material performance based on the structural response. Such combination been pursued in both macro-scale and multi-scale computations. For instance, in the macro-scale, Liu and Mahadevan [2] proposed a high-cycle multi-axial fatigue performance prediction model based on a characteristic plane computation that makes use of the structural analysis. As an example of multi-scale computation, Ghosh et al. [3] developed an adaptive multi-level methodology to model the evolution of variables at the macrostructure-level as well as track microstructural damage, in composite and porous materials. Crouch et al. [4] proposed a composite material fatigue damage prediction model using a multi-scale modeling approach to damage accumulation. Several studies have focused on improving the accuracy and efficiency of performance prediction through multi-scale modeling [5,6] and high performance computing [7].

Most current structure–material models have primarily focused on deterministic performance prediction. Consideration of uncertainty has been limited to natural variability in material properties and loading [8–10], which are modeled through random variables and random processes. However, there are many other types of uncertainty sources that need to be accounted for, in particular epistemic uncertainty resulting from lack of knowledge. These include uncertainty in

✉ Sankaran Mahadevan
sankaran.mahadevan@vanderbilt.edu

[1] Department of Civil and Environmental Engineering, Vanderbilt University, 272 Jacobs Hall, VU Mailbox: PMB 351831, Nashville, TN 37235, USA

Published online: 28 July 2017

the distribution parameters of input random variables due to limited data [11], discrepancy between the computer simulation model and the actual physics [12], and model parameter uncertainty [13]. These uncertainty sources are present in both the structural analysis model and the material damage model. Often, the structural analysis is a physics-based finite element model, incorporating multiple assumptions and approximations; on the other hand, the material damage model is often empirical, based on material test data. Each type of model is affected by different sources of uncertainty. A systematic framework is therefore required to quantify the different sources of uncertainty and to aggregate their effects on the performance prediction.

A Bayesian network approach is adopted in this paper to facilitate uncertainty quantification, aggregation and reduction. The Bayesian network will help not only in uncertainty aggregation (which is a forward problem), but also in identifying the parameters that make the highest contributions to the uncertainty of the performance prediction, and facilitate uncertainty reduction in the performance prediction through Bayesian updating with observed data.

Bayesian updating is a process of uncertainty reduction regarding unknown parameters of a model based on the fusion of prior information, mathematical model, and physical observations [14]. In current Bayesian calibration methods for uncertainty reduction, it is typically assumed that the prediction model (e.g., computational simulation model) accurately represents the actual physics or that the uncertainty between the prediction model and the actual physics can be accurately captured by adding a discrepancy term to the prediction model [15]. Thus the calibration effort typically estimates the posterior distributions of model parameters and the discrepancy term. For some problems, however, the prediction model is still quite different from the actual physics even if a discrepancy term is estimated and added [14]. In that case, Bayesian calibration based on observations using a low quality prediction model will lead to low accuracy in the model prediction even if the uncertainty in the prediction can be reduced. This implies that the challenge in model calibration is two-fold: *how to effectively reduce the uncertainty and how to maintain the accuracy of the prediction*. To tackle this challenge, the validity of the prediction model needs to be considered within Bayesian calibration.

In a structure–material performance prediction model, considering the model validity into Bayesian calibration is not straightforward. There are two main challenges that need to be solved. The first challenge is how to evaluate the validity of the model while accounting for various sources of uncertainty. The second challenge is how to include the model validity into the Bayesian calibration. To address the first challenge, this paper adopts Bayesian hypothesis testing for model validation under uncertainty [16]. To address the second challenge, an adaptive uncertainty reduction method is proposed through an iterative integration of model validation and Bayesian calibration. Thus the main contributions of this paper can be summarized as: (1) a framework for uncertainty modeling and uncertainty aggregation in structure–material performance prediction models; (2) model validation based on uncertainty aggregation; and (3) a novel adaptive method for uncertainty reduction based on iterative integration of model validation and calibration.

The remainder of the paper is organized as follows. Section 2 discusses the background of uncertainty aggregation and uncertainty reduction with respect to structure–material performance prediction. Section 3 presents the proposed uncertainty aggregation framework. Following that, Sect. 4 proposes the new adaptive uncertainty reduction approach. Section 5 illustrates the proposed framework using a composite rotorcraft hub example. Section 6 presents concluding remarks.

2 Background

2.1 Combination of physics-based and empirical models

Various types of models have been developed in the past decades to predict the structural performance based on structural analysis and material damage modeling. Figure 1 shows a schematic procedure of structure–material performance prediction. The computational structural analysis model predicts the structural response under different loading conditions and the material damage model predicts the material performance based on the structural response. The structural analysis and material damage analysis can be performed sequentially or coupled together.

The structure model is usually a physics-based finite element model as shown in Fig. 1. It incorporates modeling assumptions regarding the structural behavior, boundary conditions etc., and further employs numerical approximations (e.g., discretization) to compute the structural response. The model includes parameters (e.g., mechanical properties) which may be obtained from test data.

The material damage is often empirical (especially at the macro-level), based on curve fitting of material test data. Consider, for example, Paris' law that is commonly used to predict long crack growth in structures. The relationship between the number of cycles (N), the crack size (a), and the stress intensity factor range (ΔK) is written as [17]

$$da/dN = C\Delta K^n \left(1 - \frac{\Delta K_{th}}{\Delta K}\right)^p, \tag{1}$$

where C, n, p, and ΔK_{th} are model parameters estimated from test data that measure crack size versus number of load-

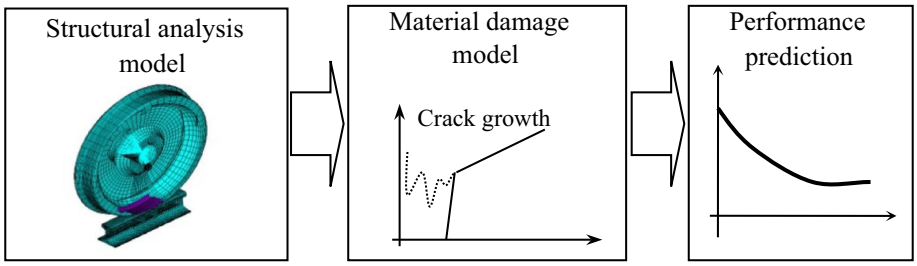


Fig. 1 Illustration of the structure–material performance prediction model

ing cycles under different loading conditions. n controls the slope of the Paris regime and p controls the curvature in the near-threshold regime. For a structure of interest as in Fig. 1, ΔK is computed by the structural analysis model and input to the material model in Eq. (1) to predict the crack growth under a loading condition of interest.

The above is a simple illustrative example. In the literature, many sophisticated models have been developed to predict the performance of structures using computational structural analysis model and material damage model. These range from macro-level fatigue damage prediction methods to handle different fatigue conditions (e.g., the characteristic plane approach for multi-axial fatigue [2]), to multi-scale modeling of material damage [18]. The focus of this paper, however, is not the details of such models, but on how to quantify and reduce the uncertainty in performance prediction *when multiple models are combined*. Our particular focus is on aggregating the uncertainty arising from multiple aleatory and epistemic sources, and on reducing the uncertainty in performance prediction. The developed methodology in this paper is applicable not only to structure–material performance prediction models, but also to other computer simulation models.

2.2 Sources of uncertainty

Model uncertainty sources can be classified into two groups: aleatory and epistemic.

- *Aleatory uncertainty* This is natural variability and is irreducible. Natural variability may be present in loading, material properties, component geometry, and boundary conditions. The variability may be across specimens, as well as over time and space.
- *Epistemic uncertainty* This type of uncertainty comes from lack of knowledge, which is reducible when more information is available. Depending on the source, it can be further classified into two groups: (1) Data uncertainty, which arises from limited, imprecise, qualitative, erroneous or missing data. It causes uncertainty in the parameters of the probability distributions used to describe the aleatory model inputs. For instance, we may not be able to precisely determine the mean and variance of the material property random variable (e.g., Young's modulus) when we only have limited experimental data. (2) Model uncertainty, which arises from three types of sources, namely, model form, model parameters, and solution approximations. Model form error arises during mathematical representation of physical reality (due to simplifying assumptions), whereas solution approximation errors arise in numerically solving the mathematical equations (e.g., discretization, reduced order modeling, surrogate modeling, and sampling). Since model parameters are often not directly measurable but only inferred from the observations of model inputs and outputs, limited or imprecise observations cause uncertainty regarding the model parameters (e.g., the fatigue limit parameter in the crack growth model).

All these sources of uncertainty affect the prediction of the structure–material performance prediction model. A systematic approach is needed to quantify the individual sources as well as aggregate their effects on the overall uncertainty in the model prediction.

3 Uncertainty aggregation

3.1 Modeling of uncertainty sources

A commonly used strategy for the modeling of aleatory uncertainty is to use random variables for time-independent uncertainty, stochastic processes for time-dependent uncertainty, and random fields for quantities with spatial variability [19,20]. Standard approaches are available for the modeling of aleatory uncertainty, such as probability distributions, time-series models, and Karhunen–Loève expansion. Details of these methods can be found in many textbooks, such as Ref. [19]. Here, we mainly focus on the modeling of epistemic uncertainty.

3.1.1 Data uncertainty

Multiple mathematical frameworks have been proposed to handle epistemic uncertainty, such as the Bayesian approach, imprecise probabilities, fuzzy sets, evidence theory, and

information gap theory. This paper follows the Bayesian framework, where the analyst's lack of knowledge regarding an unknown quantity is represented by a probability distribution. The unknown quantity may either be deterministic, or the parameter of a stochastic quantity (random variable, random process or random field). When new data $\mathbf{D}$ is available, the uncertainty regarding the unknown quantity $\boldsymbol{\theta}$ can be updated using Bayes' theorem as [21]

$$f(\boldsymbol{\theta}|\mathbf{D}) = \frac{L(\mathbf{D}|\boldsymbol{\theta}) f(\boldsymbol{\theta})}{\int_{\boldsymbol{\theta}} L(\mathbf{D}|\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}} \propto L(\mathbf{D}|\boldsymbol{\theta}) f(\boldsymbol{\theta}), \tag{2}$$

where $L(\mathbf{D}|\boldsymbol{\theta})$ is the likelihood of observing the data $\mathbf{D}$, $f(\boldsymbol{\theta})$ is the prior distribution (i.e., before observing data $\mathbf{D}$), $f(\boldsymbol{\theta}|\mathbf{D})$ is the posterior distribution (i.e., after observing data $\mathbf{D}$), and "$\propto$" stands for "proportional to". Methods for computing $L(\mathbf{D}|\boldsymbol{\theta})$ are different, depending on the modeling of $\boldsymbol{\theta}$. For example, the likelihood function $L(\mathbf{D}|\boldsymbol{\theta})$ is quite different for random variables and time-series models [22].

3.1.2 Model discrepancy

The model discrepancy between the prediction model and the experiment data may come from different sources. For instance, model form assumptions made in the mathematical model result in the discrepancy between prediction model and experiment. Further, numerical approximations such as the discretization error in the finite element simulation cause discrepancy between the prediction model and the experiments. If a surrogate model is used in place of the original physics model, then surrogate model uncertainty is also introduced into the discrepancy term. Consider a single scalar physical quantity of interest (QOI) y_{ph}; the simulation model output y_{sim}, which is the prediction of y_{ph}, is given by

$$y_{sim} = G_{sim}(\mathbf{x}), \tag{3}$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_{n_x}] \in \mathbb{R}^{n_x}$ is the vector of inputs of the computer simulation model.

The model discrepancy terms are used to correct the computer simulation model (y_{sim}) to achieve closer agreement to the actual physics (y_{ph}) as

$$y_c(\mathbf{x}) = y_{sim}(\mathbf{x}) + \varepsilon_{num}(\mathbf{x}) + \delta(\mathbf{x}), \tag{4}$$

where $y_c(\mathbf{x})$ is the corrected computer simulation model, $\delta(\mathbf{x})$ is the model form error, and $\varepsilon_{num}(\mathbf{x})$ is the numerical error due to solution approximations. The numerical error could be further decomposed to discretization error, truncation error (in reduced-order modeling), round-off error etc. The different error terms may combine in different ways in different problems, such as linear, nonlinear, nested, or iterative. A Bayesian network has been proposed to aggregate the various model error terms in a systematic manner, without imposing any assumptions such as root-mean-square (RMS) regarding their combination [23].

With more and more model error terms being properly added to Eq. (4), the corrected computer simulation model $y_c(\mathbf{x})$ will become more and more close to the true physics y_{ph}. For different sources of model uncertainty, the quantification methods are different. For example, a commonly used method for the modeling of discretization error is Richardson extrapolation [24,25]. For a given input setting $\mathbf{d}$, the discretization error in the output $y_{sim}(\mathbf{d})$ of an FEA simulation with mesh size h_1 can be quantified based on two more FEA simulations with finer mesh size h_2 and finest mesh size h_3 as follows [24,25]:

$$\varepsilon_{FEA}(\mathbf{d}) = \left[y_{sim}^{(1)}(\mathbf{d}) - y_{sim}^{(2)}(\mathbf{d}) \right] \Big/ \left(r_{mesh}^{p_c} - 1 \right), \tag{5}$$

where $y_{sim}^{(k)}(\mathbf{d})$ is the result of FEA simulation with inputs $\mathbf{d}$ and mesh size h_k, $r_{mesh} = h_2/h_1 = h_3/h_2$ is the mesh refinement ratio, and the convergence p_c is estimated as

$$p_c = \ln \left[\left(y_{sim}^{(3)}(\mathbf{d}) - y_{sim}^{(2)}(\mathbf{d}) \right) \Big/ \left(y_{sim}^{(2)}(\mathbf{d}) - y_{sim}^{(1)}(\mathbf{d}) \right) \right] \Big/ \ln r_{mesh}. \tag{6}$$

Liang and Mahadevan [26] provide methods for quantifying different types of model errors in computational simulation.

3.1.3 Model parameter uncertainty

Similar to the modeling of distribution parameter uncertainty, the Bayesian approach in Eq. (2) can be used for the quantification of physics model parameter uncertainty with input–output observations. The likelihood function for Bayesian calibration needs to be defined according to the prediction model. When model parameter uncertainty and model discrepancy are both present, a common approach for quantifying both simultaneously is to implement the Kennedy and O'Hagan framework [27] which uses a Gaussian process model to represent the model discrepancy. Depending on the objectives of the analyst, one might either calibrate an overall model discrepancy term (which will include both model form and numerical errors), or first quantify the numerical errors and correct the simulation model in order to isolate the model form error during the calibration step [28].

3.2 Uncertainty aggregation

A Bayesian network (BN) approach is used in this paper to aggregate the contributions of various sources of uncertainty towards the overall model output of interest. BN is a flexible tool for modeling the multivariate joint probability density of

n random variables $X_1,\ X_2,\ \ldots, X_n$ in terms of conditional probabilities as [21]

$$P(\mathbf{X}) = P(X_1,\ X_2,\ \ldots, X_n) = \prod_{i=1}^{n} P(X_i|\pi_i), \tag{7}$$

where π_i is the set of parent nodes of node X_i and $P(X_i|\pi_i)$ is the conditional probability density function (PDF) of node X_i for a given realization of its parents. The nodes without parents are called root nodes.

Consider a structural analysis model

$$g_s = G_S\left(\mathbf{x},\ \boldsymbol{\omega}_a,\ \boldsymbol{\theta}_{ae}^S,\ \boldsymbol{\theta}_e^S\right), \tag{8}$$

where $\mathbf{x}$ is a vector of deterministic inputs, $\boldsymbol{\omega}_a$ is a vector of aleatory inputs, $\boldsymbol{\theta}_{ae}^S$ is a vector of epistemic parameters of $\boldsymbol{\omega}_a$ due to data uncertainty, and $\boldsymbol{\theta}_e^S$ is a vector of physics model parameters with epistemic uncertainty. $g_s = G_S(\mathbf{x},\ \boldsymbol{\omega}_a,\ \boldsymbol{\theta}_{ae}^S,\ \boldsymbol{\theta}_e^S)$ can be a computational simulation model, or an empirical data-driven model.

Similarly, consider a material damage model

$$T = G_M\left(\boldsymbol{\theta}_e^M,\ g_s\right), \tag{9}$$

where $\boldsymbol{\theta}_e^M$ is a vector of model parameters with epistemic uncertainty.

In addition to these two models, according to Eq. (4), we also have the model discrepancy term $\delta_S(\mathbf{x},\ \boldsymbol{\omega}_a)$ of the physics-based structural model, and the stochastic residual term $\varepsilon_M(g_s)$ (from curve-fitting) for the empirical material model. (The latter term could be dependent on g_s as shown, in case a Gaussian process model is fit to the material data, or independent of g_s, as is common in simple regression models where zero-mean normally distributed residuals with constant variance are assumed). Here $\delta_S(\mathbf{x},\ \boldsymbol{\omega}_a)$ refers to the overall model discrepancy term which includes both numerical error and model form error. Note that the model discrepancy term is not a function of model parameters because the epistemic uncertainty parameters are integrated out in the KOH framework [27].

Often surrogate models such as Kriging model [29] and polynomial chaos model [30] are used to replace the original prediction models, in order to save computational expense, especially in inverse problems as in model calibration. The use of a surrogate model introduces surrogate model uncertainty. For example, if a Kriging surrogate model $\hat{G}_S(\mathbf{x},\ \boldsymbol{\omega}_a,\ \boldsymbol{\theta}_{ae}^S,\ \boldsymbol{\theta}_e^S)$ is used to replace $G_S(\mathbf{x},\ \boldsymbol{\omega}_a,\ \boldsymbol{\theta}_{ae}^S,\ \boldsymbol{\theta}_e^S)$, then the prediction of the surrogate model is a normal random variable for a given realization of the inputs and parameters $\mathbf{x}_{G_S} = [\mathbf{x},\ \boldsymbol{\omega}_a,\ \boldsymbol{\theta}_{ae}^S,\ \boldsymbol{\theta}_e^S]$, with mean of the prediction given as

$$\mu_{\hat{G}_S}(\mathbf{x}_{G_S}) = \mathbf{h}(\mathbf{x}_{G_S})^T\boldsymbol{\upsilon} + \mathbf{r}(\mathbf{x}_{G_S})^T\mathbf{R}^{-1}(\mathbf{g} - \mathbf{H}\boldsymbol{\upsilon}), \tag{10}$$

where $\boldsymbol{\upsilon} = [\upsilon_1,\ \upsilon_2,\ \ldots,\ \upsilon_p]^T$ is a vector of coefficients estimated from surrogate modeling, $\mathbf{h}(\mathbf{x}_{G_S}) = [h_1(\mathbf{x}_{G_S}),\ h_2(\mathbf{x}_{G_S}),\ \ldots,\ h_p(\mathbf{x}_{G_S})]^T$ is a vector of regression functions, $\mathbf{h}(\mathbf{x}_{G_S})^T\boldsymbol{\upsilon}$ is the trend of prediction, $\mathbf{r}(\mathbf{x}_{G_S}) = [R(\mathbf{x}_{G_S},\ \mathbf{x}_{G_S1}),\ R(\mathbf{x}_{G_S},\ \mathbf{x}_{G_S2}),\ \ldots,\ R(\mathbf{x}_{G_S},\ \mathbf{x}_{G_Sn_s})]$, $R(\mathbf{x}_{G_Si},\ \mathbf{x}_{G_Sj})$ is the correlation function, $\mathbf{x}_1, \ldots,\ \mathbf{x}_{n_s}$ are the training points, and $\mathbf{R}$, $\mathbf{g}$, and $\mathbf{H}$ are $R(\mathbf{x}_{G_Si},\ \mathbf{x}_{G_Sj})$, G_S, and $\mathbf{h}(\mathbf{x})$ evaluated at the training points.

The variance of the prediction is given by [38]

$$\begin{aligned}\sigma^2_{\hat{G}_S}(\mathbf{x}_{G_S}) = \sigma_\varepsilon^2\{1 &- \mathbf{r}(\mathbf{x}_{G_S})^T\mathbf{R}^{-1}\mathbf{r}(\mathbf{x}_{G_S}) \\ &+ [\mathbf{H}^T\mathbf{R}^{-1}\mathbf{r}(\mathbf{x}_{G_S}) - \mathbf{h}(\mathbf{x}_{G_S})]^T(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1} \\ &\times [\mathbf{H}^T\mathbf{R}^{-1}\mathbf{r}(\mathbf{x}_{G_S}) - \mathbf{h}(\mathbf{x}_{G_S})],\end{aligned} \tag{11}$$

where σ_ε^2 is a parameter estimated during the surrogate modeling.

Based on the above definitions, the data uncertainty, model discrepancy, and model parameter uncertainty in the structure–material performance prediction model can be aggregated into the uncertainty of the performance prediction T using the BN shown in Fig. 2.

In the above BN, the elliptical nodes represent random variables, and the square nodes represent observations. ε_{obs} is the measurement error, which is an aleatory uncertainty source, and T_{obs} is the observed structure–material performance measure (e.g., fatigue life). As shown in Eq. (7), different nodes in the BN are connected to each other through the conditional PDF to model the joint PDF of all the uncertain variables associated with both models. Forward propagation of the uncertainties through the BN using Monte Carlo simulation (MCS) helps to quantify the uncertainty in the performance prediction and thus achieve uncertainty aggregation. Therefore, the key step in uncertainty aggregation is the construction of the BN. The BN given in Fig. 2 can be constructed either using physics models which are connected through the inputs and outputs of the models or using data from which the connections are learned. In the structure–material prediction model, the BN is constructed using a combined physics and data driven models.

The BN approach provides two additional benefits. *First*, it can help us to identify the uncertainty sources that make the highest contributions on the uncertainty in the prediction. *Second*, it can significantly facilitate uncertainty reduction as will be discussed in Sect. 4.

Global sensitivity analysis (GSA) can be employed to identify the contributions of different uncertainty sources on the overall uncertainty in performance prediction. Variance-based GSA ranks the contribution of each input random variable (X) on the variance Var(Y) of an output QOI (Y). Sobol' indices are commonly used in this context and two

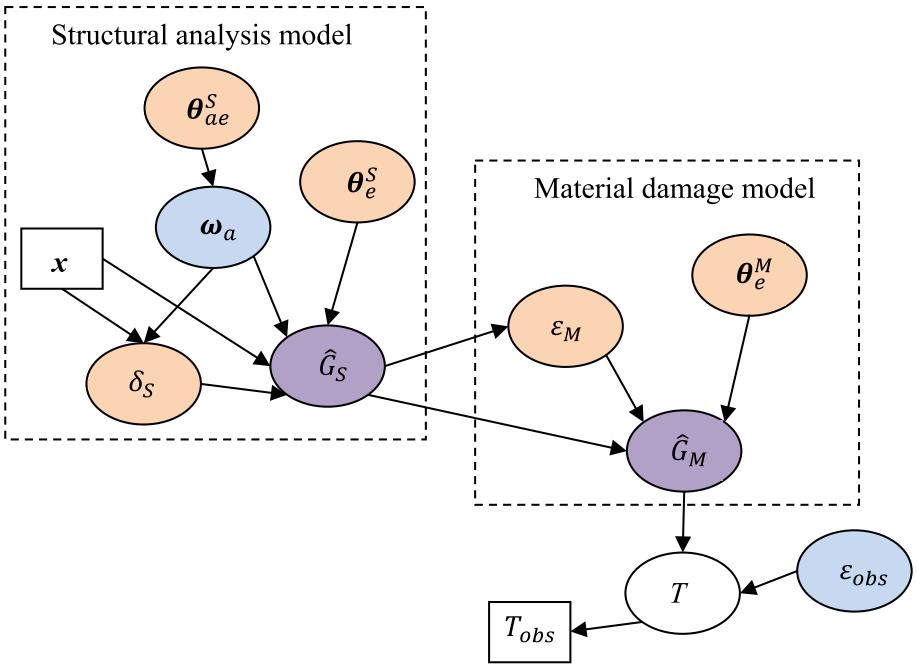


Fig. 2 Uncertainty aggregation in the structure–material performance prediction model

kinds of indices are used to quantify the uncertainty contributions: first-order indices and total effects indices. The first-order index measures the contribution of an individual variable without considering its interactions with other variables and is given by [31]

$$S_i^I = \frac{Var_{X_i}(E_{X\sim i}(Y|X_i))}{Var(Y)}, \tag{12}$$

where X_i is the ith input variable, $X_{\sim i}$ is the vector of variables excluding variable X_i, $Var(Y)$ is the variance of QOI (Y), and $E_{X\sim i}(Y|X_i)$ is the expectation by freezing X_i. The total effects index measures the contribution of an individual variable and its interactions with other variables. The total effects index is given by

$$S_i^T = 1 - \frac{Var_{X_{\sim i}}(E_{X_i}(Y|X_{\sim i}))}{Var(Y)}, \tag{13}$$

Note that the above GSA method is for a single prediction model. To apply GSA to the BN given in Fig. 2 which combines multiple models, a BN-based GSA method has been recently developed [32]. After the contributions of various uncertainty sources on the uncertainty of the output are identified, appropriate computational and experimental resources can be allocated to reduce the dominant uncertainty sources and thus reduce the uncertainty in the model prediction [33]. Here we do not focus on resource allocation; instead, we focus on how to perform effective uncertainty reduction based on the observed data already collected from experiments or field observations.

4 Uncertainty reduction

4.1 Bayesian updating for uncertainty reduction

A Bayesian approach is employed in this section for uncertainty reduction. The epistemic uncertainty sources modeled in Sect. 3.1 and as shown in Fig. 2 are updated using observation data. For the sake of illustration, we denote all the updating (calibration) parameters as $\boldsymbol{\theta}^{Cali}$, which include the parameters of both structural and material models. Suppose we have n_t observation data of the structure–material performance as $\mathbf{t}^e = [t_1^e,\ t_2^e,\ \ldots,\ t_{n_t}^e]$ as follows

$$t_i^e = G_{S-M}(\mathbf{x}(i),\ \mathbf{x}_M(i),\ \boldsymbol{\theta}^{Cali}) + \varepsilon_{obs}^*, \tag{14}$$

where $G_{S-M}(\mathbf{x}(i),\ \mathbf{x}_M(i),\ \boldsymbol{\theta}^{Cali})$ is the actual response under the ith experiment input setting $\mathbf{x}(i)$ and $\mathbf{x}_M(i)$, t_i^e is the ith observation, ε_{obs}^* is a realization of the observation noise ε_{obs}, and $\varepsilon_{obs} \sim N(0,\ \sigma_{obs}^2)$ is the observation noise.

The PDF of $\boldsymbol{\theta}^{Cali}$ can be updated using Bayesian calibration as

$$\begin{aligned} f\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}^e\right) &= \frac{L(\mathbf{t}^e|\,\boldsymbol{\theta}^{Cali})f(\boldsymbol{\theta}^{Cali})}{\int_{\boldsymbol{\theta}^{Cali}} L(\mathbf{t}^e|\,\boldsymbol{\theta}^{Cali})f(\boldsymbol{\theta}^{Cali})d\boldsymbol{\theta}^{Cali}} \\ &\propto L(\mathbf{t}^e\,|\,\boldsymbol{\theta}^{Cali})f(\boldsymbol{\theta}^{Cali}), \end{aligned} \tag{15}$$

where $f(\boldsymbol{\theta}^{Cali})$ is the prior distribution of the calibration parameters which are obtained from the uncertainty modeling of epistemic uncertainty sources in Sect. 3.1, $f(\boldsymbol{\theta}^{Cali}|\,\mathbf{t}^e)$ is the posterior distribution of $\boldsymbol{\theta}^{Cali}$ after uncertainty reduction using observations $\mathbf{t}^e$, and $L(\mathbf{t}^e|\,\boldsymbol{\theta}^{Cali})$ is the likelihood

function of observing $\mathbf{t}^e$ for given $\boldsymbol{\theta}^{Cali}$.$L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali})$ is computed by

$$L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali}) = \prod_{i=1}^{n_t} L\left(t_i^e|\boldsymbol{\theta}^{Cali}\right), \tag{16}$$

in which $L(t_i^e|\boldsymbol{\theta}^{Cali})$ is computed using the BN given in Fig. 2 by fixing the epistemic uncertainty sources at $\boldsymbol{\theta}^{Cali}$ as follows

$$\begin{aligned} &L\left(t_i^e|\boldsymbol{\theta}^{Cali}\right) \\ &= f\left(t_i^e|G_M\right) f\left(G_M|\boldsymbol{\theta}_e^M,\ \varepsilon_M,\ G_S\right) f\left(\varepsilon_M|G_S\right) \times \\ &f\left(G_S|\boldsymbol{\theta}_e^S,\ \delta_S,\ \boldsymbol{\omega}_a,\ \mathbf{x}(i)\right) f\left(\delta_S|\boldsymbol{\omega}_a,\ \mathbf{x}(i)\right) f\left(\boldsymbol{\omega}_a|\boldsymbol{\theta}_{ae}^S\right). \end{aligned} \tag{17}$$

Note that Eq. (14) is not used to compute the likelihood given in Eq. (15) because we do not know the true physics, but only the structure–material prediction model. Since the dimension of $\boldsymbol{\theta}^{Cali}$ is usually high, directly solving Eqs. (15) through (17) is computationally very expensive. To overcome this challenge, sampling-based methods have been developed to approximate the posterior distributions given by Eq. (15), such as the Markov Chain Monte Carlo (MCMC) [34] and the particle filter (PF) [35] methods. In this paper, the PF method is used. The PF represents the posterior distribution of the state variables by a set of particles that evolve and adapt recursively as new information is available. In each iteration, the particles are re-sampled according to the weights of the particles given by

$$weight(\boldsymbol{\theta}^{Cali}(i)) = \frac{L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali}(i))}{\sum_{i=1}^{n_{PF}} L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali}(i))}, \tag{18}$$

where $\boldsymbol{\theta}^{Cali}(i)$ is the ith particle generated from the prior distributions and n_{PF} is the number of samples in the PF.

By updating the prior distributions to the posterior distributions, the uncertainty in the epistemic uncertainty parameters is reduced as illustrated in Fig. 3. However, we encounter a new challenge discussed below during the uncertainty reduction. As shown in Fig. 3, if the model used for the updating does not adequately capture the physics, the updated posterior distribution may be far away from the true value of the calibration parameter even if the uncertainty in the parameter is reduced. This implies that Bayesian updating cannot always guarantee the accuracy of calibration (i.e., reduction of bias) even if it can reduce the variance. Next, we will analyze where this issue comes from and why it matters.

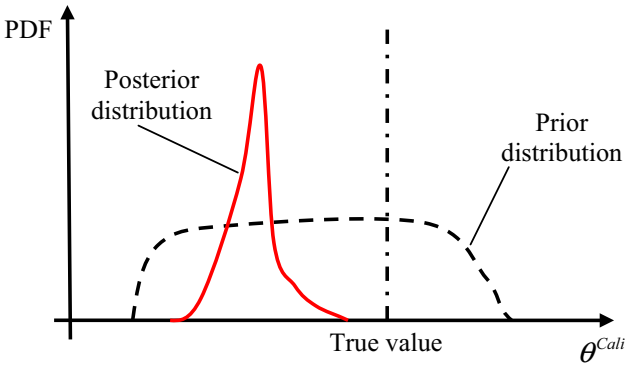


Fig. 3 Illustration of variance reduction versus bias in Bayesian calibration

4.1.1 Where does the error come from?

To identify where the error of Bayesian calibration comes from, we first look at Eq. (15). In Eq. (15), there are two terms that affect the posterior distribution, namely prior distribution ($f(\boldsymbol{\theta}^{Cali})$) and likelihood function ($L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali})$). For the Bayesian calibration example shown in Fig. 3, the prior distribution is able to cover the location of the true value. In that case, the error issue should be mainly attributed to the likelihood function.

As shown in Eq. (17), the likelihood function is computed using the BN given in Fig. 2 by fixing the calibration parameters $\boldsymbol{\theta}^{Cali}$. According to the BN and Eq. (5), we can define the corrected material damage model as

$$G_{MC} = G_M\left(\boldsymbol{\theta}_e^M,\ g_{sc}\right) + \varepsilon_M(g_{sc}), \tag{19}$$

where

$$g_{sc} = G_S(\mathbf{x},\ \boldsymbol{\omega}_a,\ \boldsymbol{\theta}_{ae},\ \boldsymbol{\theta}_e) + \delta_S, \tag{20}$$

in which g_{sc} is the corrected output of the structural analysis model.

Using the BN given in Fig. 2 to compute the likelihood function, we are trying to approximate the unknown statistical model of the actual physics $G_{S-M}(\mathbf{x},\ \mathbf{x}_M,\ \boldsymbol{\theta}^{Cali})$ given in Eq. (14) using the BN built based on the corrected structure–material prediction model given in Eqs. (19)–(20) or the node G_M in Fig. 2. This approximation may work in some design regions. But in some design region, it is quite often that

$$L_{S-M}\left(t_i^e|\boldsymbol{\theta}^{Cali}\right) = L\left(t_i^e|\boldsymbol{\theta}^{Cali}\right) + \delta_L\left(t_i^e|\boldsymbol{\theta}^{Cali}\right), \tag{21}$$

where $L_{S-M}(t_i^e|\boldsymbol{\theta}^{Cali})$ is the likelihood of observing t_i^e from the true physics if the calibration parameters equal to $\boldsymbol{\theta}^{Cali}$ and $\delta_L(t_i^e)$ is the error of the likelihood computation using the BN.

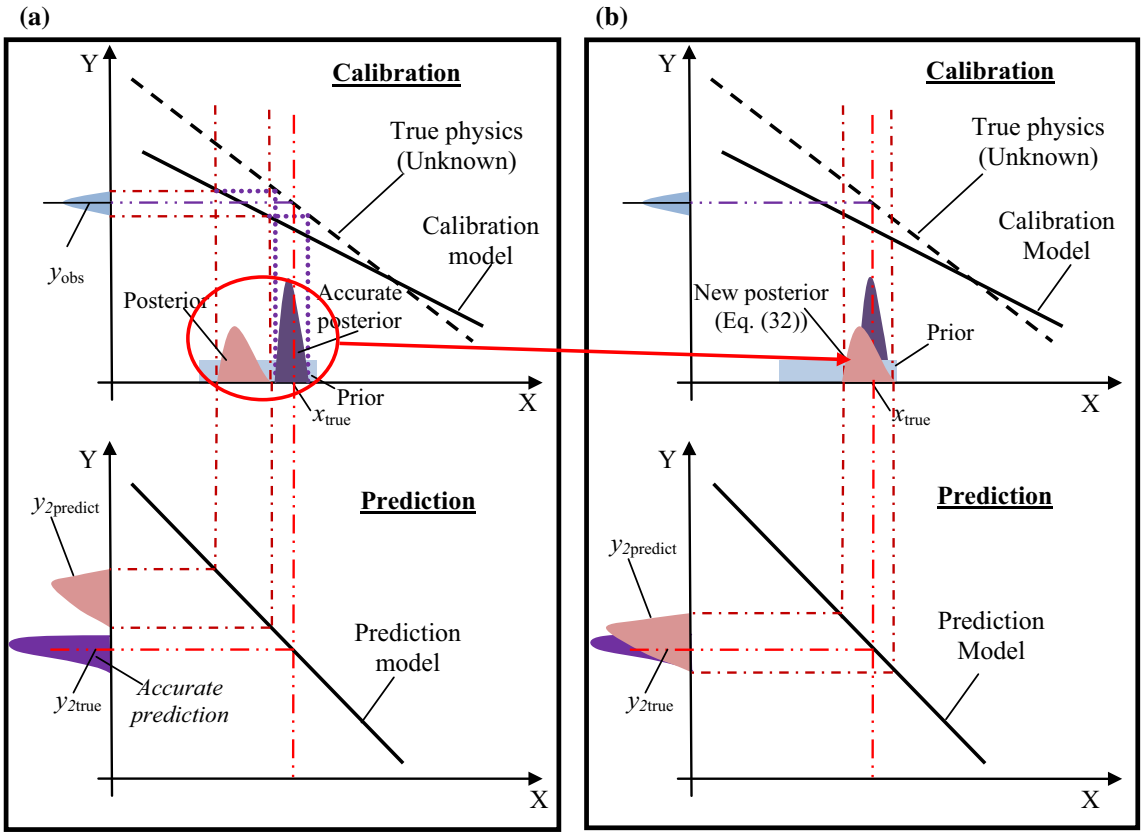


Fig. 4 Error propagation and reduction from calibration to prediction. **a** Direct Bayesian calibration and **b** adaptive Bayesian calibration

As a result, we also have error in the computed $L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali})$ as follows

$$L_{S-M}(\mathbf{t}^e|\boldsymbol{\theta}^{Cali}) = L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali}) + \delta_L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali}), \tag{22}$$

If the true physics (Eq. 13) is used for Bayesian calibration, we have

$$\begin{aligned} f_{acc}(\boldsymbol{\theta}^{Cali}|\mathbf{t}^e) &= \frac{\left[L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali}) + \delta_L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali})\right] f(\boldsymbol{\theta}^{Cali})}{\int_{\boldsymbol{\theta}^{Cali}} \left[L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali}) + \delta_L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali})\right] f(\boldsymbol{\theta}^{Cali}) d\boldsymbol{\theta}^{Cali}}, \\ &\propto \left[L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali}) + \delta_L(\mathbf{t}^e|\boldsymbol{\theta}^{Cali})\right] f(\boldsymbol{\theta}^{Cali}), \end{aligned} \tag{23}$$

where $f_{acc}(\boldsymbol{\theta}^{Cali}|\mathbf{t}^e)$ is the accurate posterior distribution obtained using the true physics.

It is apparent that $f_{acc}(\boldsymbol{\theta}^{Cali}|\mathbf{t}^e)$ and $f(\boldsymbol{\theta}^{Cali}|\mathbf{t}^e)$ will be different due to the error in the likelihood computation. The fundamental reason is that the likelihood function given in Eq. (15) is computed using prediction model while the observation data are collected from the actual physics. Even if the model discrepancy and bias correction terms are applied to correct the prediction model, it cannot be guaranteed that the statistical property of the corrected calibration model is always close to that of the actual physics. In Fig. 4a, we use a simple one-dimensional example to illustrate this point. In this figure, we have a calibration model after bias correction, which is used to approximate the actual physics. We want to calibrate x_{true} based on an observation y_{obs} obtained from the experiment (i.e., true physics). Since there is aleatory uncertainty in the experimental observation, y_{obs} can follow a probability distribution. We then solve the inverse problem based on the calibration model. As shown in the upper half of Fig. 4a, due to the difference (i.e., bias) between the calibration model and the actual physics, the obtained posterior distribution ($f(\boldsymbol{\theta}^{Cali}|\mathbf{t}^e)$) is quite away from the true value x_{true} and the accurate posterior distribution ($f_{acc}(\boldsymbol{\theta}^{Cali}|\mathbf{t}^e)$) obtained from Bayesian calibration using the actual physics.

Does the discrepancy between the posterior distribution and the true value matter? Next, we will analyze when and why it matters.

4.1.2 Why does it matter?

The discrepancy between the posterior distribution and the true value may not always matter for uncertainty quantification in the prediction. In calibration, we are solving the inverse problem corresponding to Eq. (23):

$$t_i^e = G_{MC}\left(\mathbf{x}(i),\ \mathbf{x}_M(i),\ \boldsymbol{\theta}_{MC}^{Cali}\right) + \varepsilon_{obs}^*, \tag{24}$$

where $G_{MC}(\cdot)$ is the combined structure–material prediction model, and $\boldsymbol{\theta}_{MC}^{Cali}$ are $\boldsymbol{\theta}^{Cali}$ in the calibration model. However, if we knew the true physics, we would solve the following inverse problem:

$$t_i^e = G_{S-M}\left(\mathbf{x}(i),\ \mathbf{x}_M(i),\ \boldsymbol{\theta}_{S-M}^{Cali}\right) + \varepsilon_{obs}^*, \tag{25}$$

where $\boldsymbol{\theta}_{S-M}^{Cali}$ are $\boldsymbol{\theta}^{Cali}$ corresponding to the actual physics.

If there is a one-to-one relationship between the calibration parameters and the prediction QOI, due to the link of t_i^e, we have the following relationship

$$\boldsymbol{\theta}_{MC}^{Cali} = G_{MC}^{-1}\left(\mathbf{x}(i),\ \mathbf{x}_M(i),\ G_{S-M}\left(\mathbf{x}(i),\ \mathbf{x}_M(i),\ \boldsymbol{\theta}_{S-M}^{Cali}\right)\right), \tag{26}$$

or

$$\boldsymbol{\theta}_{MC}^{Cali} = f_{map}\left(\mathbf{x}(i),\ \mathbf{x}_M(i),\ \boldsymbol{\theta}_{S-M}^{Cali}\right), \tag{27}$$

$$L_{S-M}\left(t_i^e \middle| \boldsymbol{\theta}_{S-M}^{Cali}\right) = L\left(t_i^e \middle| \boldsymbol{\theta}_{MC}^{Cali}\right), \tag{28}$$

where $f_{map}(\mathbf{x}(i),\ \mathbf{x}_M(i),\ \boldsymbol{\theta}_{S-M}^{Cali})$ stands for a function mapping $\boldsymbol{\theta}_{S-M}^{Cali}$ to $\boldsymbol{\theta}_{MC}^{Cali}$.

Note that the above equations are based on the assumption that the difference between the actual physics and the calibration model is only the mean prediction. Equations (26)–(28) illustrate that this difference is accounted for by mapping $\boldsymbol{\theta}_{S-M}^{Cali}$ in actual physics to $\boldsymbol{\theta}_{MC}^{Cali}$ in the calibration model. It is also this mapping that causes the difference between the posterior distribution obtained from the calibration model and the accurate posterior distribution obtained from the actual physics. When the posterior distribution $f\left(\boldsymbol{\theta}_{MC}^{Cali} \middle| \mathbf{t}^e\right)$ is used for uncertainty quantification in the same calibration model, we have

$$f_{pre}^{MC}(t) = \int_{\boldsymbol{\theta}_{MC}^{Cali}} L\left(t \middle| \boldsymbol{\theta}_{MC}^{Cali}\right) f\left(\boldsymbol{\theta}_{MC}^{Cali} \middle| \mathbf{t}^e\right) \mathbf{d}\boldsymbol{\theta}_{MC}^{Cali}. \tag{29}$$

Similarly, when the true posterior distribution $f(\boldsymbol{\theta}_{S-M}^{Cali} \middle| \mathbf{t}^e)$ is used for uncertainty quantification in the actual physics, we have

$$f_{pre}^{S-M}(t) = \int_{\boldsymbol{\theta}_{S-M}^{Cali}} L_{S-M}\left(t \middle| \boldsymbol{\theta}_{S-M}^{Cali}\right) f\left(\boldsymbol{\theta}_{S-M}^{Cali} \middle| \mathbf{t}^e\right) \mathbf{d}\boldsymbol{\theta}_{S-M}^{Cali}. \tag{30}$$

According to Eq. (27), the uncertainty of $\boldsymbol{\theta}_{MC}^{Cali}$ is governed by the uncertainty of $\boldsymbol{\theta}_{S-M}^{Cali}$ and thus we can rewrite Eq. (29) as

$$f_{pre}^{MC}(t) = \int_{\boldsymbol{\theta}_{S-M}^{Cali}} L_{S-M}\left(t \middle| \boldsymbol{\theta}_{S-M}^{Cali}\right) f\left(\boldsymbol{\theta}_{S-M}^{Cali} \middle| \mathbf{t}^e\right) \mathbf{d}\boldsymbol{\theta}_{S-M}^{Cali}. \tag{31}$$

Comparing Eqs. (30) and (31), we have $f_{pre}^{S-M}(t) = f_{pre}^{MC}(t)$. It means that the discrepancy between the posterior distribution and the accurate posterior distribution does not matter for uncertainty quantification if the mean prediction is the only difference between the actual physics and calibration model, there is a one-to-one relationship between the calibration parameter and response, and the posterior distribution is used for prediction of the same calibration model. As shown in upper half of Fig. 4a, when the posterior distribution from calibration is applied to the same calibration model, the predicted distribution will be close to that from the physics.

But this is not always the case. Most of the assumptions made in the above derivations (Eqs. 24–31) will not hold for most problems. Since we can only obtain $f\left(\boldsymbol{\theta}_{MC}^{Cali} \middle| \mathbf{t}^e\right)$ from Bayesian calibration which is the posterior distribution after mapping as indicated in Eq. (26), $f\left(\boldsymbol{\theta}_{MC}^{Cali} \middle| \mathbf{t}^e\right)$ may be used as inputs of another model (prediction model) for prediction. This prediction model can be a completely different model for a different structure with the same material, or the prediction model of the same structure (calibration model) but under different application conditions, such as the structural analysis model under a different load condition. Due to the fact that the mapped posterior distribution ($f\left(\boldsymbol{\theta}_{MC}^{Cali} \middle| \mathbf{t}^e\right)$) is away from the true value, prediction based on that will have large errors. As depicted in the lower half of Fig. 4b, the discrepancy of the posterior distribution results in a large error in the prediction of the prediction model. This indicates that the discrepancy between the posterior distribution and the true value *matters* when the posterior distribution is applied in a different prediction model, i.e., model of a different structure, or the model of the same structure but under different conditions.

Motivated by this, in this paper, we propose an *adaptive Bayesian calibration* method to mitigate the effects of the statistical bias between the calibration model and the actual physics on the posterior distributions obtained from Bayesian calibration. The adaptive Bayesian calibration method has two features: (1) integration of Bayesian calibration and model validation, which is an extension of the integration method presented in Ref. [28]; and (2) an adaptive scheme to selectively choose among the observation data in order to maximize the accuracy of Bayesian calibration. As shown in Fig. 4b, in the proposed method, we want to reduce the discrepancy between the posterior distribution and the true value and thus increase the accuracy of the prediction when the calibrated parameter is applied to a new prediction model (as indicated in the lower half of Fig. 4b). In the new adaptive

Bayesian calibration method, we not only reduce the uncertainty, but also maximize the accuracy of the calibration. Details of the proposed method are given in the subsequent section.

4.2 Proposed adaptive Bayesian calibration

In this section, we first give the main idea of the proposed adaptive Bayesian calibration method. Following that, we provide details of the implementation procedure.

4.2.1 Overview of the adaptive Bayesian calibration method

Considering that the discrepancy between the posterior distribution and the true value of the calibration parameter is caused by the error in the likelihood function $L(t_i^e \,|\, \boldsymbol{\theta}^{Cali})$, we can improve the accuracy of Bayesian calibration by improving the accuracy of the likelihood function. Since the likelihood function is affected by the statistical bias between the prediction model and the actual physics as illustrated in Fig. 4a and the bias varies over the prediction domain, the accuracy of the likelihood function will also vary over the prediction domain. If the likelihood function is accurate in certain prediction regions, Bayesian calibration based on that using Eq. (15) is able to reduce the uncertainty accurately. Otherwise, the uncertainty is reduced incorrectly, i.e., bias of the posterior distribution is increased even if the variance is reduced. Since the physics model uncertainty is unknown, it is hard to exactly determine whether the likelihood computation is accurate or not. To address this issue, we consider the probability that the likelihood computation is accurate.

Based on the above analyses, according to the range of the collected observation data, we first discretize the performance prediction domain into several segments $[t_i^L,\ t_i^U]$, where t_i^L is the lower bound of the ith segment, t_i^U is the upper bound of the ith segment, $i = 1,\ 2,\ \ldots,\ N_T$, and N_T is the number of segments. Defining $\mathbf{t}^e \in [t_i^L,\ t_i^U]$ as $\mathbf{t}_i^e$, the corrected posterior distribution of $\boldsymbol{\theta}^{Cali}$ after considering the probability that the likelihood computation is accurate can be written as

$$f_{new}\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_i^e\right) = \Pr(\mathrm{L}_i)\, f\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_i^e,\ \mathrm{L}_i\right) + [1-\Pr(\mathrm{L}_i)]\, f(\boldsymbol{\theta}^{Cali}), \tag{32}$$

where $\Pr(\mathrm{L}_i)$ is the probability that the likelihood function computation is accurate for $[t_i^L,\ t_i^U]$, $f(\boldsymbol{\theta}^{Cali}\,|\,\mathbf{t}_i^e,\ \mathrm{L}_i)$ is the posterior distribution obtained from Bayesian calibration using $\mathbf{t}_i^e$ and Eq. (15), $f(\boldsymbol{\theta}^{Cali})$ is the prior distribution, and $f_{new}(\boldsymbol{\theta}^{Cali}\,|\,\mathbf{t}_i^e)$ is the new posterior distribution.

Equation (32) is a form of Bayesian model averaging which uses the probability that the likelihood computation is accurate as a weighting factor; as a result, $f_{new}(\boldsymbol{\theta}^{Cali}\,|\,\mathbf{t}_i^e)$is the distribution of $\boldsymbol{\theta}^{Cali}$ provided to the prediction model, and not simply the posterior distribution $f(\boldsymbol{\theta}^{Cali}\,|\,\mathbf{t}_i^e,\ \mathrm{L}_i)$. If the probability that the likelihood function is accurate is high, this formula assigns higher weight to the posterior distribution; otherwise the prior distribution gets a higher weight, since the posterior distribution obtained from the wrong likelihood will bring wrong information regarding the calibration parameters. The integrated Bayesian calibration method given in Eq. (31) enables us to use only the correct information (i.e., accurate likelihood) to reduce the uncertainty in our simulation and thus enhance the accuracy of Bayesian calibration.

In addition, since the range of observation data is divided into multiple segments, the posterior distribution obtained from the observation data in one segment can be employed as the prior when using data from another segment for subsequent calibration. Based on this idea, we rewrite Eq. (32) as follows

$$f_{new}\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_i^e\right) = \Pr(\mathrm{L}_i)\, f\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_i^e,\ \mathrm{L}_i\right) + [1-\Pr(\mathrm{L}_i)]\, f_{new}\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_j^e\right), \tag{33}$$

where $f_{new}(\boldsymbol{\theta}^{Cali}\,|\,\mathbf{t}_j^e)$ is the posterior distribution obtained using Eq. (31) based on the observation data of the previous segment $[t_j^L,\ t_j^U]$.

When $\Pr(\mathrm{L}_i)$ is very small (e.g., $\Pr(\mathrm{L}_i) < 0.05$), it means that the posterior distribution $f(\boldsymbol{\theta}^{Cali}\,|\,\mathbf{t}_i^e,\ \mathrm{L}_i)$ is not useful at all. Since Bayesian calibration is computationally expensive, to save computational effort, we rewrite Eq. (33) as

$$f_{new}\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_i^e\right) = \begin{cases} \Pr(\mathrm{L}_i)\, f\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_i^e,\ \mathrm{L}_i\right) \\ \quad +[1-\Pr(\mathrm{L}_i)] \\ \quad f_{new}(\boldsymbol{\theta}^{Cali}\,|\,\mathbf{t}_j^e), & \text{if } \Pr(\mathrm{L}_i) \geq 0.05 \\ f_{new}\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_j^e\right), & \text{otherwise}. \end{cases} \tag{34}$$

A key step in implementing the adaptive Bayesian calibration method given in Eq. (34) is the computation of $\Pr(\mathrm{L}_i)$. However, estimating $\Pr(\mathrm{L}_i)$ is very difficult given the fact that the actual physics is unknown. $L(t_i^e\,|\,\boldsymbol{\theta}^{Cali})$ is computed using the PDF of the structure–material performance prediction at t_i^e; since the PDF is the first derivative of the cumulative density function (CDF) of the prediction, we assume that the PDF computation (and thus the likelihood computation) is accurate if the CDF computation from the prediction model is accurate. Based on this assumption, we have

$$\Pr(\mathrm{L}_i) \approx \Pr(C_i), \tag{35}$$

where C_i is the event that the CDF computation over segment $[t_i^L,\ t_i^U]$ is accurate and $\Pr(\cdot)$ is the probability operator.

Equations (32) and (34) are then rewritten as

$$f_{new}\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_i^e\right) = \Pr(C_i)\, f\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_i^e,\ C_i\right) + [1-\Pr(C_i)]\, f(\boldsymbol{\theta}^{Cali}), \tag{36}$$

$$f_{new}\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_i^e\right) = \begin{cases} \Pr(C_i)\, f\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_i^e,\ \mathrm{L}_i\right) \\ \quad +[1-\Pr(C_i)] \\ \quad f_{new}\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_j^e\right), & \text{if } \Pr(C_i) \geq 0.05 \cdot \\ f_{new}\left(\boldsymbol{\theta}^{Cali}\middle|\,\mathbf{t}_j^e\right), & \text{otherwise} \end{cases} \tag{37}$$

Model validation can be employed to estimate $\Pr(C_i)$. In the following section, we discuss how to estimate $\Pr(C_i)$ using model validation; then in Sect. 4.2.3 we discuss how to implement the adaptive Bayesian calibration in detail.

4.2.2 *Model validation for estimation of* $\Pr(C_i)$

Model validation is a process of quantifying the agreement between the prediction model and experiment observations for the intended use of the model [36,37]. Several model validation metrics have been studied, include classical and Bayesian hypothesis testing [38], reliability metric [37,39], area metric [40], and others. Bayesian hypothesis testing (BHT) and the reliability metric give the model validation result as a probability measure, and thus are suitable for implementing Eqs. (36) and (37); BHT is illustrated here.

In the BHT method, we check the plausibility of two hypotheses: the null hypothesis (H_0) and the alternative hypothesis (H_1). For given observation data $\mathbf{t}_i^e$, the likelihood ratio of the two hypotheses, known as Bayes factor (B), is computed as [37]

$$B = \frac{\Pr\left(\mathbf{t}_i^e\middle|\,H_0\right)}{\Pr\left(\mathbf{t}_i^e\middle|\,H_1\right)}, \tag{38}$$

The Bayes factor is originally developed to compare the data support for two models. It can be extended to assess the validity of a model (M) as follows

$$B = \frac{\Pr\left(\mathbf{t}_i^e\middle|\,M \text{ is valid}\right)}{\Pr\left(\mathbf{t}_i^e\middle|\,M \text{ is not valid}\right)}, \tag{39}$$

To evaluate the accuracy of the CDF estimate over a certain segment of the output quantity of interest (e.g., life interval in the context of life prediction), Bayes factor has been further extended by Zhang and Mahadevan as [16]

$$B(C_i) = \frac{P\left(\mathbf{t}_i^e\,|H_0 : C_{itrue} = C_i\right)}{P\left(\mathbf{t}_i^e\,|H_1 : C_{itrue} \neq C_i\right)}, \tag{40}$$

where C_{itrue} is the true CDF over $[t_i^L,\ t_i^U]$, and C_i is the predicted CDF.

For n_t independent experiments, with the CDF over $[t_i^L,\ t_i^U]$ as C_i, if k_i experiment data are observed in the segment $[t_i^L,\ t_i^U]$, the probability of observing the data is given by

$$P(k_i; C_i, n_t) = \binom{n_t}{k_i} C_i^{k_i}(1-C_i)^{n_t-k_i}. \tag{41}$$

If the CDF over $[t_i^L,\ t_i^U]$ is not C_i, it can be any value over interval [0, 1]. The probability of observing the data is then computed by

$$\begin{aligned} &P(k_i; C_{itrue} \in [0,\ 1], n_t) \\ &= \int_0^1 \binom{n_t}{k_i} C_{itrue}^{k_i}(1-C_{itrue})^{n_t-k_i} dC_{itrue} \\ &= \binom{n_t}{k_i} \frac{k_i!(n_t-k_i)!}{(n_t+1)!}. \end{aligned} \tag{42}$$

Combining Eqs. (40)–(42), we have

$$\begin{aligned} B(C_i) &= \frac{P\left(t^e\,|H_0 : C_{itrue} = C_i\right)}{P\left(t^e\,|H_1 : C_{itrue} \neq C_i\right)} \\ &= \frac{C_i^{k_i}(1-C_i)^{n_t-k_i}(n_t+1)!}{k_i!(n_t-k_i)!}, \\ &= C_i^{k_i}(1-C_i)^{n_t-k_i}(n_t+1)\binom{n_t}{k_i}. \end{aligned} \tag{43}$$

In Eqs. (40)–(43), C_i is computed based on the CDF of the prediction model given in Fig. 2 by only considering aleatory uncertainty as follows

$$C_i = F_T\left(t_i^U\right) - F_T\left(t_i^L\right), \tag{44}$$

where $F_T(t)$ is the CDF of the life distribution prediction from the model.

As discussed in Sect. 3, we have epistemic uncertainty sources $\boldsymbol{\theta}^{Cali}$ present in the structure–material model. For a given realization of the epistemic uncertainty sources $\boldsymbol{\theta}^{Cali}$, $C_i\,|\,\boldsymbol{\theta}^{Cali}$ is given by

$$C_i\,|\,\boldsymbol{\theta}^{Cali} = F_T\left(t_i^U\right)\Big|\,\boldsymbol{\theta}^{Cali} - F_T\left(t_i^L\right)\Big|\,\boldsymbol{\theta}^{Cali}, \tag{45}$$

in which $F_T(t_i^U)\big|\,\boldsymbol{\theta}^{Cali}$ is the CDF of the life distribution by fixing the epistemic uncertainty sources at $\boldsymbol{\theta}^{Cali}$ and propagating the aleatory uncertainty through the BN model.

The probability of observing k_i experiment data in the segment $[t_i^L,\ t_i^U]$ becomes

$$P\left(H_0:\ \text{CDF is accurate in interval } \left[t_i^L,\ t_i^U\right]\right) = \binom{n_t}{k_i}\int_{\boldsymbol{\theta}^{Cali}}\left(C_i^{k_i}\,\middle|\,\boldsymbol{\theta}^{Cali}\right)(1-C_i\,|\,\boldsymbol{\theta}^{Cali})^{n_t-k_i}d\boldsymbol{\theta}^{Cali}. \tag{46}$$

The Bayes factor of the predicted CDF over $[t_i^L,\ t_i^U]$ after considering epistemic uncertainty is given by

$$B_i = \frac{P\left(\mathbf{t}_i^e\,|H_0:\text{CDF estimate is accurate in }\left[t_i^L,\ t_i^U\right]\right)}{P\left(\mathbf{t}_i^e\,|H_1:\text{CDF estimate is not accurate in }\left[t_i^L,\ t_i^U\right]\right)}, = \int_{\boldsymbol{\theta}^{Cali}}\left(C_i^{k_i}\,\middle|\,\boldsymbol{\theta}^{Cali}\right)(1-C_i\,|\,\boldsymbol{\theta}^{Cali})^{n_t-k_i}d\boldsymbol{\theta}^{Cali} \times\frac{(n_t+1)!}{k_i!(n_t-k_i)!}. \tag{47}$$

where B_i is the Bayes factor corresponding to the CDF estimate over $[t_i^L,\ t_i^U]$ after considering epistemic uncertainty.

Note that in the above derivation of the Bayes factor, only the number of observed experiment data in the segment $[t_i^L,\ t_i^U]$ is used. Eq. (47) can therefore be rewritten as

$$B_i = \frac{P\left(k_i\text{ experiment data observed in }\left[t_i^L,\ t_i^U\right]\,|H_0:\text{CDF estimate is accurate in }\left[t_i^L,\ t_i^U\right]\right)}{P\left(k_i\text{ experiment data observed in }\left[t_i^L,\ t_i^U\right]\,|H_1:\text{CDF estimate is not accurate in }\left[t_i^L,\ t_i^U\right]\right)}, = \int_{\boldsymbol{\theta}^{Cali}}\left(C_i^{k_i}\,\middle|\,\boldsymbol{\theta}^{Cali}\right)(1-C_i\,|\,\boldsymbol{\theta}^{Cali})^{n_t-k_i}d\boldsymbol{\theta}^{Cali}\frac{(n_t+1)!}{k_i!(n_t-k_i)!}. \tag{48}$$

Based on the above Bayes factor, the posterior distribution of $P(H_0:$ CDF estimate is accurate in $[t_i^L,\ t_i^U]|k_i$ experiment data observed in $[t_i^L,\ t_i^U])$ can be computed as [41]

$$P\left(H_0:\text{CDF estimate is accurate in }\left[t_i^L,\ t_i^U\right]\,\middle|\, k_i\text{ experiment data observed in }\left[t_i^L,\ t_i^U\right]\right) = \frac{B_i}{B_i+1}. \tag{49}$$

Note that the above expression is obtained based on equal prior probabilities of 0.5 for H_0 and H_1 (i.e., no prior information regarding model validity). We then approximate $\Pr(C_i)$ as

$$\Pr\left(C_i\,|\,k_i\right) \approx P\left(H_0:\text{CDF estimate is accurate in }\left[t_i^L,\ t_i^U\right]\,\middle|\, k_i\text{ experiment data observed in }\left[t_i^L,\ t_i^U\right]\right). \tag{50}$$

The above approximation may not be accurate since the probability that the CDF estimate is accurate in $[t_i^L,\ t_i^U]$ only provides an overall evaluation of the CDF accuracy in that interval. If the number of experimental data is large enough, we are able to obtain a better approximation based on other validation methods, such as the area metric method which compares the CDF curves directly. How to further improve the accuracy of the approximation given in Eq. (49) is worth studying in future research.

Substituting Eq. (50) into Eq. (36) and (37), we have

$$f_{new}\left(\boldsymbol{\theta}^{Cali}\,\middle|\,\mathbf{t}_i^e\right) = \Pr(C_i\,|\,k_i)\,f\left(\boldsymbol{\theta}^{Cali}\,\middle|\,\mathbf{t}_i^e,\ C_i\right) + [1-\Pr(C_i\,|\,k_i)]f(\boldsymbol{\theta}^{Cali}), \tag{51}$$

$$f_{new}\left(\boldsymbol{\theta}^{Cali}\,\middle|\,\mathbf{t}_i^e\right) = \begin{cases} \Pr\left(C_i\,|\,k_i\right) f\left(\boldsymbol{\theta}^{Cali}\,\middle|\,\mathbf{t}_i^e,\ C_i\right) \\ \quad +[1-\Pr\left(C_i\,|\,k_i\right)] \\ \quad f_{new}\left(\boldsymbol{\theta}^{Cali}\,\middle|\,\mathbf{t}_j^e\right), & \text{if } \Pr\left(C_i\,|\,k_i\right)\ge 0.05 \\ f_{new}\left(\boldsymbol{\theta}^{Cali}\,\middle|\,\mathbf{t}_j^e\right), & \text{otherwise} \end{cases}, \tag{52}$$

where $\Pr(C_i\,|\,k_i)$ is the probability that the CDF estimate is accurate given that experiment data are observed in $[t_i^L,\ t_i^U]$.

Next, we discuss how to implement the proposed integrated Bayesian calibration method based on the discussions given in Sects. 4.2.1 and 4.2.2.

4.2.3 Implementation procedure

Suppose n_t observations are collected and denoted as $\mathbf{t}^e = [t_1^e,\ t_2^e,\ \ldots,\ t_{n_t}^e]$. As discussed in Sect. 4.2.1, we first discretize the response prediction domain into N_T segments as $[t_i^L,\ t_i^U], i = 1,\ 2,\ \ldots,\ N_T$. For each segment, we perform the integrated Bayesian calibration using observation data $\mathbf{t}_i^e,\ i = 1,\ 2,\ \ldots,\ N_T$. As mentioned in Sect. 4.2.1, the posterior distributions obtained from one segment can be used as the prior distributions of another segment. If the accuracy of the likelihood function is not considered in Bayesian calibration, the sequence of calibrations may not be very important. However, when the accuracy of the likelihood function is included in the Bayesian calibration (i.e., the adaptive Bayesian calibration method), the accuracy of the estimated $\Pr(C_i)$ will be affected by the epistemic uncertainty sources as shown in Eq. (48). The accuracy of $\Pr(C_i)$ will then affect the accuracy of Bayesian calibration results. To minimize the effect of the epistemic

Table 1 Pseudo code of the adaptive Bayesian calibration method

Step	Description
1	Discretize the observation data domain into N_T segments, $[t_i^L, t_i^U]$, $i=1,2,\cdots,N_T$.
2	Denote the observation data belong to $[t_i^L, t_i^U]$ as $\mathbf{t}_i^e$, $i=1,2,\cdots,N_T$.
3	**Set** $N_T^{new}=N_T$.
4	**For** i=1: N_T
5	Compute the Bayes factor of each segment based on the prior distribution of $\boldsymbol{\theta}^{Cali}$ and the aleatory uncertainty information using the BN given in Fig. 2, the method presented in Sec. 4.2.2, and Eq. (48). Denote the obtained Bayes factor as $B_i, i=1,2,\cdots,N_T^{new}$.
6	Identify the segment with the highest Bayes factor by $i_{\max}=\underset{i=1,2,\cdots,N_T^{new}}{\arg\max}\{B_i\}$
7	Perform Bayesian calibration using the Eqs. (15)-(18) and observation data $\mathbf{t}_{i_{\max}}^e$ and obtain the posterior samples of $f(\boldsymbol{\theta}^{Cali}\vert\mathbf{t}_{i_{\max}}^e, C_{i_{\max}})$.
8	Compute $\Pr(C_i\vert k_i)$ using Eqs. (49) and (50) and $B_{i_{\max}}$.
9	Obtain the samples of $f_{new}(\boldsymbol{\theta}^{Cali}\vert\mathbf{t}_i^e)$ using weighted sampling method based on the samples of $f(\boldsymbol{\theta}^{Cali}\vert\mathbf{t}_{i_{\max}}^e, C_{i_{\max}})$, the prior distribution samples, and Eqs. (51) and (52).
10	Update the prior distributions of $\boldsymbol{\theta}^{Cali}$ as $f_{new}(\boldsymbol{\theta}^{Cali}\vert\mathbf{t}_i^e)$.
11	Remove $\mathbf{t}_i^e$ from the observation data and remove segment $[t_{i_{\max}}^L, t_{i_{\max}}^U]$.
12	$N_T^{new}=N_T^{new}-1$.
13	**End**
14	Obtain the posterior distribution $\boldsymbol{\theta}^{Cali}$ from the adaptive Bayesian calibration method.

uncertainty on the estimation of $\Pr(C_i)$, we first perform the adaptive Bayesian calibration for the segment with the highest Bayes factor, which gives the highest credibility in the estimation of $\Pr(C_i)$. After the epistemic uncertainty is reduced, we perform the adaptive Bayesian calibration again for a new segment with the highest Bayes factor among the remaining intervals. This process continues until all the segments that have Bayes factor above the desired threshold are included.

In Table 1, we provide a pseudo code of the adaptive Bayesian calibration method.

From the implementation procedure presented in Table 1, it can be found that the Bayes factors are updated adaptively over the iterations and the Bayesian calibration is performed adaptively. By implementing the proposed adaptive Bayesian calibration method, we can achieve the purpose of uncertainty reduction of the structure–material performance prediction model while also maximizing the accuracy of calibration. Next, we will use the composite rotorcraft hub component example to illustrate the proposed framework.

5 Composite rotorcraft hub component life prediction

5.1 Structure–material model

A composite rotorcraft hub component shown in Fig. 5 is used to illustrate the proposed framework. The rotorcraft hub is made of a laminated composite to reduce weight, drag, and number of parts [42]. The rotor hub yoke is subjected to axial tensile load caused by the rotation of the rotor, and a cyclic bending load as a result of the interaction of the rotor passage with the fuselage or other static structures and aerodynamic gust loads. During flight, the hub could fail due to delamination between individual plies at the thick-to-taper transition resin pocket caused by fatigue due to centrifugal and bending loads. To estimate the fatigue life of the component, a structure–material performance assessment model is developed [42].

As depicted in Fig. 6, during the analysis, a constant axial load (P_h) is applied at the end of the component to simulate the constant centrifugal load. A cyclic load (P_v) is also

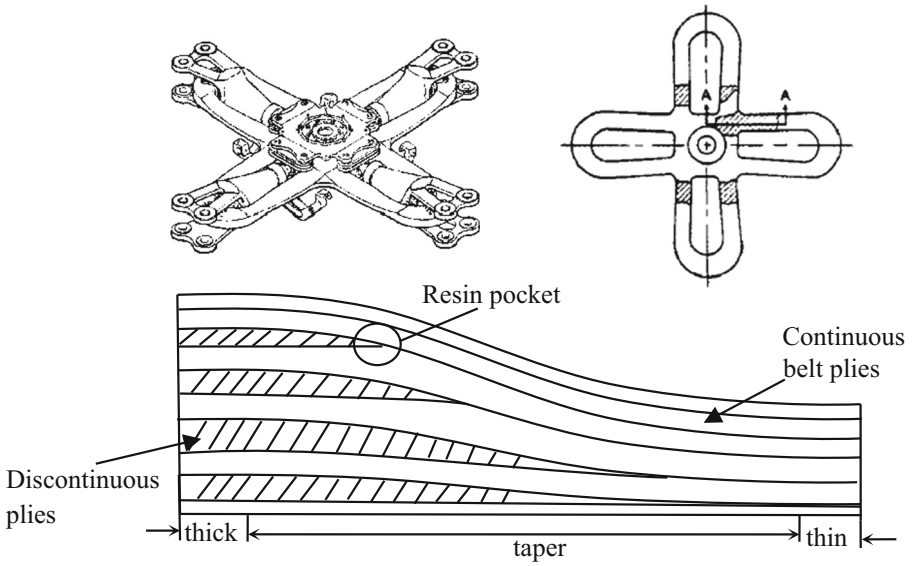


Fig. 5 Half of the symmetric section of tapered composite component [43]

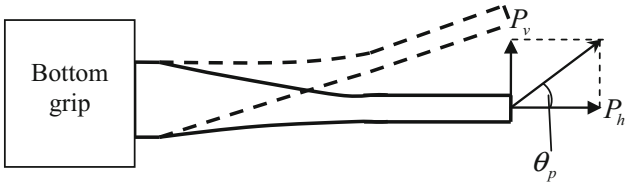


Fig. 6 Illustration of the load condition of the component

applied to the component, which induces an angular displacement, simulating the flexural bending in the yoke [42]. The angle between the axial load and the cyclic load is $\boldsymbol{\theta}_p$. Due to the combined axial and cyclic load, the component may fail due to delamination at the thick to taper transition caused by an initial tension crack at the internal ply drop-offs.

It is assumed that delamination onset occurs when the strain energy release rate G_e exceeds a critical value G_{crit}, which is derived from material coupon delamination tests [42,44]. A 2D ABAQUS structural analysis model for the helicopter rotor hub is first built to assess the fatigue delamination using the virtual crack closure technique (VCCT) as shown in Fig. 7. We first construct a finite element analysis (FEA) model for hub specimen (as shown in Fig. 7a, b). We then compute the strain energy release rate associated with the fatigue delamination of the component under given geometry, material property, and loads using VCCT. Figure 7c shows the stress contour of the analysis result. Figure 8 gives the results of strain energy release rate. The finite element mesh is refined around the crack tip to accurately compute the energy release rate.

Based on the composite material coupon delamination testing, an empirical material damage model is developed to estimate the relationship between the fatigue life and G_e as follows [42]

$$G_e = A_G - B_G \log_{10}(N_F), \tag{53}$$

where G_e is the critical energy release rate, N_F is the number of load cycles, and A_G and B_G are two parameters estimated based on experimental data.

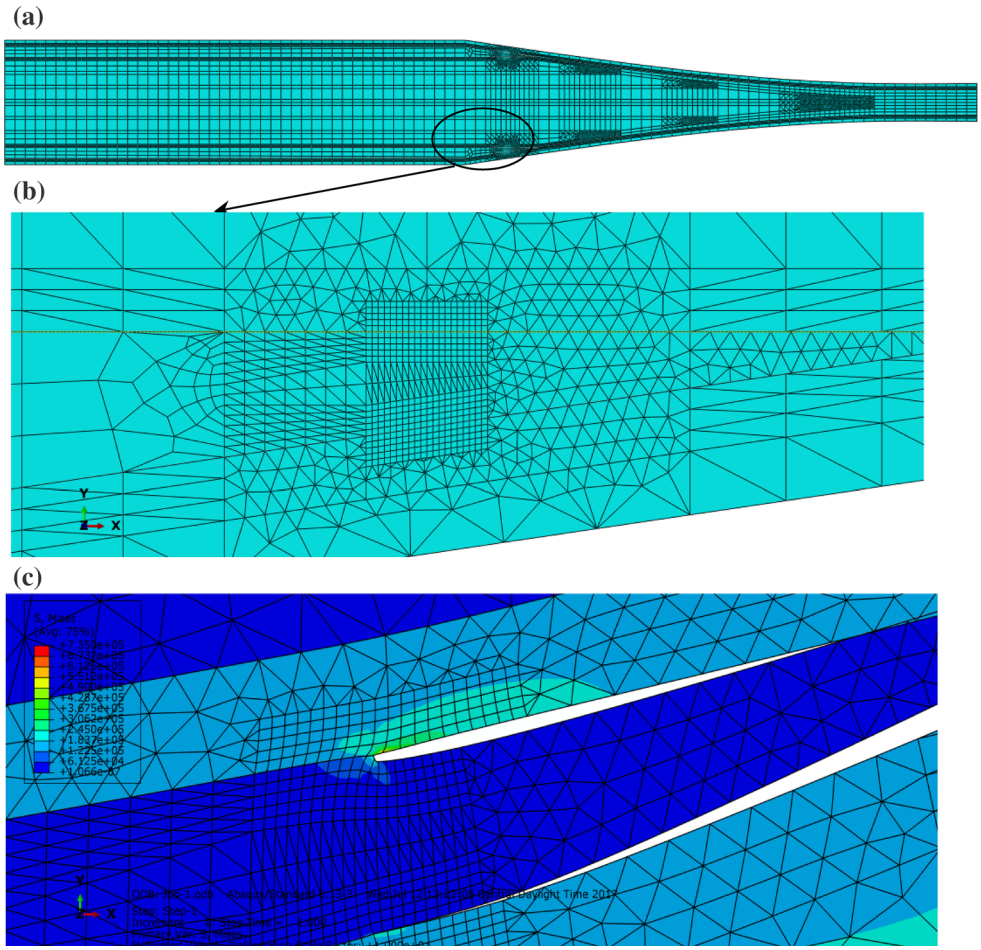
Due to the uncertainty in the experimental data, A_G and B_G are estimated as $A_G \sim N(439.3,\ 25^2)$ and $B_G \sim N(56.48,\ 3^2)$, where $N(\mu,\ \sigma^2)$ represents a normal distribution with mean μ and standard deviation σ. There are also other sources of uncertainty present in the above structure–material life prediction model. Next, we aggregate these uncertainty sources into the uncertainty of the fatigue life.

5.2 Uncertainty aggregation

Before performing uncertainty aggregation, we model the various sources of uncertainty using the method presented in Sect. 3.1. Based on collected data, we model aleatory variables P_l and θ_p as normal distributions and we have $P_h = P_l \sin(\theta_P)$ and $P_v = P_l \cos(\theta_P)$. Due to limited data, we have epistemic uncertainty regarding the standard deviations of P_l and θ_p. Since the finite element model is used predict G_e, we also have model discrepancy of G_e. In the KOH framework, we usually model the discrepancy term as a Gaussian process (GP) model which is a function of input variables. In a GP model, the model discrepancy term is model by

$$\delta_G(\mathbf{x}) = \mathbf{h}(\mathbf{x})^T \boldsymbol{\upsilon} + \varepsilon(\mathbf{x}), \tag{54}$$

where $\boldsymbol{\upsilon}$ is a vector of unknown coefficients, $\mathbf{h}(\mathbf{x})$ is a vector of regression functions, $\mathbf{h}(\mathbf{x})^T \boldsymbol{\upsilon}$ is the trend of prediction, and



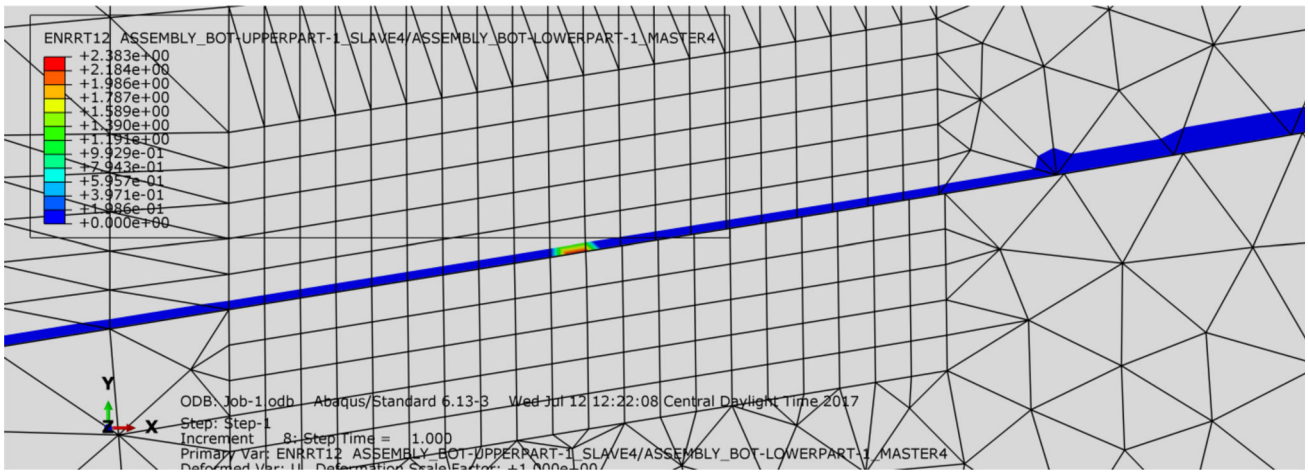


Fig. 7 Simulation of a laminated composite rotorcraft hub component. **a** FEA model of laminated composite rotorcraft hub component, **b** refined mesh and **c** stress contour

Fig. 8 Result of energy release rate obtained from ABAQUS simulation

$\varepsilon(\mathbf{x})$ is assumed to be a Gaussian process with zero mean and covariance $Cov[\varepsilon(\mathbf{x}_i),\ \varepsilon(\mathbf{x}_j)]$.

The covariance between two points $\mathbf{x}_i$ and $\mathbf{x}_j$ is given by

$$Cov[\varepsilon(\mathbf{x}_i),\ \varepsilon(\mathbf{x}_j)] = \sigma_\varepsilon^2 R(\mathbf{x}_i,\ \mathbf{x}_j), \tag{55}$$

where $R(\mathbf{x}_i,\ \mathbf{x}_j)$ is the correlation function and σ_ε^2 is the variance of $\varepsilon(\mathbf{x})$.

Ling et al. [14] investigated different options for the stochastic second term in Eq. (54), such as constant, normal distribution, input-dependent normal distribution, stationary

Gaussian process, and non-stationary Gaussian process. (The more sophisticated options were found to reduce the bias but increase the variance in the calibration posteriors; however, in prediction, the performance of the different options depended on the difference between the calibration and prediction configurations. The adaptive calibration method in Sect. 4 addresses this issue). In this example, for the sake of illustration, we model the discrepancy term as

$$\delta_G = 0.5P_l + 1.2\theta_p + 0.07P\theta_p + \varepsilon_G, \tag{56}$$

using a linear trend function with an interaction term, and the stochastic term is modeled simply as a normal variable, i.e., $\varepsilon_G \sim N(0,\ \sigma_\delta^2)$. To account for the error of the assumed discrepancy term, we have uncertainty in σ_δ, which is updated using calibration data. The coefficients of the trend function may also be uncertain and calibrated using observation data; however, for simplicity, we assume that these coefficients are known constants from previous knowledge.

Table 2 Various sources of uncertainty in the composite rotorcraft hub problem

Variable	Parameter 1	Parameter 2	Distribution type
E_{11} (Msi)	6.7	0.1	Normal
P_l (kips)	27	σ_p	Normal
θ_p(°)	10	σ_{θ_p}	Normal
σ_p (kips)	3.5	6	Uniform
σ_{θ_p}(°)	1	1.2	Uniform
σ_δ	4	7	Uniform
A_G	439.3	25	Normal
B_G	56.48	3	Normal
ε_{obs} (cycles)	0	20	Normal

In addition to the uncertainty sources in the model and the model inputs, we also assume aleatory uncertainty in the material property E_{11} and the observation error ε_{obs}. Table 2 summarizes various sources of aleatory and epistemic uncertainty in this problem. In this table, "parameter 1" and "parameter 2" denote the mean and standard deviation for a normal distribution, and the lower and upper bounds for a uniform distribution.

Since the finite element analysis model is computationally expensive for uncertainty aggregation, we build a surrogate model for the critical energy release rate using the Kriging surrogate modeling method [45]. Due to the surrogate model, surrogate model uncertainty is also introduced in the prediction. For given fixed inputs, the surrogate model prediction is a random variable modeled as a normal distribution. Based on that, we connect various sources of uncertainty, the structural analysis model, the material damage model using the BN method discussed in Sect. 3.2. Figure 9 shows the constructed BN for the composite rotorcraft hub example.

We then perform uncertainty aggregation using the BN and the uncertainty information given in Table 2. Figure 10 shows the comparison of fatigue life distribution of the rotorcraft hub obtained by only considering the aleatory uncertainty sources and by aggregating all the uncertainty sources in Table 2. It shows that the uncertainty in the prediction will be increase after considering both aleatory and epistemic uncertainty.

Next we perform global sensitivity analysis to quantify the contributions of different aleatory and epistemic uncertainty sources towards the uncertainty in the life prediction. Table 3

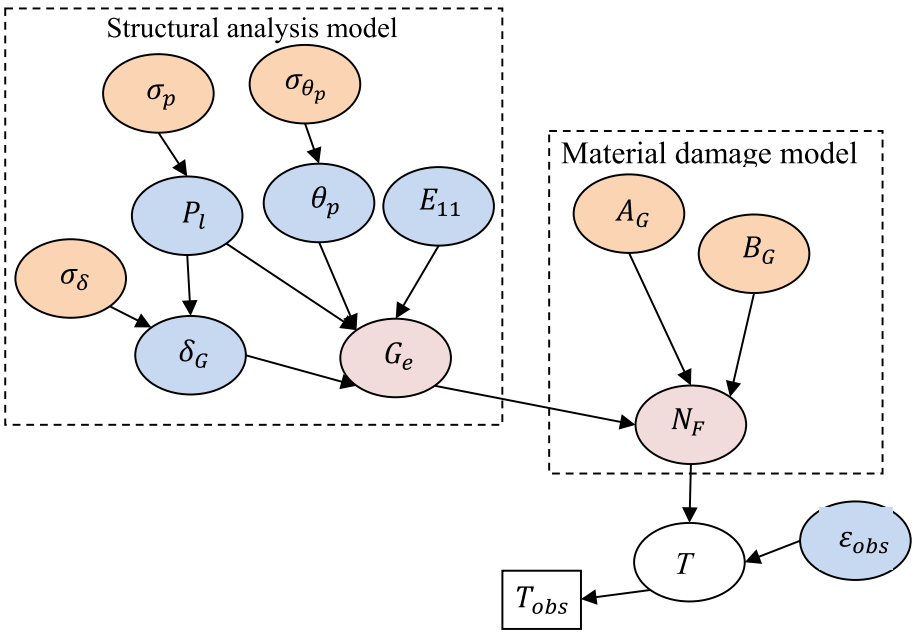


Fig. 9 Bayesian network for the fatigue life prediction model

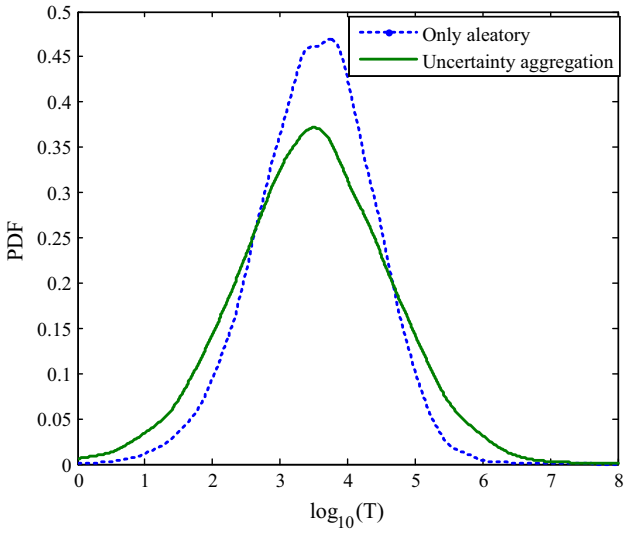


Fig. 10 Fatigue life distribution obtained from uncertainty aggregation

gives the first-order Sobol indices for the uncertainty sources. The results show that the aleatory uncertainty sources P_l and θ_p make the highest contribution. Among the epistemic uncertainty sources, σ_p and P_l make the highest contributions to the life prediction uncertainty.

Next, we will illustrate how to reduce the uncertainty in the epistemic uncertainty sources when fatigue life observations are available.

5.3 Uncertainty reduction

Based on the BN given in Fig. 9, we can perform uncertainty reduction when observations of fatigue life are available. In order to verify the effectiveness of the proposed adaptive Bayesian calibration method, we generate synthetic data for fatigue life observations using the structure–material life prediction model. Based on the fatigue life data, we divide the life domain $\log_{10}(T)$ into nine segments: [1.5, 2], [2, 2.5], [2.5, 3], [3, 3.5], [3.5, 4], [4, 4.5], [4.5, 5], [5, 5.5], [5.5, 6]. 25 observations are generated from the structure–material life prediction model by fixing the epistemic uncertainty at assumed true values. After that, we add five random observations in the life intervals [1.5, 2] and [5.5, 6]. By doing so, we induce statistical bias in these two life intervals between the prediction model and the model used to generate the synthetic data. But during the analysis, we assume that we do not know whether the models have bias in these regions. Table 4 gives the Bayes factor update history during the iterations. There are totally eight iterations. After the eighth iteration the Bayes factors for intervals [1.5, 2] and [5.5, 6] are very low; this is as expected since we generated biased data for these two intervals. The maximum Bayes factor of each iteration is also highlighted in Table 4 using bold font. Note that the number of Bayes factors decreases over iterations because each life interval is removed from the validation analysis after it is used for calibration (Step 14 of Table 1). The results show that the proposed method is able to effectively identify the life intervals where the calibration model is poor (i.e., low Bayes factors). The low prediction quality of the model in intervals 1 and 9 synthetically introduced in this example is reflective of common observations in crack growth experiments, where noisy data is observed in short crack and unstable long crack regimes, at the ends of the Paris regime. In such regimes, simply collecting more test data or adding model discrepancy during calibration may not be enough; substantial refinement in physics modeling may be needed to achieve satisfactory prediction accuracy.

Figure 11 compares the posterior distributions obtained from direct Bayesian calibration and the proposed adaptive Bayesian calibration method. The direct Bayesian calibration method simply performs Bayesian calibration with all the data at once, without considering the accuracy of likeli-

Table 3 First-order Sobol indices of uncertainty sources (before uncertainty reduction)

Variable	σ_p	σ_{θ_p}	σ_δ	A_G	B_G	E_{11}	P_l	θ_p	ε_G	ε_{obs}
S_i^I	0.233	0.032	0.01	0.09	0.1702	0.018	0.591	0.253	0.011	0.02

Table 4 Update history of the Bayes factor over iterations

Iteration	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
1	0.074	4.05	6.69	4.62	5.86	6.99	4.09	**10.77**	0.33
2	0.004	**7.92**	7.63	3.17	5.70	6.31	1.33	–	1.39
3	0.03	–	6.95	4.29	6.05	**6.96**	2.89	–	0.47
4	0.002	–	**7.62**	3.30	5.24	–	1.15	–	1.23
5	0.045	–	–	4.92	**5.36**	–	3.85	–	0.18
6	0.013	–	–	**5.13**	–	–	3.43	–	0.09
7	0.015	–	–	–	–	–	**5.67**	–	0.01
8	0.01	–	–	–	–	–	–	–	**0.016**

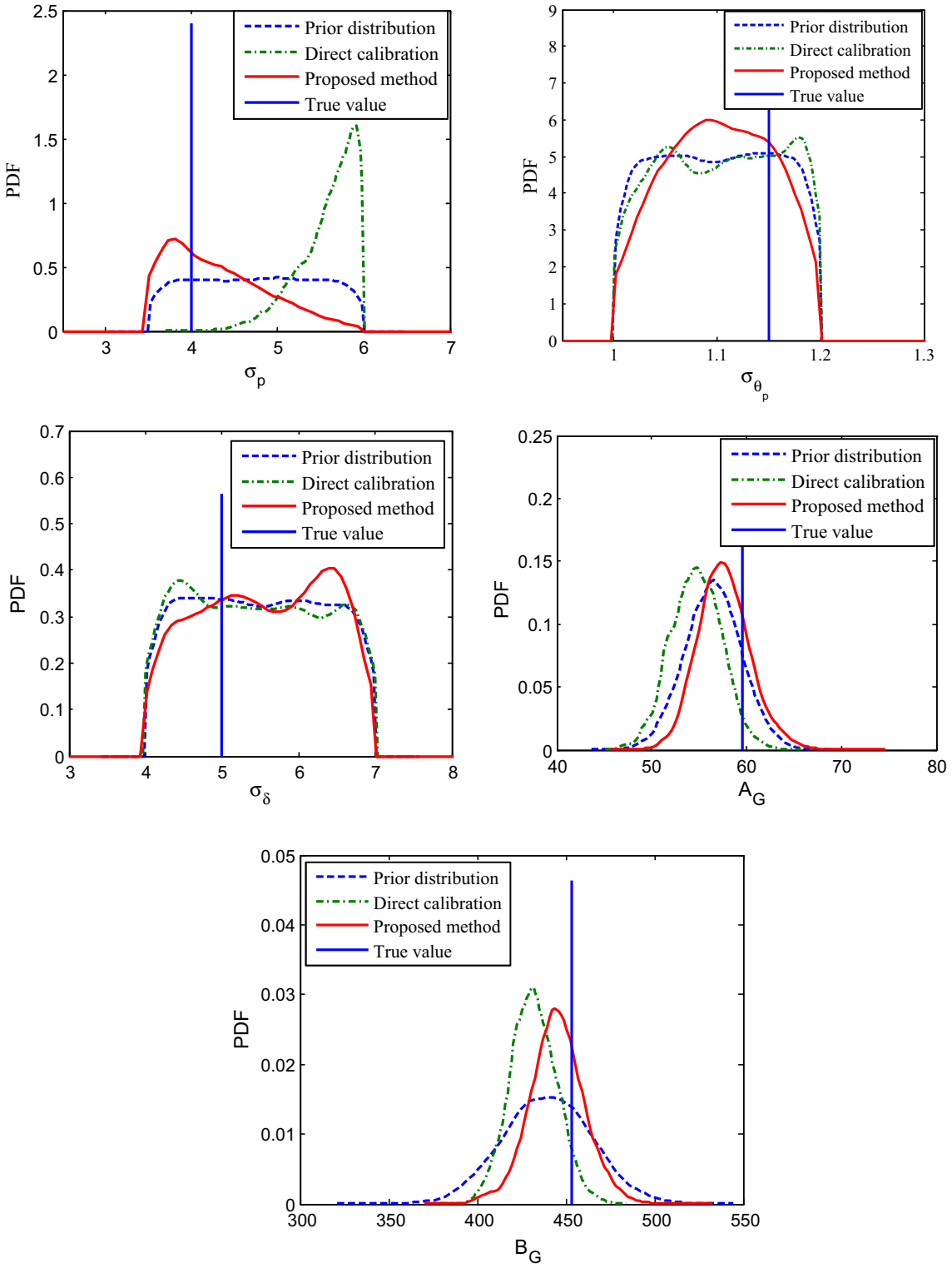


Fig. 11 Comparison of posterior distributions obtained from different methods

hood function. The results in Fig. 11 indicate that both direct calibration and the proposed adaptive calibration can reduce the uncertainty of the epistemic parameters. However, the proposed method can not only reduce the epistemic uncertainty, but also maximize the accuracy of the calibration, which is indicated by closer maximum a posteriori (MAP)

Table 5 First order Sobol indices of uncertainty sources (after uncertainty reduction)

Variable	σ_p	σ_{θ_p}	σ_δ	A_G	B_G	E_{11}	P_l	θ_p	ε_G	ε_{obs}
S_i^I	0.218	0.053	0.003	0.045	0.052	0.035	0.612	0.278	0.012	0.03

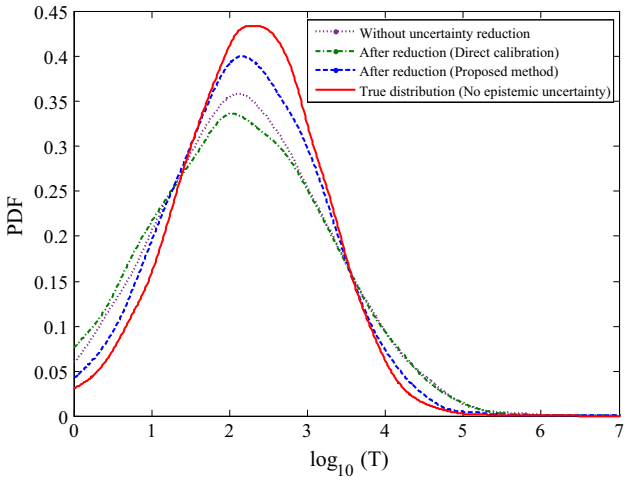


Fig. 12 Comparisons of predictions after uncertainty reduction

estimates to the true values. This demonstrates the effectiveness of the proposed method. After the uncertainty reduction, we perform global sensitivity analysis again. Table 5 gives the first order Sobol indices of the uncertainty sources after uncertainty reduction. It shows that the sensitivities of the epistemic uncertainty sources are reduced after uncertainty reduction. As a result, the contributions of aleatory uncertainty sources are found to be higher.

5.4 Prediction after uncertainty reduction

After the uncertainty in the calibration parameters is reduced using either direct Bayesian calibration or the proposed adaptive Bayesian calibration, we can perform structure–material performance prediction. In the calibration scenario, we assume that the mean value of the load P_l is 27 kips, whereas in the prediction scenario we assume the mean value of P_l as 35 kips. (Note that in this example, we can carry the discrepancy term to the prediction, since only the load *value* has changed and the discrepancy term is a function of the input. However, if the prediction structure—defined by geometry, material, loading configuration, and boundary conditions—were different from the experimental setup used for calibration, then we would not be able to carry the discrepancy term to the prediction. The integration calibration approach of Eq. (32) partially alleviates this problem, since the model validation result which is a measure of discrepancy in the calibration model is carried to the prediction through the integrated posterior; however, we still do not know the discrepancy in the prediction model).

Figure 12 shows the comparison of the predicted life distributions. The results show that compared with the life distribution without uncertainty reduction, the prediction uncertainty is increased with the direct calibration posterior whereas the prediction uncertainty is reduced with the adaptive calibration posterior. More importantly, the predicted distribution with the adaptive calibration posterior is closer to the true life distribution, which demonstrates the effectiveness of the proposed adaptive Bayesian calibration method.

In summary, the results of the composite rotorcraft hub example demonstrate that the proposed adaptive method that integrates calibration and validation can reduce both the uncertainty and the bias in calibration and prediction.

6 Conclusion

Structural and material models are often combined to predict the performance of the structure–material system. Various sources of uncertainty such as data uncertainty, model uncertainty, and natural variability are present in both types of models. To quantify the effects of various sources of uncertainty on the prediction QOI, this paper first developed a Bayesian network approach for uncertainty aggregation. Contributions of different types of uncertainty sources on the uncertainty of QOI are analyzed based on global sensitivity analysis with the Bayesian network. When experimental observations of QOI are available, the uncertainty regarding the epistemic sources (model parameters, distribution parameters of model inputs, and model discrepancy) can be reduced through Bayesian updating. However, the Bayesian updating may introduce errors when the calibration model cannot accurately represent the actual physics of the experiment. To overcome this challenge, this paper proposes an iterative integrated approach, where the accuracy of the prediction model is incorporated within the updating procedure. The observation data on QOI is divided into segments, and the segments are selected adaptively to reduce the uncertainty based on Bayesian calibration and model validation. The results of a composite rotorcraft hub component show that the proposed framework can effectively perform uncertainty aggregation and uncertainty reduction by making optimal use of the data to maximize the accuracy in the result.

Although a structure–material life prediction model is used in this paper for illustration, the proposed uncertainty aggregation and uncertainty reduction framework is not limited to this type of problems. It is a general approach that

can be applied to other prediction models where uncertainty sources are present. Extension of the proposed integrated Bayesian calibration method to multi-level and multi-physics models can be investigated in future work. Another interesting and useful direction of research is test resource allocation, i.e., finding the optimum number of tests and designing their settings (test design) in order to collect the most valuable data for uncertainty reduction. Previous work in this direction has considered only direct Bayesian calibration [15,33]; the incorporation of the proposed adaptive calibration approach within test resource allocation can be studied in the future.

Acknowledgements This study was supported by funds from the National Science Foundation (Grant No. 1404823, CDSE Program) and the Air Force Office of Scientific Research (Grant No. FA9550-15-1-0018, Program Officer: Dr. David Stargel). The support is gratefully acknowledged.

References

1. Fish J, Yu Q (2002) Computational mechanics of fatigue and life predictions for composite materials and structures. Comput Methods Appl Mech Eng 191(43):4827–4849
2. Liu Y, Mahadevan S (2005) Multiaxial high-cycle fatigue criterion and life prediction for metals. Int J Fatigue 27(7):790–800
3. Ghosh S, Lee K, Raghavan P (2001) A multi-level computational model for multi-scale damage analysis in composite and porous materials. Int J Solids Struct 38(14):2335–2385
4. Crouch R, Oskay C, Clay S (2012) Multiscale modeling of damage accumulation in carbon fiber reinforced polymers subjected to fatigue. In: Proceedings of 53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA, p 1615
5. Xia Z, Curtin W, Peters P (2001) Multiscale modeling of failure in metal matrix composites. Acta Mater 49(2):273–287
6. Rannou J, Limodin N, Réthoré J, Gravouil A, Ludwig W, Baïetto-Dubourg M-C, Buffière J-Y, Combescure A, Hild F, Roux S (2010) Three dimensional experimental and numerical multiscale analysis of a fatigue crack. Comput Methods Appl Mech Eng 199(21):1307–1325
7. McDowell D, Dunne F (2010) Microstructure-sensitive computational modeling of fatigue crack formation. Int J Fatigue 32(9):1521–1542
8. Bogdanor MJ, Oskay C, Clay SB (2015) Multiscale modeling of failure in composites under model parameter uncertainty. Comput Mech 56(3):389–404
9. Hu Z, Du X, Conrad D, Twohy R, Walmsley M (2014) Fatigue reliability analysis for structures with known loading trend. Struct Multidiscip Optim 50(1):9–23
10. Li C, Mahadevan S (2016) Relative contributions of aleatory and epistemic uncertainty sources in time series prediction. Int J Fatigue 82:474–486
11. Hu Z, Mahadevan S, Du X (2016) Uncertainty quantification of time-dependent reliability analysis in the presence of parametric uncertainty. ASCE-ASME J Risk Uncertain Eng Syst Part B Mech Eng 2(3):031005
12. Nannapaneni S, Hu Z, Mahadevan S (2016) Uncertainty quantification in reliability estimation with limit state surrogates. Struct Multidiscip Optim 54(6):1509–1526
13. Hu Z, Mahadevan S (2016) Uncertainty quantification in prediction of material properties during additive manufacturing. Scr Mater 135:135–140
14. Ling Y, Mullins J, Mahadevan S (2014) Selection of model discrepancy priors in Bayesian calibration. J Comput Phys 276:665–680
15. Hu Z, Ao D, Mahadevan S (2017) Calibration experimental design considering field response and model uncertainty. Comput Methods Appl Mech Eng 318:92–119
16. Zhang R, Mahadevan S (2003) Bayesian methodology for reliability model acceptance. Reliab Eng Syst Saf 80(1):95–103
17. Wheeler OE (1972) Spectrum loading and crack growth. J Basic Eng 94(1):181–186
18. Oskay C, Fish J (2004) Multiscale modeling of fatigue for ductile materials. Int J Multiscale Comput Eng 2(3):329–353
19. Haldar A, Mahadevan S (2000) Probability, reliability, and statistical methods in engineering design. Wiley, New York
20. Devathi H, Hu Z, Mahadevan S (2016) Snap-through buckling reliability analysis under spatiotemporal variability and epistemic uncertainty. AIAA J 54:3981–3993
21. Mahadevan S, Zhang R, Smith N (2001) Bayesian networks for system reliability reassessment. Struct Saf 23(3):231–251
22. Hu Z, Mahadevan S, Du X (2015) Uncertainty quantification in time-dependent reliability analysis. In: Proceedings of ASME 2015 international design engineering technical conferences and computers and information in engineering conference. American Society of Mechanical Engineers, pp V02BT03A062–V002BT003A062
23. Sankararaman S, Ling Y, Shantz C, Mahadevan S (2011) Uncertainty quantification in fatigue crack growth prognosis. Int J Progn Health Manag 2(1):15
24. Richardson LF (1911) The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam. Philos Trans R Soc Lond Ser A 210:307–357
25. Celik I, Karatekin O (1997) Numerical experiments on application of Richardson extrapolation with nonuniform grids. J Fluids Eng 119(3):584–590
26. Liang B, Mahadevan S (2011) Error and uncertainty quantification and sensitivity analysis in mechanics computational models. Int J Uncertain Quant 1(2):147–161
27. Kennedy MC, O'Hagan A (2001) Bayesian calibration of computer models. J R Stat Soc Ser B (Stat Methodol) 63(3):425–464
28. Sankararaman S, Mahadevan S (2015) Integration of model verification, validation, and calibration for uncertainty quantification in engineering systems. Reliab Eng Syst Saf 138:194–209
29. Hu Z, Du X (2015) Mixed efficient global optimization for time-dependent reliability analysis. J Mech Des 137(5):051401
30. Xiu D, Karniadakis GE (2002) The Wiener–Askey polynomial chaos for stochastic differential equations. SIAM J Sci Comput 24(2):619–644
31. Li C, Mahadevan S (2016) An efficient modularized sample-based method to estimate the first-order Sobolx index. Reliab Eng Syst Saf 153:110–121
32. Li C, Mahadevan S (2016) Global sensitivity analysis for a Bayesian network. In: Proceedings of ASME 2016 international design engineering technical conferences and computers and information in engineering conference. American Society of Mechanical Engineers, pp V01AT02A023–V001AT002A023
33. Li C, Mahadevan S (2016) Robust test resource allocation using global sensitivity analysis. In: Proceedings of 18th AIAA non-deterministic approaches conference, p 0952
34. Gilks WR (2005) Markov chain monte carlo. Wiley Online Library, London
35. Arulampalam MS, Maskell S, Gordon N, Clapp T (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans Signal Process 50(2):174–188

36. Committee AS (1998) AIAA guide for the verification and validation of computational fluid dynamics simulations (G-077-1998). AIAA
37. Ling Y, Mahadevan S (2013) Quantitative model validation techniques: new insights. Reliab Eng Syst Saf 111:217–231
38. Rebba R, Mahadevan S, Huang S (2006) Validation and error estimation of computational models. Reliab Eng Syst Saf 91(10):1390–1397
39. Rebba R, Mahadevan S (2008) Computational methods for model reliability assessment. Reliab Eng Syst Saf 93(8):1197–1207
40. Ferson S, Oberkampf WL, Ginzburg L (2008) Model validation and predictive capability for the thermal challenge problem. Comput Methods Appl Mech Eng 197(29):2408–2430
41. Sankararaman S, Ling Y, Mahadevan S (2011) Uncertainty quantification and model validation of fatigue crack growth prediction. Eng Fract Mech 78(7):1487–1504
42. Mahadevan S, Dey A, Tryon R, Wang Y, Rousseau C (2001) Reliability analysis of rotorcraft composite structures. J Aerosp Eng 14(4):140–146
43. Ao D, Hu Z, Mahadevan S (2016) Validation of surrogate model-based life prediction for a composite rotorcraft hub component. In: Proceedings of the American Society for composites: thirty-first technical conference
44. Murri GB, Salpekar SA, O'Brien TK (1991) Fatigue delamination onset prediction in unidirectional tapered laminates. In: Composite materials: fatigue and fracture (third volume). ASTM International, Philadelphia
45. Hu Z, Mahadevan S (2016) Global sensitivity analysis-enhanced surrogate (GSAS) modeling for reliability analysis. Struct Multidiscip Optim 53(3):501–521