AIAA JOURNAL
Vol. 49, No. 5, May 2011

Quantifying Multiple Types of Uncertainty in Physics-Based Simulation Using Bayesian Model Averaging

Inseok Park* and Ramana V. Grandhi†
Wright State University, Dayton, Ohio 45435

DOI: 10.2514/1.J050741

Given experimental data measured from an engineering system, response predictions by a stochastic simulation model involve both parametric uncertainty and random errors. Also, model form uncertainty arises when two or more simulation models predict the responses of an engineering system because it is beyond capability to identify the best approximating model among the considered model set. In this research, a methodology is developed to quantify model probability using measured deviations between experimental data and model predictions of the data under a Bayesian statistical framework. Model averaging is used to combine the predictions of a system response by a model set into a single prediction. Then, a nonlinear spring-mass system is used to demonstrate the process for implementing model averaging. Finally, the methodology is applied to the engineering benefits of a laser peening process, and a confidence band for a residual stress field is established to indicate the reliability of the composite prediction of the stress field.

Nomenclature

D	=	set of experimental data
d	=	single experimental data
f	=	model operator
g	=	nonlinear spring-force function, N/cm
h	=	probability density function
K	=	number of model set
L	=	number of random samples from joint distribution of random input parameters
M	=	physical model
N	=	number of xperimental data set
u	=	displacement, cm
X	=	set of deterministic input parameters
y	=	system response to be predicted
ε	=	prediction error
Θ	=	set of random input parameters
μ	=	mass

Subscripts

k	=	index of models
l	=	index of random samples on random input parameters
n	=	index of experimental data

I. Introduction

THE structural analysis of engineering systems is becoming more dependent on the computer simulation of complex physical phenomena with advances in nonlinear modeling processes. A simulation model can vary, depending on the underlying physics considered and on the manner in which a mathematical model is converted into a simulation model. In other words, we may have two or more different models to simulate an identical engineering system. The consideration of different simulation models creates model form uncertainty (uncertainty involved in selecting the best approximating model from a set of models) in addition to the uncertainty in random input parameters within the considered model set. Uncertainty in model form can be significant when predictions from different models vary considerably.

While input parameter uncertainty has been investigated and is well documented, model form uncertainty has not been extensively researched in the engineering field when compared with other fields such as statistics, economics, and environmental science. Alvin et al. [1] used Bayesian model averaging (BMA) to estimate the model form uncertainty in the frequency predictions of a component mounting bracket resulting from the use of three stochastic simulation models where the elastic material modulus is considered to be a random input parameter. The three simulation models had different levels of simplifying assumptions in their partial differential equation form, as well as different spatial meshes and different model orders (i.e., number of discrete solution variables). Model probability, assigned to each model to quantify model form uncertainty, was simply assumed to be uniform across the simulation models considered. Zio and Apostolakis [2] used the adjustment factor approach to estimate the model form uncertainty in the response predictions regarding the cumulative release of a radionuclide to the water table given by six different models. The six models differ by some fundamental hypotheses on groundwater flow and transport mechanisms. Model probabilities were evaluated based on expert opinions. Zhang and Mahadevan [3] estimated the failure probabilities for the butt welds of a steel bridge using two competing crack growth models (the Forman and the Weertman crack growth models). They made a reliability analysis of fatigue life by averaging the estimated failure probabilities weighted by model probabilities. The uncertainty in crack size measurement was quantified to evaluate model probabilities using Bayes's theorem.

Zouaoui and Wilson [4] used BMA to quantify the model form uncertainty in the prediction of a message delay in a computer communication network. Although simulation models were used to predict a message delay, the model probabilities were quantified, not by using any simulation model but by using three different types of distributions that represent the uncertainty in an input variable (i.e., a message length). The distributions were assumed to be of exponential, normal, and lognormal forms. McFarland and Bichon [5] also used BMA to incorporate probability distribution model form uncertainty into the estimation of failure probability of a bistable microelectromechanical systems device. As in the work of Zouaoui and Wilson [4], model probabilities were assigned to the three types

Presented as Paper 2010-2512 at the 12th AIAA Non-Deterministic Approaches Conference, Orlando, FL, 12–15 April 2010; received 8 July 2010; revision received 22 November 2010; accepted for publication 9 December 2010.

*Ph.D. Candidate, Department of Mechanical and Materials Engineering. Member AIAA.

†Distinguished Professor, Department of Mechanical and Materials Engineering. Fellow AIAA.

of distributions (normal, lognormal, and Weibull) representing the uncertainty in an input variable (i.e., edge bias on beam widths).

In the research mentioned previously, model probability was not quantified using experimental data on system responses. The quantified model probabilities could not provide effective measures of model form uncertainty in terms of predictions of system responses. In this research, a methodology to evaluate model likelihood, which measures how well a model is supported by experimental data relative to other models, using measured differences of experimental and model outcomes under a Bayesian statistical framework is developed in order to make an informed estimation of the probabilities of a stochastic simulation model set. To incorporate model form uncertainty, prediction by each stochastic simulation model, which involves both parametric uncertainty and random error, is averaged using the evaluated model probability as weight. The proposed methodology is then applied to the engineering problem of a laser peening (LP) process: a time-dependent high-impact process.

This paper is organized as follows: The definition of model form uncertainty is presented in Sec. II. The proposed methodology is discussed in Sec. III. Model averaging to combine model predictions is discussed in Sec. IV. Model averaging is illustrated with the numerical problem of a nonlinear spring-mass system, and the proposed methodology is applied to the computer simulation of a LP process in Sec. V.

II. Definition of Model Form Uncertainty

A model represents a particular aspect of the real physical system in a form we can interpret. We can predict system responses with the help of a physical model. The generation of a model entails an idealization of a real system of interest because, in general, a system is too complicated to be modeled without idealization. Therefore, a model is not the representation of a real system but the representation of an ideal system. Different models can be generated to represent an identical system when different sets of assumptions are made in the process of modeling the system. In addition to the simplifying assumptions, a model can vary depending on the decisions made in the modeling process with regard to the executor's preferences, the decision maker's requirements, and economic considerations. For instance, to analyze a structure, a modeler can build a number of finite element (FE) models that are different in terms of element type, geometry, shape function, mesh size, material behavior, expected operating load, and boundary conditions, such as various Saturn V FE models [6].

Given two or more approximating models for a system, the problem may occur of whether it is feasible to select the model that represents the system with the highest fidelity from the set of considered models. Generally, it is beyond our ability to select the best approximating model from a set of models, because we cannot with certainty estimate the relative degrees of discrepancy between a physical system and a model set approximating it due to lack of complete knowledge about the system; here, the best model is defined as the one that makes the closest predictions of system responses of interest among a set of possibilities. In this instance, uncertainty is involved in selecting the best approximating one from a set of models. This uncertainty is called model form uncertainty to distinguish it from the parametric uncertainty associated with each model. Model form uncertainty is categorized as epistemic uncertainty [7], since it derives from our lack of knowledge. Model form uncertainty is represented by model probabilities assigned to a model set.

Given the data, it is practical to make an educated guess that a certain model is the best from a set of models using model selection criteria, such as the Akaike information criterion or the Bayesian information criterion [8]. Model selection criteria select a particular model that better fits into observed experimental data and has a smaller number of regression parameters when compared with the other models in a model set as the best model. However, the relative goodness of fit of a model set may vary depending on the sample size of experimental data. Therefore, a model regarded as the best of a model set might not be the best if additional experimental data support another model. Because of this possibility, model form uncertainty cannot be eliminated even after observing experimental data. Ignoring model form uncertainty often leads to underestimating uncertainty in prediction of a response [9,10].

III. Quantification of Model Probability

The methodology to quantify model probability is developed in this section.

A. Bayes's Theorem for Quantifying Model Probability

Model probability is defined as the degree of belief that a model is the best approximating one among a model set [11]. Given experimental data D, Bayes's theorem presents a way to update the prior probability of model M_k into posterior probability of M_k by

$$Pr(M_k|D) = \frac{Pr(D|M_k)Pr(M_k)}{\sum_{m=1}^{K} Pr(D|M_m)Pr(M_m)}, \qquad k = 1, \ldots, K \quad (1)$$

where K is the number of models. Prior model probability $Pr(M_k)$ is the probability of model M_k evaluated before observing experimental data D. When given neither expert opinion nor engineering information about the models considered, prior model probability $Pr(M_k)$ is usually given a uniform probability, $Pr(M_k) = 1/K$. In Eq. (1), $Pr(D|M_k)$, which is broadly known as $L(M_k|D)$, is called the likelihood of model M_k given experimental data D, because the first argument D is known, while the second argument M_k is not held fixed [12]. The numerical value of the likelihood $Pr(D|M_k)$ is of no importance; all that matters is the ratio of two different likelihoods: $Pr(D|M_p)$ and $Pr(D|M_q)$, $p \neq q$, given the same data D. The posterior probability $Pr(M_k|D)$ of model M_k is only proportional to prior probability $Pr(M_k)$ and model likelihood $Pr(D|M_k)$ because the denominator in Eq. (1) is common to all models.

B. Formulation to Involve a Random Error into Model Prediction

Model likelihood $Pr(D|M_k)$ is evaluated for each model by measuring the degree of agreement between experimental data and predictions of the data by each model. For this purpose, the probabilistic relationship between experimental data and model predictions involving uncertainty should be described [13]. There are various formulations developed to describe the probabilistic relationship. Usually, bias function and measurement error are included as parts of the probabilistic relationship when matching model predictions with experimental data. The bias function captures systematic discrepancies between the true responses and predictions by a model [14,15]. The measurement error is usually assumed to be an independent normal variable with the mean of zero. Kennedy and O'Hagan [16] used Bayesian statistical methodology to quantify the uncertainty in the bias function modeled by a Gaussian process. Xiong et al. [17] treated the bias function as a deterministic regression model. Additionally, the fundamental concepts and methodologies to quantify random prediction errors associated with large-scale simulation models are being actively developed by professional societies, including the AIAA [18] and the American Society of Mechanical Engineers [19], especially aiming at validating simulation models.

In this research, a formulation that combines bias function associated with an approximating model and measurement error on data is used to describe the probabilistic relationship between experimental data and model predictions [20]. The formulation is represented by

$$y_k = f_k(X_k) + \varepsilon_k, \qquad \varepsilon_k \sim \mathcal{N}(0, \sigma_k^2) \quad (2)$$

X_k is a set of deterministic parameters within model M_k, and ε_k is a random variable that encompasses both bias associated with model prediction $f_k(X_k)$ and measurement error on experimental data. The normal random variable ε_k is assumed to be independent and identically distributed (IID) with zero mean and unknown variance σ_k^2. The variable ε_k results from the inadequate model form, imperfect

numerical implementation, inaccurate input parameters, and measurement error. The variable ε_k is considered to involve uncertainty even if the uncertainties in the model forms and parameters under consideration are perfectly quantified, because we cannot take all relevant models and parameters into consideration [21].

The use of ε_k with zero mean reflects the fact that a model claims that its prediction $f_k(X_k)$ is the most probable value. The reason that the bias function is not included separately in the probabilistic relationship in this research is that a separate incorporation of the bias function results in shifting a model prediction from the initially predicted value $f_k(X_k)$. Although this can reduce the bias involved in model prediction, model probabilities will be evaluated depending on the shifted predictions. Also, including an unknown bias function separately may result in the problem of degrees of freedom given that there are generally quite limited amounts of experimental data such as the engineering applications. Increasing the number of unknown parameters usually leads to a reduction in the accuracy of estimation, because it costs a degree of freedom to estimate each unknown parameter. If a more complicated formulation is desired, the covariance matrix of errors may be estimated. However, errors are assumed to be independent in this research due to insufficient experimental data to estimate the correlation among errors.

C. Predictive Distribution of a Response Under Each Model

Consider that a set of models predicts a system response. Since the mean of random error ε_k is zero, Eq. (2) is rewritten as

$$h_Y(y|M_k, D) = \mathcal{N}[f_k(X_k), \sigma_k^2] \tag{3}$$

The unknown variable σ_k^2 can be determined by using the differences between a set of experimental data and the predictions of the data by each model M_k. Consider that experimental data $D = \{d_1, \ldots, d_N\}$ are observed. The measured prediction errors of model M_k given experimental data D are

$$\varepsilon_{k_n} = d_n - f_k(X_{k_n}), \qquad n = 1, \ldots, N \tag{4}$$

where $f_k(X_{k_n})$ is the prediction of a single experimental data d_n by model M_k. Because a measured prediction error ε_{k_n} is considered to be a realization of an IID normal variable $\varepsilon_k \sim \mathcal{N}(0, \sigma_k^2)$, σ_k^2 is calculated by Eq. (5) based on the maximum likelihood estimation approach:

$$\sigma_k^2 = \frac{\sum_{n=1}^{N} \varepsilon_{k_n}^2}{N} \tag{5}$$

Now, consider that input parameters within model M_k consist of a set of random parameters Θ_k as well as a set of deterministic parameters X_k. Uncertainty in Θ_k is represented by a joint probability density function $h_{\Theta_k}(\theta_k|M_k)$ under model M_k. For this research, a set of random parameters Θ_k are considered to be independent of experimental data D. Therefore, $h_{\Theta_k}(\theta_k|M_k)$ is not updated after observing data D. Incorporating a random prediction error of model M_k, as well as parametric uncertainty in Θ_k, the predictive distribution of response y under M_k given data D is represented by

$$h_Y(y|M_k, D) = \int_{\theta_k} h_{\Theta_k}(\theta_k|M_k) h_Y(y|M_k, \theta_k, X_k, D)\, \mathrm{d}\theta_k \tag{6}$$

$$\text{and} \quad h_Y(y|M_k, \theta_k, X_k, D) = \mathcal{N}[f_k(\theta_k, X_k), (\sigma_k)^2_{\theta_k}] \tag{7}$$

Equation (6) [10] represents the marginal distribution of response y where the random variable set Θ_k is marginalized out. Because prediction error ε_k is assumed to be normally distributed, $h_Y(y|M_k, \theta_k, X_k, D)$ is represented as a normal density function given $\Theta_k = \theta_k$ in Eq. (7). However, due to the existence of the collection of random parameters Θ_k, the mean $f_k(\theta_k, X_k)$ and variance of $h_Y(y|M_k, \theta_k, X_k, D)$ are probabilistically distributed, unlike $h_Y(y|M_k)$ in Eq. (3).

It is impossible to analytically solve Eq. (6) for most engineering problems where $f_k(\theta_k, X_k)$ and $(\sigma_k)^2_{\theta_k}$ cannot not be explicitly represented by the probability density functions of Θ_k. Using the Monte Carlo integration technique, the integral in Eq. (6) is numerically solved using data points randomly sampled from $h_{\Theta_k}(\theta_k|M_k)$, as expressed by

$$h_Y(y|M_k, D) \approx \frac{1}{L} \sum_{l=1}^{L} \mathcal{N}[f_k(\hat{\Theta}_{k_l}, X_k), (\sigma_k)^2_{\hat{\Theta}_{k_l}}] \tag{8}$$

$\hat{\Theta}_{k_l}$ is the lth data point randomly sampled from $h_{\Theta_k}(\theta_k|M_k)$, and L is the number of sampled data points. Given a realization $\hat{\Theta}_{k_l}$ of Θ_k, $f_k(\hat{\Theta}_{k_l}, X_k)$ and $(\sigma_k)^2_{\hat{\Theta}_{k_l}}$ are deterministic values. The mean $f_k(\hat{\Theta}_{k_l}, X_k)$ is the prediction from model M_k given $\Theta_k = \hat{\Theta}_{k_l}$. The variance $(\sigma_k)^2_{\hat{\Theta}_{k_l}}$ can be calculated using Eqs. (4) and (5); $f_k(X_{k_n})$ in (4) is replaced with $f_k(\hat{\Theta}_{k_l}, X_{k_n})$.

D. Evaluation of Model Likelihood

The likelihood of model M_k given experimental data D is represented by [10]

$$Pr(D|M_k) = \int_{\theta_k} h_{\Theta_k}(\theta_k|M_k) h_{\Theta_k}(D|M_k, \theta_k, X_k)\, \mathrm{d}\theta_k \tag{9}$$

Consider that experimental data $D = \{d_1, \ldots, d_N\}$ are observed, and the predictions of data D are given as $f_k(\theta_k, X_{k_1}), \ldots, f_k(\theta_k, X_{k_N})$ by model M_k. Given the assumption that experimental data D are independent of one another, $h_{\Theta_k}(D|M_k, \theta_k, X_k)$ is represented by

$$\begin{aligned} h_{\Theta_k}(D|M_k, \theta_k, X_k) &= \prod_{n=1}^{N} h_{\Theta_k}(d_n|M_k, \theta_k, X_{k_n}) \\ &= \prod_{n=1}^{N} \mathcal{N}[d_n; f_k(\theta_k, X_{k_n}), (\sigma_k)^2_{\theta_k}] \end{aligned} \tag{10}$$

$\mathcal{N}[d_n; f_k(\theta_k, X_{k_n}), (\sigma_k)^2_{\theta_k}]$ is the probability density of $\mathcal{N}[f_k(\theta_k, X_{k_n}), (\sigma_k)^2_{\theta_k}]$ for the experimental data d_n. To analytically calculate the integral in Eq. (9), $\mathcal{N}[d_n; f_k(\theta_k, X_{k_n}), (\sigma_k)^2_{\theta_k}]$ should be represented as a function (likelihood function) of θ_k. However, the explicit expression of $\mathcal{N}[d_n; f_k(\theta_k, X_{k_n}), (\sigma_k)^2_{\theta_k}]$ cannot usually be acquired in engineering problems.

The Monte Carlo integration technique is used to find an approximate solution for Eq. (9), depending on data points $\hat{\Theta}_{k_l}$, $l = 1, \ldots, L$, randomly sampled from $h_{\Theta_k}(\theta_k|M_k)$. Given $\Theta_k = \hat{\Theta}_{k_l}$, $h_{\Theta_k}(D|M_k, \theta_k, X_k)$ in Eq. (10) is represented by

$$\begin{aligned} &h_{\hat{\Theta}_{k_l}}(D|M_k, \hat{\Theta}_{k_l}, X_k) \\ &\quad = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi(\sigma_k)^2_{\hat{\Theta}_{k_l}}}} \exp\left(-\frac{\{d_n - f_k(\hat{\Theta}_{k_l}, X_{k_n})\}^2}{2(\sigma_k)^2_{\hat{\Theta}_{k_l}}}\right) \end{aligned} \tag{11}$$

Because

$$(\sigma_k)^2_{\hat{\Theta}_{k_l}} = \frac{\sum_{n=1}^{N} \{d_n - f_k(\hat{\Theta}_{k_l}, X_{k_n})\}^2}{N}$$

by Eq. (5), Eq. (11) is simplified into

$$h_{\hat{\Theta}_{k_l}}(D|M_k, \hat{\Theta}_{k_l}, X_k) = \left(\frac{1}{2\pi(\sigma_k)^2_{\hat{\Theta}_{k_l}}}\right)^{N/2} \exp\left(-\frac{N}{2}\right) \tag{12}$$

When

$$h_{\hat{\Theta}_{k_l}}(D|M_k, \hat{\Theta}_{k_l}, X_k), \ldots, h_{\hat{\Theta}_{k_L}}(D|M_k, \hat{\Theta}_{k_L}, X_k)$$

are calculated, the integral in Eq. (9) is approximated by

$$Pr(D|M_k) \approx \frac{1}{L}\sum_{l=1}^{L}\left(\frac{1}{2\pi(\sigma_k)^2_{\hat{\Theta}_{kl}}}\right)^{N/2}\exp\left(-\frac{N}{2}\right) \tag{13}$$

IV. Model Averaging

The model averaging technique averages the predictions of a response by a model set using model probabilities as weights. The predictions of a response made by a set of models are weighted by the probabilities assigned to the set of models and then integrated into the composite prediction of the response, as in

$$h_Y(y|D) = \sum_{k=1}^{K} Pr(M_k|D)h_Y(y|M_k, D) \tag{14}$$

As shown in Eqs. (6) and (7), the posterior probability density function $h_Y(y|M_k, D)$ represents the uncertainty in the prediction of response y by model M_k due to both uncertainty in the input parameter set Θ_k and uncertainty in prediction error ε_k. The composite prediction $h_Y(y|D)$ incorporates model form uncertainty as well as both the uncertainties, because $h_Y(y|D)$ is obtained by averaging over $h_Y(y|M_k, D)$ weighted by probability value $Pr(M_k|D)$.

If predictive distribution $h_Y(y|M_k, D)$ estimated by model M_k has the mean $E(y|M_k, D)$ and the variance $\mathrm{Var}(y|M_k, D)$, the mean and variance of composite predictive distribution $h_Y(y|D)$ are calculated by [9]

$$E(y|D) = \sum_{k=1}^{K} Pr(M_k|D)E(y|M_k, D) \tag{15}$$

and

$$\mathrm{Var}(y|D) = \sum_{k=1}^{K} Pr(M_k|D)\mathrm{Var}(y|M_k, D) + \sum_{k=1}^{K} Pr(M_k|D)]E(y|M_k, D) - E(y|D)]^2 \tag{16}$$

As shown in Eq. (15), the mean of $h_Y(y|D)$ is the average of means of predictive distributions under a model set, using model probabilities as weights. The variance of $h_Y(y|D)$ is decomposed into the sum of two terms. as shown in Eq. (16). The first term, called within-model variance, represents the average degree of uncertainty in each model prediction of response y. The second term, called between-model variance, represents the degree of model form uncertainty in prediction of response y.

V. Demonstration Problems

The model averaging technique is demonstrated using a problem of natural frequency prediction in a nonlinear spring-mass system. Next, the methodology suggested to quantify the model form uncertainty associated with physics-based simulations using experimental data is applied to the FE analysis of a LP process.

A. Nonlinear Spring-Mass System

1. Problem Description

The free vibration of a single-degree-of-freedom spring-mass system with a nonlinear spring is represented by

$$\mu\ddot{u} + g(u) = 0 \tag{17}$$

where μ is a mass, and $g(u)$ is a nonlinear spring-force function of displacement u.

Using different spring-force functions, different mathematical models are created to represent the free vibration of a nonlinear spring-mass system. Three spring-force functions considered for this problem are expressed by

$$g_1(u) = \alpha u^{1/3} \tag{18}$$

$$g_2(u) = \beta u + \gamma u^3 \tag{19}$$

and

$$g_3(u) = \delta u + \frac{\zeta u}{\sqrt{1+u^2}} \tag{20}$$

where α, β, γ, δ, and ζ are the parameters used to specify the spring-force functions. The three nonlinear spring-force functions are graphically represented in Fig. 1, given the parameter values of $\alpha = 0.65$ N/cm$^{4/3}$, $\beta = 1$ N/cm^2, $\gamma = -0.35$ N/cm^4, $\delta = 1$ N/cm^2, and $\zeta = -0.5$ N/cm. The nonlinear spring is described as the stiffest by Eq. (18) and as the most flexible by Eq. (20).

2. Natural Frequency Prediction of a Spring-Mass System Using Three Mathematical Models

The natural frequency predictions by the three mathematical models depend on the following equations [22]:

$$\omega = 1.070451\alpha^{1/2}u(0)^{-1/3} \tag{21}$$

$$\omega = \sqrt{\beta + \frac{3}{4}\gamma u(0)^2} \tag{22}$$

and

$$\omega = \sqrt{\frac{\int_0^{\pi/2}\{\delta\cos^2 t + [\zeta\cos^2 t/\sqrt{1+u(0)^2\cos^2 t}]\}\,\mathrm{d}t}{\int_0^{\pi/2}\sin^2 t\,\mathrm{d}t}} \tag{23}$$

The mass and initial velocity of the system are given as $\mu = 1$ kg and $\mathrm{d}u/\mathrm{d}t(0) = 0$ cm/s. The initial displacement is regarded as a random input parameter and represented by a normal distribution,

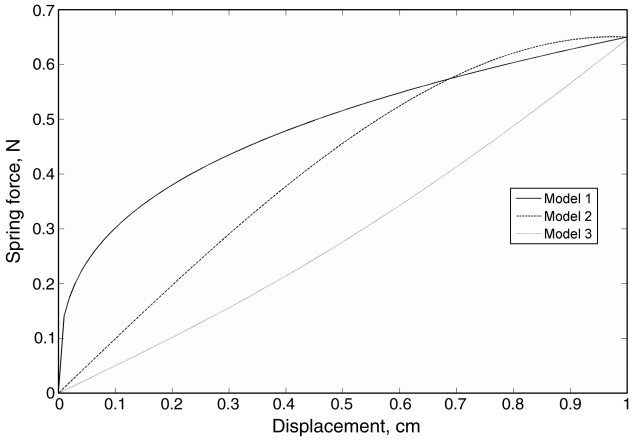

Fig. 1 Nonlinear spring force in three different models of a spring-mass system.

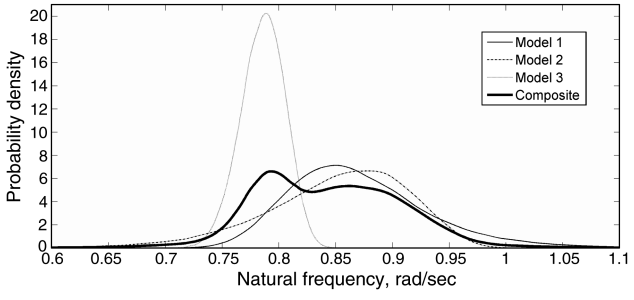

Fig. 2 Composite predictive distribution of natural frequency together with predictive distributions under three models.

Table 1 Model probabilities, means and variances of natural frequency predictions by three mathematical models and the composite prediction for a spring-mass system

	Model probability	Mean, rad/s	Variance, rad^2/s^2
Model 1	0.3	0.871	4.00×10^{-3}
Model 2	0.5	0.850	4.20×10^{-3}
Model 3	0.2	0.786	3.74×10^{-4}
Composite	1.0	0.844	$3.37 \times 10^{-3} + 9.17 \times 10^{-4} = 4.29 \times 10^{-3}$

$u(0) = \mathcal{N}(1\text{ cm}, 0.04\text{ cm}^2)$. The randomness in the initial displacement results in parametric uncertainty in the prediction of the natural frequency by each model. To construct a predictive distribution of the natural frequency under each model, 100,000 random samples are drawn from the normal distribution that represents the randomness in the initial displacement. Figure 2 shows the kernel density estimates for the predictive distributions under the three models based on the 100,000 random samples. The means and variances of the three predictive distributions are shown in Table 1. The mean of natural frequency prediction by model 1 (0.8712 rad/s) is the largest among the three models, since model 1 describes the nonlinear spring of the system as the stiffest. The mean of model 3 prediction (0.7858 rad/s) is the smallest, since model 3 describes the spring as the most flexible. The variances of model 1 and model 2 predictions of the natural frequency (4.00×10^{-3} rad^2/s^2 and 4.20×10^{-3} rad^2/s^2, respectively) are little different from each other. The variance of model 3 prediction (3.74×10^{-4} rad^2/s^2) is considerably smaller compared with the variances of model 1 and model 2 predictions.

3. *Quantification of Model Probabilities*

The probabilities of the three models must be quantified to weigh the predictive distributions from the three models. The model probabilities are assumed for this problem because no information is available to evaluate them (Table 1). Model 2 has the largest effect on the integration of the three model predictions, because model 2 has the highest model probability.

4. *Combination of Predictive Distributions*

Given the model probabilities and the predictive distributions estimated by the three models, the composite predictive distribution of the natural frequency can be constructed using Eq. (14). As shown in Fig. 2, the composite predictive distribution has two modes due to the effect of model 3 prediction having a significantly large mode. Using Eqs. (15) and (16), the mean and variance of the composite prediction are calculated as shown in Table 1. The mean of the composite prediction is almost the same as the mean of the prediction by model 2, because the weighted means of the three model predictions are almost symmetrical around the mean of model 2 prediction. The variance of the composite prediction (4.29×10^{-3} rad^2/s^2) is the sum of the within-model variance (3.37×10^{-3} rad^2/s^2), which measures the average degree of parametric uncertainty in each model prediction. and the between-model variance (9.17×10^{-4} rad^2/s^2), which measures the degree of model form uncertainty in the composite prediction. The dominance of the within-model variance over the between-model variance indicates that the parametric uncertainty associated with the three models considered is more significant than the model form uncertainty.

B. Finite Element Simulation for Laser Peening Process

1. *Problem Description*

LP is an advanced surface enhancement technique that has been shown to increase the fatigue life of metallic components. LP has also been shown to increase the corrosion and fretting properties of metals. During the LP process, laser energy is converted into shock waves at the surface that induce compressive residual stresses. Fatigue life is improved as the induced compressive residual stresses inhibit the formation of cracks. A detailed description of a LP process can be found in [23,24].

In simulating the LP process, obtaining an accurate description of material behavior is a challenging task because of the high strain rates experienced by the material. During the LP process, the strain rates experienced by a material can reach as high as $10^6/s$. In such high strain-rate processes, different material models are available to describe the elastic–plastic behavior. In a recent paper [25], three material models were considered to describe the material behavior: the Johnson–Cook (JC) model, the Zerilli–Armstrong (ZA) model, and the Khan–Huang–Liang (KHL) model.

Each material model creates a different FE simulation to predict the residual stress field induced by a LP process. The simulation requires intensive computer effort due to modeling of the material behavior under high-pressure shock waves with plastic deformation and time marching numerical procedures. A schematic illustration of the LP FE model for this problem is shown in Fig. 3. In this problem, the peak value of a pressure pulse induced by a laser shot is considered to be a random input parameter. The parametric uncertainty in the peak pressure value is represented by a normal distribution with the mean of 5.5 GPa and the standard deviation of 0.275 GPa. To quantify uncertainty in prediction of the residual stress field by each model, 800 random samples are drawn from the normal distribution representing the uncertainty in the peak pressure value. Even if a comparatively small number of random data are sampled due to a high computational cost, a set of 800 sampled data (FE simulations) proves to be sufficient for representing the parametric uncertainty with acceptable accuracy after a convergence study. The means of the predictions of the residual stress field by the three FE models are shown with observed experimental data [26] in Fig. 4. Basing a prediction of the residual stress field on a single model can cause unreliable results because it is beyond our capability to know which of the three models makes a prediction closest to the true residual stress field. The significant differences between the means of

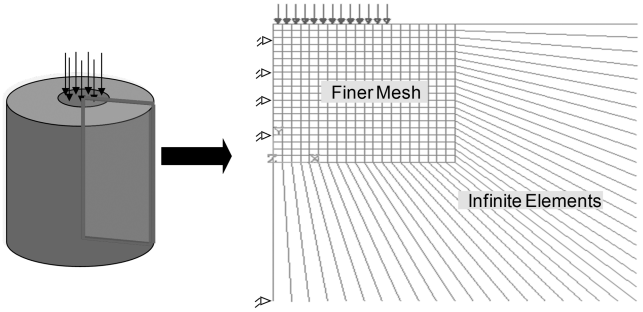

Fig. 3 Representative axisymmetric FE mesh for the LP simulation model.

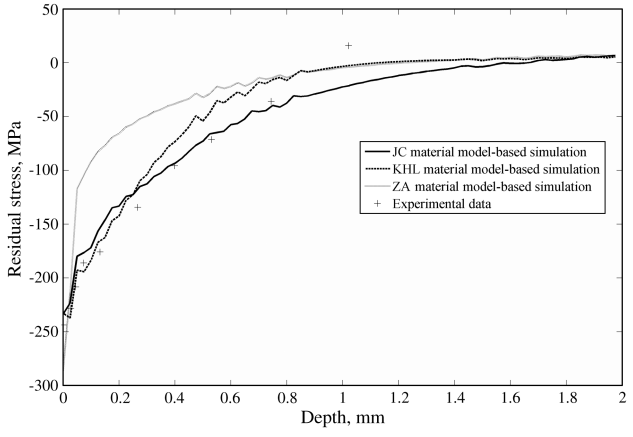

Fig. 4 Residual stress comparison between experimental data measured at a distance of 1 mm from the center of a cylindrical LP component, and means of the predictions of the data by three FE models.

Table 2 Prior probabilities, likelihoods, and posterior probabilities of three FE models for an axisymmetric LP component

	Prior model probability	Model likelihood	Posterior model probability
JC-based FE model	3.333×10^{-1}	4.305×10^{-19}	8.075×10^{-1}
ZA-based FE model	3.333×10^{-1}	1.026×10^{-19}	1.925×10^{-1}
KHL-based FE model	3.333×10^{-1}	8.450×10^{-25}	1.585×10^{-6}

the three simulation outcomes imply that the model form uncertainty in the prediction of the residual stress field might be considerable. The measured differences between the experimental data observed at 10 points of different depths and the predictions of the data by the models considered are assumed to be independent of one another, and they are used as information to estimate the probabilities of the three FE models.

2. *Quantification Model Form Uncertainty in Finite Element Simulation for Laser Peening Process*

For this problem, the prior probabilities of the three FE models are assumed to be uniform because of the lack of information to quantify them. The evaluation of model likelihoods is required to update the prior model probabilities into the posterior model probabilities using Bayes's theorem, shown in Eq. (1). The prediction of a residual stress at any depth by a FE model involves an unknown random error as well as the parametric uncertainty arising from the randomness in the peak pressure pulse. Given each of the 800 samples randomly drawn from the distribution for the peak pressure, the variance of random prediction error associated with each FE model is computed using the differences measured between the observed data and the predictions of the data made by each model based on each random sample by Eq. (5). Using Eq. (8), the quantified random error is incorporated into the prediction of a residual stress made by each model. The likelihoods of the considered three FE models given the observed experimental data are computed using the number of the observed data, and the calculated variances of random prediction errors as shown in Eq. (13).

Using model averaging, shown in Eq. (14), the predictions of a residual stress at any depth by the three models involving both parametric uncertainty and random prediction errors are combined into a single composite prediction. The composite predictive distribution of a residual stress incorporates the uncertainties arising from the uncertainty associated with the use of the three different FE models, the randomness in the peak pressure value, and the random errors involved in the predictions of the residual stress by the models considered. The mean and variance of the composite prediction of a residual stress are calculated using Eqs. (15) and (16).

3. *Results of Quantifying Model Form Uncertainty and Combining Model Predictions*

The posterior probabilities of the three FE models are shown in Table 2 together with the prior model probabilities and model likelihoods. As shown in Eq. (1), equal prior model probabilities cancel out in the calculation of posterior model probabilities. The posterior probability of a model is the ratio of the likelihood for the model to the sum of likelihoods for all the considered models. Therefore, a posterior model probability indicates the measure of how well a model is supported by the experimental data relative to the other models. The posterior probability of the FE model that includes the ZA material model (1.585×10^{-6}) is significantly smaller than the two FE models that include the JC and the KHL material models (8.075×10^{-1} and 1.925×10^{-1}, respectively), because the first FE model is poorly supported by the observed experimental data when compared with the last two FE models. In other words, the first FE model has considerably less chance of being the best approximating one than the last two FE models. It can be inferred from this fact that the JC and the KHL material models are much more effective in the simulation of the LP component than the ZA material model. The FE model that includes the JC material model (the JC-based FE model) has a considerable degree of likelihood of being the best model when compared with the other two models. However, it cannot be concluded that the uncertainty involved in the selection of the best model is negligible, because the amount of the experimental data is comparatively small.

Table 3 Mean and standard deviation of predictions by three FE models and the composite prediction of the residual stress on the surface at 1 mm distance from center

	Mean, MPa	Standard deviation, MPa
JC-based FE model	−233.732	29.586
ZA-based FE model	−232.719	66.308
KHL-based FE model	−285.596	96.769
Composite	−233.537	39.413

Figure 5 shows the composite predictive distribution of the residual stress on the surface at 1 mm distance from the center of the LP component and the predictive distribution of the stress estimated by each model. The mean and standard deviation of the composite prediction of the residual stress are shown in Table 3 along with the mean and standard deviation of prediction of the stress by each model. The mean of the composite prediction (−233.54 MPa) is slightly different from the mean of the JC-based FE model prediction (−233.73 MPa) because the effect of the JC-based FE model prediction on the mean of the composite prediction is significant due to the comparatively large value of the probability of the JC-based FE model. The standard deviation of the composite prediction (39.41 MPa) is larger than the standard deviation of the JC-based FE model prediction (29.59 MPa), mainly due to the considerably large standard deviation of the KHL-based FE model prediction (66.31 MPa). The standard deviation of the composite prediction indicates the total degree of uncertainty in the prediction of the residual stress due to model form uncertainty, parametric uncertainty, and random prediction errors. Figure 6 shows the mean of the

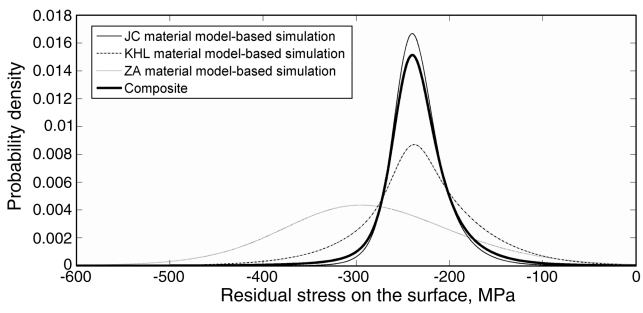

Fig. 5 Composite predictive distribution of the residual stress on the surface along with predictive distributions estimated by three FE models.

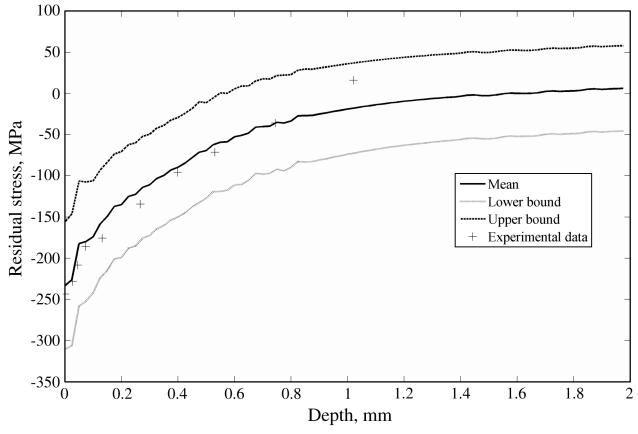

Fig. 6 Mean of composite prediction and 95% confidence band of residual stress field.

composite prediction of the residual stress field that indicates the expected value of the true residual stress at every depth.

4. Establishment of Confidence Band for Residual Stress Field

In addition to the expected value of the true residual stress at every depth represented by the mean of the composite prediction, an interval estimate of the true stress is made to indicate the reliability of the composite prediction. For this problem, a 95% confidence interval for a residual stress at every depth is established by choosing an interval where the probability of being below the interval is the same as the probability of being above [27]. A 95% confidence interval represents the interval that is expected to include the true residual stress with a probability of 0.95. By connecting the upper and lower endpoints at all the depths, an upper and a lower bound curve are drawn, respectively (Fig. 6). The width of the 95% confidence band may be thought of as the measure of the degree of uncertainty in the composite prediction of the residual stress field due to the three types of uncertainty. As shown in Fig. 6, all the observed experimental data are included within the established confidence band. Therefore, residual stresses not yet observed are expected with a high degree of reliability to fall within the established confidence band.

VI. Conclusions

In this research, a methodology to evaluate model likelihood using the measured differences between observed experimental data and model predictions of the data involving both parametric uncertainty and random errors is developed to quantify model form uncertainty. The expression used to describe the probabilistic relationship between experimental data and model predictions makes the evaluation of model likelihood easy to implement. By incorporating model form uncertainty, the prediction of a system response is made unconditionally, independent of the specific model considered; the variance in the prediction that would be missing if a single model was considered is incorporated into the composite prediction.

The model averaging technique is illustrated with the combination of the predictions of the fundamental natural frequency for a nonlinear spring-mass system made by the three mathematical models. The proposed methodology is applied to the FE simulation of a LP process. Model form uncertainty is quantified by evaluating model probability using the limited experimental data on residual stresses and the stochastic predictions of the data from the considered FE models. The composite prediction of a residual stress field incorporates model form uncertainty, parametric uncertainty, and random errors involved in model predictions. Given stochastic predictions by a model set and observed experimental data, the proposed methodology can be applied to any problem of model form uncertainty quantification, regardless of the number of considered models and experimental data. However, model probability is quantified based on observed experimental data alone, not using prior information about a model set from the physical approach (either expert opinion or partial engineering knowledge).

Acknowledgment

The authors acknowledge the support of this research work through contract FA8650-04-D-3446, delivery order no. 25, sponsored by Wright-Patterson Air Force Base, Ohio.

References

[1] Alvin, K. F., Oberkampf, W. L., Diegert, K. V., and Rutherford, B. M., "Uncertainty Quantification in Computational Structural Dynamics: A New Paradigm for Model Validation," *Proceedings of the 16th International Modal Analysis Conference*, Santa Barbara, CA, Soc. of Photo-Optical Instrumentation Engineers, Bellingham, WA, 1998, pp. 1191–1198.

[2] Zio, E., and Apostolakis, G., "Two Methods for the Structured Assessment of Model Uncertainty by Experts in Performance Assessments of Radioactive Waste Repositories," *Reliability Engineering and System Safety*, Vol. 54, Nos. 2–3, 1996, pp. 225–241. doi:10.1016/S0951-8320(96)00078-6

[3] Zhang, R., and Mahadevan, S., "Model Uncertainty and Bayesian Updating in Reliability-Based Inspection," *Structural Safety*, Vol. 22, No. 2, 2000, pp. 145–160. doi:10.1016/S0167-4730(00)00005-9

[4] Zouaoui, F., and Wilson, J. R., "Accounting for Input Model and Parameter Uncertainty in Simulation," *Proceedings of the 33nd Conference on Winter Simulation*, Arlington, VA, IEEE Publ., Piscataway, NJ, 2001, pp. 290–299.

[5] McFarland, J. M., and Bichon, B. J., "Bayesian Model Averaging for Reliability Analysis with Probability Distribution Model Form Uncertainty," 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Palm Springs, CA, AIAA Paper 2009-2231, 2009.

[6] Craig, R., and Kurdila, A., *Fundamentals of Structural Dynamics*, Wiley, Hoboken, NJ, 2006, pp. 4–5.

[7] Oberkampf, W. L., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., and Ferson, S., "Challenge Problems: Uncertainty in System Response Given Uncertain Parameters," *Reliability Engineering and System Safety*, Vol. 85, No. 1, 2004, pp. 11–19. doi:10.1016/j.ress.2004.03.002

[8] Claeskens, G., and Hjort, N. L., *Model Selection and Model Averaging*, Cambridge Univ. Press, Cambridge, England, U.K., 2008, pp. 22–98.

[9] Draper, D., "Assessment and Propagation of Model Uncertainty," *Journal of the Royal Statistical Society Series B (Methodological)*, Vol. 57, No. 1, 1995, pp. 45–97.

[10] Raftery, A. E., "Approximate Bayes Factors and Accounting for Model Uncertainty in Generalized Linear Models," *Biometrika*, Vol. 83, No. 2, 1996, pp. 251–266. doi:10.1093/biomet/83.2.251

[11] Park, I., Amarchinta, H. K., and Grandhi, R. V., "A Bayesian Approach for Quantification of Model Uncertainty," *Reliability Engineering and System Safety*, Vol. 95, No. 7, 2010, pp. 777–785. doi:10.1016/j.ress.2010.02.015

[12] Robert, C. P., *The Bayesian Choice: A Decision Theoretic Motivation*, Springer-Verlag, New York, 1994, pp. 13–21.

[13] Bayarri, M. J., Berger, J. O., Paulo, R., Sacks, J., Cafeo, J. A., Cavendish, J., Lin, C. H., and Tu, J., "A Framework for Validation of Computer Models," *Technometrics*, Vol. 49, No. 2, 2007, pp. 138–154. doi:10.1198/004017007000000092

[14] Riley, M. E., and Grandhi, R. V., "Quantification of Model-Form and Predictive Uncertainty for Multi-Physics Simulation," *Computers and Structures* (in press). doi:10.1016/j.compstruc.2010.10.004

[15] Soundappan, P., Nikolaidis, E., and Dheenadayalan, B., "Targeted Testing Methods for Reliability Validation of Components and Systems," *International Journal of Reliability and Safety*, Vol. 4, No. 1, 2010, pp. 16–45. doi:10.1504/IJRS.2010.029563

[16] Kennedy, M. C., and O'Hagan, A., "Bayesian Calibration of Computer Models," *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, Vol. 63, No. 3, 2001, pp. 425–464. doi:10.1111/1467-9868.00294

[17] Xiong, Y.,Chen, W., Tsui, K. L., and Apley, D. W., "A Better Understanding of Model Updating Strategies in Validating Engineering Models," *Computer Methods in Applied Mechanics and Engineering*, Vol. 198, Nos. 15–16, 2009, pp. 1327–1337. doi:10.1016/j.cma.2008.11.023

[18] *Guide for the Verification and Validation of Computational Fluid Dynamics Simulations (AIAA-G-077)*, AIAA Standards Series, AIAA, Reston, VA, 1998, pp. 10–14.

[19] *V&V 10: Guide for Verification and Validation in Computational Solid Mechanics*, American Soc. of Mechanical Engineers, New York, 2006, pp. 16–19.

[20] McFarland, J. M., Mahadevan, S., Swiler, L., and Giunta, A., "Bayesian Calibration of the QASPR Simulation," 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Honolulu, HI, AIAA Paper 2007-1952, 2007.

[21] Draper, D., "Model Uncertainty in Stochastic and Deterministic Systems," *Proceedings of the 12th International Workshop on Statistical Modeling*, Biel/Bienne, Switzerland, Statistical Modelling Soc., Amsterdam, 1997, pp. 43–59.

[22] He, J. H., "Variational Approach for Nonlinear Oscillators," *Chaos, Solitons, and Fractals*, Vol. 34, No. 5, 2007, pp. 1430–1439. doi:10.1016/j.chaos.2006.10.026

[23] Ding, K., and Ye, L., *Laser Shock Peening Performance and Process Simulation*, Woodhead Publ. Limited, Cambridge, England, U.K., 2006, pp. 7–46.

[24] Singh, G., and Grandhi, R. V., "Mixed-Variable Optimization Strategy

Employing Multifidelity Simulation and Surrogate Models," *AIAA Journal*, Vol. 48, No. 1, 2010, pp. 215–223.
doi:10.2514/1.43469

[25] Amarchinta, H. K., Grandhi, R. V., Clauer, A. H., Langer, K., and Stargel, D., "Simulation of Residual Stress Induced by a Laser Peening Process Through Inverse Optimization Of Material Models," *Journal of Materials Processing Technology*, Vol. 210, No. 14, 2010, pp. 1997–2006.
doi:10.1016/j.jmatprotec.2010.07.015

[26] Nam, T., "Finite Element Analysis of Residual Stress Field Induced by Laser Shock Peening," Ph.D. Dissertation, Mechanical Engineering Department, Ohio State Univ., Columbus, OH, 2002.

[27] Chen, W., Baghdasaryan, L., Buranathiti, T., and Cao, J., "Model Validation via Uncertainty Propagation and Data Transformations," *AIAA Journal*, Vol. 42, No. 7, 2004, pp. 1406–1415.
doi:10.2514/1.491

A. Messac
Associate Editor