INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2011; **87**:262–272
Published online 28 September 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nme.3038

Control of modeling error in calibration and validation processes for predictive stochastic models

J. Tinsley Oden*, † and Serge Prudhomme

Institute for Computational Engineering and Sciences, The University of Texas at Austin, 1 University Station C0200, Austin, TX 78712, U.S.A.

SUMMARY

The idea of adaptive control of modeling error is expanded to include ideas of statistical calibration, validation, and uncertainty quantification. Copyright © 2010 John Wiley & Sons, Ltd.

Received 5 February 2010; Revised 31 May 2010; Accepted 2 June 2010

KEY WORDS: error estimation; adaptive modeling; Bayesian approaches; quantities of interest

The first author would like to dedicate this work to the memory of his dear and long-time friend, Olek Zienkiewicz, a founder of IJNME and a great scholar whose work profoundly changed the way numerical methods in engineering are done.

1. INTRODUCTION

Our view of how computer modeling and simulation impact the world of science and engineering has evolved in recent years to one in which the principal factors affecting predictability of a simulation and their interaction with one another are more clearly distinguishable than ever before. Among these factors are (1) the selection of a mathematical model that provides an abstraction of the physical events of interest; (2) the identification of appropriate parameters that define the model; (3) the use of physical observations and measurements to calibrate and validate the model; (4) the development of a computational model through discretization of the mathematical model; (5) the identification of the specific goals of the simulation, the *quantities of interest* or *target outputs*; and (6) the quantification of the uncertainty in the predictions.

An obvious property of each of these factors is that they all are imprecise and lead to error and uncertainty in predictions: the heuristic act of selecting a mathematical abstraction of physical reality can be the most critical source of error, which we call modeling error; the parameters of the model are rarely known precisely and can generally be determined only in some rough statistical sense, and the same is true of observations and measurements that are obtained through imperfect systems and devices and which are generally limited to only a few features of the response of the system; the corruption of the model due to discretization leads to discretization (or 'numerical') error, and all these errors propagate through the target outputs, which represent quantities with significant uncertainty.

*Correspondence to: J. Tinsley Oden, Institute for Computational Engineering and Sciences, The University of Texas at Austin, 1 University Station C0200, Austin, TX 78712, U.S.A.
†E-mail: oden@ices.utexas.edu

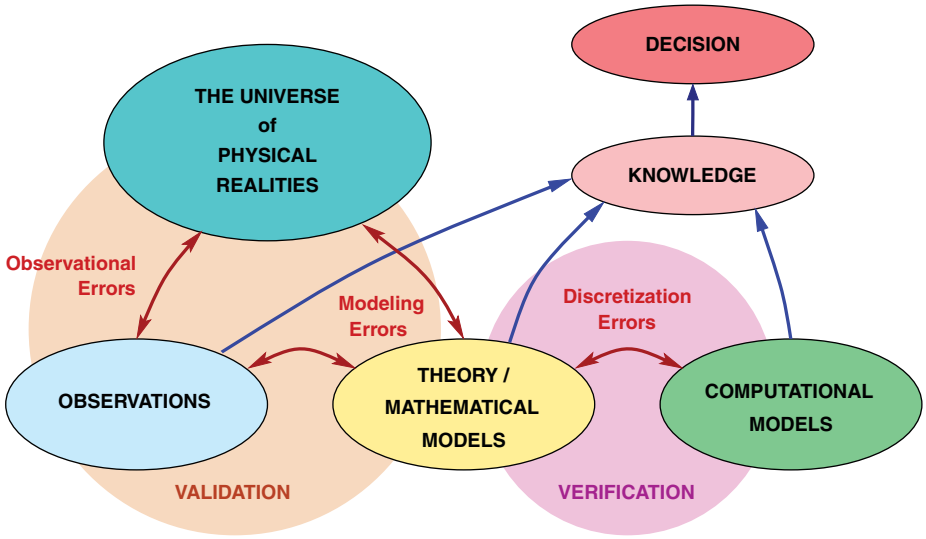

Figure 1. The imperfect paths to knowledge.

These various sources of error in the broad landscape of scientific discovery are depicted symbolically in Figure 1 [1]. There it is indicated that the issue of modeling error (the selection of the mathematical model that one hopes can best produce reliable quantities of interest) is the goal of *validation processes*. The assessment of the effect of discretization error on the quantities of interest is the province of *verification processes*, and is a problem of *a posteriori* error estimation. Selecting the best characterization of model parameters and observational data is an issue of statistical calibration, which is an exercise in the statistical inverse theory.

A broad framework for embracing all these components of modeling physical events and of coping with error and uncertainty is provided by the statistical inverse theory based on Bayesian (or Bayesian like) statistics. We describe one such framework in this paper. The estimation and control of modeling error that we have explored in previous papers (e.g. [2–6]) pertains to *relative error*, in which results produced by a sequence of surrogate models are compared with that capable of being delivered by a grand, master, base model, which itself may be in error with reality. In the present work, we improve this setting by bringing into play statistical features of the base model based on enriching the likelihood probability density that relates model parameters to observations.

Following this introductory section, we lay down in Section 2 an abstract mathematical framework for modeling systems with uncertain parameters and calibration measurements. This is configured into a statistical inverse framework provided by a Bayesian approach. The generalization of the models to stochastic systems is taken up in Section 3. Model calibration, validation, and prediction with quantifiable uncertainty are discussed in Section 4. We then take up the statistical aspects of modeling error in Section 5.

2. PRELIMINARIES

2.1. *The abstract mathematical model*

We shall first need to lay down a collection of preliminary concepts and notations that provide suitable abstractions of the idea of a family of mathematical models of a physical event of interest. We are interested in modeling a physical system to predict its behavior under a particular scenario of conditions. Based on our experience, available information, and possibly heuristic arguments, we select a mathematical theory we believe can describe the event for a particular scenario of a

DOI: 10.1002/nme

solution domain, boundary conditions, initial conditions, etc. This can be expressed as the abstract problem of finding a function u in a space $\mathscr{U}$ of trial functions, such that:

$$A(\boldsymbol{m}, S, u(\boldsymbol{m}, S)) = 0. \tag{1}$$

Here $A(\cdot)$ represents an abstract operator defined by the theory we propose to depict the event of interest (e.g. $A(\cdot)$ may represent the equations of compressible fluid or gas dynamics, elastodynamics, quantum mechanics, molecular dynamics, etc. depending on the physical event, along with appropriate boundary and initial conditions), $\boldsymbol{m}$ is a vector of parameters defining specific models in the class, S is the set of scenario parameters such as the solution domain, boundary and initial data, and $u(\boldsymbol{m}, S)$ is the solution of the equations describing the model for particular choices of parameters $\boldsymbol{m}$ and S. Thus, once a theory and its associated mathematical characterization ($A(\cdot)$) are selected, and a particular scenario S is selected, we may traverse over infinitely many models in this class by varying the model parameters $\boldsymbol{m}$. The parameter set $\boldsymbol{m}$ is assumed to belong to a model differentiable manifold M, which we discuss fully in Section 2.2 below.

Example 1

The differentiation between the model class and the various scenarios is important. As an example, the model class could correspond to linear elastostatics of isotropic, homogeneous bodies. Then

$$A(\boldsymbol{m}, S, u(\boldsymbol{m}, S)) = 0 \sim \begin{cases} -\dfrac{\partial \sigma_{ij}}{\partial x_j}(\boldsymbol{x}) = f_i(\boldsymbol{x}), \quad 1 \leqslant i, j \leqslant n, \ \ \boldsymbol{x} \in \Omega \subset \mathbb{R}^n, \\ \sigma_{ij}(\boldsymbol{x}) = \dfrac{E}{1+\nu}\varepsilon_{ij}(\boldsymbol{x}) + \dfrac{E\nu}{(1+\nu)(1-2\nu)}\varepsilon_{kk}(\boldsymbol{x})\delta_{ij}, \\ 2\varepsilon_{ij}(\boldsymbol{x}) = \dfrac{\partial u_i}{\partial x_j}(\boldsymbol{x}) + \dfrac{\partial u_j}{\partial x_i}(\boldsymbol{x}), \\ \sigma_{ij}(\boldsymbol{x})\ n_i(\boldsymbol{x}) = g_i(\boldsymbol{x}), \quad \boldsymbol{x} \in \Gamma \subset \partial\Omega, \\ u_i(\boldsymbol{x}) = 0, \quad \boldsymbol{x} \in \partial\Omega \backslash \Gamma. \end{cases} \tag{2}$$

The parameter set is taken to be $\boldsymbol{m} = (\log E/E_0, \log \nu/\nu_0)$ where E_0 and ν_0 are arbitrary constant moduli. There are many possible scenarios:

$$S = S_1 \sim \text{a simple tension or torsion test}, \tag{3}$$

$$S = S_2 \sim \text{a two-dimensional square plate}, \tag{4}$$

$$S = S_3 \sim \text{a three-dimensional body}. \tag{5}$$

Thus, whether a quantity such as a body force $\boldsymbol{f}$, surface tractions $\boldsymbol{g}$, etc. are designated as the parameters of the model (and included in $\boldsymbol{m}$) or as scenario data S is generally a matter of subjective choice.

For reasons that will become clear later, it is important at this point to recast (1) into a function space setting. For each $\boldsymbol{m}$ and S, we presume there exists a topological vector space (TVS) of trial functions $\mathscr{U}_S$ and a TVS of test functions $\mathscr{V}_S$, and that $A(\boldsymbol{m}, S, \cdot)$ maps $\mathscr{U}_S$ into the dual $\mathscr{V}_S'$ of $\mathscr{V}_S$. Then (1) can be expressed as the condition

$$\mathscr{R}(\boldsymbol{m}, S, u(\boldsymbol{m}, S); v) = 0 \quad \forall v \in \mathscr{V}_S, \tag{6}$$

where $\mathscr{R}(\cdot;\cdot)$ is the continuous linear functional,

$$\mathscr{R}(\boldsymbol{m}, S, u(\boldsymbol{m}, S); v) = \langle A(\boldsymbol{m}, S, u(\boldsymbol{m}, S)); v \rangle \tag{7}$$

and $\langle \cdot;\cdot \rangle$ denotes the duality pairing on $\mathscr{V}_S' \times \mathscr{V}_S$. We shall refer to $\mathscr{R}(\cdot;\cdot)$ as the *residual functional* and to problem (6) as the *forward problem*. Equality (6) thus states that of all the trial functions

DOI: 10.1002/nme

in $\mathscr{U}_S$, those for which the residual functional vanishes for all test functions in $\mathscr{V}_S$ are solutions of (1) (for fixed $\boldsymbol{m}$ and S).

2.2. Quantities of interest

Central to the goal of predictive simulation is the specification of *quantities of interest*, 'QoI's', the target features of a physical event that we identify as critical to making decisions about the outputs of a model of physical or engineered systems. These are the goals of the simulation. The QoI's are not, in general, the solution of some general set of equations, but they may be characterized as functionals of the solutions of a system of governing equations.

We shall assume that each QoI is a continuous differentiable functional Q_S on the space of trial functions $\mathscr{U}_S$ with variational (Gâteaux) derivative $Q'_S : \mathscr{U}_S \times \mathscr{U}_S \longrightarrow \mathbb{R}$:

$$Q'_S(u; v) = \lim_{\theta \to 0} \theta^{-1} [Q_S(u + \theta v) - Q_S(u)]. \tag{8}$$

Here $Q'_S(u; \cdot)$ is a linear functional on $\mathscr{U}_S$ for a given scenario S.

2.3. The adjoint problem

Corresponding to each quantity of interest Q, we can formulate the dual or adjoint problem to (1)

$$\mathscr{R}'(\boldsymbol{m}, S, u(\boldsymbol{m}, S); v, p(\boldsymbol{m}, S)) = Q'_S(u(\boldsymbol{m}, S); v) \quad \forall v \in \mathscr{V}_S, \tag{9}$$

where $\mathscr{R}'(\cdot; \cdot, \cdot) : \mathscr{U}_S \times \mathscr{U}_S \times \mathscr{V}_S \longrightarrow \mathbb{R}$ is the functional:

$$\mathscr{R}'(\boldsymbol{m}, S, u; v, p) = \lim_{\theta \to 0} \theta^{-1} [\mathscr{R}(\boldsymbol{m}, S, u + \theta v; p) - \mathscr{R}(\boldsymbol{m}, S, u; p)]. \tag{10}$$

In (9), $p(\boldsymbol{m}, S)$ is the adjoint or dual solution or generalized Green's function corresponding to Q_S and problem (6). Problem (9) is linear in p with coefficients possibly dependent on the solution $u(\boldsymbol{m}, S)$ to (6).

2.4. Statistical inverse analysis: calibration of the model

Up to this point, the mathematical models (6) and corresponding adjoint models (9) have been treated as deterministic. But this, in general, is unacceptable, as there is generally significant uncertainty in the values and variables in the parameters $\boldsymbol{m}$. Moreover, our initial knowledge of the parameters may be incomplete and may not reflect the actual physical setting of the events we wish to study. Our knowledge of the system can be increased with a judicious choice of physical experiments for certain scenarios S that provide sets of observational data $\boldsymbol{d}$ that can be used to calibrate the model. But the observational data themselves have uncertainties.

To address these difficulties, we will call upon Bayesian methods of statistical inverse analysis (see, e.g. [7, 8] or the variant proposed by Tarantola [9]). Forthwith, for the model manifold M, we define a probability space $(M, \mathscr{F}_M, P_M)$, with probability measure P_M, and we assume that one can define the random parameters $\boldsymbol{m}$ by a *prior probability density function* (pdf) $\rho_M(\boldsymbol{m})$, $dP_M = \rho_M(\boldsymbol{m}) d\boldsymbol{m}$. Likewise, we define a data manifold $\mathscr{D}$ on which a probability measure P_D and a corresponding data pdf $\rho_D(\boldsymbol{d})$ can be defined as well as a corresponding homogeneous pdf, $\mu_D(\boldsymbol{d})$ (see [9]). The pdf's ρ_M and ρ_D are obviously independent, so we may define the *joint prior probability density* as

$$\rho(\boldsymbol{m}, \boldsymbol{d}) = \rho_M(\boldsymbol{m}) \rho_D(\boldsymbol{d}). \tag{11}$$

Here, and hereafter, we employ the notation of Tarantola [9].

The information provided by the mathematical theory we use to describe the physical events (in the case of model calibration, these events are the observations $\boldsymbol{d}$) is embodied in the *likelihood probability*, which is characterized by a conditional pdf, $\theta(\boldsymbol{d}|\boldsymbol{m})$ that defines the theoretical

DOI: 10.1002/nme

probability density of an observation $\boldsymbol{d}$ given parameters $\boldsymbol{m}$. Finally, the *joint posterior probability density*, $\sigma(\boldsymbol{m},\boldsymbol{d})$, is defined as the conjunction of the prior and the likelihood probabilities:

$$\sigma(\boldsymbol{m},\boldsymbol{d})=k\frac{\rho(\boldsymbol{m},\boldsymbol{d})}{\mu_D(\boldsymbol{d})}\theta(\boldsymbol{d}|\boldsymbol{m}). \tag{12}$$

Here k is a normalization constant. The marginal probability density,

$$\sigma_M(\boldsymbol{m})=\int_{\mathscr{D}}\sigma(\boldsymbol{m},\boldsymbol{d})\,\mathrm{d}\boldsymbol{d} \tag{13}$$

is the solution to the statistical inverse problem. It defines the calibrated pdf for the model parameters $\boldsymbol{m}$. Introducing (11) and (12) into (13) gives

$$\sigma_M(\boldsymbol{m})=k\rho_M(\boldsymbol{m})L(\boldsymbol{m}), \tag{14}$$

where $L(\boldsymbol{m})$ is the *likelihood function*

$$L(\boldsymbol{m})=\int_{\mathscr{D}}\frac{\rho_D(\boldsymbol{d})}{\mu_D(\boldsymbol{d})}\theta(\boldsymbol{d}|\boldsymbol{m})\,\mathrm{d}\boldsymbol{d}. \tag{15}$$

Given the prior pdf's, the homogeneous pdf's, and the likelihood conditional pdf, we can generate the posterior pdf in (14) using standard sampling methods. See [7–10].

3. STOCHASTIC SYSTEMS

We now consider generalizations of the forward and adjoint problems to stochastic systems. As noted in Section 2.4, the model parameters $\boldsymbol{m}$ define points in a sample set, also denoted M, in a probability space

$$\mathscr{P}_M=(M,\mathscr{F}_M,P_M), \tag{16}$$

where $\mathscr{F}_M$ is a σ-algebra on M and P_M is a probability measure on $\mathscr{F}_M$. A random vector $\boldsymbol{d}(\boldsymbol{m})$ is a function from M into $\mathbb{R}^n$. We equip the space $\mathscr{P}_M$ with an $L^2(M)$—inner product $(\cdot,\cdot)_M$ so that if $\boldsymbol{\alpha}(\boldsymbol{m})$ and $\boldsymbol{\beta}(\boldsymbol{m})$ are two random vectors,

$$(\boldsymbol{\alpha}(\boldsymbol{m}),\boldsymbol{\beta}(\boldsymbol{m}))_M=\int_M\boldsymbol{\alpha}(\boldsymbol{m})\cdot\boldsymbol{\beta}(\boldsymbol{m})\,\mathrm{d}P_M. \tag{17}$$

Hereafter, we assume that $\mathrm{d}P_M$ is defined by a probability density σ_M, so that $\mathrm{d}P_M=\sigma_M(\boldsymbol{\alpha}(\boldsymbol{m}),\boldsymbol{\beta}(\boldsymbol{m}))\,\mathrm{d}\boldsymbol{\alpha}\mathrm{d}\boldsymbol{\beta}$.

Returning now to (6), we introduce the TVS's of random vector fields for a fixed scenario S,

$$\mathscr{X}=\mathscr{P}_M\times\mathscr{U}_S,\quad \mathscr{Y}=\mathscr{P}_M\times\mathscr{V}_S. \tag{18}$$

Let A denote a map from $\mathscr{X}$ to $\mathscr{Y}$ which depends on parameters $\boldsymbol{m}$ characterized by the pdf $\sigma_M(\boldsymbol{m})$. We can now introduce the stochastic forward problem

$$\mathscr{R}(\sigma_M(\boldsymbol{m}),S,u(\sigma_M(\boldsymbol{m}),S);v)=0\quad\forall v\in\mathscr{Y}, \tag{19}$$

where $\mathscr{R}(\cdot)$ is the residual,

$$\mathscr{R}(\sigma_M(\boldsymbol{m}),S,u(\sigma_M(\boldsymbol{m}),S);v)=\int_M\langle A(\sigma_M(\boldsymbol{m}),S,u(\sigma_M(\boldsymbol{m}),S);v\rangle\,\mathrm{d}P_M. \tag{20}$$

The QoI is defined by

$$Q:\mathscr{X}\longrightarrow\mathbb{R},\quad Q(u(\sigma_M(\boldsymbol{m}),S))=q_S(\boldsymbol{m}). \tag{21}$$

DOI: 10.1002/nme

Let v and w be arbitrary random fields in $\mathscr{X}$ and $\mathscr{Y}$, respectively. We define:

$$\mathscr{R}'(\sigma_M(\boldsymbol{m}), S, u(\sigma_M(\boldsymbol{m}), S); v, w) = \lim_{\theta\to 0} \theta^{-1}[\mathscr{R}(\sigma_M(\boldsymbol{m}), S, u(\sigma_M(\boldsymbol{m}), S)+\theta v; w) - \mathscr{R}(\sigma_M(\boldsymbol{m}), S, u(\sigma_M(\boldsymbol{m}), S); w)], \quad (22)$$

$$Q'(u(\sigma_M(\boldsymbol{m}), S); v) = \lim_{\theta\to 0} \theta^{-1}[Q(u(\sigma_M(\boldsymbol{m}), S)+\theta v) - Q(u(\sigma_M(\boldsymbol{m}), S))]. \quad (23)$$

The stochastic adjoint problem is then to find $p(\sigma_M(\boldsymbol{m}))$ such that:

$$\mathscr{R}'(\sigma_M(\boldsymbol{m}), S, u(\sigma_M(\boldsymbol{m}), S); v, p(\sigma_M(\boldsymbol{m})) = Q'(u(\sigma_M(\boldsymbol{m}), S); v) \quad \forall v \in \mathscr{Y}. \quad (24)$$

Here we have introduced a somewhat heavier but more descriptive notation. The parameters $\boldsymbol{m}$ are now replaced with the posterior pdf σ_M determined by the calibration defined in (14). The solution $u(\sigma_M(\boldsymbol{m}), S)$ for a given scenario S is a random field, as is the adjoint solution $p(\sigma_M(\boldsymbol{m}))$. The scenario S is arbitrary for the moment and assumed to be deterministic, but extensions to random scenarios are straightforward. The output values of the QoI's, $Q(\cdot)$ are random fields. As noted earlier, the posterior pdf's $\sigma_M(\boldsymbol{m})$ in (14) are evaluated by standard sampling methods such as the Markov Chain Monte Carlo (MCMC) or Hasting methods (see e.g. [7, 10]), which have been implemented in open-source codes [10]. The general framework represented in (19) can now be applied to the key steps in the calibration and validation of the model.

4. THE PREDICTION PYRAMID: MODEL CALIBRATION, VALIDATION, AND PREDICTION

The contemporary concept of a predictive computational model of a physical event presumes that observational data is acquired that enables the calibration of the model for simple, possibly laboratory, scenarios of the theory, and other experimental observations in more complex settings which are part of the process of validation of the model. These three processes, calibration, validation, and prediction, make up levels of the *prediction pyramid*, illustrated in Figure 2 (cf. [11]).

At the peak of the pyramid is the most sophisticated and complete model scenario (denoted S_p in the figure) where we hope to predict the ultimate QoI, denoted Q_p, with a model that survives our validation process. At a lower level of the pyramid, we consider simplified validation scenarios and we target validation observables $\boldsymbol{d}_v$ (or, equivalently, validation QoI's). These validation levels may involve submodels of the general model solved with validation scenarios S_v. At the lowest level of the pyramid we consider simple (possibly component based) scenarios for the purpose of

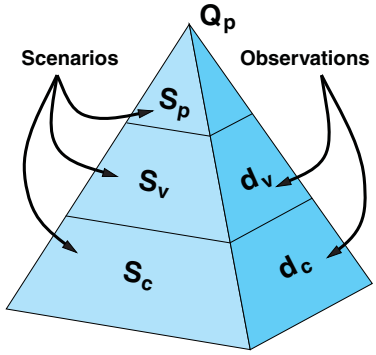

Figure 2. The prediction pyramid.

DOI: 10.1002/nme

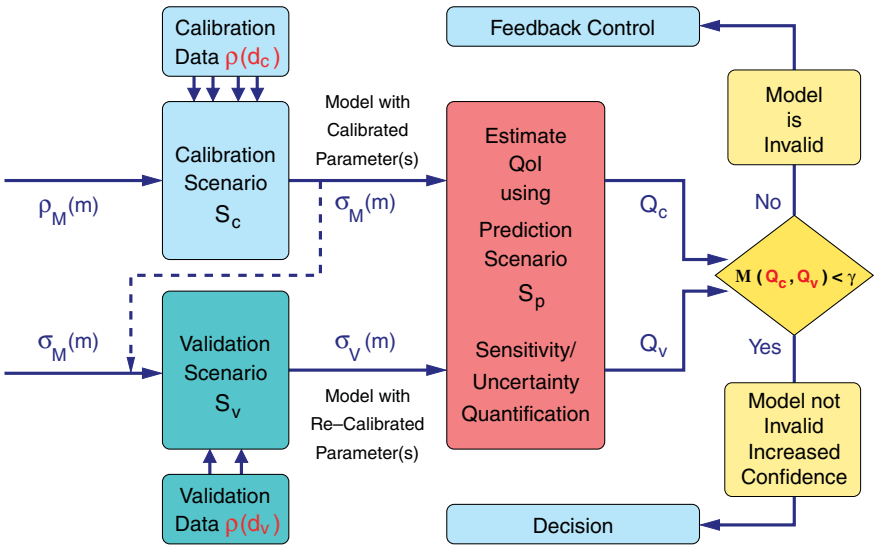

Figure 3. The flow of a calibration, validation, and prediction processes. Prior pdf's of model parameters and pdf of observational data are used to calibrate the model for a given calibration scenario S_c. Inverse analysis yields the posterior pdf $\sigma_M(\boldsymbol{m})$. This is the prior for the more elaborate validation process which involves a new scenario S_v and validation data with prior $\rho(\boldsymbol{d}_v)$. Inverse analysis yields the posterior $\sigma_V(\boldsymbol{m})$. The stochastic forward problems are solved using calibrated and validated data and the corresponding prediction QoI's are compared. If they meet a preset tolerance, the model is not invalidated and the calibrated model is used to make the predictions. Otherwise the model is declared invalid.

model calibration, and we perform experiments that provide observables represented by the data $\boldsymbol{d}_c \in \mathscr{D}$, which can be regarded as QoI's for the calibration process. There may be many steps of increasing complexity represented by layers in the validation levels of the pyramid, but we shall consider a single validation process here for simplicity.

The calibration, validation, and prediction processes are illustrated in Figure 3 (cf. [11, 12]) and involve the following steps:

Calibration: For a selected theoretical framework, defined by the calibration scenario S_c, we

(1) provide the prior pdf $\rho_M(\boldsymbol{m})$,
(2) determine the pdf $\rho_D(\boldsymbol{d}_c)$ (and $\mu_D(\boldsymbol{d}_c)$) from the given experimental observables,
(3) introduce the likelihood pdf $\theta(\boldsymbol{d}_c|\boldsymbol{m})$ for the calibration scenario based on the theory $(A(\cdot,\cdot,\cdot))$,
(4) solve the inverse problem for the posterior pdf σ_M as in (14), i.e.

$$\sigma_M(\boldsymbol{m}) = k\rho_M(\boldsymbol{m})L_c(\boldsymbol{m}),$$

where $L_c(\boldsymbol{m})$ is the likelihood function (15) computed using the likelihood pdf $\theta(\boldsymbol{d}_c|\boldsymbol{m})$ for the calibration problem, using sampling methods such as those described in [7, 8, 10].

The calibrated model for the prediction problem can now be written,

$$\mathscr{R}(\sigma_M(\boldsymbol{m}), S_p, u(\sigma_M(\boldsymbol{m}), S_p); v) = 0 \quad \forall\, v \in \mathscr{Y}. \tag{25}$$

Validation: A validation scenario S_v is now selected together with validation observations $\boldsymbol{d}_v$, which are hoped to reflect the ability of the model to deliver acceptable predictions of the prediction QoI, Q_p. The data manifold for the validation experiments is denoted $\mathscr{D}_v$ and the pdf for the validation data is denoted $\rho_V(\boldsymbol{d}_v)$, with homogeneous pdf $\mu_V(\boldsymbol{d}_v)$. Then

(1) the posterior pdf of the calibration process is used as the prior pdf for the validation process,
(2) we introduce the likelihood pdf for the validation process, $\theta(\boldsymbol{d}_v|\boldsymbol{m})$,

DOI: 10.1002/nme

(3) we solve the inverse problem for the posterior pdf $\sigma_V(\boldsymbol{m})$ for the validation process, i.e.

$$\sigma_V(\boldsymbol{m}) = k\sigma_M(\boldsymbol{m})L_v(\boldsymbol{m}),$$

where k is a normalization constant and $L_v(\boldsymbol{m})$ is the validation likelihood function

$$L_v(\boldsymbol{m}) = \int_{\mathscr{D}_v} \frac{\rho_V(\boldsymbol{d}_v)}{\mu_V(\boldsymbol{d}_v)} \theta(\boldsymbol{d}_v|\boldsymbol{m})\,\mathrm{d}\boldsymbol{d}_v. \tag{26}$$

(4) the model for the prediction problem produced by the validation process (which as yet may be an invalid model) may be written

$$\mathscr{R}(\sigma_V(\boldsymbol{m}), S_p, u(\sigma_V(\boldsymbol{m}), S_p); v) = 0 \quad \forall v \in \mathscr{Y}. \tag{27}$$

The mathematical model chosen for the prediction (1) can never be validated; it can, at best, be not-invalidated for the specific validation experiments we performed. The determination of a criterion for accepting a model as 'not-invalidated' is a subjective decision that requires the acceptance of a metric to compare prediction QoI's produced by the calibration and the validation processes, and a tolerance that we establish as an acceptable measure of the predictability of the model. Thus, our validation process involves comparing $Q_p(u(\sigma_M(\boldsymbol{m}), S_p))$ and $Q_p(u(\sigma_V(\boldsymbol{m}), S_p))$. Let $\mathbb{D}$ denote a metric on the space of random functions containing $Q_p(u(\sigma_M(\boldsymbol{m}), S_p))$ and $Q_p(u(\sigma_V(\boldsymbol{m}), S_p))$ and let γ_{tol} denote a preset tolerance. We will declare the model **not-invalid** if

$$\mathbb{D}(Q_p(u(\sigma_M(\boldsymbol{m}), S_p)), Q_p(u(\sigma_V(\boldsymbol{m}), S_p))) < \gamma_{\mathrm{tol}}. \tag{28}$$

There are many metrics that could be used in (28). One reasonable possibility is suggested in Figure 4. Let us use the more concise notation,

$$\left.\begin{aligned} Q_C(\boldsymbol{m}) &= Q_p(u(\sigma_M(\boldsymbol{m}), S_p) \\ Q_V(\boldsymbol{m}) &= Q_p(u(\sigma_V(\boldsymbol{m}), S_p) \end{aligned}\right\}. \tag{29}$$

We can always compute the cumulative pdfs [11]

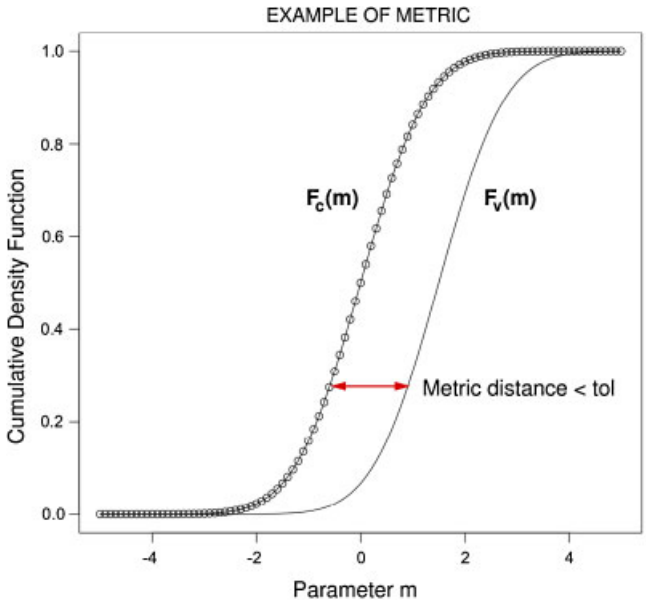

Figure 4. Illustration of the metric $\mathbb{D}$ defined in (31).

DOI: 10.1002/nme

$$\left.\begin{aligned} F_C(\boldsymbol{m}) &= \int_{-\infty}^{\boldsymbol{m}} \hat{\sigma}_C(\boldsymbol{m}')\,\mathrm{d}m' \\ F_V(\boldsymbol{m}) &= \int_{-\infty}^{\boldsymbol{m}} \hat{\sigma}_V(\boldsymbol{m}')\,\mathrm{d}m' \end{aligned}\right\}, \tag{30}$$

where $\hat{\sigma}_C$ and $\hat{\sigma}_V$ are the pdf's corresponding to $Q_C(\boldsymbol{m})$ and $Q_V(\boldsymbol{m})$, respectively. Then a comparison of Q_C and Q_V is afforded by

$$\mathbb{D}(Q_C, Q_V) = \sup_{y\in[0,1]} |F_C^{-1}(y) - F_V^{-1}(y)|. \tag{31}$$

Uncertainty Quantification. With a (subjective) choice of a tolerance γ_{tol}, one can check (28) (using (31)) to determine if the criterion for not-invalidity is satisfied. If (28) holds, the model has passed the tests laid down in the validation process, and based on these tests, it has not been found to be invalid. It is commonplace to call this model 'validated', but of course, no model can ever be 'validated', it can only be declared not invalid for a specific set of validation tests.

Let us suppose that the model has passed the validation criterion (28) for a chosen metric $\mathbb{D}$ and tolerance γ_{tol}. How can the uncertainty in the predicted quantity of interest Q_p be quantified? Denote

$$q_p(\boldsymbol{m}) = Q_p(u(\sigma_M(\boldsymbol{m}), S_p). \tag{32}$$

Then the following statistical properties of $q_p(\boldsymbol{m})$ can be calculated: its means and centered moments,

$$\mathbb{E}[q_p] = \int_M q_p(\boldsymbol{m})\ \sigma_M(\boldsymbol{m})\,\mathrm{d}m, \tag{33}$$

$$\mu_k[q_p] = \mathbb{E}[q_p - \mathbb{E}[q_p]]^k = \int_M [q_p(\boldsymbol{m}) - \mathbb{E}[q_p]]^k \sigma_M(\boldsymbol{m})\,\mathrm{d}m, \quad k = 2, 3, \ldots \tag{34}$$

and particularly the variance, for $k=2$, i.e.

$$\mathrm{Var}[q_p] = \mu_2[q_p] = \mathbb{E}[q_p - \mathbb{E}[q_p]]^2. \tag{35}$$

These numbers, (33)–(35), quantify the uncertainty in the quantity of interest.

5. ESTIMATION OF MODELING ERROR

The notion of modeling error introduced in [2, 4, 6] is that of determining the relative error between QoI's evaluated using a possibly incorrect surrogate model compared to those delivered by a more general and complete base model. The base model is thus used as a datum with respect to which all other models are compared. Let $\mathscr{U}^*_{S_p}$ and $\mathscr{V}^*_{S_p}$ be two topological vector spaces and let $\mathscr{R}^*$, defined on $\mathscr{U}^*_{S_p} \times \mathscr{V}^*_{S_p}$ for a given set of parameters $\boldsymbol{m}^*$ and scenario S_p, denote a semilinear functional describing the forward base problem:

$$\mathscr{R}^*(\boldsymbol{m}^*, S_p, u^*; v) = 0 \quad \forall v \in \mathscr{V}^*_{S_p}. \tag{36}$$

For simplicity, we assume that the base problem is deterministic with parameters $\boldsymbol{m}^* \in M^*$ fixed, so that $u^* = u^*(\boldsymbol{m}^*, S_p)$, but our results are easily extended to cases in which the base forward problem is endowed with, for example, a lognormal distribution of model parameters.

The following proposition is proved in [2].

Proposition 1
Let the functional $\mathscr{R}^*$ in (36) and the functional Q_p of the prediction scenario be thrice Gâteaux differentiable (in the sense of (8)) and let u_0 be an arbitrary element of $\mathscr{U}^*_{S_p}$. Then

$$Q_p(u^*) - Q_p(u_0) = \mathscr{R}^*(\boldsymbol{m}^*, S_p, u_0; p^*) + \Delta(\boldsymbol{m}^*, u^* - u_0), \tag{37}$$

DOI: 10.1002/nme

where $p^* = p^*(\boldsymbol{m}^*) \in \mathscr{V}^*_{S_p}$ is the solution of the adjoint problem

$$\mathscr{R}^{*\prime}(\boldsymbol{m}^*, u^*; v, p^*) = Q'_p(u^*; v) \qquad \forall v \in \mathscr{U}^*_{S_p} \tag{38}$$

and $\Delta(\cdot,\cdot)$ is the remainder:

$$\Delta(\boldsymbol{m}^*, u^* - u_0) = \int_0^1 \mathscr{F}(\boldsymbol{m}^*, s u_0 + (1-s)u^*, s)\,\mathrm{d}s \tag{39}$$

and the integrand $\mathscr{F}(\boldsymbol{m}^*, \cdot, \cdot)$ is a function of quadratic and cubic terms in the error $e_0 = u^* - u_0$.

Hereafter, we neglect the remainder $\Delta(\cdot,\cdot)$ in the equality (37) and approximate the error in the QoI as

$$\mathscr{E}_{\mathrm{QoI}} = Q_p(u^*) - Q_p(u_0) \approx \mathscr{R}^*(\boldsymbol{m}^*, S_p, u_0; p^*). \tag{40}$$

Obviously, the approximation (40) is legitimate only if u_0 is close to u^* in some sense. In general, the validity of the *a posteriori* modeling error estimate (40) must be checked with each choice of $\mathscr{R}^*(\boldsymbol{m}^*, S_p, \cdot; \cdot)$, either numerically for different choices of u_0 or analytically when possible.

Beyond this point, our ideas are in a very preliminary stage and much additional work remains to be done. We do know that, in principle, the generation of sequences of surrogate models to systematically reduce the relative modeling error $\mathscr{E}_{\mathrm{QoI}}$ can be accomplished via goal-oriented adaptive modeling, and is discussed at length in earlier papers (e.g. [3, 13]). The target quantity of interest evaluated using the base model is denoted as $Q_p(u^*)$, while that of these calibration and validation models are $Q_{p_c}(\boldsymbol{m}) = Q_p(\pi u(\sigma_M(\boldsymbol{m}), S_p))$ and $Q_{p_v}(\boldsymbol{m}) = Q_p(\pi u(\sigma_V(\boldsymbol{m}), S_p))$, π being a map from $\mathscr{U}_{S_p}$ (in fact $\mathscr{X}_{S_p}$ since $\boldsymbol{m}$ is random) into $\mathscr{U}^*_{S_p}$. The operator π is, at this point, ill-defined and many possibilities exist. For example, one can compute the expected value $E[\boldsymbol{m}]$, then $u(E[\boldsymbol{m}], S_p)$ is deterministic and can be prolongated to a function $u_0 \in \mathscr{U}^*_{S_p}$. Remarkably, the error in the quantity of interest between calibrated parameters and validation parameters is

$$\begin{aligned} Q_{p_c}(\boldsymbol{m}) - Q_{p_v}(\boldsymbol{m}) &= Q_{p_c}(\boldsymbol{m}) - Q_p(u^*) + Q_p(u^*) - Q_{p_v}(\boldsymbol{m}) \\ &= [Q_p(u^*) - Q_p(\pi u(\sigma_V(\boldsymbol{m}), S_p))] - [Q_p(u^*) - Q_p(\pi u(\sigma_M(\boldsymbol{m}), S_p))] \\ &\approx \mathscr{R}^*(\boldsymbol{m}^*, S_p, \pi u(\sigma_V(\boldsymbol{m}), S_p); p^*) - \mathscr{R}^*(\boldsymbol{m}^*, S_p, \pi u(\sigma_M(\boldsymbol{m}), S_p); p^*). \end{aligned} \tag{41}$$

The right-hand side defines a (presumed) computable random field that provides a measure of the validity of the model when appropriately quantified.

Feedback Control. The essentially open problem of controlling the model, via the likelihood probability and the parameter priors, to improve a model that is initially judged to be invalid, can be viewed as a feedback control problem with complex sampling. In the special case where the surrogate and base models have same parameters, one can update the base model such that

$$\left.\begin{aligned} \sigma^*_M(\boldsymbol{m}) &= k_1\ \sigma_M(\boldsymbol{m}) \int_{\mathscr{D}} \rho_D(\boldsymbol{d}_c)\ \mu_D^{-1}(\boldsymbol{d}_c)\ \theta^*(\boldsymbol{d}_c|\boldsymbol{m})\,\mathrm{d}\boldsymbol{d}_c \\ \sigma^*_V(\boldsymbol{m}) &= k_2\ \sigma^*_M(\boldsymbol{m}) \int_{\mathscr{D}_v} \rho_V(\boldsymbol{d}_v)\ \mu_V^{-1}(\boldsymbol{d}_v)\ \theta^*(\boldsymbol{d}_v|\boldsymbol{m})\,\mathrm{d}\boldsymbol{d}_v \end{aligned}\right\}, \tag{42}$$

denote the posterior parameter pdf's determined using the likelihood pdf θ^* for the calibration and validation processes, respectively, based on a grand base model, for which the forward problem is of the form,

$$\mathscr{R}^*(\sigma^*_M(\boldsymbol{m}), S_p, u^*(\sigma^*_M(\boldsymbol{m}), S_p); v) = 0 \quad \forall v \in \mathscr{V}^*. \tag{43}$$

In fact, in (42), the set $\boldsymbol{m}$ could also include additional entries to reflect the possible increase in the number of parameters required in the base model. The new QoI is then

$$Q^*_{p_c}(\boldsymbol{m}) \equiv Q_p(u^*(\sigma^*_M(\boldsymbol{m}), S_p)) \tag{44}$$

DOI: 10.1002/nme

and the adjoint problem is,

$$\mathscr{R}^{*\prime}(\sigma_M^*(\boldsymbol{m}), S_p, u^*(\sigma_M^*(\boldsymbol{m}), S_p); v, p_c^*) = Q_p^*(v) \quad \forall v \in \mathscr{X}^*. \tag{45}$$

Likewise, we can formulate forward and adjoint problems based on the validation pdf $\sigma_V^*(\boldsymbol{m})$ and the QoI $Q_{p_v}^*(\boldsymbol{m}) \equiv Q_p(u^*(\sigma_V^*(\boldsymbol{m}), S_p))$. Thus

$$\mathbb{D}(Q_{p_c}^*, Q_{p_v}^*) = \mathbb{D}(Q_{p_c}(\boldsymbol{m}) + \mathscr{R}_c^*(\boldsymbol{m}), Q_{p_v}(\boldsymbol{m}) + \mathscr{R}_v^*(\boldsymbol{m})), \tag{46}$$

where $\mathscr{R}_c^*(\boldsymbol{m})$ is given by (43) and $\mathscr{R}_v^*(\boldsymbol{m})$ is the corresponding function associated with the validation pdf. The problem is thus reduced to choosing a sequence of parameters $\boldsymbol{m}$ so that

$$\mathbb{D}(Q_{p_c}^*, Q_{p_v}^*) \leqslant \gamma_{\text{tol}}. \tag{47}$$

If this is not satisfied, the forward model must again be enriched.

ACKNOWLEDGEMENTS

We have benefitted from numerous discussions with colleagues over recent years on the subjects of verification, validation, statistical inverse analysis, and uncertainty quantification. In particular, we thank Ivo Babuška, Omar Ghattas, Raúl Tempone, Ernesto Prudencio, and Robert Moser for helping us develop our current but still developing understanding of those subjects. We also gratefully acknowledge support of our work under the DOE Multiscale Mathematics Contract DE-FG02-05ER25701, the DOE PSAAP Contract DE-FC52-08NA28615, and the KAUST Contract US00003, all with ICES.

REFERENCES

1. Oden JT. A brief view of verification, validation, and uncertainty quantification. A presentation to the Board of Mathematical Sciences and their Applications, National Research Council, Washington, DC, October 30 2009.
2. Oden JT, Prudhomme S. Estimation of modeling error in computational mechanics. *Journal of Computational Physics* 2002; **182**:496–515.
3. Oden JT, Prudhomme S, Romkes A, Bauman PT. Multi-scale modeling of physical phenomena: adaptive control of models. *SIAM Journal on Scientific Computing* 2006; **28**(6):2359–2389.
4. Oden JT, Vemaganti K. Estimation of local modeling error and goal-oriented modeling of heterogeneous materials 1. Error estimates and adaptive algorithms. *Journal of Computational Physics* 2000; **164**:22–47.
5. Prudhomme S, Bauman PT, Oden JT. Error control for molecular statics problems. *International Journal for Multiscale Computational Engineering* 2006; **4**:647–662.
6. Vemaganti K, Oden JT. Estimation of local modeling error and goal-oriented modeling of heterogeneous materials; part ii: a computational environment for adaptive modeling of heterogeneous elastic solids. *Computer Methods in Applied Mechanics and Engineering* 2001; **190**:6089–6124.
7. Calvetti D, Somersalo E. *Introduction to Bayesian Scientific Computing*. Springer: New York, 2007.
8. Kaipo J, Somersalo E. *Statistical and Computational Inverse Problems*. Springer: New York, 2005.
9. Tarantola A. *Inverse Problems*: *Theory and Methods for Parameter Estimation*. SIAM: Philadelphia, 2004.
10. Prudencio E, Schulz K. The parallel C++ statistical library QUESO: quantification of uncertainty for estimation, simulation and optimization. *ICES Report*, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, 2010; in preparation.
11. Babuška I, Nobile F, Tempone R. A systematic approach to model validation based on Bayesian updates and prediction-related rejection criteria. *Computer Methods in Applied Mechanics and Engineering* 2008; **197**: 2517–2539.
12. Oden JT, Hawkins A, Prudhomme S. General diffuse-interface theories and an approach to predictive tumor growth modeling. *Mathematical Models and Methods in Applied Sciences* 2010; **20**(3):477–517.
13. Bauman PT, Oden JT, Prudhomme S. Adaptive multiscale modeling of polymeric materials with Arlequin coupling and goals algorithms. *Computer Methods in Applied Mechanics and Engineering* 2009; **198**:799–818.

DOI: 10.1002/nme