Comput. Methods Appl. Mech. Engrg. 198 (2009) 1327–1337

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

A better understanding of model updating strategies in validating engineering models

Ying Xiong [a], Wei Chen [a,*], Kwok-Leung Tsui [b], Daniel W. Apley [c]

[a] *Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, USA*
[b] *School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA*
[c] *Department of Industrial Engineering and Management Sciences, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA*

ARTICLE INFO

Article history:
Received 17 July 2008
Received in revised form 20 November 2008
Accepted 27 November 2008
Available online 24 December 2008

Keywords:
Model updating
Model validation
Bayesian approach
Maximum likelihood estimation
Thermal challenge problem

ABSTRACT

Our objective in this work is to provide a better understanding of the various model updating strategies that utilize mathematical means to update a computer model based on both physical and computer observations. We examine different model updating formulations, e.g. calibration and bias-correction, as well as different solution methods. Traditional approaches to calibration treat certain computer model parameters as fixed over the physical experiment, but unknown, and the objective is to infer values for the so-called calibration parameters that provide a better match between the physical and computer data. In many practical applications, however, certain computer model parameters vary from trial to trial over the physical experiment, in which case there is no single calibrated value for a parameter. We pay particular attention to this situation and develop a maximum likelihood estimation (MLE) approach for estimating the distributional properties of the randomly varying parameters which, in a sense, calibrates them to provide the best agreement between physical and computer observations. Furthermore, we employ the newly developed u-pooling method (by Ferson et al.) as a validation metric to assess the accuracy of an updated model over a region of interest. Using the benchmark thermal challenge problem as an example, we study several possible model updating formulations using the proposed methodology. The effectiveness of the various formulations is examined. The benefits and limitations of using the MLE method versus a Bayesian approach are presented. Our study also provides insights into the potential benefits and limitations of using model updating for improving the predictive capability of a model.

1. Introduction

Computer models have been widely used in engineering design and analysis to simulate complex physical phenomena. The accuracy or adequacy of a computer model can be assessed by means of model validation, which refers to the process of determining the degree to which a computational simulation is an accurate representation of the real world from the perspective of the intended uses of the model [1]. While there exists no unified approach to model validation, it is increasingly recognized that model validation is not merely a process of assessing the accuracy of a computer model, but should also help improve the model based on the validation results.

Strategies for model improvement roughly fall into two categories: *model refinement* and *model updating*. *Model refinement* involves changing the physical principles in modeling or using other means to build a more sophisticated model that better represents the physics of the problem by, for example, using a non-linear finite element method to replace a linear method, correcting and refining boundary conditions, or introducing microscale modeling in addition to macroscale modeling, etc. *Model updating*, on the other hand, utilizes mathematical means (e.g. calibrating model parameters and bias-correction) to match model predictions with the physical observations. While model refinement is desirable for fundamentally improving the predictive capability, the practical feasibility of refinement is often restricted by available knowledge and computing resources. In contrast, model updating is a cheaper means that can be practical and useful if done correctly. Here, *predictive capability* refers to the capability of making accurate predictions in domains (or locations) where no physical data are available.

While various model updating strategies (formulations and solution methods) exist, there is a lack of understanding of the effectiveness and efficiency of these methods. It is our interest in this work to examine various model updating strategies to achieve a better understanding of their merits. We are particularly interested in the role that model updating plays versus model validation and prediction. A detailed review is provided in Section 3. In summary, conventional calibration approaches [2] assume

* Corresponding author. Tel.: +1 847 491 7019; fax: +1 847 491 3915.
E-mail address: weichen@northwestern.edu (W. Chen).

0045-7825/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2008.11.023

Nomenclature

$\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$	n controllable input variables	$y^{m'}(\mathbf{x}, \Theta)$	updated model
$\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_m\}$	m uncontrollable input variables	$y^{pred}(\mathbf{x})$	predictive model
$y^e(\mathbf{x})$	physical experiments	$\boldsymbol{\Theta}$	model updating parameters
$y^m(\mathbf{x})$ or $y^m(\mathbf{x}, \boldsymbol{\theta})$	computer model	$L(\boldsymbol{\Theta})$	likelihood function
$\delta(\mathbf{x})$	bias function	$F_{\mathbf{x}_i^e}(y_i^e)$	cumulative distribution function (CDF) at y_i^e
ε	experimental error		

calibration parameters are fixed and estimated, typically using least squares to match the model with the physical observations. This type of approach for model updating is inconsistent with the primary concerns of model validation and prediction in which various uncertainties should be accounted for either explicitly or implicitly. Examples of such uncertainties include experimental error, lack of data, uncertain model parameters, and model uncertainty (systematic model inadequacy). The more recent Bayesian style calibration, also named calibration under uncertainty (CUU) or stochastic calibration, treats calibration parameters as unknown entities that are fixed over the course of the physical experiment. Initial lack of knowledge of the parameters is represented by assigning prior distributions to them, and, given the experimental data, this lack of knowledge is revised by updating their distributions (from priors to posteriors) based on the observed data through Bayesian analysis [3,4]. However, as we discuss in a more thorough examination in Section 3.2, several limitations of applying the Bayesian calibration approaches are identified in this work.

One limitation of the aforementioned calibration approaches is that the calibration parameters are assumed to remain fixed over the entire course of the physical experiment and beyond. In contrast, it is frequently the case that some parameters vary randomly over the physical experiment, perhaps due to manufacturing variation, variation in raw materials, variation in environmental or usage conditions, etc. This violates the assumptions under which the Bayesian or regression-based calibration analyses are derived. In this situation, rather than assuming fixed parameters and updating their posterior distributions to represent our lack of knowledge of them, it is more reasonable to treat them a randomly varying and estimate their distributional properties by integrating the physical data with the model. In essence, the distributional properties (e.g. the mean and variance of the randomly varying parameters) become the calibration parameters for the model, and the objective is to identify values for them that provide the best agreement with the observed distributional properties (e.g. the dispersion [5]) of the physical experimental data. In this paper, we present a maximum likelihood estimation (MLE) [6] approach for accomplishing this. The MLE method is used to estimate a set of unknown parameters (heretofore called model updating parameters) associated with several modeling updating formulations, which include the distributional properties of parameters that vary randomly over the experiment, as well as well as more traditional fixed calibration parameters and quantities associated with bias-correction and random experimental error.

The remainder of the paper is organized as follows. In Section 2, we discuss the role that model updating plays versus model validation and prediction. In Section 3, we examine the existing model updating formulations under two categories, namely, model bias-correction and calibration. The popular Bayesian approach is described and its limitations are highlighted. In Section 4, we describe our proposed MLE based model updating approach, together with the introduction of the u-pooling validation metric. In Section 5, a benchmark thermal challenge problem adopted by the Sandia Validation workshop [7,8] is used as an example to illustrate the proposed approach, draw important conclusions, and portray these conclusions in relation to conclusions from prior studies. Section 6 is the closure with a summary of the features of the proposed method, the relative merits of different approaches, the insights gained, and future research directions.

2. Role of model updating vs. model validation

In this work, model updating is viewed as a process that continuously improves a computer model through mathematical means based on the results from newly added physical experiments, until the updated model satisfies the validation requirement or the resource is exhausted. Therefore, even though model updating is interrelated with model validation, it is viewed as a separate activity that occurs before "validation". As shown in Fig. 1, the model updating procedure integrates the computer simulation model y^m with the physical experiment data $\mathbf{y}^e$ to yield an updated model $y^{m'}(\cdot)$. This updated model is then subject to a validation procedure that utilizes additional physical experiments $\mathbf{y}^e$ in the intended region of interest for validation. As noted from this diagram, unlike many contemporary model validation works, model validation in this work is used to evaluate an evolved, updated model $y^{m'}(\cdot)$, rather than the original computer model $y^m(\cdot)$. Besides, the updated model $y^{m'}(\cdot)$ is the one used for making future prediction $y^{pred}(\cdot)$ with the consideration of various sources of uncertainties. For implementing model updating and validation in a computationally efficient manner, it is indicated in Fig. 1 that a metamodel (surrounded by a dashed box) may be used to substitute the original computer model if it is expensive to compute.

As more details are provided in the remaining sections,*model updating* utilizes mathematical means (e.g. calibrating model parameters, bias-correction) to match model predictions with the physical observations. Model updating provides not only the formulation of an *updated model*, but also the characterization of model updating parameters $\boldsymbol{\Theta}$, together with the associated assumptions. As noted, the model updating procedure, during which $y^m(\cdot)$ is treated as a black-box, is largely driven by the ob-

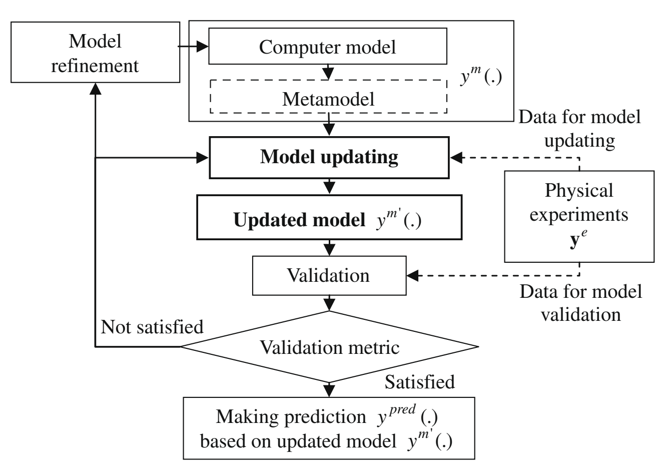

Fig. 1. Relationship of model updating, model refinement, and model validation.