Reliability Engineering and System Safety 129 (2014) 46–56

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

A Bayesian statistical method for quantifying model form uncertainty and two model combination methods

Inseok Park [a,*], Ramana V. Grandhi [b]

[a] Honda R&D Americas, Inc. Ohio Center, 21001 State Route 739, Raymond, OH 43067, USA
[b] Department of Mechanical and Materials Engineering, Wright State University, 3640 Colonel Glenn Hwy Dayton, OH 45435, USA

ARTICLE INFO

Article history:
Received 27 December 2010
Received in revised form
24 April 2014
Accepted 27 April 2014
Available online 5 May 2014

Keywords:
Model form uncertainty
Model probability
Maximum likelihood estimation
Bayesian inference
Model combination

ABSTRACT

Apart from parametric uncertainty, model form uncertainty as well as prediction error may be involved in the analysis of engineering system. Model form uncertainty, inherently existing in selecting the best approximation from a model set cannot be ignored, especially when the predictions by competing models show significant differences. In this research, a methodology based on maximum likelihood estimation is presented to quantify model form uncertainty using the measured differences of experimental and model outcomes, and is compared with a fully Bayesian estimation to demonstrate its effectiveness. While a method called the adjustment factor approach is utilized to propagate model form uncertainty alone into the prediction of a system response, a method called model averaging is utilized to incorporate both model form uncertainty and prediction error into it. A numerical problem of concrete creep is used to demonstrate the processes for quantifying model form uncertainty and implementing the adjustment factor approach and model averaging. Finally, the presented methodology is applied to characterize the engineering benefits of a laser peening process.

1. Introduction

Structural analysis of complex physical phenomena is becoming more dependent on computer simulation as nonlinear modeling methods advance. A simulation model can vary depending on the underlying physics and engineering and the manner in which a mathematical model is converted into a simulation model. This implies that we may have two or more different simulation models to analyze an identical engineering system. In statistics, a variety of model selection criteria have been developed from different perspectives with the aim to select the best model from a plausible regression model set such as Akaike Information Criterion (AIC) [1], Bayesian Information Criterion (BIC) [2], Mallows' Cp [3] and minimum description length [4]. Generally, a statistical model selection process is implemented with the concept that the best model balance goodness of fit with simplicity (usually measured by counting the number of regression parameters) better than the other considered models.

In general, uncertainty exists in the process of selecting the best model from a model set. Uncertainty involved in model selection, called *model form uncertainty* in this paper because of its existence in model form, is due to lack of confidence in selecting the best model. Ignoring model form uncertainty is problematic because it may lead to underestimating the variability of predictions or making erroneous predictions [5].

It has been argued that a simple way to account for model form uncertainty is by model combination which takes all the predictions by a model set into account (detailed in Section 3) [6]. Model combination produces predictions that incorporate the (epistemic) variation inherent in the statistical model selection process as well as the (aleatory and/or epistemic) variation conditional on each model. Model combination aims to predict unknown responses more reliably than each model in a set rather than better represent the physics of a real system or update predictions of each model given measured experimental data. The model selection criteria to select the best model from a plausible regression model set can be adapted to the task of model combination. Unlike the statistical regression models that fit observed experimental data, mathematical or simulation models used in the engineering field (specifically functional forms in the models) are fundamentally created based on scientific and engineering knowledge. Some of analytical models such as semi-empirical models have their functional forms derived from both theoretical knowledge and empirical data.

Recently, an increasing number of papers on the model form uncertainty quantification have been published in the engineering field while there exists plenty of literature on the big topic of V &

* Corresponding author. Tel: +1 937 877 8354.
E-mail address: IPark@oh.hra.com (I. Park).

http://dx.doi.org/10.1016/j.ress.2014.04.023
0951-8320/

V, which broadly covers the model form UQ [7,8]. Alvin et al. [9] used BMA to estimate the model form uncertainty in the frequency prediction of a component mounting bracket resulting from the use of three stochastic simulation models having the parametric uncertainty in elastic material modulus. The three simulation models had different levels of simplifying assumptions in their PDE forms, as well as different spatial meshes and different discrete solution variables. Model probability, which is assigned to each model to quantify model form uncertainty (the detailed description of model probability is given in Section 2.1), was simply assumed to be uniformly distributed over the considered simulation models. Zio and Apostolakis [10] used the *adjustment factor approach*—which is being described in Section 3.2—to estimate the model form uncertainty in the response predictions regarding the cumulative release of a radionuclide to water table given by six different models. The six models differ by some fundamental hypotheses on the groundwater flow and transport mechnisms. Model probabilities were evaluated based on expert opinions. Zhang and Mahadevan [11] estimated the failure probabilities for the butt welds of a steel bridge using two competing crack growth models (the Foreman and the Weertman crack growth models). They made a reliability analysis of fatigue life by averaging the estimated failure probabilities weighted by the probabilities of the crack models. The uncertainty in crack size measurement was quantified to evaluate model probabilities using Bayes' theorem.

Zouaoui and Wilson [12] used BMA to quantify the model form uncertainty in prediction of a system response (a message delay in a computer communication network). Although simulation models were used to predict a message delay, model probabilities were not assigned to the simulation models but to three different types of distributions that represent the uncertainty in an input variable (i.e. a message length). The distributions for a message length were assumed to be of exponential, normal and lognormal forms. McFarland and Bichon [13] used BMA to incorporate probability distribution model form uncertainty into the estimation of failure probability of a bistable MEMS device. As in the work of Zouaoui and Wilson, model probabilities were assigned to the three types of distributions (normal, lognormal and Weibull) that represent the uncertainty in an input variable (i.e. edge bias on beam widths).

In the above-mentioned research, evaluating model probability did not rely on a statistical analysis of the degree of agreement between the physically observed data on system responses and the simulated model predictions of the data. In recent years, attention of researchers also in the engineering field has increasingly been drawn to the statistical evaluation of the discrepancies between the test and the simulated data for the model probability quantification. To evaluate model probability using the measured differences of physically observed and simulated data in a practical and effective way, Park et al. [14] developed a statistical approach (detailed in Section 2) based on the fundamental idea behind BIC. The variance in prediction errors of each model is estimated using the Maximum Likelihood Estimation (MLE), and the best point estimate of the variance plays a critical role in the quantification of model probability. However, the composite prediction that only incorporates model form uncertainty was shown to underestimate the uncertainty in response predictions.

The present research accounts for both model form uncertainty and (unknown) prediction error involved in each model to obtain highly reliable prediction of a system response. The two types of uncertainties are merged into response prediction using a model combination technique called model averaging.

The MLE based method for quantifying model form uncertainty is presented with a fully Bayesian estimation for doing it in Section 2. Two model combination methods, model averaging and the adjustment factor approach, are described and compared in Section 3. In Section 4.1, the presented uncertainty quantification method and the model combination methods are illustrated with a numerical problem of concrete creep. Model form uncertainty and prediction errors associated with the FE analysis of a laser peening process are quantified and are incorporated into composite prediction in Section 4.2.

2. Quantification of model probability

2.1. Bayes' theorem for quantification of model probability

Model probability is defined as the degree of belief that a model is the best approximation among a set of possibilities; here, the best approximating model is defined as the model that predicts system responses of interest, usually unknown, more accurately than the other models in a model set. Consider a set of models denoted by M_1, M_2, …, M_K and experimental data D. Given experimental data D, Bayes' theorem presents a way to update prior probability of model M_k into posterior probability of M_k by

$$\Pr(M_k|D) = \frac{\Pr(M_k) \times L(M_k|D)}{\sum_{l=1}^{K} \Pr(M_l) \times L(M_l|D)}, k = 1, \ldots, K \tag{1}$$

$Pr(M_k)$ is prior probability of model M_k, the degree of belief that model M_k is the best approximation assessed prior to observing experimental data D. Zio and Apostolakis investigated the formal process of eliciting and interpreting expert judgments to quantify prior model probability $Pr(M_k)$ [10]. The quantification of $Pr(M_k)$ using a corpus of knowledge is arbitrary in nature because logically rigorous relations do not exist between the knowledge about a model set and prior model probability. Prior model probability $Pr(M_k)$ is often given a uniform value to avoid the difficulty of numerically specifying prior knowledge. $L(M_k|D)$ is called the likelihood of model M_k given experimental data D. $L(M_k|D)$ is deeply related to the goodness of fit of model M_k to experimental data D relative to the other models in the model set. Since uniform prior model probability is assumed for this research, the only concern is the evaluation of model likelihood $L(M_k|D)$. A methodology to evaluate model likelihood given experimental data is presented in Section 2.2.

2.2. Evaluation of model likelihood

2.2.1. Unknown prediction error

Because a (mathematical or simulation) model is just an approximation to a real physical system, it unavoidably involves an error in its prediction of a response; here prediction error is the difference between model prediction and experimental data. A prediction error is unknown unless the corresponding response is measured from the considered physical system. Probability distribution is generally used to describe an unknown prediction error; more specifically, (aleatory and epistemic) uncertainty in prediction error is mathematically characterized by a probability distribution. In statistics, prediction errors are usually assumed to be normally distributed with zero mean and a constant variance [15]. It is often stated that all variations in observed experimental data that cannot be explained by the considered model are included in the error term [16]. The statistical theory for dealing with random prediction error is well-understood and allows for constructing easily interpretable statistical intervals for predictions. However, this way of describing prediction error does not accommodate any framework to discern between the sources causing prediction error.