Reliability Engineering and System Safety 133 (2015) 79–86

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Evidence-based quantification of uncertainties induced via simulation-based modeling

Matthew E. Riley*

Department of Mechanical Engineering, Perimeter Dr. MS 0902, University of Idaho, Moscow, ID 83844-0902, United States

ARTICLE INFO

Article history:
Received 7 February 2014
Received in revised form
19 August 2014
Accepted 31 August 2014
Available online 6 September 2014

Keywords:
Model-form uncertainty
Epistemic uncertainty
Reliability

ABSTRACT

The quantification of uncertainties in simulation-based modeling traditionally focuses upon quantifying uncertainties in the parameters input into the model, referred to as parametric uncertainties. Often neglected in such an approach are the uncertainties induced by the modeling process itself. This deficiency is often due to a lack of information regarding the problem or the models considered, which could theoretically be reduced through the introduction of additional data. Because of the nature of this epistemic uncertainty, traditional probabilistic frameworks utilized for the quantification of uncertainties are not necessarily applicable to quantify the uncertainties induced in the modeling process itself. This work develops and utilizes a methodology – incorporating aspects of Dempster–Shafer Theory and Bayesian model averaging – to quantify uncertainties of all forms for simulation-based modeling problems. The approach expands upon classical parametric uncertainty approaches, allowing for the quantification of modeling-induced uncertainties as well, ultimately providing bounds on classical probability without the loss of epistemic generality. The approach is demonstrated on two different simulation-based modeling problems: the computation of the natural frequency of a simple two degree of freedom non-linear spring mass system and the calculation of the flutter velocity coefficient for the AGARD 445.6 wing given a subset of commercially available modeling choices.

1. Introduction

Simulation-based modeling refers to the general process of estimating a series of output responses of interest in a virtual environment by integrating methods and approaches across potentially numerous disciplines. Such models will generally take one of two forms: physics-based models—such as the classic finite element formulations or constituent physics models – or mathematical models – such as empirical data fits, surrogate models, or numerical approximations. Shared among all forms of computational models is their often inexact representation of the physical scenario that is being modeled. This inaccuracy arises from multiples sources such as an incomplete or imprecise representation of the underlying physics of the problem—as is often experienced in physics-based models – or numerical and round-off errors – as is often experienced in complex mathematical models. As a result of the assumptions made in the construction of the model, it is not uncommon for multiple models to produce conflicting estimates of an output response of interest given the same set of input parameters. In such a case, a decision is often made to select a single model among the set being considered that is thought to best describe the system [1]. However, in the early stages of design, there is often an incomplete model set, inducing an uncertainty in the selection of the best model among the model set being considered, referred to as model-form or model-selection uncertainty [2,3]. For a simulation based model to yield a complete estimation of a set of output responses, it is critical that all potential uncertainties, including those induced by the modeling process itself, be quantified and accounted for in the prediction of the output response of interest.

While the quantification of uncertainties in simulation-based design, including those epistemic in nature, has been explored in depth in the literature, most methods focus primarily upon the quantification of parametric uncertainties in the modeling process—inherent natural variations of either the input variables to the models or the model parameters themselves [4–7]. In such approaches, a representation of the uncertainty present in the model's estimation of the output response(s) of interest is often achieved by developing a representation of the uncertainty inherent to each of the parameters, be it probabilistic or non-probabilistic in nature, and then propagating such information through the simulation-based models in a rigorous manner. While such approaches have been shown to be quite effective and efficient for the quantification of uncertainties in the input parameters of the model, they often fail to completely address the potential inaccuracy of the models themselves. Such an inaccuracy

* Tel.: +1 208 885 9029.
E-mail address: riley@uidaho.edu

http://dx.doi.org/10.1016/j.ress.2014.08.016
0951-8320/

in the modeling process originates from one of two sources: the uncertainty associated with the selection of the most accurate model given multiple potential models—referred to as model-form uncertainty in this work but also referred to as model uncertainty in other works [1,8] – and the uncertainty associated with the discrepancy between the model of interest and the true physical value of interest – referred to as predictive uncertainty [9] but also referred to as model inadequacy [10] and model-form uncertainty [11] in other works. Such a deficiency can become critical, as it has been observed in complex problems that the variabilities induced by the modeling process itself, both model-form and predictive uncertainties, can become a significant, if not greater source of the overall uncertainty in the simulation-based modeling process than the parametric uncertainty alone [10].

While numerous approaches have been developed and demonstrated in the literature for the quantification of parametric uncertainties, these approaches are not necessarily directly viable or feasible for the quantification of model-form or predictive uncertainties. Such a deficiency is primarily due to the different forms and origins of the types of uncertainties, with parametric uncertainty often being intrinsic to the problem while model-form and predictive uncertainties are induced through the modeling process itself [4]. Additionally, quantification of predictive uncertainty will often require the introduction of additional information regarding the theoretical true scenario of interest, often in the form of experimental data [10]. Such data is often not abundant and readily available in early stages of design, making the quantification of predictive uncertainties difficult in early design. Instead, there is often an incomplete set of information available to the designer regarding both the constituent models within a given model set and the true physical scenario of interest [3]. Examples of this deficiency can include situations where multiple models are made available to the designer to represent the same given scenario, often run at a limited number of parameter sets or sparse experimental data sets. This incomplete set of information is not indicative of the lack of a correct model, as it is entirely possible that any of the given models being considered be the "truth model", but is instead due to the inavailability of such information at the particular stage of design. As this deficiency of knowledge could theoretically be reduced and ultimately resolved through the introduction of additional data regarding the physical scenario of interest – whether that be additional experimental data, more precise experimental data, or additional models – such uncertainty is epistemic in nature [4].

It has been established that the model-form uncertainty is to be treated as epistemic in nature, but it is common to encounter parametric uncertainty that is aleatory. In order to maintain epistemic generality with regards to the model-form uncertainty, nested loop approaches are often utilized to quantify the aleatory uncertainty within an inner loop and the epistemic uncertainty within an outer loop [12]. Prior work on nested aleatory/epistemic loops has focused upon the quantification of multiple types of parametric uncertainties, where some parameters are aleatory while others are epistemic in nature [13]. While a similar approach is fundamentally feasible to extend to model-form uncertainties, a framework must first be established to deal with the fundamental difference between parametric and model-form uncertainties.

This work proposes a methodology for quantifying the uncertainty in both the selection of the "best" model in the presence of multiple available models as well as the uncertainty inherent to the model itself by treating both uncertainties as epistemic in nature. Prior work in the quantification of modeling-induced uncertainties has often treated the uncertainties as probabilistic or probabilistic-like in nature, developing probability distributions for the models of interest and integrating them with classical model probabilities to develop a distribution of some output response(s) of interest [8,9,14,15]. One potential drawback of such an approach is that by representing model likelihoods in a probabilistic setting, there is often an accompanying assumption that the model set being considered is complete. However, such is often not the case in preliminary design stages where the available models and data is often only a small subset of the full knowledge base. Additionally, in a probabilistic framework, the quantification of such uncertainties is often hindered by the restrictions of the laws of classical probability, meaning that additional assumptions must be made regarding the problem to satisfy requirements of the probability theory and approaches utilized to quantify the uncertainties, not necessarily due to physical data driving the assumptions. By quantifying the modeling-induced uncertainties as purely epistemic in nature, they can be quantified in a non-probabilistic setting, freeing the designer of the restrictions that accompany classical probability theory as well as alleviating the implicit assumptions of a fully-realized data set.

2. Quantification of modeling-induced uncertainties

While methods exist in the literature exist for quantifying a subset of the sources of uncertainty, few exist for the quantification of uncertainty from all potential sources. This deficiency arises from the fact that there are often constraining assumptions that are made when quantifying uncertainty of a particular form that limit the capability for handling uncertainties of the other forms. These assumptions can include assumptions in the treatment of the different sources of uncertainty as aleatory or epistemic in nature or the availability of data and models. Riley and Grandhi summarized the current state of uncertainty quantification approaches that consider the contribution of uncertainties from all three sources as it applies to simulation-based models in the aerospace field [15].

A common feature among much of the prior work is the probabilistic representation of modeling induced uncertainties, even in the early stages of design. Due to the epistemic nature of modeling-induced uncertainties at this stage in design, though, the quantification of such uncertainties is often hindered by the restrictions of the laws of classical probability, meaning that additional assumptions must be made regarding the problem or the uncertainty to satisfy requirements of the algorithms and approaches utilize. Such assumptions can potentially either introduce an additional level of uncertainty into the problem or assign a false level of confidence in a particular result. One potential remedy for such deficiencies is to quantify the model-form and predictive uncertainties as purely epistemic in nature, eliminating the confining assumptions necessary for classical probability theory. Instead, only the current state of knowledge will be used to determine the uncertainty in the system.

The fundamental basis for using the current state of knowledge as the determining factor in system-level modeling-induced was established in Allaire and Willcox's work using a maximum entropy estimate to the represent the level of uncertainty in a multi-fidelity modeling problem [16]. The authors utilized a probabilistic representation of the uncertainty in the problem by modeling the information entropy, showing that the introduction additional data into the framework serves to reduce the overall level of information entropy, reducing the uncertainty associated with the prediction of the output response of interest. Park and Grandhi expanded upon this principal, utilizing Dempster–Shafer's Theory of Evidence to quantify the effects of model selection through interval analysis [2]. By utilizing the disjunctive rule of combination, parametric uncertainties that were epistemic in nature were able to be integrated with the uncertainty in the selection of the most accurate model among the model set being