Received 28 March 2016
Revised 22 November 2016
Accepted 11 April 2017

Model assessment in scientific computing

Considering robustness to uncertainty in input parameters

Saurabh Prabhu and Sez Atamturktur
Glenn Department of Civil Engineering, Clemson University, Clemson, South Carolina, USA, and

Scott Cogan
Department of Applied Mechanics, University of Franche-Comté, Besançon, France

Abstract

Purpose – This paper aims to focus on the assessment of the ability of computer models with imperfect functional forms and uncertain input parameters to represent reality.

Design/methodology/approach – In this assessment, both the agreement between a model's predictions and available experiments and the robustness of this agreement to uncertainty have been evaluated. The concept of satisfying boundaries to represent input parameter sets that yield model predictions with acceptable fidelity to observed experiments has been introduced.

Findings – Satisfying boundaries provide several useful indicators for model assessment, and when calculated for varying fidelity thresholds and input parameter uncertainties, reveal the trade-off between the robustness to uncertainty in model parameters, the threshold for satisfactory fidelity and the probability of satisfying the given fidelity threshold. Using a controlled case-study example, important modeling decisions such as acceptable level of uncertainty, fidelity requirements and resource allocation for additional experiments are shown.

Originality/value – Traditional methods of model assessment are solely based on fidelity to experiments, leading to a single parameter set that is considered fidelity-optimal, which essentially represents the values which yield the optimal compensation between various sources of errors and uncertainties. Rather than maximizing fidelity, this study advocates for basing model assessment on the model's ability to satisfy a required fidelity (or error tolerance). Evaluating the trade-off between error tolerance, parameter uncertainty and probability of satisfying this predefined error threshold provides us with a powerful tool for model assessment and resource allocation.

Keywords Uncertainty quantification, Model calibration, Model validation, Predictive modelling, Satisfying boundary, Scientific computing

Paper type Research paper

1. Introduction

Computer models must provide sufficient realism to be of use in aiding our understanding of and the having the ability to probe the reality of interest (Mankin *et al.*, 1977). Model prediction accuracy and precision are therefore necessary only to the extent that they contribute to the determination of the answers for the questions asked

The authors thank Parker Shields, Ismail Farajpour and Greg Roche for their contribution during the earlier stages of this work, as well as Godfrey Kimball for his editorial help.

Engineering Computations
Vol. 34 No. 5, 2017
pp. 1700-1723

0264-4401
DOI 10.1108/EC-03-2016-0109

of the model (Kiviat, 1967; Marks, 2007; Klein and Herskovitz, 2005). With this in mind, rather than searching for a single, optimal model[1] that best reproduces the experiments, we present an approach comparing the abilities of competing models in satisfying desired fidelity threshold levels.

Three aspects of a model are important in this comparison:

(1) The domain in which the problem is evaluated, defined by the *control parameters* that dictate the environmental or operational conditions of the system.

(2) The mathematical representation of the underlying processes, also referred to as *model form*, defined in accordance with the identified domain.

(3) The *input parameters* that characterize the properties of the system of interest, defined in accordance with the mathematical representations.

Control parameters define this applicability domain and are measurable and controllable in an experiment. In contrast, input parameters are variables that can be calibrated/updated within physically plausible bounds to match model predictions to observations. While control parameters have a physical meaning that is measurable, input parameters may or may not have a physical meaning. The domain of interest defined by control parameters is a critical step, as the model's internal structure must be representative of the behavior of the system within this domain (Oberkampf *et al.*, 2007; Atamturktur *et al.*, 2010). However, in this paper, we will assume that the problem domain is correctly defined and emphasize the model form and its associated parameter values (Aspects 2 and 3).

The second aspect declares that no model form is a perfect representation of reality, and the last aspect focuses on the parameterization of a given model form. Virtually all models concerning non-trivial, real-life systems that include input parameters with imprecise values elicit the following question: Should we use parameter values that suitably compensate for other forms of errors and uncertainties in the model (e.g. model form imperfectness, experimental uncertainty) or those that most accurately depict the real parameters? The former is typically what is achieved by calibrating model parameters against experiments (Atamturktur *et al.*, 2015a; Atamturktur *et al.*, 2015b); the latter is only possible if the parameter has a physical, quantifiable meaning (i.e. density of a material, geometric dimension, etc.) that can be measured. In the former case, the physical significance of the parameter is no longer retained if the parameter corrects for other sources of error (Link and Hanke, 1999) and the model is only suitable for replicating the observed data. Such a model updating where the real error source is not consistent with the assumed error source is termed *inconsistent model updating* (Govers and Link, 2010). The result is multiple models that yield the same fidelity to observations and with physically meaningless parameter values. This has prompted development of several updating procedures considering measurement errors (Mares *et al.*, 2006; Steenackers and Guillaume, 2006; Hua *et al.*, 2008; Khodaparast *et al.*, 2008; Govers and Link, 2010). In these studies, measurement error is considered in finite element model updating using modal parameters such as natural frequency and mode shapes. In such settings, measurement error can be quantified by repeated experiments. However, in many practical problems, the measurement error is not easily quantifiable owing to the cost constraints in performing repeated experiments. In problems involving naturally occurring phenomenon like earthquakes, the event repetition rate itself is very low. Moreover, quantification of modeling error is extremely difficult. The problem of inconsistency in updating procedure thus remains unavoidable in many practical applications.

Thus, the question remains how one should address this problem if a model's parameters have no physical meaning, or worse, if those with physical meanings are wrongly excluded during model idealization? As seen, many complications arise when inferring suitable values for imprecisely known input parameters of imperfect model forms using experimental measurements. However, no universally accepted approach exists to help in their determination.

In this paper, we argue that the selected mathematical representation (i.e. model form) should not only provide sufficiently accurate predictions of observable reality, but that it should do so given the uncertainty in its own parameter values (Deraemaeker *et al.*, 2004; Mares *et al.*, 2006). We define the capability of a model form in accommodating parametric uncertainty while satisfying a predefined fidelity threshold as a model's robustness (Gupta and Rosenhead, 1968; Rosenhead *et al.*, 1972; Bertsimas and Sim, 2004). A robust model predicts experimental outputs within acceptable fidelity bounds despite variations in its input parameter values (Stevens *et al.*, 2015). Based on the concept of robustness, we develop and illustrate indicators, as well as the trade-off between these indicators, to aid in the assessment of computer models.

This paper is organized as follows. Section 2 of this paper introduces the concept of satisfying boundaries for evaluating fidelity and robustness of model predictions. An optimization-based approach to obtain the satisfying boundaries for a given model is presented in Section 3. Section 4 describes a controlled academic case-study application, which is used later in Section 5 to identify the satisfying boundaries and discuss the trade-offs between the model's fidelity to experiments and its robustness to uncertainty considering alternate model forms. In Section 6, existing model assessment criteria are discussed and shown to be special cases of the satisfying boundary approach. Finally, Section 7 concludes with a discussion of the limitations of the proposed approach and the future direction for extension.

2. Concepts and methodological perspectives

Let us consider a model with uncertain parameters and assume that the model is a proper, uniformly continuous function, where an admissible, finite change in the input parameters yields a small, finite change in the model's output. Such functions are micro-continuous, where $x \approx y \Rightarrow f(x) \approx f(y)$. Hence, unstable systems which may respond to small perturbations in their inputs with disproportionately large changes in their output are left out of the scope of our discussion. The implication of confining our scope to proper uniformly micro-continuous functions is that for a compact space of input parameters, the model yields a compact (i.e. closed and bounded) output space (see Figure 1).

Within the compact output space, it is perhaps tempting to identify the model realization that is closest to experimental observations. This approach is often used to identify parameter values that yield the best agreement with available experiments (a process widely known as model calibration). However, calibrating the model input parameters to achieve such so-called *best-fidelity* solutions may yield misleading results, as experimental observations are neither collected with certainty nor available in sufficient quantity. Moreover, the agreement (or lack thereof) between the predictions and observations is often represented in an averaged, aggregated manner and the input parameters to be calibrated tend to compensate for each other, leading to multiple equally plausible inverse solutions. Finally, the discrepancy bias in model predictions is often non-negligible, resulting in parameters being calibrated to incorrect values to account for the model bias. In practical applications, these aspects unavoidably allow multiple combinations of input parameter values to yield predictions of similar agreement with experiments. Hence, fidelity is an

incomplete and potentially misleading criterion for model assessment. Rather than selecting the single best solution, a more viable approach is to evaluate all alternative solutions that satisfy a minimum fidelity threshold (March and Simon, 1958; Simon, 1959). Following this idea, we identify the model predictions, which satisfy a certain *error tolerance* when compared with experimental observations. The measure of model adequacy for its intended use is hence reflected in this error tolerance (see Mayes, 2009 for a discussion on determining such error tolerance criteria).

We introduce the concept of a *satisfying boundary* that encompasses all model input parameter realizations that satisfy this prescribed error tolerance. This boundary exists as an n-dimensional hypervolume where n is the number of model input parameters to be evaluated. The volume encompassed by this boundary is reminiscent of Starr's domain criterion (Starr, 1966; Schneller and Sphicas, 1983). The satisfying boundary can be derived using an appropriate optimization algorithm to search the output space for the hypersurface at which the prediction error is equal to the error tolerance. The shape and size of the satisfying boundary solely depend upon the model form and available experiments, meaning that it requires no knowledge of the distributions of the uncertain input parameters (Figure 1). This is important because there are many engineering and science problems in which the probabilistic knowledge of parameters involved is either incomplete or entirely unavailable. For well-posed problems, among two alternative model forms with the same set of uncertain parameters, the model with a larger satisfying boundary, i.e. larger volume of the n-dimensional hypervolume, should be preferred, as it can accommodate higher levels of uncertainty while meeting the error tolerance requirement[2].

2.1 Probability of satisfying

To assess the model's capability to reproduce experiments and tolerate uncertainties, we consider two independent sources of information:

(1) a satisfying boundary that is intrinsic to the model form; and

(2) the parametric uncertainty that is intrinsic to the model inputs.

For a given level of uncertainty in input parameters, a model's ability to reproduce experiments can be quantified as the probability of the model predicting within the acceptable error tolerance. Accepting the principle of indifference[3], the probability of satisfying can be estimated as the ratio of the number of parameter combinations within the

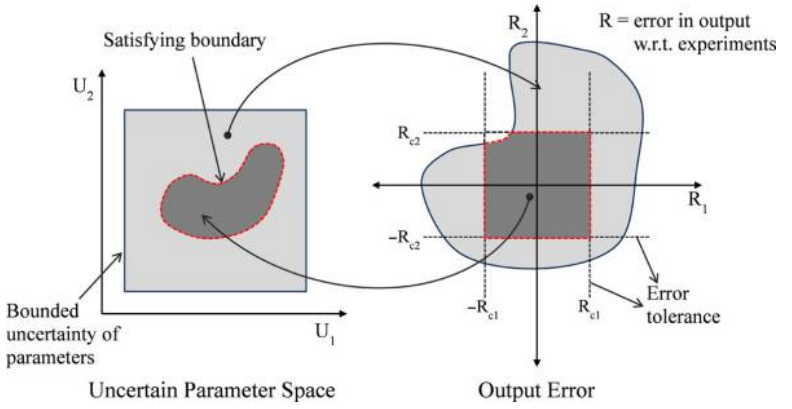


Figure 1. Conceptual figure showing the mapping from uncertain parameter space to output space and the compact satisfying boundary

satisfying boundary to the total number of parameter combinations sampled with the given uncertainty distributions:

$$P_s = \Pr\left(g(\theta) \le b^*\right) = \Pr\left(\theta \notin F_t\right) = \int_{\Theta} I_f(\theta) q(\theta) d\theta \cong \frac{1}{N} \sum_{j=1}^{N} I_f(\theta_j) \quad (1)$$

where, $g(\theta)$ is the performance function and b^* is the critical threshold. $I_f(\theta)$ is a binary indicator which is equal to 1 if $g(\theta) \le b^*$ and 0 otherwise. For high dimensional problems, Monte Carlo simulation (MCS) can be used, in which case, P_s can be estimated as:

$$P_s \cong \frac{1}{N} \sum_{i=1}^{N} I_f(\theta) \quad (2)$$

where N is the number of random samples. The knowledge of probability of satisfying helps determine the influence of parameter uncertainty on the model's outputs of interest. For instance, for a given error tolerance, if 100 per cent of the parameter value sets within the bounded uncertainty space are effectively contained within the satisfying boundary, parameter uncertainty can be deemed inconsequential, as the model predictions satisfy the error tolerance requirement regardless of lack of precise values for the parameters. Thus, the model form is deemed robust and can make reliable predictions. However, when this is not the case, we can resort to quantifying the probability that the model output will satisfy the error tolerance given the uncertainty in its parameters. For the desired error tolerance, if the probability of satisfying is inordinately low, it may indicate a problem with either the model form or the uncertainty distribution assumed on the parameter values. To rectify this, we may invest in developing a better model form or better define our parameters. Such a case is illustrated in Figure 2, where the probability of satisfying the error tolerance is 0 per cent. As the uncertain parameter space expands, some model instances may begin to satisfy the error tolerance, leading to a counter-intuitive increase in the probability of satisfying with increasing parameter uncertainty. Such behavior can point to flaws in either the model form or the distribution of the parameter values.

Typically, parametric uncertainties can be reduced by additional data collection or further analysis. However, it is first necessary to verify that a reduction in parameter

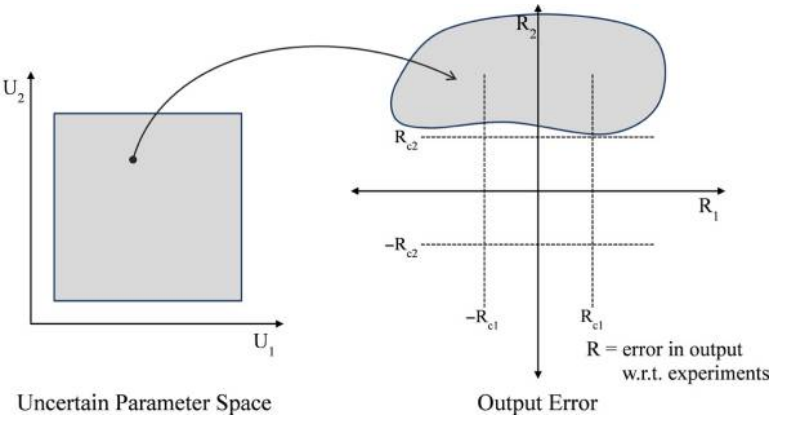


Figure 2.
Conceptual figure showing the absence of a satisfying boundary as output errors fall outside the error tolerance

uncertainty will indeed improve the model's ability to reproduce experiments (i.e. in our definition, the probability of satisfying a prescribed error tolerance). Hence, in our evaluation, rather than requiring that a single uncertainty bound be specified for an input parameter, an array of differing levels of uncertainty bounds is studied to observe how a model's probability of satisfying the fidelity tolerance is affected. Such a trade-off analysis reveals the value of the efforts to reduce uncertainty.

2.2 Error tolerance

Until now, we have assumed that a normative definition for an error tolerance is known and such a tolerance may not be well established; hence, the model developer may first need to determine what tolerance is acceptable. However, determining an error tolerance (i.e. the maximum allowable prediction error: How good is good enough?) in turn obscures the differences in the behaviors of alternative models for errors less than this predefined tolerance. As an example, what would it mean if a model shows the largest robustness for an error tolerance of 10 per cent, but which is the least robust for a smaller error tolerance of 9 per cent? This is shown in Figure 3, where Models 1 and 2 appear to provide the same satisfying boundary for the given error tolerance. However, the robustness of Model 1 is superior to Model 2 for error levels that are less than the predefined tolerance. Hence, it is meaningful to evaluate the satisfying boundary for varying levels of error tolerance. A satisfying boundary, by definition, monotonically increases in size as the error tolerance becomes less stringent and encompasses more input parameter sets (refer Figure 4). Given the dependency among the three criteria, i.e. the required error tolerance, the parameter uncertainty and the model's probability of satisfying, we propose that model assessment must be based on an exploration of the trade-off between these criteria.

Note that in Figure 3, the gray band represents the bounded experimental uncertainty, which may originate from a variety of sources such as measurement noise, unit-to-unit variability and operator-to-operator variability. This experimental uncertainty can be reflected in the analysis by defining the error tolerance with respect to the uncertainty bounds (instead of the mean) or by making the error tolerance itself uncertain. The latter option will be discussed in Section 6.3.

3. Toward the derivation of satisfying boundary

The satisfying boundary can be constructed by sampling the parameter space and generating hypervolumes (in n-dimensional input space) encompassing instances that satisfy the error tolerance. Once the boundary is defined, low-cost sampling techniques can be adapted to determine the ratio of the parameter samples that fall within the

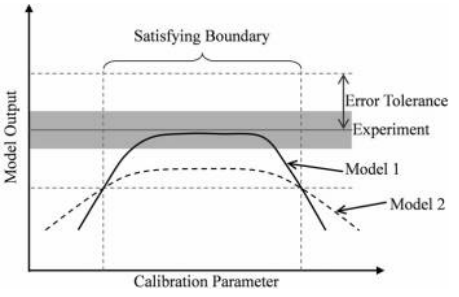


Figure 3.
Illustration of an inability to distinguish between multiple models within the same satisfying boundary

satisfying boundary to the total number of samples drawn with the uncertainty distributions of the input parameters, i.e. the probability of satisfying. It is then possible, for example, to use the *inpolygon* and *inhull* (D'Ericco, 2006) functions in MATLAB to determine if a point sampled in the input parameter domain lies inside or outside the boundary.

Let us now consider predicted response $y_p = g(\theta)$, where $g(\theta)$ is the performance function and $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ is the uncertain random input variable representing the input parameters. Assume that $g(\theta)$ is proper continuous in the n-dimensional parameter space, which means that if the input variables are a compact set, the response is also a compact set.

Here, the analyst is assumed to have prior knowledge (i.e. the bounds for plausible values of the parameters) regarding the uncertainty bounds for θ_i: $(i = 1, 2, \ldots, n)$; however, the probability distributions of these parameters are assumed to be unknown at this point. The model error R can be defined as the deviation of these model predictions from experimental observations y_o, which can also be assumed to be a function of the input parameters as given in equation 2:

$$R = \|y_o - y_p\| = f(\theta); \qquad \theta = (\theta_1, \theta_2, \ldots, \theta_n) \tag{3}$$

where R represents the norm of the error between the model predictions y_p and the mean of experimental observations y_o. In equation 3, $\|\cdot\|$ indicates a suitable metric (such as a Euclidean or Bhattacharya distance). Under the given error tolerance R_t, the worst-case input parameters $\hat{\theta}$ satisfy the following conditions:

$$|f(\hat{\theta}) - R_t| \leq \varepsilon \tag{4}$$

where ε is a tolerance used as the optimization stopping criteria.

The set of these input parameters $\hat{\theta}$ forms the satisfying boundary for a given error tolerance. If the performance function is computationally expensive, polynomial interpolation between the MCS samples can be used to estimate the satisfying boundary. Alternatively, the expensive performance function can be replaced by a surrogate model, allowing the use of optimization techniques that rely on intensive sampling. Assuming that either the performance function is cheaply calculated or a surrogate model is developed, a simplified random search algorithm for obtaining the satisfying boundary is presented below.

By defining an objective function in the following form and minimizing the z-value, the points on the satisfying boundary can be determined using the formula below:

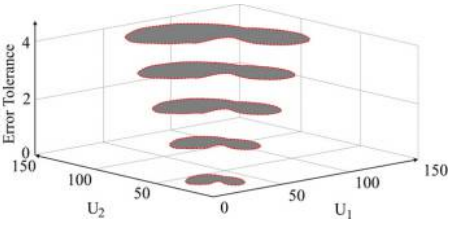


Figure 4.
A schematic representation of the satisfying boundary monotonically increasing with increasing error tolerance in the predictions (i.e. decreasing fidelity)

$$min\ z = min\,|f(\hat{\theta}) - R_t| \tag{5}$$

For defining the satisfying boundary, no prior knowledge of the distribution of parameters is required; however, the bounds $max(\theta_i)$ and $min(\theta_i)$ must be defined on the parameters to confine the search space. These bounds can correspond to the extreme values of the input parameters that define the problem domain. Most optimization approaches require an initial point for the algorithm. Here we used the center of the input parameter domain defined by the uncertainty bounds as the starting point. Subsequent points are chosen via a suitable optimization algorithm that searches for an optimal point around the previously selected point. The number of points I that define the satisfying boundary should be decided based upon the computational demand of the model, but must be sufficiently high to accurately delineate the boundary. Because our model was fast running, we set the number of points to 2,000 based on trial and error. For a computationally expensive model, a stopping criterion can be based on the change in the area of the polygon encompassed by the boundary on additional iterations. For higher-dimensional problems, the stopping criteria can be based on the hypervolume calculation, although the added cost of this calculation may not be trivial. Several efficient algorithms have been previously developed for calculating compact multi-dimensional volumes (see for instance, Schneller and Sphicas, 1983; Eiselt and Langley, 1990; Vetschera, 1997; While *et al.*, 2006; Bradstreet *et al.*, 2008). A flow chart describing the algorithm's step-wise process is shown in Figure 5.

The general process of applying the satisfying boundaries approach to model assessment is as follows:

(1) Generate N samples of input parameters $\boldsymbol{\theta}_j = (\theta_{1,j},\ \theta_{2,j}, \ldots,\ \theta_{n,j})$; $(j = 1, 2, \ldots, N)$ from uniform distribution between upper and lower bounds prescribed on each parameter.

(2) Identify satisfying boundaries corresponding to increasing levels of error tolerance R_t.

- 2.1. Compute $y_{p,j} = g(\boldsymbol{\theta}_j)$ for each of the N sample points.
- 2.2. Compute error $R_j = \|y_o - y_{p,j}\|$.

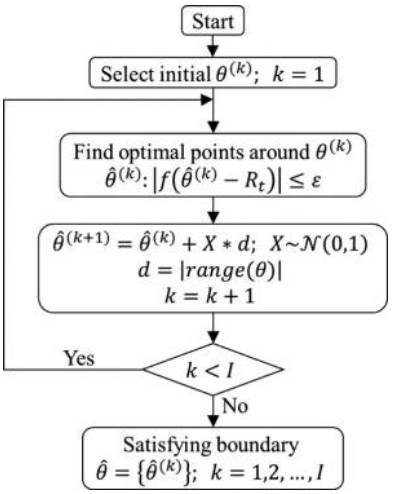


Figure 5. Flowchart of optimization algorithm used to define the satisfying boundary in the input parameter domain

- 2.3. Estimate boundary points in the input space $\hat{\theta}^{(k)}$: $\left|R_t - \left|y_o - g(\hat{\theta}^{(k)})\right|\right| \leq \varepsilon$; $k = 1, \ldots, I$.
- 2.4. Let S be the volume enclosed by $\hat{\theta} = \left\{\hat{\theta}^{(1)}, \hat{\theta}^{(2)}, \ldots, \hat{\theta}^{(I)}\right\}$.

(3) Generate $M \gg N$ samples of input parameter $\boldsymbol{\theta}_j = (\theta_{1,j}, \theta_{2,j}, \ldots, \theta_{n,j})$; ($j = 1, 2, \ldots, M$) from their respective prescribed distributions.

(4) Estimate probability of satisfying $P_s = \frac{1}{M}\sum_{j=1}^{M} I_S(\theta_j)$, where $I_S(\theta_j) = 1$ if $\theta_j \in S$, and 0 otherwise.

Several problem-specific decisions must be incorporated to improve efficiency of the process. The boundary points can be estimated by interpolation on the initial sample set in Step 2.3 or by using a search-based optimization algorithm. Interpolation requires sufficiently dense sampling for accurate estimation of boundary points. When the performance function is highly computationally intensive, it may not be realistic to achieve sufficiently dense samples to justify interpolation. Neither will a search-based algorithm be feasible given the required number of function evaluations. In this case, an expensive performance function can be replaced by an inexpensive surrogate model in Step 1. Surrogate modeling techniques such as response surface methodology (Box and Draper, 1987; Khuri and Mukhopadhyay, 2010) or Gaussian process regression (aka Kriging) (Neal, 1997; Rasmussen, 2006; Forrester *et al.*, 2007; Kleijnen, 2009) are widely used for both linear and non-linear functions (Barton and Meckesheimer, 2006; Wang and Shan, 2007 for review of surrogate modeling techniques). The choice of uncertain input parameters, output variables and error tolerance requires some model exploration, which is guided by Step 1. Sensitivity analysis (Homma and Saltelli, 1996; Saltelli *et al.*, 2000) allows estimation of fractional contribution of uncertainty in input parameters to output variations.

4. Proof-of-concept demonstration: steel moment resisting frame

4.1 Description of frame structure

The concepts introduced in the previous sections are demonstrated on the 2-D steel frame shown in Figure 6. In this academic example, the frame is constructed with vertical columns that rest on fixed supports with beams that are semi-rigidly connected to the columns at both levels. In steel frame structures, the variability of the material properties, member sizes and construction practices results in highly uncertain connection stiffness values (Sakurai *et al.*, 2001; Díaz *et al.*, 2012). Hence, in our proof-of-concept example, the connection stiffness of the two first story joints is represented using linear rotational springs with uncertain and independent stiffness constants. All members of the portal frame are assigned uniform dimensions and material properties which are provided in Table I. Static, horizontal loads are applied to the portal frame as shown, with the members oriented to bend about their strong axes.

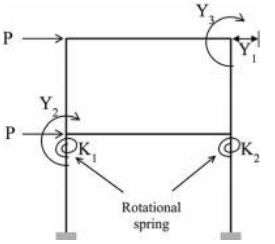


Figure 6.
Single-bay, two-story portal frame with rotational springs at the top of the first story columns

4.2 Synthetic experiments and satisfying boundary

There are several methods for developing the mathematical representation for engineering systems: a variety of simplifying assumptions may be established or idealizations may occur (Judge and McIntosh, 2000; Hossein and Hossein, 2011) which collectively lead to multiple competing model forms. In this section, we show how the probability of satisfying the error tolerance can be used to assess and compare three alternative model forms with varying levels of model imperfectness (i.e. prediction bias).

The exact model, constructed with the Timoshenko beam theory (Cowper, 1966; Hutchinson, 2001), is used to synthesize the experimental data. The data consist of both translation (Y_1) and rotation (Y_2) response under the loading conditions shown in Figure 6. The exact model not only accurately accounts for the effects of axial, shear and flexural deformations but also uses the so-called true values of K_1 and K_2 of 10 kip-in/rad (1.13 KN-m/rad) (Table I). The corresponding exact Y_1 and Y_2 are 0.095 in and 0.54×10^{-3} rad, respectively, for an applied load $P = 1$ kip (4.45 KN).

An example satisfying boundary is shown in Figure 7 for an error tolerance of 2.5 per cent in the outputs Y_1 and Y_2, identified using the optimization algorithm described in Section 3. Because the frame model is computationally inexpensive, sampling is also a plausible option for this example. In Figure 7, 40,000 instances of (K_1, K_2) are sampled within their uncertainty bounds, which in this case, is set to be between 5 and 15 kip-in/rad. Contours in the input space that join extremities of the sampled point clouds agree well with the satisfying boundary obtained using the previously discussed optimization algorithm.

Property description	Beams and columns
All member lengths (in)	72 in (1.83 m)
Cross-sectional area (in^2)	4.44 (28.4 cm^2)
Moment of inertia (in^4)	48.0 (1998 cm^4)
Young's modulus (ksi)	29,000 (200 MPa)

Table I.
Input values for the portal frame

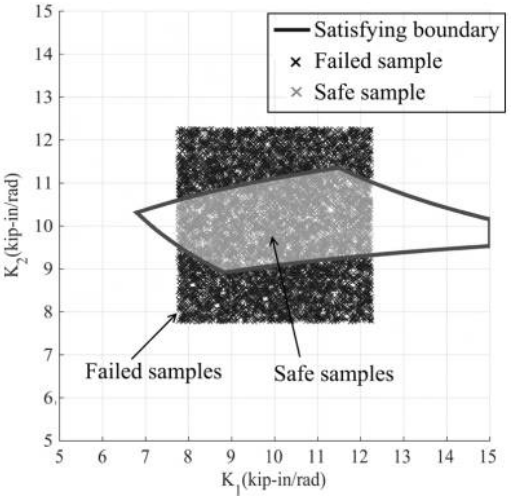


Figure 7.
The satisfying boundary for parameters Y_1 and Y_2 in the input parameter space for an error tolerance of 2.5 per cent

4.3 Competing model forms

Alongside the exact model, two inexact (biased) model forms are evaluated:

(1) one that underestimates; and

(2) one that overestimates the shear area by 25 per cent, thus offering a decision-maker three alternative options.

Hence, these two inexact models inaccurately account for the shear deformations, while K_1 and K_2 will be uncertain in all three models.

5. Satisfying boundaries for model assessment: steel moment resisting frame

5.1 Exact model with uncertain input parameters

For combinations of K_1 and K_2 within their predefined bounds, the model predictions are compared to the experimental data to calculate the percentage prediction error in the two outputs Y_1 and Y_2. The corresponding prediction errors R_1 and R_2 are the percentage differences with respect to the experiments. Figure 8 illustrates the relationship between the error in the model output and the two uncertain parameters.

Subsequently, satisfying boundaries are generated for varying error tolerances, R_t, shown in Figure 9. R_t varies from 0 to 5 per cent prediction error in steps of 0.5 per cent. In Figure 9, each contour corresponds to an error tolerance level R_t such that all instances of K_1 and K_2 that lie within the contour satisfy that error tolerance. As expected of a continuous system, the satisfying boundaries are nested with their size increasing as error tolerance increases. Similarly, the size of the boundaries decreases and approaches a point as the error tolerance approaches zero [Figure 9(c)]. The model form used in the development of this figure was *exact*, which is why the true parameter values are encompassed within the satisfying boundaries. If compensations are in effect on the satisfying boundaries, we would obtain a curve rather than a single point. This ill-conditioned behavior can be remedied by combining outputs as seen in Figure 9(c), where the outputs Y_1 and Y_2 are combined to generate satisfying boundaries such that the boundaries encompass parameter combinations that satisfy the error tolerance in both output parameters. It is now observed that as the error tolerance approaches 0 per cent, the satisfying boundary converges to a point (instead of a curve), which in this case corresponds to the exact parameter values.

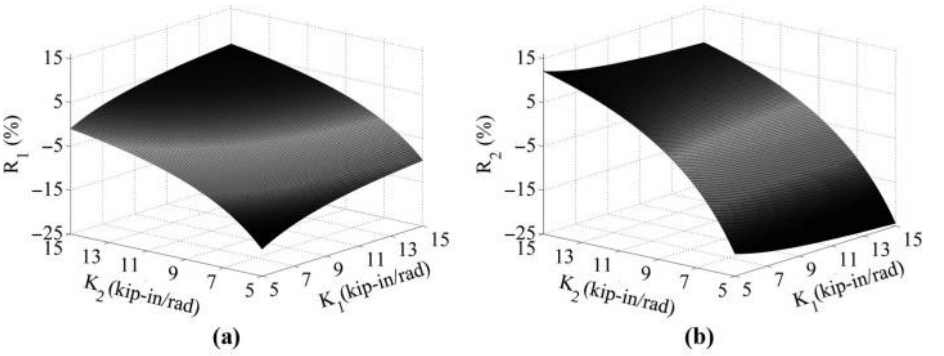


Figure 8.
3-D representation of prediction errors

Notes: (a) R_1; and (b) R_2 in the exact model

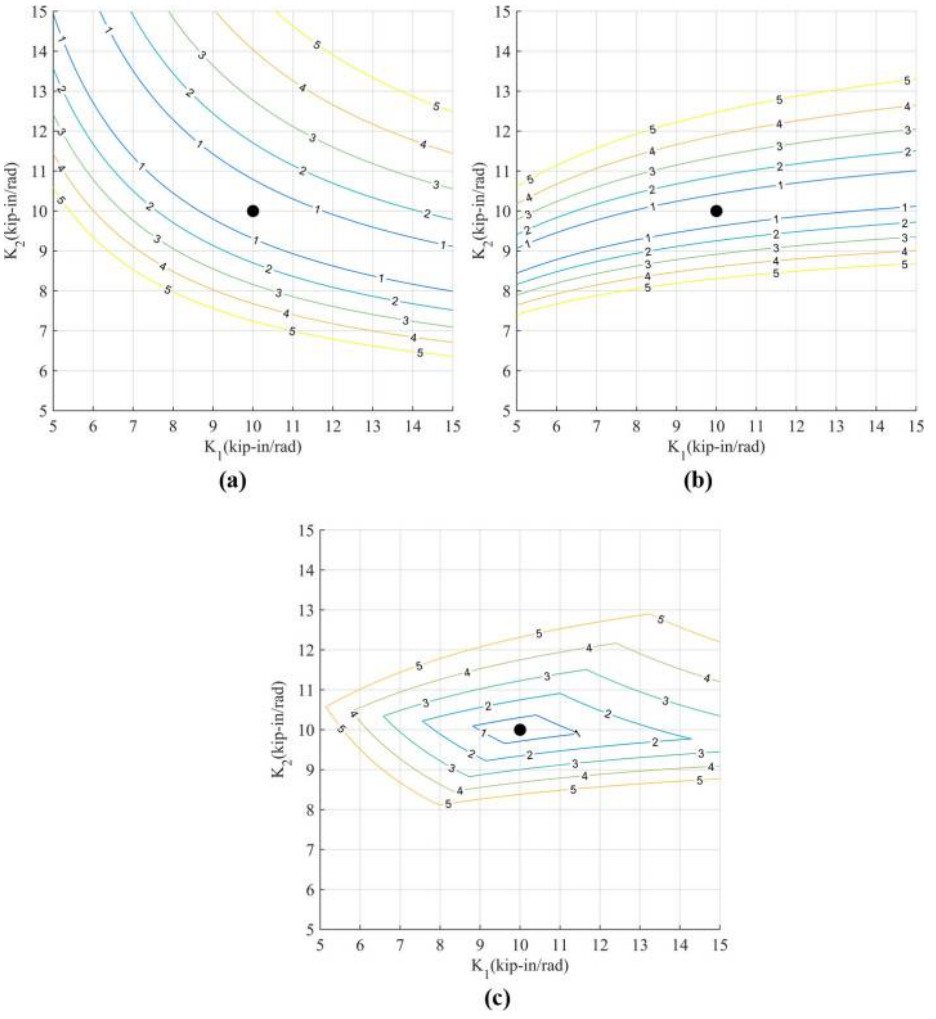


Figure 9.
Nested sets of satisfying boundaries for increasing levels of error tolerance for the exact frame model for (a) output Y_1, (b) Y_2 and (c) Y_1 and Y_2 combined. The black dot represents the location of the "true" parameter values (those that were used to generate the synthetic experiments)

5.2 Inexact models with uncertain input parameters

The two inexact finite element models underestimate and overestimate the shear area of the beam and column elements (i.e. 75 and 125 per cent shear areas, respectively), resulting in biased model predictions. The satisfying boundaries for the two inexact models are shown in Figures 10 and 11.

As seen in Figures 10 and 11, the misrepresentation of shear deformation causes the satisfying boundaries to shift such that the true parameter values represented by the black dot (K_1 = K_2 = 10 kip-in/rad) are no longer encompassed by the satisfying boundary that corresponds to small error tolerances, reflecting the inherent bias in model predictions.

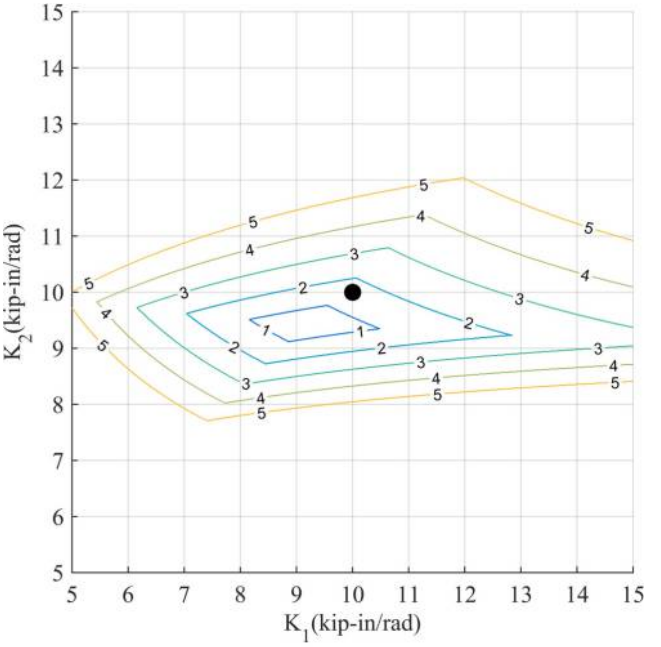


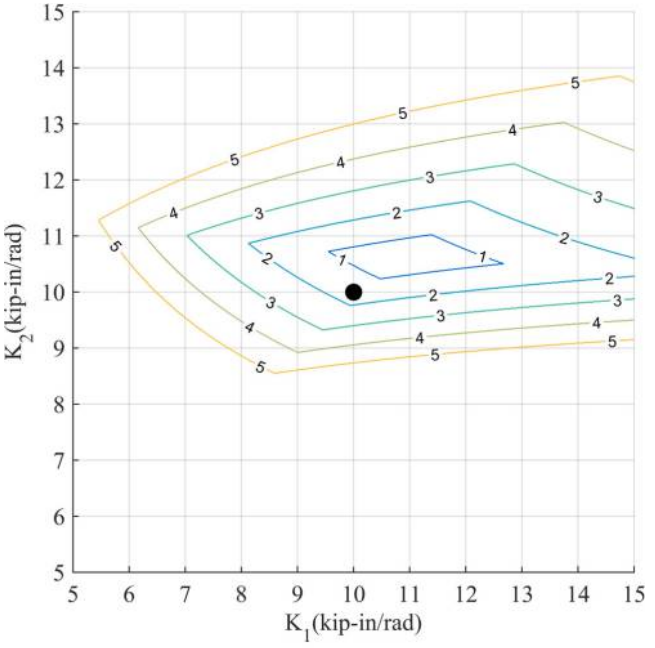
Figure 10.
Nested sets of satisfying boundaries for increasing levels of error tolerance for the biased frame model (25 per cent lower shear area) for outputs Y_1 and Y_2 combined

Note: The black dot represents the location of the "true" parameter values (those that were used while generating synthetic experiments)

5.3 Utilizing the satisfying boundaries

The satisfying boundaries for the three competing models (one exact and two inexact) discussed in the previous section can be used to evaluate the probability of satisfying predefined error tolerances within predefined uncertainty distributions of the two parameters. In this evaluation, bounded uniform uncertainty is assumed for the two parameters and is allowed to gradually increase from 0.5 to 5 kip-in/rad in steps of 0.5 kip-in/rad. An uncertainty of 5 kip-in/rad means that the parameter value can vary between 7.5 and 12.5 kip-in/rad. This evaluation is also repeated for increasing levels of error tolerance from 0 to 5 per cent, in steps of 0.5 per cent. Figure 12 displays the relationship between the varying levels of error tolerance in model predictions, the parameter uncertainty and the subsequent probability that the model satisfies this predefined error tolerance. Because two outputs, namely, the translation at the top story (Y_1) and the rotation at the first story (Y_2) of the steel frame, are considered, the joint probability of satisfying the error tolerance in both outputs is calculated (see Figure 12).

The 3-D plot in Figure 12 can be used as a diagnostic tool for the assessment and comparison of multiple computer models. It can be observed from Figure 12 that as the error tolerance increases so does the probability of satisfying the error tolerance. The rate of increase depends upon the model form. The light-shaded region in Figure 12 represents the unsuitability of the model for its intended use (as defined by the error tolerance) given the available knowledge (as defined by the parametric uncertainty). On the other hand, the darker-shaded regions in Figure 12 represent the situation where the model is a better fit for



Note: The black dot represents the location of the "true" parameter values (those that were used while generating synthetic experiments)

Figure 11.
Nested sets of satisfying boundaries for increasing levels of error tolerance for the biased frame model (25 per cent higher shear area) for outputs Y_1 and Y_2 combined

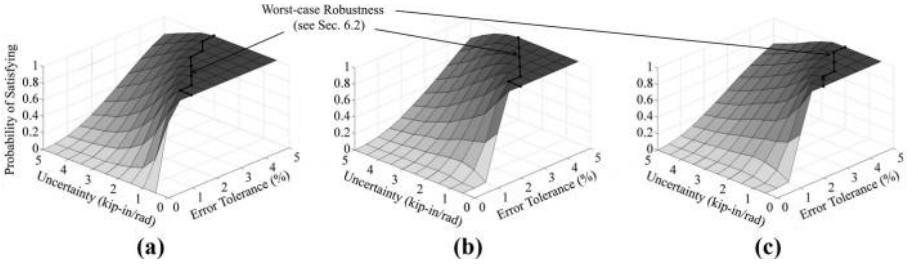


Figure 12.
Three-dimensional plot showing trade-off between probability of satisfying, error tolerance and parametric uncertainty for the (a) accurate model, (b) inaccurate model with 25 per cent underestimated shear area and (c) inaccurate model with 25 per cent overestimated shear area

its intended use. This region represents the lower values of the experimental uncertainty and the higher values of the error tolerance. The larger dark regions in the model with the exact form [Figure 12(a)] indicate a generally higher probability of satisfying at higher uncertainty and lower error tolerance compared to the two inexact models [Figure 12(b) and (c)].

The relationship between the probability of satisfying the error tolerance and the error tolerance itself for 30 per cent uncertainty in both K_1 and K_2 is shown in Figure 13. As can be seen, for an error tolerance of 0 per cent, the corresponding probability of satisfying this tolerance is 0 per cent, meaning that no model form can accommodate the given level of uncertainty and satisfy the required error tolerance. Only by increasing the error tolerance does the probability of satisfying increase. It can also be observed in Figure 13 that the exact model consistently yields a higher probability of satisfying the error tolerance compared to the two inexact models. This is of course because our uncertainty bounds for parameter values were centered on the so-called true values. Furthermore, the exact model's probability of satisfying the error tolerance increases more rapidly with error tolerance (higher slope) than the two imperfect models.

The probability of satisfying the error tolerance of 3 per cent as a function of uncertainty in the input parameters is plotted in Figure 14. The model developer may use this plot to observe a potential improvement in the probability of satisfying the desired error tolerance by reducing the uncertainty in the input parameters. For instance, if the developer of the exact model wants to ensure at least 90 per cent probability of satisfying the 3 per cent error tolerance, then resources must be allocated to ensure that the uncertainty in the input parameters is lower than 2 kip-in/rad. Also note in Figure 14 that for uncertainty levels of 0.5kip-in/rad, all models yield predictions that are 100 per cent within the error tolerance. Hence, from this figure, we can deduce that allocating resources for reducing uncertainty below 0.5 kip-in/rad is not justifiable. These increased levels of parameter uncertainty however lead to a reduction in the probability of satisfying, as expected, during which the role of bias once again becomes important. For very high levels of parameter uncertainty, all three models converge to unacceptably low probabilities of satisfying the error tolerance. As seen, Figure 14, similar to Figure 13, can be used as a comparative tool and aid in model selection. For instance, a model developer may establish a minimum probability of satisfying requirement and subsequently evaluate which model performs best given the varying degrees of parameter uncertainty.

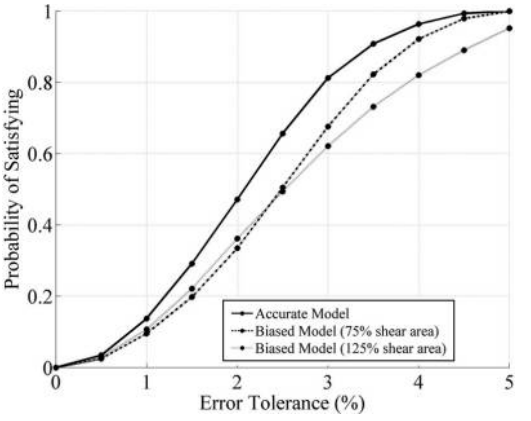
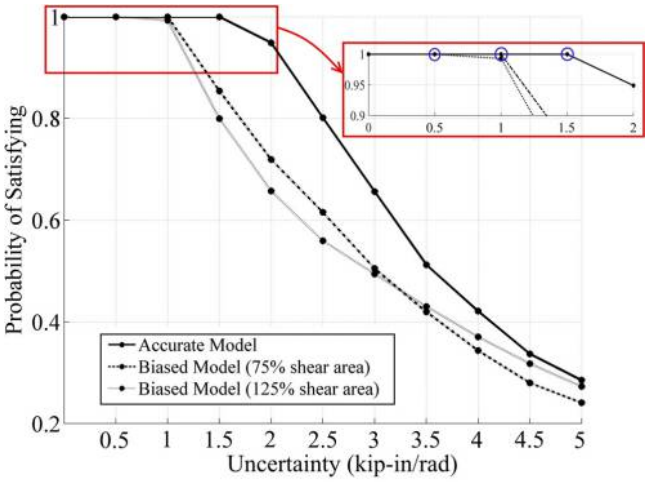


Figure 13.
Probability of satisfying as a function of error tolerance for a 3 kip-in/rad uncertainty (30 per cent uniform uncertainty) in K_1 and K_2

It is also possible to evaluate the relationship between the fidelity of model predictions and parameter uncertainty for a given probability of satisfying (shown for 80 per cent in Figure 15). The two inexact models are inadmissible when the parameter uncertainty is less than 1.8 kip-in/rad. This uncertainty is caused by the offset of the biased model's satisfying boundaries (recall Figures 10 and 11), resulting in the parameter spaces corresponding to low uncertainty falling entirely outside these satisfying boundaries. This concept, demonstrated earlier in Figure 2, supplies a means for diagnosing the fundamental flaws in either the form of this model or the values associated with the parameters of these model forms.



Note: The encircled markers show worst-case robustness

Figure 14.
Probability of satisfying as a function of parameter uncertainty for a constant error tolerance ($R_t = 3$ per cent)

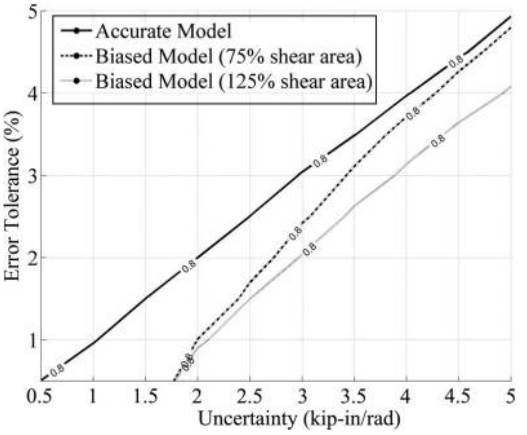


Figure 15.
Prediction error as a function of parameter uncertainty for a constant probability of satisfying ($P_s = 80$ per cent)

6. Model selection criteria based on satisfying boundaries

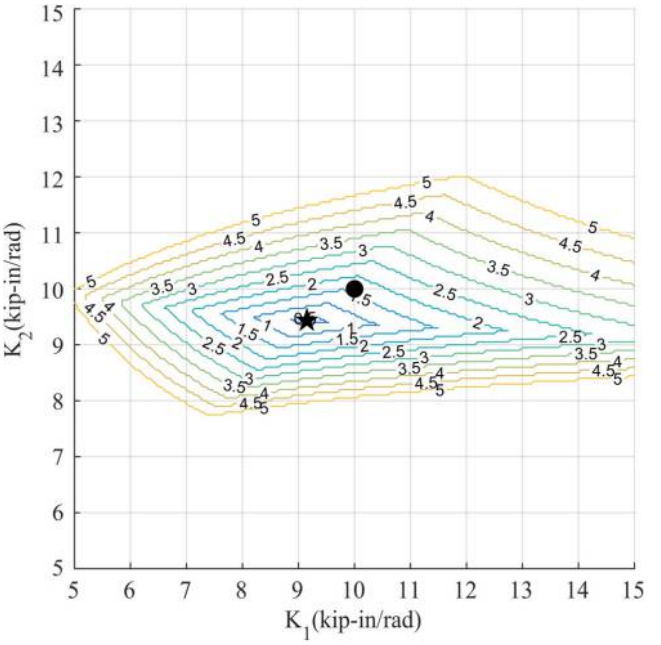
The concept of satisfying boundaries readily presents useful model selection criteria, specifically the optimal deterministic model, info-gap robustness, model distinguishability and value of information. In this section, we will discuss other, lower-dimensional model selection criteria that can be deduced from the three-dimensional plots presented in Section 5.

6.1 Deterministically optimized model

The special case of R_t = 0 per cent corresponds to the optimal deterministic design. A preference of a single model over another based only on its performance at a point is generally not recommended. The satisfying boundaries for varying error tolerances, as shown Figure 16, are specified on both outputs Y_1 and Y_2 for the biased frame model (25 per cent lower shear area). In this figure, the star represents the deterministic optimal parameter values which differ from the supposed true parameter values used while generating the synthetic experiments. The compensation of the deterministic optimal parameter values causes this difference in the model form error, masking the model's deficiency and making it appear as if it is in good agreement with experiments.

6.2 Worst-case robustness analysis

A special case of P_s = 1 reflects the maximum level of uncertainty that can be tolerated while still ensuring that the error tolerance is not exceeded. This is similar to the previously



Notes: The black dot represents the location of the exact parameter values, whereas the star represents the deterministic optimal parameters

Figure 16.
Nested sets of satisfying boundaries for increasing levels of error tolerance for the biased frame model (25 per cent lower shear area)

established Info-gap robustness (Ben-Haim, 2006), which requires the definition of a nominal parameter value for calculating the maximum allowable deviation from the nominal value while still satisfying a predefined error threshold. Worst-case robustness $\hat{\alpha}$ can be formulated as an optimization problem as follows:

For a given error threshold R_t:

$$\begin{aligned} &\text{maximize } \alpha \\ &S.T. \quad |\max\{R(U)\} - R_t| \leq \varepsilon; U \epsilon \Upsilon(\mathrm{U}_0, \alpha) \end{aligned} \tag{6}$$

where $\Upsilon(\mathrm{U}_0, \alpha)$ is the uncertainty distribution centered on an assumed nominal parameter set U_0 and α is the parameter uncertainty. This worst-case robustness is highlighted in Figure 14 for different error tolerance levels. Also recall Figure 14 where the maximum uncertainty (X-axis) corresponding to $P_s = 1$ reflects the worst-case robustness of the three competing models. For an error tolerance of 3 per cent, the worst-case robustness for the accurate model and the two biased models is 1.5, 1 and 0.5. Schematically speaking, the worst-case robustness for a model with uniform uncertainty distribution in all parameters is the half of the edge-length of the largest bounded uncertainty envelop that may be placed within the satisfying boundary as conceptualized for a two-parameter model in Figure 17.

6.3 Model distinguishability

Model distinguishability (Walter *et al.*, 1984), a fundamental characteristic of input–output spaces, quantifies the extent to which models can be ranked based on their fidelity-to-data alone given the presence of measurement errors. High model distinguishability indicates that model behaviors are sufficiently distinct from one another and that regions of fidelity-equivalent solutions are relatively small. Low distinguishability indicates that there are large regions in the parameter space with nearly equivalent output errors, indicating the insufficiency of the fidelity-to-data in selecting a model from a set of indistinguishable models (recall the phenomena of non-uniqueness discussed earlier in the introduction).

Model distinguishability presents a means for incorporating the experimental uncertainty into the evaluation. If α_e is the measurement noise, the indistinguishable models

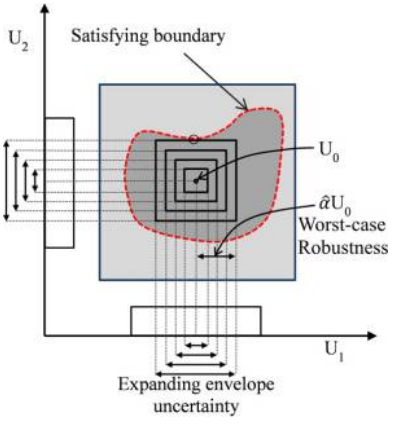


Figure 17.
Conceptual figure showing the Info-gap robustness $\hat{\alpha}$ with respect to the satisfying boundary in the input space

are a set of all models that satisfy the specified error tolerance R_t within a tolerance of α_e. The problem can be stated as:

For a given error tolerance R_t,

Find a set of input parameters U_{id}:

$$S.T. \qquad |R(U_{id}) - R_t| \leq \alpha_e \tag{7}$$

The notion of model distinguishability, illustrated in Figure 18, is based on the nested sets of satisfying boundaries for the biased frame model (25 per cent lower shear area) for the mean output error $R(U) = \frac{1}{2}(R_1 + R_2)$. The colored zones correspond to sets of indistinguishable models for measurement noise of 0.5 per cent and output errors of 0.5, 1.5 and 2.5 per cent, etc. Models with these zones are fidelity-equivalent and cannot otherwise be ranked without additional information (e.g. parameter constraints or new experiments).

6.4 Value of added information

Adding more system responses to assess the predictive ability of a model entails additional experiments or additional sensors or instruments. Investigating the changes in the satisfying boundary resulting from the inclusion or exclusion of system responses provides a means for quantifying the value of information added by including new experimentally measured response features, which can guide the need and resource allocation for additional experimentation or instrumentation.

The *global* satisfying boundary, the intersection of the individual satisfying boundaries of each output, as shown in Figure 19, is affected as new response features are added. In Figure 19(a), an error tolerance of $R_t = 2.5$ per cent is specified only on the output Y_1 (refer

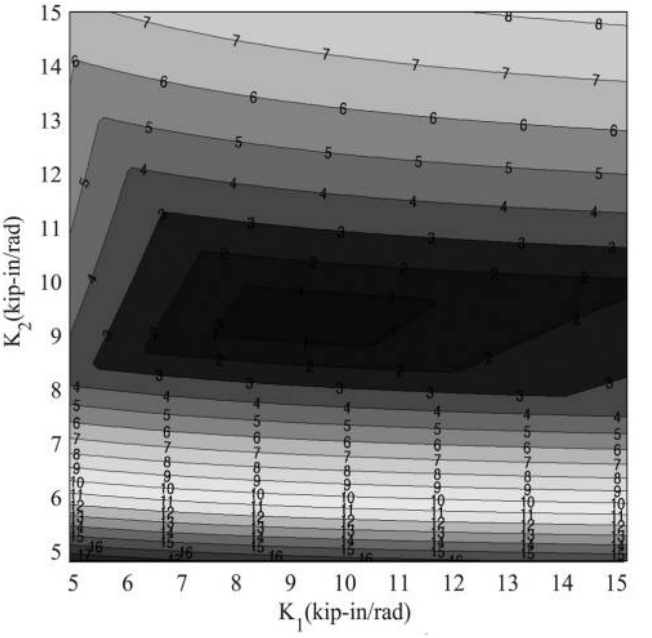


Figure 18.
Model distinguishability for the biased frame model (25 per cent lower shear area) assuming measurement noise level of 1 per cent and the mean output error $R = \frac{1}{2}(R_1 + R_2) = 0.5\%,\ 1.5\%,\ 2.5\%,$ etc.

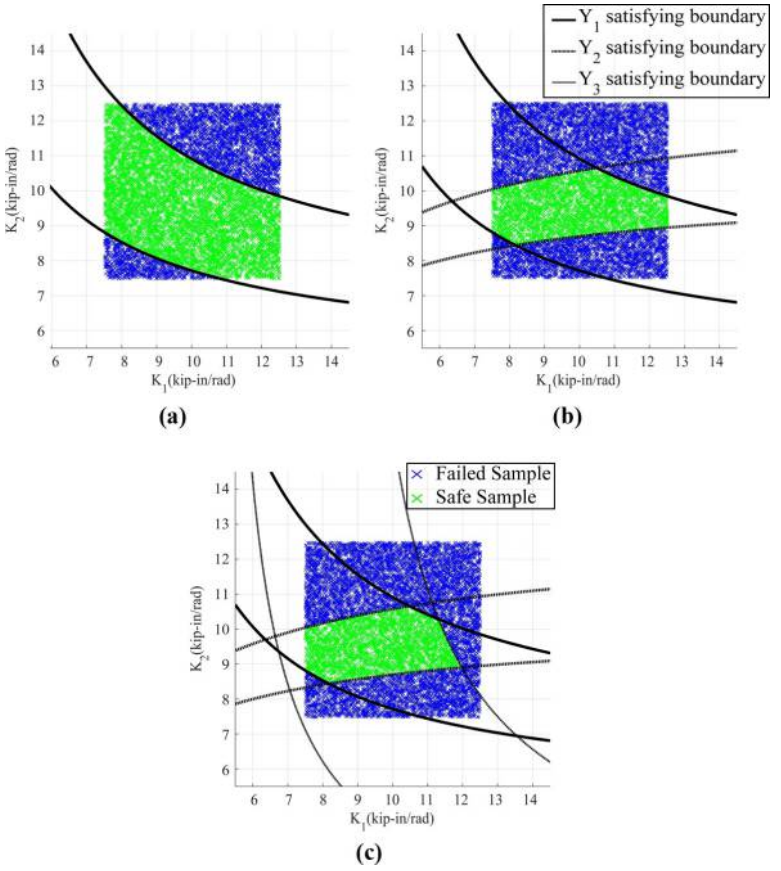


Figure 19.
Satisfying boundaries (R_t = 2.5 per cent) of three model outputs and their intersection which is the safe region

Figure 6). In Figure 19(b) and (c), satisfying boundaries corresponding to outputs Y_2 and Y_3 are added with the same error tolerance. The combined satisfying boundary is the intersection of the individual satisfying boundaries of each output, leading to lesser instances of satisfaction for (K_1, K_2), effectively further confining the satisfying boundary and reinforcing the uniqueness of the solution with the addition of more outputs. Also, the additional system response may be relatively insensitive to uncertainty in the input parameters, in which case, the additional information will not have a drastic (if any) effect on the size of the satisfying boundary. In other situations, satisfying boundaries of multiple outputs failing to intersect may reveal an inherent deficiency in the predictive ability of the model.

7. Conclusion

In numerical modeling, uncertainties occur due to imprecisely known input parameter values just as biases arise from our imperfect understanding of the underlying physics. Here a method has been presented to enhance the comparison of alternative model forms to answer questions posed given the availability of information regarding their parameter values. This evaluation is based on the trade-off between three criteria: the desired fidelity of

model predictions to experimental observations, the robustness which quantifies the ability of the model to maintain fidelity under uncertainty in its input parameters and lastly the probability of satisfying fidelity criteria.

In this study, we evaluated the case with bounded uncertainty in model input parameters represented as nested sets. The effect of growing uncertainty on the satisfying boundary is evaluated to quantify the trade-off in the three criteria. If better characterization of parameter distributions is available, it may be incorporated into estimating the probability of satisfying. Note, however, that construction of the satisfying boundary itself does not require any distribution to be specified on the input parameters.

For a given level of uncertainty, only one of these two criteria must be known or defined, from which the third can then be determined. Hence, knowledge of the trade-offs between these two criteria can provide a decision-maker with useful insight in selecting the most useful or appropriate values based upon the intended application of the model. Also, four model selection criteria closely related to the notion of a satisfying boundary, specifically the deterministic optimal solution, the robust optimal solution, model distinguishability and the value of added experimental outcomes, were also used to enhance the visibility of the important characteristics of the design space.

The primary impediment of the proposed approach to model evaluation might be the requirement of random sampling which can limit feasibility in application to many practical applications with computationally costly performance function evaluations. Interpolation and surrogate modeling techniques are suggested under such conditions. Application to a full-scale structure, which entail higher complexity (e.g. higher number of uncertain parameters, experimental uncertainty) will be investigated in future studies in the hopes of determining the implications of such complications on satisfying boundaries. The authors continue to investigate the expansion of the approach to other types of engineering problems by improving the mathematical derivation of the satisfying boundaries.

Notes

1. In this context, a model is defined collectively by the prediction domain, model form and input parameters.
2. However, if the problem is ill-posed (i.e. the number of unknown inputs is greater than the number of known outputs), then the obtained satisfying boundary would contain all possible non-unique solutions, rendering it essentially meaningless.
3. Starr (1966) interpreted the principle of indifference to mean an infinite set of any possible probability distribution to be equally likely. In our implementation, it would mean any possible bounded probability distribution that sum to one. In case the parameter distribution is not known deterministically, one can compare the average probability of satisfying of competing models considering all possible parameter distributions.

References

Atamturktur, S., Hemez, F. and Unal, C. (2010), *Calibration under Uncertainty for Finite Element Models of Masonry Monuments*, (No. LA-14414), Los Alamos National Laboratory (LANL), Los Alamos, NM.

Atamturktur, S., Hegenderfer, J., Williams, B., Egeberg, M., Ricardo, L. and Unal, C. (2015a), "A resource allocation framework for experiment-based validation of numerical models", *Journal of Mechanics of Advanced Materials and Structures*, Vol. 22 No. 8, pp. 641-654.

Atamturktur, S., Liu, Z., Cogan, S. and Juang, C.H. (2015b), "Calibration of imprecise and inaccurate numerical models considering fidelity and robustness: a multi-objective optimization based approach", *Structural and Multi-disciplinary Optimization*, Vol. 51 No. 3, pp. 659-671.

Barton, R.R. and Meckesheimer, M. (2006), "Metamodel-based simulation optimization", *Handbooks in Operations Research and Management Science*, Elsevier, New York, Vol. 13, pp. 535-574.

Ben-Haim, Y. (2006), *Info-Gap Decision Theory: Decisions under Severe Uncertainty*. Academic Press, Massachusetts.

Bertsimas, D. and Sim, M. (2004), "The price of robustness", *Operations Research*, Vol. 52 No. 1, pp. 35-53.

Box, G.E. and Draper, N.R. (1987), *Empirical Model-Building and Response Surfaces*, Wiley, New York, NY, Vol. 424.

Bradstreet, L., While, L. and Barone, L. (2008), "A fast incremental hypervolume algorithm", *IEEE Transactions on Evolutionary Computation*, Vol. 12 No. 6, pp. 714-723.

Cowper, G.R. (1966), "The shear coefficient in Timoshenko's beam theory", *Journal of Applied Mechanics*, Vol. 33 No. 2, pp. 335-340.

Deraemaeker, A., Ladevèze, P. and Romeuf, T. (2004), "Model validation in the presence of uncertain experimental data", *Engineering Computations*, Vol. 21 No. 8, pp. 808-833.

Díaz, C., Victoria, M., Querin, O.M. and Martí, P. (2012), "Optimum design of semi-rigid connections using metamodels", *Journal of Constructional Steel Research*, Vol. 78, pp. 97-106.

D'Ericco, J. (2006), Inhull, available at url: www.mathworks.com/matlabcentral/fileexchange/10226-inhull (accessed 27 February 2016).

Eiselt, H.A. and Langley, A. (1990), "Some extensions of domain criteria in decision making under uncertainty", *Decision Sciences*, Vol. 21 No. 1, pp. 138-153.

Forrester, A.I., Sóbester, A. and Keane, A.J. (2007), "Multi-fidelity optimization via surrogate modeling", *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *The Royal Society*, *London*, Vol. 463 No. 2088, pp. 3251-3269.

Govers, Y. and Link, M. (2010), "Stochastic model updating – covariance matrix adjustment from uncertain experimental modal data", *Mechanical Systems and Signal Processing*, Vol. 24 No. 3, pp. 696-706.

Gupta, S.K. and Rosenhead, J. (1968), "Robustness in sequential investment decisions", *Management Science*, Vol. 15 No. 2, pp. B18-B29.

Homma, T. and Saltelli, A. (1996), "Importance measures in global sensitivity analysis of nonlinear models", *Reliability Engineering & System Safety*, Vol. 52 No. 1, pp. 1-17.

Hossein, A. and Hossein, A. (2011), "A robust data mining approach for formulation of geotechnical engineering systems", *Engineering Computations*, Vol. 28 No. 3, pp. 242-274.

Hua, X.G., Ni, Y.Q., Chen, Z.Q. and Ko, J.M. (2008), "An improved perturbation method for stochastic finite element model updating", *International Journal for Numerical Methods in Engineering*, Vol. 73 No. 13, pp. 1845-1864.

Hutchinson, J.R. (2001), "Shear coefficients for Timoshenko beam theory", *Transactions-American Society of Mechanical Engineers Journal of Applied Mechanics*, Vol. 68 No. 1, pp. 87-92.

Judge, P.G. and McIntosh, S.W. (2000), "Non-uniqueness of atmospheric modeling", *Physics of the Solar Corona and Transition Region*, Springer, Netherlands, pp. 331-350.

Khodaparast, H.H., Mottershead, J.E. and Friswell, M.I. (2008), "Perturbation methods for the estimation of parameter variability in stochastic model updating", *Mechanical Systems and Signal Processing*, Vol. 22 No. 8, pp. 1751-1773.

Khuri, A.I. and Mukhopadhyay, S. (2010), "Response surface methodology", *Wiley Interdisciplinary Reviews: Computational Statistics*, Vol. 2 No. 2, pp. 128-149.

Kiviat, P.J. (1967), *Digital Computer Simulation: Modeling Concepts*, (No. RM-5378-PR), Rand Corporation, California.

Kleijnen, J.P. (2009), "Kriging metamodeling in simulation: a review", *European Journal of Operational Research*, Vol. 192 No. 3, pp. 707-716.

Klein, E.E. and Herskovitz, P.J. (2005), "Philosophical foundations of computer simulation validation", *Simulation & Gaming*, Vol. 36 No. 3, pp. 303-329.

Link, M. and Hanke, G. (1999), "Model quality assessment and model updating", in *Modal Analysis and Testing*, Springer, Netherlands, pp. 305-324.

Mankin, J.B., O'Neill, R.V., Shugart, H.H. and Rust, B.W. (1977), "The importance of validation in ecosystem analysis", *New Directions in the Analysis of Ecological Systems*, Vol. 1, pp. 63-71.

Mares, C., Mottershead, J.E. and Friswell, M.I. (2006), "Stochastic model updating: part 1 – theory and simulated example", *Mechanical Systems and Signal Processing*, Vol. 20 No. 7, pp. 1674-1695.

March, J.G. and Simon, H.A. (1958), *Organizations*, Wiley, New york, NY.

Marks, R.E. (2007), "Validating simulation models: a general framework and four applied examples", *Computational Economics*, Vol. 30 No. 3, pp. 265-290.

Mayes, R.L. (2009), "Developing adequacy criterion for model validation based on requirements", *Proceedings of the IMAC-XXVII, Orlando, FL*, Springer, New York.

Neal, R.M. (1997), "Monte Carlo implementation of Gaussian process models for Bayesian regression and classification", Technical Report No. 9702, Department of Statistics, University of Toronto, Toronto, ON, available at: https://arxiv.org/pdf/physics/9701026.pdf

Oberkampf, W.L., Pilch, M. and Trucano, T.G. (2007), *Predictive Capability Maturity Model for Computational Modeling and Simulation*, No. SAND2007-5948, Sandia National Laboratories, Albuquerque, NM.

Rasmussen, C.E. (2006), *Gaussian Processes for Machine Learning*, MIT Press, Cambridge.

Rosenhead, J., Elton, M. and Gupta, S.K. (1972), "Robustness and optimality as criteria for strategic decisions", *Operational Research Quarterly*, Vol. 23 No. 4, pp. 413-431.

Saltelli, A., Chan, K. and Scott, E.M. (Eds), (2000), *Sensitivity Analysis*, Wiley, New York, NY, Vol. 1.

Sakurai, S., Ellingwood, B.R. and Kushiyama, S. (2001), "Probabilistic study of the behavior of steel frames with partially restrained connections", *Engineering Structures*, Vol. 23 No. 11, pp. 1410-1417.

Schneller, G.O. and Sphicas, G.P. (1983), "Decision making under uncertainty: starr's domain criterion", *Theory and Decision*, Vol. 15 No. 4, pp. 321-336.

Simon, H.A. (1959), "Theories of decision-making in economics and behavioral science", *The American Economic Review*, Vol. 49 No. 3, pp. 253-283.

Starr, M.K. (1966), "A discussion of some normative criteria for decision-making under uncertainty", *Industrial Management Review*, Vol. 8 No. 1, pp. 71.

Steenackers, G. and Guillaume, P. (2006), "Finite element model updating taking into account the uncertainty on the modal parameters estimates", *Journal of Sound and Vibration*, Vol. 296 No. 4, pp. 919-934.

Stevens, G., Van Buren, K., Wheeler, E. and Atamturktur, S. (2015), "Evaluating the fidelity and robustness of calibrated numerical model predictions: an application on a wind turbine blade", *Engineering Computations*, Vol. 32 No. 3, pp. 621-642.

Vetschera, R. (1997), "A recursive algorithm for volume-based sensitivity analysis of linear decision models", *Computers & Operations Research*, Vol. 24 No. 5, pp. 477-491.

Walter, E., Lecourtier, Y. and Happel, J. (1984), "On the structural output distinguishability of parametric models, and its relations with structural identifiability", *IEEE Transactions on Automatic Control*, Vol. 29 No. 1, pp. 56-57.

Wang, G.G. and Shan, S. (2007), "Review of metamodeling techniques in support of engineering design optimization", *Journal of Mechanical Design*, Vol. 129 No. 4, pp. 370-380.

While, L., Hingston, P., Barone, L. and Huband, S. (2006), "A faster algorithm for calculating hypervolume", *IEEE Transactions on Evolutionary Computation*, Vol. 10 No. 1, pp. 29-38.

Corresponding author
Sez Atamturktur can be contacted at: sez@clemson.edu