Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 311 (2008) 1371–1390
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

A decision-theoretic method for surrogate model selection

R.V. Field Jr.*

Applied Mechanics Development, Sandia National Laboratories, Albuquerque, NM 87185-0346, USA

Received 14 February 2007; received in revised form 24 September 2007; accepted 12 October 2007
Available online 26 November 2007

Abstract

The use of surrogate models to approximate computationally expensive simulation models, e.g., large comprehensive finite element models, is widespread. Typical uses of surrogate models include design, optimization, sensitivity analysis and/or uncertainty quantification. A surrogate model is defined by a postulated functional form, and values for the surrogate model parameters are estimated using results from a limited number of solutions to the comprehensive model. In general, there may be multiple surrogate models, each defined by possibly a different functional form, consistent with the limited data from the comprehensive model. We refer to each as a candidate surrogate model. Methods are developed and applied to select the optimal surrogate model from the collection of candidate surrogate models. One approach is to select the surrogate model that best fits the data provided by the comprehensive model, regardless of its intended use. The proposed approach applies techniques from decision theory, where postulated utility functions are used to account for the model use within the selection process. Three applications are presented to illustrate the methods. These include surrogate model selection for the purpose of: (1) estimating the minimum of a deterministic function, (2) the design under uncertainty of a simple oscillator, and (3) the uncertainty quantification of a complex engineering system subject to a severe shock and vibration environment.

1. Introduction

Most systems in science and engineering can be described by an input/output relationship of the type shown in Fig. 1, where input **x** and operator **f** are, in general, vector valued. Typically, **f** is defined by a collection of differential, integral, and/or algebraic equations with (possibly) random coefficients. The objective is to calculate properties of an output vector, **y**. For example, **f** can be a finite element (FE) model of a spacecraft that maps an applied pressure field, **x**, to the displacement response, **y**, of an internal component. Properties of **y**, e.g., the maximum in time, can then be calculated.

Mathematical models for the system shown in Fig. 1 are developed for one or more reasons, what we refer to as the *model use*. For example, we may use the models described above to select the appropriate stiffness and/or location of the internal component attachment point such that maximum in time of its response to the prescribed load is less than some critical value. In this case, we say the model use is design.

*Tel.: +1 505 284 4060; fax: +1 505 844 9297.
E-mail address: rvfield@sandia.gov

0022-460X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2007.10.030

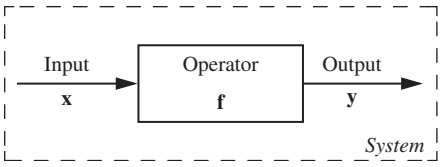

Fig. 1. Model for a system as an input/output relationship.

Real physical systems such as the example described above are often very complex. The models developed to study such systems can therefore involve a large number of equations that can only be solved numerically with a computer, requiring many hours to obtain an accurate solution. We refer to models of this type as *comprehensive models* for the system. Circumstances may require a simplified approximation for the comprehensive model; we refer to this approximation as a *surrogate model* for the system. To illustrate, consider the case where multiple solutions of the large FE model for the spacecraft are necessary. This can occur, for example, if the applied pressure field is random in space and/or time, and Monte Carlo simulation is used as the method for analysis. Optimization and sensitivity analyses also often require multiple model solutions to estimate, for example, the gradient of the output to changes in one or more design variables.

Surrogate models are typically based on a limited number of calculations from the comprehensive models they approximate. Because of this, there may be more than one surrogate model that is consistent with the available information. We refer to the collection of these models as the *collection of candidate surrogate models* for the system. Typical surrogate models include, but are not limited to, polynomials functions [1], radial basis functions (RBFs) [2,3], Kriging interpolation [4,5], and multivariate adaptive regression splines [6]. One type of surrogate model for non-Gaussian random variables and stochastic processes is the polynomial chaos approximation [7,8].

Most often, one surrogate model from the collection of candidate surrogate models is selected and used for analysis. Classical methods to select a surrogate, such as the approaches discussed in Ref. [1], Chapter 2, and/or Ref. [9], typically do not consider the model use. It has been demonstrated [10] that this limitation may render classical methods for surrogate model selection inappropriate for some applications. Herein, we also consider a decision-theoretic method, introduced in Refs. [10,11], to select a surrogate model for the system. Decision theory provides a representation for possible models for the system, a set of options for action, and a utility function that quantifies the decision maker's preferences for each action taken, under each possible model [12]. One criticism of the method is the possible subjectivity of the utility function since the construction of such a function requires a fair understanding of the consequences of unsatisfactory system behavior. We assume these consequences to be well-specified for all examples considered. These methods for model selection have been applied to turbulence models for re-entry random vibration [13] and the monitoring of vehicles in the vicinity of critical national assets [14].

For illustration, we herein apply both methods for surrogate model selection to three applications with different model uses. We choose optimal surrogate models to estimate the minimum of a deterministic function, for design under uncertainty of a simple oscillator, and for uncertainty quantification of a complex engineering system, for applications 1, 2, and 3, respectively.

2. The model selection problem

Consider the following input/output relationship motivated by Fig. 1

$$\mathbf{y} = \mathbf{f}(\mathbf{x}), \tag{1}$$

where $\mathbf{f}: \mathbb{R}^d \to \mathbb{R}^l$ is a deterministic, measurable mapping, and $\mathbf{x} = (x_1, \ldots, x_d)^{\mathrm{T}}$ and $\mathbf{y} = (y_1, \ldots, y_l)^{\mathrm{T}}$ are $\mathbb{R}^d$- and $\mathbb{R}^l$-valued vectors, respectively. Vectors $\mathbf{x}$ and $\mathbf{y}$ may be deterministic or random; for the latter case, we replace $\mathbf{x}$ and $\mathbf{y}$ with $\mathbf{X}$ and $\mathbf{Y}$, respectively. We assume: (i) $\mathbf{f}$ is the comprehensive model for a physical system developed for a specific purpose, i.e., the model use, (ii) the functional form for $\mathbf{f}$ is not explicitly known, but given a value for $\mathbf{x}$, we can calculate the corresponding value for $\mathbf{f}(\mathbf{x})$, and (iii) limited information on $\mathbf{f}$ is available.

The limited information on $\mathbf{f}$ is of two types: (a) calibration data, denoted by $(\mathbf{z}_i, \mathbf{w}_i)$, $i = 1, \ldots, n$, where $\mathbf{z}_i \in \mathbb{R}^d$ and $\mathbf{w}_i = \mathbf{f}(\mathbf{z}_i)$, and (b) prior knowledge, i.e., any information, other than data, on the underlying physics of the system shown in Fig. 1. Prior knowledge is made up by the opinions and theories of experts, as well as any literature on the subject; the fact that $\mathbf{f}$ is nonnegative is one example of prior knowledge. We refer to items (a) and (b) collectively as the available information on $\mathbf{f}$. Note that by assumptions (ii) and (iii), the effects of any solution error are not included.

2.1. Candidate models

There may be more than one surrogate model for $\mathbf{f}$ that is consistent with the available information. Define

$$\mathscr{G} = \{\mathbf{g}_1, \mathbf{g}_2, \ldots\}, \tag{2}$$

where, for each j, $\mathbf{g}_j: \mathbb{R}^d \to \mathbb{R}^l$ denotes a surrogate model for $\mathbf{f}$. We refer to $\mathscr{G}$ as the collection of surrogate models for $\mathbf{f}$. Each $\mathbf{g}_j \in \mathscr{G}$ must be consistent with the available information on $\mathbf{f}$, meaning that: (i) any surrogate model that violates the prior knowledge on $\mathbf{f}$ is excluded from the collection, and (ii) given a functional form for $\mathbf{g}_j$, we estimate values for the coefficients of $\mathbf{g}_j$ using the calibration data. For example, consider the case where each $\mathbf{g}_j$ is a polynomial function of the coordinates of $\mathbf{x} \in \mathbb{R}^d$ up to and including order j; the coefficients of surrogate model $\mathbf{g}_j$ can be estimated using, for example, the method of least squares.

The objective is to select the optimal surrogate model for $\mathbf{f}$, denoted $\mathbf{g}^\star \in \mathscr{G}$; this requires a procedure to rank or order the members of $\mathscr{G}$. One way to order the collection of candidate surrogate models is to assess the accuracy of each $\mathbf{g}_j \in \mathscr{G}$ at various values for $\mathbf{x} \in \mathbb{R}^d$. A comparison of the surrogate models at the calibration data alone may be inadequate since, in many cases, two candidate surrogate models may give the same results, i.e., $\mathbf{g}_i(\mathbf{z}_k) = \mathbf{g}_j(\mathbf{z}_k)$, for $i \neq j$ and $k = 1, \ldots, n$. We therefore introduce validation data, denoted by $(\mathbf{z}'_k, \mathbf{w}'_k)$, $k = 1, \ldots, m$, where $\mathbf{z}'_k \in \mathbb{R}^d$ and $\mathbf{w}'_k \in \mathbb{R}^l$.

Validation data may originate from two sources: (i) solutions of $\mathbf{f}$ at values for $\mathbf{x}$ that do not coincide with the calibration data, i.e., $\mathbf{w}'_k = \mathbf{f}(\mathbf{z}'_k)$, $\mathbf{z}'_k \neq \mathbf{z}_i$, $i = 1, \ldots, n$, $k = 1, \ldots, m$, and/or (ii) experimental observations of the system shown in Fig. 1. As the name implies, validation data may not be used for surrogate model calibration. The concept of using validation data to assess the accuracy of a surrogate model is common; see, for example, Ref. [9]. Cross validation techniques may also be used [15], but are not considered here.

Define

$$p_j \propto \left(\lambda \sum_{k=1}^{m} \|\mathbf{w}'_k - \mathbf{g}_j(\mathbf{z}'_k)\|^2 + (1-\lambda) \sum_{i=1}^{n} \|\mathbf{w}_i - \mathbf{g}_j(\mathbf{z}_i)\|^2 \right)^{-1/2}, \tag{3}$$

where $\propto$ is used to imply that p_j is proportional to the RHS of Eq. (3), $\lambda \in (0, 1)$ is a deterministic constant, and $\|\boldsymbol{\zeta}\|^2 = \sum_{i=1}^{d} \zeta_i^2$ denotes the square of the 2-norm of vector $\boldsymbol{\zeta} \in \mathbb{R}^d$. We scale Eq. (3) such that $\sum_j p_j = 1$ and interpret p_j to be the probability that surrogate model $\mathbf{g}_j \in \mathscr{G}$ is true. The values for $p_1, p_2, \ldots$ therefore define an ordering for the members of $\mathscr{G}$.

2.2. Optimal model by classical method

One technique to select $\mathbf{g}^\star \in \mathscr{G}$ is to consider only the model probabilities defined by Eq. (3). We refer to this approach as the classical method for surrogate model selection and note that the purpose or use of the model, e.g., design or optimization, is not considered in the selection process. Hence, with this method, the optimal model is independent of the model use.

By the classical method, surrogate model $\mathbf{g}_i \in \mathscr{G}$ is optimal and denoted by $\mathbf{g}^\star$ if, and only if

$$p_i \geqslant p_j, \quad j = 1, 2, \ldots \ . \tag{4}$$

We note that by Eq. (3) if validation data is unavailable, p_j depends on the calibration data alone, and $\mathbf{g}^\star \in \mathscr{G}$ is that model which minimizes the mean squared error between $\mathbf{w}_i$ and $\mathbf{g}^\star(\mathbf{z}_i)$. If, in addition, we have $\mathbf{w}_i = \mathbf{g}_j(\mathbf{z}_i)$, $i = 1, \ldots, n$, $j = 1, \ldots$, meaning that each $\mathbf{g}_j \in \mathscr{G}$ interpolates the calibration data, we cannot rank

any surrogate model higher than any other; in this case the models are assumed equally likely, i.e., $p_1 = p_2 = \cdots$.

2.3. Optimal model by decision-theoretic method

Consistent with the approach developed in Refs. [10,11] we propose to instead assess the utility of each candidate surrogate model for the intended model use, then rank the members of $\mathscr{G}$ according to their expected utility. Let $U(\mathbf{g}_i, \mathbf{g}_j) \geqslant 0$ denote the utility associated with using surrogate model $\mathbf{g}_i \in \mathscr{G}$ for the intended use, assuming surrogate $\mathbf{g}_j \in \mathscr{G}$ is the best surrogate model available. We construct U such that $U(\mathbf{g}_i, \mathbf{g}_j) \geqslant U(\mathbf{g}_k, \mathbf{g}_j)$ if, and only if, the consequences of using model $\mathbf{g}_i$ are preferable to the consequences of using model $\mathbf{g}_k$. By definition, the utility function is a random variable, and

$$u_i = \mathrm{E}[U(\mathbf{g}_i, \mathscr{G})] = \sum_j U(\mathbf{g}_i, \mathbf{g}_j) p_j \tag{5}$$

denotes the expected utility of surrogate model $\mathbf{g}_i$.

By the decision-theoretic method, surrogate model $\mathbf{g}_i \in \mathscr{G}$ is optimal and denoted by $\mathbf{g}^\star$ if, and only if

$$u_i \leqslant u_j, \quad j = 1, 2, \ldots . \tag{6}$$

The utility, U, is sometimes referred to as the "opportunity loss" (see Ref. [16], p. 60) so that the solution to Eq. (6) agrees with intuition, i.e., $\mathbf{g}^\star \in \mathscr{G}$ minimizes the expected loss. For the special case where $U(\mathbf{g}_i, \mathbf{g}_j) = 1 - \delta_{ij}$, where $\delta_{ij} = 1$ for $i = j$ and zero otherwise, the optimal surrogate model by the classical method (Eq. (4)) is recovered (see Ref. [17], p. 23).

The utility function, U, depends on the model use, meaning that different surrogate models may be selected for a different model use, even when the comprehensive model, $\mathbf{f}$ defined by Eq. (1), remains unchanged. To illustrate this, we apply the decision-theoretic method for surrogate model selection to the design of a simple oscillator in Section 4. Two different design metrics are considered, and the method selects different surrogates depending on the design metric considered.

The overall process of surrogate model section as applied herein is outlined by Fig. 2. Beginning with the available information, i.e., the calibration and validation data, a collection of candidate surrogate models can be defined. Next, we identify an appropriate utility function based on the model use. Finally, we select the optimal model from the collection. As mentioned, classical methods do not consider the model use; the use of the dashed line in Fig. 2 is meant to convey this.

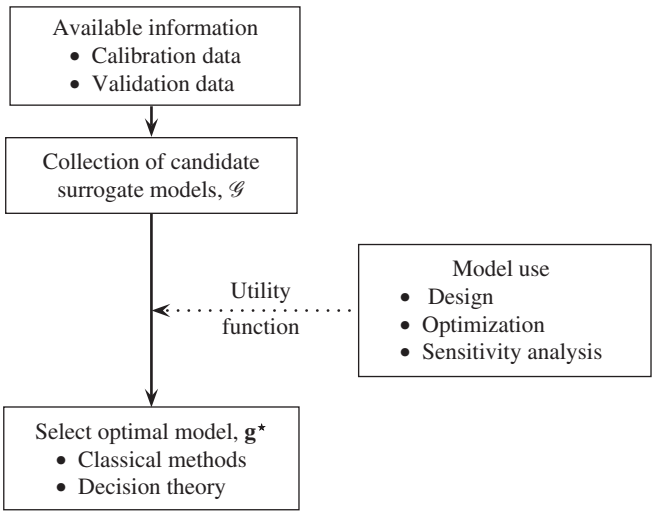

Fig. 2. The process of surrogate model selection (adapted from Ref. [10]).

3. Surrogate models

Three classes of surrogate models are briefly reviewed. Detailed descriptions of these models can be found in, for example, Ref. [1], Chapter 2, and Ref. [9]. The purpose here is not to present an exhaustive list of the numerous types of surrogate models used in practice, but rather to present an overview of a few simple relevant models so as to illustrate the concept of surrogate model selection. We restrict our discussion to the case of scalar output $\mathbf{y}$; we therefore replace $\mathbf{f}$, $\mathbf{g}$, and $\mathbf{y}$ defined in Section 2 with f, g, and y, respectively.

We consider surrogate models for $y = f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$, of the following type:

$$g(\mathbf{x}; \mathbb{H}_d^r) = \sum_{j=1}^{r} c_j h_j(\mathbf{x}) = \mathbf{c}^{\mathrm{T}} \mathbf{h}(\mathbf{x}), \tag{7}$$

where $\mathbf{c} = (c_1, \ldots, c_r)^{\mathrm{T}}$ denotes a deterministic vector of coefficients that must be determined, and $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \ldots, h_r(\mathbf{x}))^{\mathrm{T}}$ denotes an array of deterministic vector-valued basis functions. We explicitly write g as a function of the collection $\mathbb{H}_d^r = \{h_1(\mathbf{x}), \ldots, h_r(\mathbf{x}); \mathbf{x} \in \mathbb{R}^d\}$ to denote the dependence of the surrogate model on the choice of basis.

The method of least-squares can be used to solve for the coefficients of Eq. (7), i.e.,

$$\mathbf{c} = (\mathbf{a}^{\mathrm{T}}\mathbf{a})^{-1}\mathbf{a}^{\mathrm{T}}\mathbf{w}, \tag{8}$$

where $\mathbf{a}$ is an $n \times r$ matrix with elements $a_{ij} = h_j(\mathbf{z}_i)$, $\mathbf{w} = (w_1, \ldots, w_n)^{\mathrm{T}}$, and $(\mathbf{z}_i, w_i)$, $i = 1, \ldots, n$, denotes the calibration data defined in Section 2. For the special case when $r = n$ and $\mathbf{a}$ has full rank, Eq. (8) reduces to

$$\mathbf{c} = \mathbf{a}^{-1}\mathbf{w}. \tag{9}$$

Many of the surrogate models used in practice assume the calibration data, $(\mathbf{z}_i, w_i)$, $i = 1, \ldots, n$, satisfy the following criteria:

$$\begin{aligned} &\sum_{j=1}^{n} w_j = \sum_{j=1}^{n} z_{k,j} = 0, \ k = 1, \ldots, d, \quad \text{and} \\ &\sum_{j=1}^{n} w_j^2 = \sum_{j=1}^{n} z_{k,j}^2 = 1, \ k = 1, \ldots, d, \end{aligned} \tag{10}$$

where $z_{k,j}$ denotes the kth coordinate of vector $\mathbf{z}_j$. When necessary, we will make the same assumption; the extension to the case where the calibration data do not satisfy Eq. (10) is straightforward.

As noted, the surrogate model for f defined by Eq. (7) depends on the choice of basis, $\mathbb{H}_d^r$. We next discuss some properties of g for three different choices for $\mathbb{H}_d^r$, where each choice is commonly used in practice. The bases studied include polynomial, exponential, and indicator functions of $\mathbf{x}$, and are discussed in Sections 3.1–3.3, respectively.

3.1. Polynomial basis

Let $\mathbb{P}_d^r = \{h_1(\mathbf{x}), \ldots h_r(\mathbf{x}), \ \mathbf{x} \in \mathbb{R}^d\}$ denote the collection of multidimensional polynomials of $\mathbf{x}$ up to, and including, order q, where

$$h_j(\mathbf{x}) = x_1^{q_1} x_2^{q_2} \cdots x_d^{q_d}, \tag{11}$$

each $q_i \in \{0, 1, \ldots, q\}$, and $\sum_{i=1}^{d} q_i \leqslant q$. It follows that r, the number of terms in Eq. (7), is given by $r = \prod_{j=1}^{q}(1 + d/j)$. For example, consider the case of $d = q = 2$ so that $r = (1 + 2)(1 + 2/2) = 6$. The corresponding collection of basis functions is given by $\mathbb{P}_2^6 = \{h_1(x_1, x_2), \ldots, h_6(x_1, x_2)\}$, where

$$\begin{aligned} h_1(x_1, x_2) &= 1, \\ h_2(x_1, x_2) &= x_1, \\ h_3(x_1, x_2) &= x_2, \end{aligned}$$

$$h_4(x_1,x_2)=x_1^2,$$
$$h_5(x_1,x_2)=x_1x_2,$$
$$h_6(x_1,x_2)=x_2^2. \tag{12}$$

The polynomial basis, perhaps the most frequently used approximation for dealing with functions on bounded domains, has some interesting properties. First, in general $g(\mathbf{z}_i;\mathbb{P}_d^r)\neq w_i$, meaning that a surrogate model defined on $\mathbb{P}_d^r$ does not necessarily interpolate f. Second, by Weierstrass's theorem, as $r\to\infty$, any continuous f can be approximated on a finite interval with arbitrary precision (see Ref. [18], p. 159).

3.2. Exponential basis

Assume $r=n$ and let $\mathbb{E}_d^n=\{h_1(\mathbf{x}),\ldots,h_n(\mathbf{x}),\ \mathbf{x}\in\mathbb{R}^d\}$ denote a collection of exponential functions of $\mathbf{x}$, where

$$h_j(\mathbf{x})=\exp(-\theta_j\|\mathbf{x}-\mathbf{z}_j\|^2),\quad j=1,\ldots,n \tag{13}$$

and $\theta_j>0$ is a deterministic parameter. An alternative collection of basis functions, denoted by $\tilde{\mathbb{E}}_d^n=\{\tilde{h}_1(\mathbf{x}),\ldots,\tilde{h}_n(\mathbf{x})\}$, can be considered, where

$$\tilde{h}_j(\mathbf{x})=h_j(\mathbf{x})+\frac{1-\mathbf{1}^{\mathrm{T}}\mathbf{a}^{-1}\mathbf{h}(\mathbf{x})}{\mathbf{1}^{\mathrm{T}}\mathbf{a}^{-1}\mathbf{1}}. \tag{14}$$

$\mathbf{1}=(1,1,\ldots,1)^{\mathrm{T}}$ denotes an $n\times 1$ vector of ones, and $\mathbf{h}(\mathbf{x})$ is defined by Eq. (7).

The bases defined by Eqs. (13) and (14) are similar. Because $r=n$, g interpolates f for either basis, meaning that $g(\mathbf{z}_i;\mathbb{E}_d^n)=g(\mathbf{z}_i;\tilde{\mathbb{E}}_d^n)=w_i$, $i=1,\ldots,n$. The second term on the RHS of Eq. (14) is zero for $\mathbf{x}=\mathbf{z}_i$, $i=1,\ldots,n$; it follows that $\tilde{h}_j(\mathbf{z}_i)=h_j(\mathbf{z}_i)$, $i,j=1,\ldots,n$, and the coefficients of Eq. (7) are therefore identical under basis $\mathbb{E}_d^n$ and $\tilde{\mathbb{E}}_d^n$. The surrogate models defined above have special names in the literature. Eq. (13) is a particular type of radial basis function (RBF), and under basis $\tilde{\mathbb{E}}_d^n$, Eq. (7) is a Kriging approximation for f [3,5]. The RBF and Kriging approximations are commonly used in practice as surrogate models for large, complex FE models (see, for example, Ref. [19]).

3.3. Indicator basis

Assume $r=n$ and let $\mathbb{I}_d^n=\{h_1(\mathbf{x}),\ldots,h_n(\mathbf{x}),\ \mathbf{x}\in\mathbb{R}^d\}$ denote a collection of basis functions, where

$$h_j(\mathbf{x})=1(\mathbf{x}\in\rho_j) \tag{15}$$

and $\rho_1,\ldots,\rho_r\subseteq\mathbb{R}^d$ denote non-overlapping subsets of $\mathbb{R}^d$. The function $1(A)=1$ if event A is true, and $1(A)=0$ otherwise, is referred to as an indicator function. We consider the special case where each ρ_j is a rectangle in $\mathbb{R}^d$ with center $\mathbf{x}=\mathbf{z}_j$ and size $\phi_1\times\cdots\times\phi_d$, i.e.,

$$\rho_j=\overset{d}{\underset{k=1}{\times}}[z_{k,j}-\phi_k/2,z_{k,j}+\phi_k/2],\quad j=1,\ldots,n \tag{16}$$

and

$$\phi_k=\min_{z_{k,i}\neq z_{k,j}}|z_{k,i}-z_{k,j}|. \tag{17}$$

By Eqs. (16) and (17), $\mathbf{a}=\mathbf{i}_n$, where $\mathbf{i}_n$ denotes the $n\times n$ identity matrix, so that, by Eq. (9), $\mathbf{c}=\mathbf{w}$. More sophisticated techniques are available to select both the number of basis functions, r, and the subsets, $\rho_1,\ldots,\rho_r$; one popular approach is the Multivariate Adaptive Regression Spline (MARS) [6].

4. Applications

Three applications are provided to demonstrate the methods for surrogate model selection. Various types of model use are considered, including: (i) deterministic prediction, (ii) design under uncertainty, and (iii) uncertainty quantification. For (i), we consider a collection of surrogate models for f, a deterministic,

fourth-order polynomial function of two variables, for the purpose of estimating the minimum of f. For (ii), we consider the design of a simple oscillator subject to a time-varying forcing function, where one of the system parameters is modeled as a random variable. Two different design metrics are considered, and the optimal surrogate model is selected for each metric. For (iii), f is a highly nonlinear transient dynamics FE model of a flexible structure penetrating a hard target, and surrogate models are needed in order to quantify the effects of uncertainty in certain initial conditions on system response. Applications (i), (ii), and (iii) are discussed in Sections 4.1, 4.2, and 4.3, respectively.

4.1. Deterministic prediction

Let

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \quad \mathbf{x} \in D, \tag{18}$$

where $D = [-2, 2] \times [-2, 2]$, be the comprehensive model of interest. This particular example has been extensively studied by the optimization community and is commonly known as the Rosenbrock test function [20,21]. The function f is illustrated in Fig. 3; a contour plot is also shown.

The objectives are: (i) select the optimal surrogate model for f, and (ii) use the optimal surrogate to estimate

$$\eta = \min_{\mathbf{x} \in D} f(\mathbf{x}). \tag{19}$$

The exact solution to Eq. (19) is $\eta = 0$ at $\mathbf{x} = (1, 1)^{\mathrm{T}}$, as denoted by the solid circle in Fig. 3.

4.1.1. Candidate models

The available information on f is limited to n calibration data, denoted by $(\mathbf{z}_i, w_i)$, $i = 1, \ldots, n$. We assume m validation data, denoted by $(\mathbf{z}'_j, w'_j)$, $j = 1, \ldots, m$, where $w'_j = f(\mathbf{z}'_j)$ and $\mathbf{z}'_j \neq \mathbf{z}_i$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. The values for $\mathbf{z}_i$ and $\mathbf{z}'_j$ are illustrated by Fig. 4, where domain D is discretized into 64 non-overlapping $1/2 \times 1/2$ regions, defined by the 81 nodes in the figure. As denoted by the black squares in Fig. 4, $m = 5$ of the 81 nodes are reserved for surrogate model validation. The calibration points are selected at random from the remaining 76 nodes. The values for $\mathbf{z}_1, \ldots, \mathbf{z}_n$ for the case of $n = 4$ and 10 are illustrated by Fig. 4.

Six candidate surrogate models for f are considered, i.e.,

$$\begin{aligned} \mathscr{G} &= \{g_1, \ldots, g_6\} \\ &= \{g(\mathbf{x}; \mathbb{P}_2^6), g(\mathbf{x}; \mathbb{P}_2^{10}), g(\mathbf{x}; \mathbb{P}_2^{15}), g(\mathbf{x}; \mathbb{E}_2^n), g(\mathbf{x}; \tilde{\mathbb{E}}_2^n), g(\mathbf{x}; \mathbb{I}_2^n)\}, \end{aligned} \tag{20}$$

where the functional form for each surrogate is defined by Eq. (7), and the parameters for each surrogate are estimated from the calibration data shown in Fig. 4. By Eq. (20), g_1, g_2, and g_3 are second-, third-, and fourth-order polynomial functions of $\mathbf{x} \in \mathbb{R}^2$, respectively. Surrogate models g_4 and g_5 are defined on the exponential basis defined by Eqs. (13) and (14), respectively, with $\theta_j = 1$, $j = 1, \ldots, n$; g_6 is defined on the indicator basis discussed in Section 3.3.

We refer to $g_3 \in \mathscr{G}$ as the "true surrogate model" for f because, by Eq. (18), f is a fourth-order polynomial in $\mathbf{x}$. Contours of surrogates g_1, g_3, g_4, and g_6 are shown in Fig. 5 for the case of $n = 15$. These results

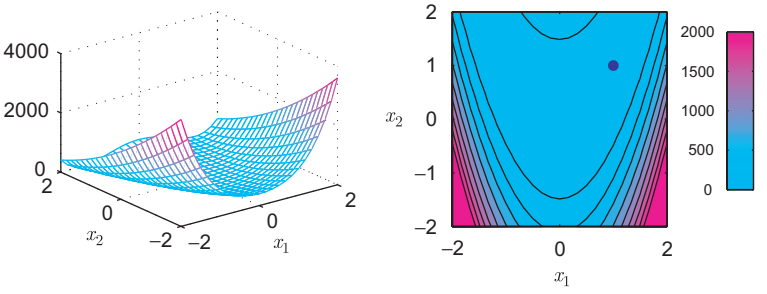

Fig. 3. The Rosenbrock function, $f(x_1, x_2)$, considered in application #1.

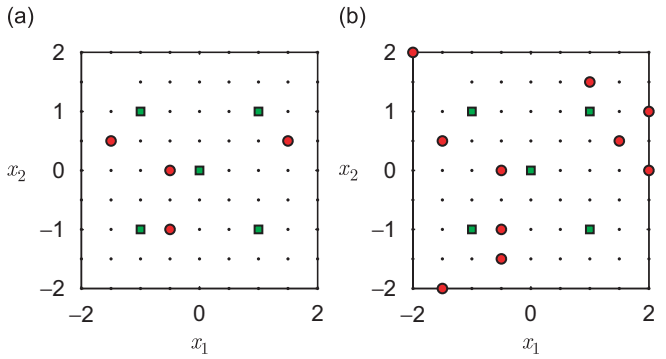

Fig. 4. Locations of n calibration data (red circles) and m validation data (green squares) for application #1 assuming: (a) $(m,n)=(5,4)$, and (b) $(m,n)=(5,10)$.

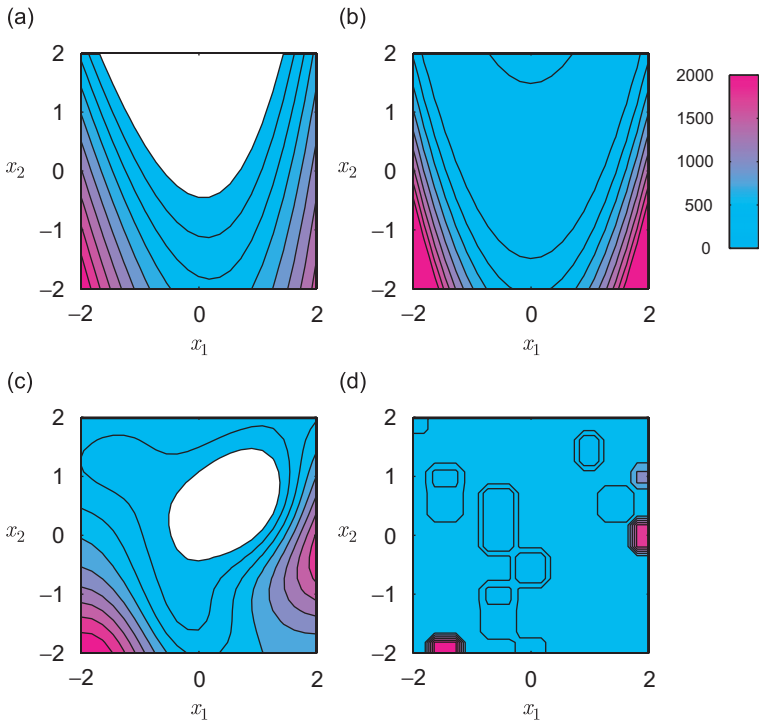

Fig. 5. Contours of four of the six candidate surrogate models for f assuming $n=15$: (a) g_1, (b) g_3, (c) g_4, and (d) g_6.

demonstrate that, for fixed n, the surrogate models can be very different from one another, but each is consistent with the available information. Further, as illustrated by Fig. 6, a surrogate model can change dramatically with n, the amount of available information. For very few calibration data, as illustrated by Fig. 6a, surrogate g_4 is very different from the exact function f illustrated by Fig. 3; as the number of calibration data grows large, g_4 looks very much like f (see Fig. 6d).

4.1.2. Optimal model

We apply the classical and decision-theoretic methods to select optimal surrogate models for f, then use these models to estimate $\eta=\min f(\mathbf{x})$. Further, we study the evolution of the optimal model and corresponding estimates for η as the number of calibration data, n, increases.

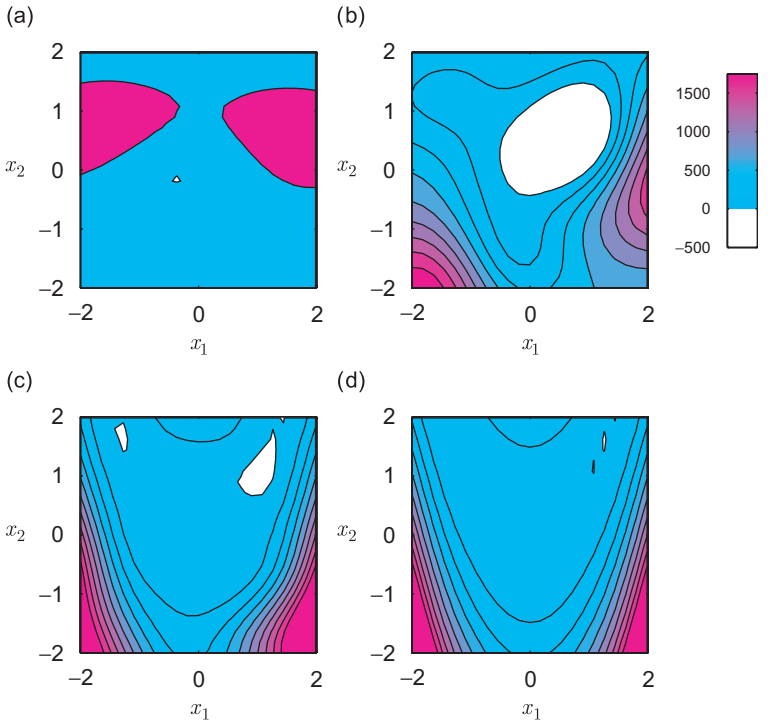

Fig. 6. Contours of surrogate model g_4 assuming: (a) $n = 4$, (b) $n = 15$, (c) $n = 38$, and (d) $n = 76$.

As discussed in Section 2, the decision-theoretic method for model selection requires a utility function. Let

$$\xi_i = \min_{\mathbf{x}\in D} g_i(\mathbf{x}), \tag{21}$$

so that ξ_i denotes an approximation for η defined by Eq. (19) using surrogate $g_i \in \mathcal{G}$. Assuming surrogate $g_j \in \mathcal{G}$ is true, we say $g_i \in \mathcal{G}$ is conservative if it over-predicts the minimum, i.e., if $\xi_i \geqslant \xi_j$. We assume conservative models are preferable to non-conservative models and use this assumption as the basis for our interpretation of surrogate model utility.

An appropriate value for the utility of surrogate model $g_i \in \mathcal{G}$, assuming surrogate $g_j \in \mathcal{G}$ is true, is given by

$$U(g_i, g_j) = \tilde{U}(\xi_i, \xi_j) = \begin{cases} \beta_1(\xi_i - \xi_j)^2 & \text{if } \xi_i \geqslant \xi_j, \\ \beta_2(\xi_i - \xi_j)^2 & \text{if } \xi_i < \xi_j, \end{cases} \tag{22}$$

where $\beta_2 \geqslant \beta_1 \geqslant 0$ are deterministic parameters, and we replace U with $\tilde{U}$ to denote that the utility function can be expressed as a function of ξ_i and ξ_j alone. By Eq. (22), non-conservative models are assigned a large utility; overly conservative models are also subject to penalty. We note that definitions of model utility are problem dependent; alternative definitions can be used.

The surrogate model probabilities, $p_1, \ldots, p_6$, are illustrated by Fig. 7(a) for $4 \leqslant n \leqslant 20$; shown in Fig. 7(b) are the expected utilities of each surrogate model, $u_1, \ldots, u_6$. Parameters $\lambda = 1/2$, $\beta_1 = 1$, and $\beta_2 = 10$ were used for calculations. The optimal surrogate model, $g^\star \in \mathcal{G}$, using the classical method (by Eq. (4)) and decision-theoretic method (by Eq. (6)) are shown in Fig. 8 for $4 \leqslant n \leqslant 20$. For $n < 15$, different surrogate models for f are optimal under the two methods for model selection. For $n \geqslant 15$, $p_3 = 1$ so that the true model, g_3, is selected by both methods; this is because g_3 is a fourth-order polynomial requiring $r = 15$ coefficients (see Section 3.1).

Recall that only the decision-theoretic method includes the model use and, therefore, assigns a large utility to those models that provide non-conservative predictions of Eq. (19). To illustrate this, let

$$\xi^\star = \min_{\mathbf{x}\in D} g^\star(\mathbf{x}) \tag{23}$$

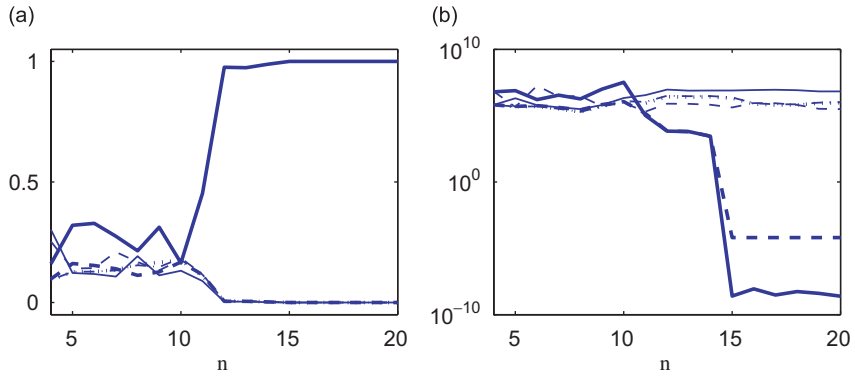

Fig. 7. Model probabilities (a) and expected utilities (b) for each surrogate: g_1 (thin solid), g_2 (thin dashed), g_3 (thick solid), g_4 (dash–dot), g_5 (dotted), and g_6 (thick dashed).

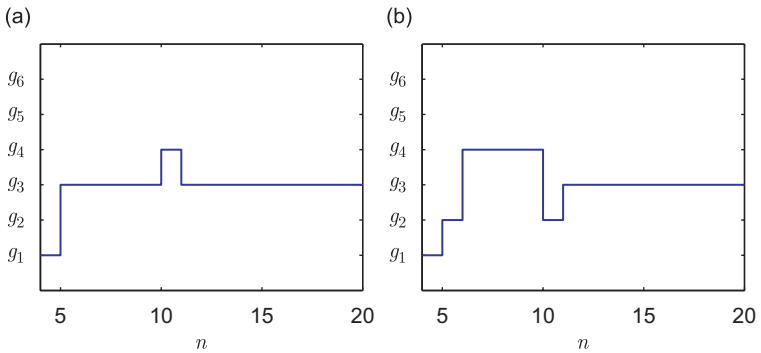

Fig. 8. Optimal surrogate model, $g^\star \in \mathcal{G}$, for application #1 using classical method (a) and decision-theoretic method (b).

denote the prediction of Eq. (19) under the optimal model, $g^\star \in \mathcal{G}$. Values for $\xi^\star$ using the classical method and proposed decision-theoretic method are shown in Fig. 9. Note that with the classical method, a non-conservative model, i.e., one that under-predicts the minimum, may be selected for $n<15$. For $n \geqslant 15$, $\xi^\star = \min_{\mathbf{x}\in D} f(\mathbf{x}) = 0$ using both methods for surrogate model selection.

The results of this section illustrate two important properties. First, when data is limited the two methods can deliver different results; only the decision-theoretic method accounts for the model use and ensures a conservative estimate for η defined by Eq. (19). Second, if one of the surrogate models exactly duplicates the comprehensive model, it will be selected by both methods considered. We remark that the results presented depend on the: (i) location and number of the validation data, (ii) order in which the calibration data is selected (see Fig. 4), and (iii) utility function. More sophisticated methods are available to select locations for calibration and validation data (see, for example, Ref. [22], Section 4.3); these methods can easily be included in the model selection framework presented here. A discussion on the sensitivity of the optimal model by the decision-theoretic method to changes in the utility function are discussed in Ref. [10].

4.2. Design under uncertainty

We next consider the 2 degree-of-freedom oscillator shown in Fig. 10, a model commonly used for applications in structural dynamics, where ζ denotes the forcing function, X denotes the value for a spring constant, m_1 and m_2 denote the values for the two masses, and v denotes the relative displacement of the two masses. We assume: (i) input $\zeta(t)$ is a perfectly known and deterministic function of time, t, (ii) the value for one of the spring constants is known and fixed, (iii) X is a uniform random variable on [1500,2500], and

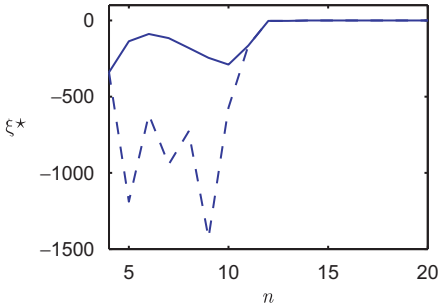

Fig. 9. Optimal approximations for $\min f$ vs. n under decision-theoretic method (solid) and classical method (dashed).

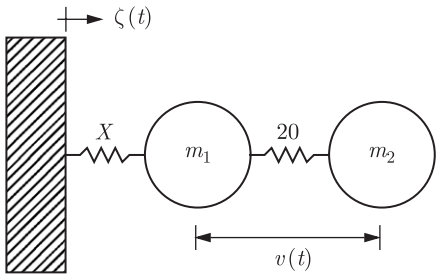

Fig. 10. Two degree-of-freedom oscillator considered in application #2.

(iv) the values for the two masses are deterministic design parameters that must satisfy the following constraints:

$$m_1 + m_2 = 100, \tag{24a}$$

$$\frac{1}{99} \leqslant \frac{m_2}{m_1} \leqslant 1. \tag{24b}$$

The objectives are to: (i) select the optimal surrogate model for the comprehensive system model, and (ii) select values for m_1 and m_2, such that certain properties of v satisfy a prescribed set of conditions.

Let $\mathbf{m}$, $\mathbf{d}$, and $\mathbf{k}$ denote the 2×2 mass, damping, and stiffness matrices, respectively, of the two degree-of-freedom oscillator shown in Fig. 10, i.e.,

$$\mathbf{m} = \begin{bmatrix} \dfrac{100}{1+\delta} & 0 \\ 0 & \dfrac{100\delta}{1+\delta} \end{bmatrix} \quad \text{and} \quad \mathbf{k} = \begin{bmatrix} X+20 & -20 \\ -20 & -20 \end{bmatrix}, \tag{25}$$

where $\delta = m_2/m_1$ is the ratio of the two masses, and $\mathbf{d}$ is such that the system is classically damped with a constant damping ratio of 4% for each mode. The relative displacement of the two masses, assuming zero initial conditions, is g,

$$v(t, X, \delta) = \mathbf{c} \int_0^t \exp[\mathbf{a}(t-\tau)]\, \mathbf{b}\, \zeta(\tau)\, \mathrm{d}\tau, \tag{26}$$

where

$$\mathbf{a} = \begin{bmatrix} \mathbf{0} & \mathbf{i} \\ -\mathbf{m}^{-1}\mathbf{k} & -\mathbf{m}^{-1}\mathbf{d} \end{bmatrix}, \quad \mathbf{b} = [0 \quad 0 \quad -1 \quad -1]^{\mathrm{T}}, \quad \mathbf{c} = [-1 \quad 1 \quad 0 \quad 0], \tag{27}$$

and $\mathbf{i}$ denotes the 2×2 identity matrix. We consider two metrics of system performance, given by

$$P(Y \leqslant \bar{y}), \tag{28a}$$

and

$$\mathrm{E}[Y], \tag{28b}$$

where $\bar{y} \geqslant 0$ is a prescribed deterministic parameter, and

$$Y = f(X, \delta) = \max_{t \geqslant 0} | v(t, X, \delta) | \tag{29}$$

is the output of interest. We write $Y = f(X, \delta)$ in Eq. (29) to be consistent with the general input/output relationship described by Fig. 1, and we replace $\mathbf{x}$ with (X, δ) and $\mathbf{y}$ with Y; capital letters for X and Y are used to denote that these two quantities are random variables.

To estimate the performance metrics defined by Eq. (28), it is convenient to approximate Eq. (29) with a surrogate model; approximation may become necessary when we consider nonlinear systems or linear systems with many degrees-of-freedom. Herein, we employ methods for surrogate model selection for the purpose of design under uncertainty. We consider two cases:

$$\text{Case \#1 : select } \delta \in [1/99, 1] \quad \text{such that } P(Y \leqslant \bar{y}) = \bar{q}, \tag{30a}$$

$$\text{Case \#2 : select } \delta \in [1/99, 1] \quad \text{such that } \mathrm{E}[Y] = \bar{r}, \tag{30b}$$

where $\bar{q}$ and $\bar{r}$ denote prescribed deterministic parameters that define the design constraints. The model use for Case #1 is different than for Case #2 because, by Eq. (30), the design constraints are different. Optimal surrogate models for Case #1 and Case #2 are discussed in Sections 4.2.2.1 and 4.2.2.2, respectively.

4.2.1. Candidate models

The available information on f is limited to n calibration data, denoted by $(\mathbf{z}_i, w_i)$, $i = 1, \ldots, n$, and m validation data, denoted by $(\mathbf{z}'_k, w'_k)$, $k = 1, \ldots, m$. We assume the latter is given by simulated experimental observations of the system shown in Fig. 10. For calculations, we model each experimental observation as the solution of the comprehensive model at one of the calibration points, subject to additive noise, i.e., for $k = 1, \ldots, m$, w'_k is one sample of random variable

$$f(\mathbf{z}'_k) + E_k, \tag{31}$$

where $\{E_k\}$ denotes a sequence of zero-mean iid Gaussian random variables with variance σ^2, each $\mathbf{z}'_k$ coincides with one of $\mathbf{z}_i$, $i = 1, \ldots, n$, and $m \leqslant n$. The data are illustrated by Fig. 11 for $m = 5$ and $n = 10$. The values for $\mathbf{z}'_1, \ldots, \mathbf{z}'_5$, the validation data, are shown in Fig. 11(a), while the values for $\mathbf{z}_1, \ldots, \mathbf{z}_{10}$, the calibration data, are shown in Fig. 11(b). For calculations, the forcing function, $\zeta(t)$, is given by one sample of Gaussian white noise with intensity $10,000/\pi$ (see Ref. [23], p. 29).

We consider 6 candidate surrogate models for f, i.e.,

$$\begin{aligned} \mathscr{G} &= \{g_1, \ldots, g_6\} \\ &= \{g(X, \delta; \mathbb{P}_2^6), g(X, \delta; \mathbb{P}_2^{10}), g(X, \delta; \mathbb{P}_2^{15}), g(X, \delta; \mathbb{E}_2^n), g(X, \delta; \tilde{\mathbb{E}}_2^n), g(X, \delta; \mathbb{I}_2^n)\}, \end{aligned} \tag{32}$$

where the functional form for each surrogate is defined by Eq. (7), and the parameters for each surrogate are estimated from the calibration data illustrated by Fig. 11(b). Note the functional form for each surrogate considered is identical to the functional form considered in Section 4.1.1. Unlike the example of Section 4.1, there is no true surrogate model for f. Contours of surrogates g_1, g_2, g_4, and g_6 are shown in Fig. 12 for the case of $n = 15$, illustrating that the candidate surrogate models for f are very different, but each is consistent with the available calibration data.

4.2.2. Optimal model

We first apply the classical method for model selection; results using the decision-theoretic method, which depend on the model use, are discussed in Sections 4.2.2.1 and 4.2.2.2. The surrogate model probabilities, $p_1, \ldots, p_6$, are illustrated by Fig. 13(a) for $5 \leqslant n \leqslant 81$; the optimal surrogate model for f using the classical

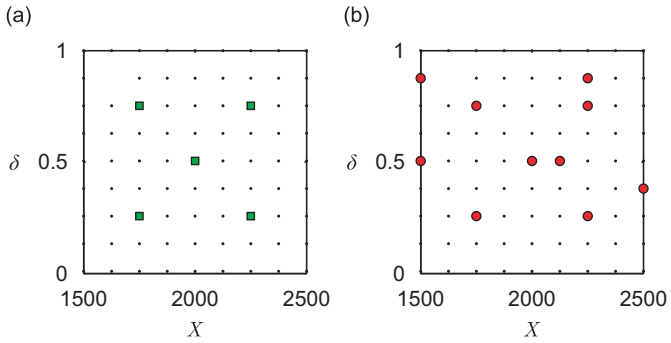

Fig. 11. Locations of $m = 5$ validation data (a) and $n = 10$ calibration data (b) for application #2.

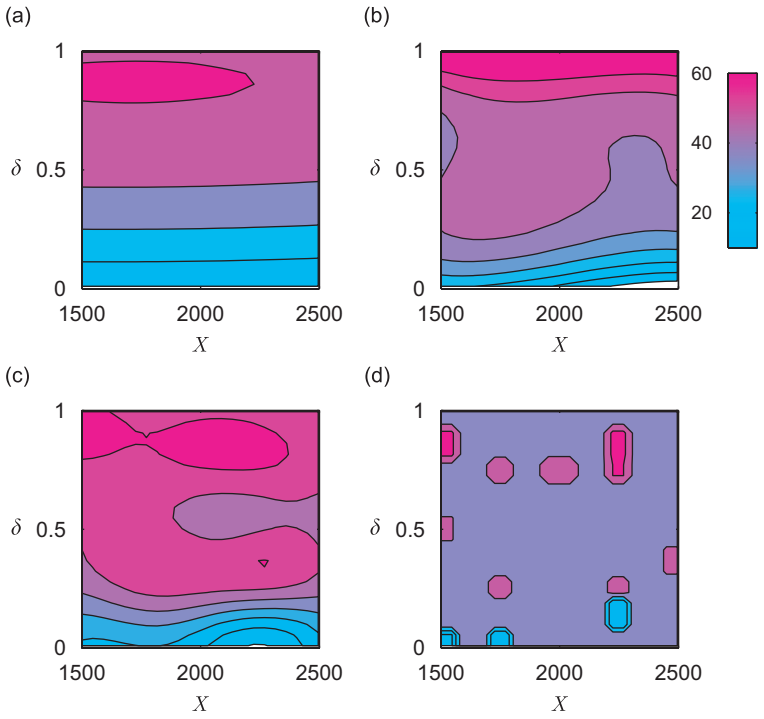

Fig. 12. Contours of four of the six candidate surrogate models for f assuming $n = 15$: (a) g_1, (b) g_2, (c) g_4, and (d) g_6.

method, i.e., by Eq. (4), is shown in Fig. 13(b) as a function of n. Parameters $\lambda = 1/2$ and $\sigma^2 = 10$ were used for calculations. For $n < 16$, any surrogate model can be selected. For $n \geqslant 16$, values for p_1, p_2, and p_3, which correspond to the polynomial models, approach zero, while values for p_4, p_5, and p_6 are nearly identical and approach $1/3$. Hence, among models $\{g_4, g_5, g_6\}$ there is no strong preference of one over another for $n \geqslant 16$; this is further demonstrated by the oscillatory behavior of $g^\star$ illustrated by Fig. 13(b).

4.2.2.1. Case #1. As discussed in Section 2, the decision-theoretic method for model selection requires a utility function; we next develop such a function to quantify the utility of each surrogate that is consistent with the model use (Eq. (30a)).

Assuming it exists, we define a_i such that

$$P(g_i(X, a_i) \leqslant \bar{y}) = \bar{q}, \tag{33}$$

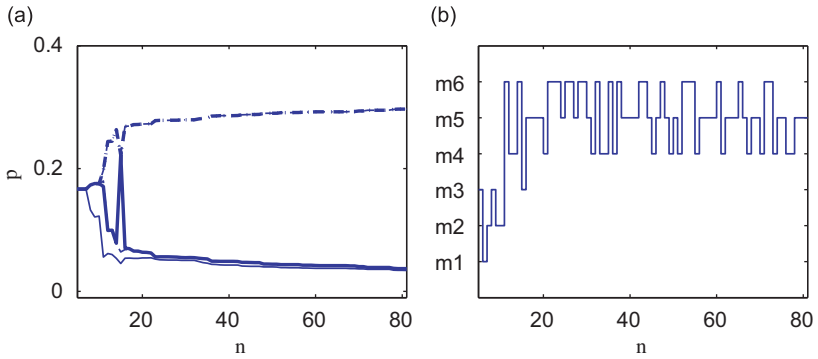

Fig. 13. Model selection by classical method for application #2. Model probabilities (a) for each surrogate: g_1 (thin solid), g_2 (thin dashed), g_3 (thick solid), g_4 (dash–dot), g_5 (dotted), and g_6 (thick dashed). Optimal model (b), $g^\star \in \mathscr{G}$.

so that a_i is the value for δ that satisfies design condition #1, defined by Eq. (30a), under surrogate model g_i. In the case of a non-unique solution, we choose the minimum a_i that satisfies Eq. (33). The performance of design a_i, assuming model $g_j \in \mathscr{G}$ is true, is given by

$$\xi_{ij} = P(g_j(X, a_i) \leqslant \bar{y}), \tag{34}$$

which may or may not equal the required reliability, $\bar{q}$. We say design a_i is conservative if the reliability exceeds $\bar{q}$, and non-conservative otherwise. We assume models that favor conservative designs are favorable to models that favor non-conservative designs and use this assumption as the basis for our interpretation of surrogate model utility. A similar approach for conservative estimates of probability distributions is presented by Picheny [24].

An appropriate value for the utility of surrogate model $g_i \in \mathscr{G}$, if surrogate model $g_j \in \mathscr{G}$ is true, is given by

$$U(g_i, g_j) = \tilde{U}(a_i, a_j) = \begin{cases} \beta_1(a_i - a_j)^2 & \text{if } \xi_{ij} \geqslant \bar{q}, \\ \beta_2(a_i - a_j)^2 & \text{if } \xi_{ij} < \bar{q}, \end{cases} \tag{35}$$

where $\beta_2 \geqslant \beta_1 \geqslant 0$ are deterministic parameters, and we replace U with $\tilde{U}$ to denote that the utility function can be expressed as a function of a_i and a_j. By Eq. (35), models that favor non-conservative designs are assigned a large utility; models that favor overly conservative designs are also subject to penalty.

The expected utility of each surrogate model, denoted by $u_1, \ldots, u_6$, is illustrated by Fig. 14(a) for $5 \leqslant n \leqslant 81$; the corresponding optimal surrogate model for each value for n is shown in Fig. 14(b). Parameters $\beta_1 = 1$, $\beta_2 = 100$, $\bar{q} = 0.9$, and $\bar{y} = 50$ were used for calculations. The results are different than those shown in Fig. 13 because the optimal model under the decision-theoretic method depends on the model use. For example, surrogate model g_1, a second-order polynomial, is often optimal because it results in a conservative design of the system, i.e., values for design parameter δ such that $P(Y \leqslant \bar{y}) \geqslant \bar{q}$. Models selected by the classical method provide no such guarantee on design performance.

4.2.2.2. Case #2. By Eq. (30), the model use for Case #2 is different than for Case #1; our definition for the utility function must therefore reflect this. Assuming it exists, we define a_i such that

$$\mathrm{E}[g_i(X, a_i)] = \bar{r}, \tag{36}$$

so that a_i is the value for δ that satisfies design condition #2, defined by Eq. (30b), under surrogate model g_i. As before, in the case of a non-unique solution, we choose the minimum a_i that satisfies Eq. (36). The performance of design a_i, assuming model $g_j \in \mathscr{G}$ is true, is given by

$$\xi_{ij} = \mathrm{E}[g_j(X, a_i)], \tag{37}$$

which may or may not be equal to $\bar{r}$, the design requirement. We say design a_i is non-conservative if the mean value exceeds $\bar{r}$, and conservative otherwise; we assume models that favor conservative designs are favorable to models that favor non-conservative designs.

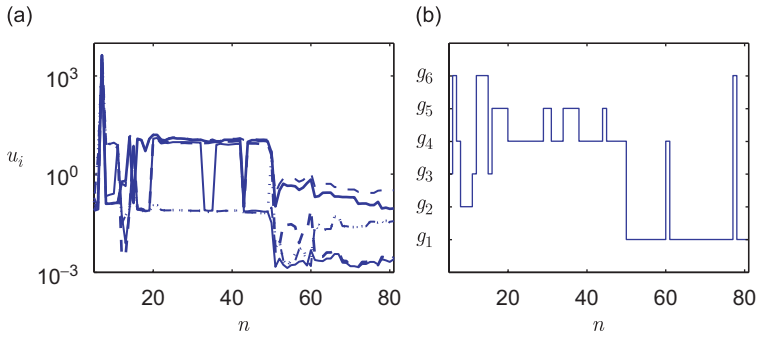

Fig. 14. Model selection by decision-theoretic method for application #2, case #1. Expected utilities (a) for each surrogate: g_1 (thin solid), g_2 (thin dashed), g_3 (thick solid), g_4 (dash–dot), g_5 (dotted), and g_6 (thick dashed). Optimal model (b), $g^\star \in \mathcal{G}$.

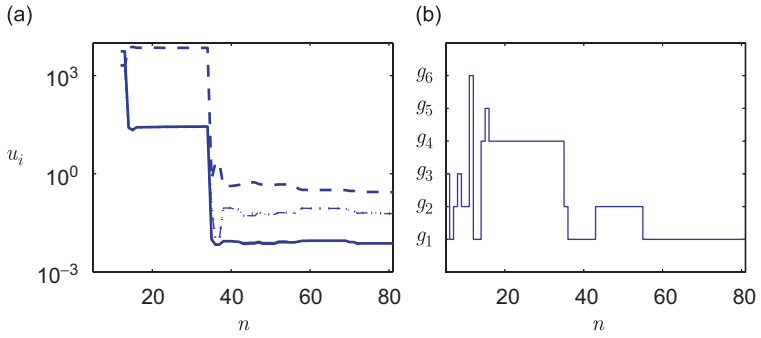

Fig. 15. Model selection by decision-theoretic method for application #2, case #2. Expected utilities (a) for each surrogate: g_1 (thin solid), g_2 (thin dashed), g_3 (thick solid), g_4 (dash–dot), g_5 (dotted), and g_6 (thick dashed). Optimal model (b), $g^\star \in \mathcal{G}$.

An appropriate value for the utility of surrogate model $g_i \in \mathcal{G}$, if surrogate model $g_j \in \mathcal{G}$ is true, is given by

$$U(g_i, g_j) = \tilde{U}(a_i, a_j) = \begin{cases} \beta_1 (a_i - a_j)^2 & \text{if } \xi_{ij} \leqslant \bar{r}, \\ \beta_2 (a_i - a_j)^2 & \text{if } \xi_{ij} > \bar{r}, \end{cases} \tag{38}$$

where $\beta_2 \geqslant \beta_1 \geqslant 0$ denote deterministic parameters. The expected utility of each surrogate model, denoted by $u_1, \ldots, u_6$, is illustrated by Fig. 15(a) for $5 \leqslant n \leqslant 81$; the corresponding optimal surrogate model for each value of n is shown in Fig. 15(b). Parameter $\bar{r} = 30$ was used for calculations. The expected utilities are undefined for $n < 12$ since no value for δ exists to satisfy the design requirement given by Eq. (30b); $g^\star$ is therefore identical to results using the classical method (Fig. 13) for $n < 12$. Together, Figs. 14 and 15 illustrate that, in general, different surrogate models are optimal for the two different design constraints defined by Eq. Eq. (30). This feature further demonstrates that the optimal surrogate by the decision-theoretic method depends on the model use.

4.3. Uncertainty quantification

Device penetration into media such as metal and soil is an application of some engineering interest [25,26]. Often, these devices contain internal components which must survive the severe environment that accompanies the penetration event. In addition, these systems must be robust to perturbations in operating conditions, some of which can only be described to within some level of uncertainty.

A schematic of a penetration system just prior to impact with a soil target is illustrated by Fig. 16(a), where X_1 denotes the orientation angle of the penetrator with respect to the velocity vector, $\boldsymbol{v}$, and X_2 is a parameter characterizing target hardness. A considerable amount of uncertainty exists in our knowledge of X_1 and

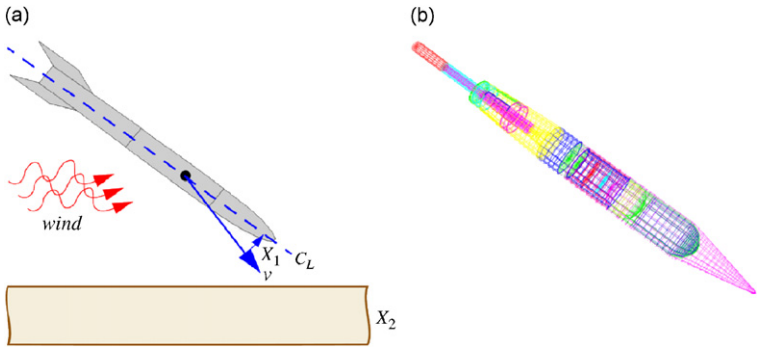

Fig. 16. Penetration system: (a) schematic, and (b) finite element model.

X_2 due to unknown wind conditions and target properties. Accordingly, we model X_1 as a Gaussian random variable with zero mean and unit variance, and X_2 as a lognormal random variable with a mean and standard deviation of 40 and 15, respectively, to represent this uncertainty. Further, we assume X_1 and X_2 to be independent; for additional discussion of these uncertainties, see Ref. [27].

Component response to the penetration event is quantified by the shock response spectrum (SRS) of lateral acceleration at the centroid of a critical internal component; this component is denoted by symbol C. Of particular interest is the peak value of the SRS of lateral acceleration, normalized by a constant and denoted by Y. Our objective is to assess

$$p_F = P(Y > 0.9), \tag{39}$$

the probability that output Y will exceed 0.9, a known, specified threshold. We refer to p_F as the probability of system failure.

A comprehensive FE model for the penetrator, comprised of approximately 60,000 degrees-of-freedom, is illustrated by Fig. 16(b). PRONTO3D [28], a nonlinear transient dynamics FE code developed at Sandia National Laboratories, coupled with a spherical cavity expansion representation for the soil–structure interaction [25,29], is used to predict the transient acceleration response of C during the penetration event. Filtering routines are then used to compute the frequency-domain SRS and corresponding value for Y. However, the calculation of output Y for a single value of X_1 and X_2 using this comprehensive FE model requires over 33 hours of compute time. This renders traditional methods to estimate p_F defined by Eq. (39) by Monte Carlo sampling infeasible; surrogate models for the comprehensive FE model illustrated by Fig. 16(b) are therefore needed.

4.3.1. Candidate models

Results from $m = 5$ independent flight tests of the penetrator are available and define the validation data for the problem. Various environmental conditions were monitored during each test, including values for X_1, the orientation angle of the penetrator at impact, and X_2, a parameter quantifying target hardness. The measured values for X_1 and X_2, as well as normalized measurements of component SRS, are illustrated by Figs. 17a and 18a, respectively. Next, we suppose there is sufficient computer resources to perform $n = 11$ evaluations of the comprehensive FE model for the penetration system, where each evaluation is defined by distinct fixed values for random variables X_1 and X_2. Design of experiment methods (see, for example, Ref. [30]) can be used to select values for X_1 and X_2; the values considered herein are illustrated by Fig. 17(b). The corresponding normalized predictions for component SRS using the comprehensive FE model are illustrated by Fig. 18(b). Together, Figs. 17b and 18b illustrate the available calibration data.

We consider 5 candidate surrogate models for f, given by

$$\begin{aligned}\mathscr{G} &= \{g_1, g_2, g_4, g_5, g_6\} \\ &= \{g(X_1, X_2; \mathbb{P}_2^6), g(X_1, X_2; \mathbb{P}_2^{10}), g(X_1, X_2; \mathbb{E}_2^{11}), g(X_1, X_2; \tilde{\mathbb{E}}_2^{11}), g(X_1, X_2; \mathbb{I}_2^{11})\}\end{aligned} \tag{40}$$

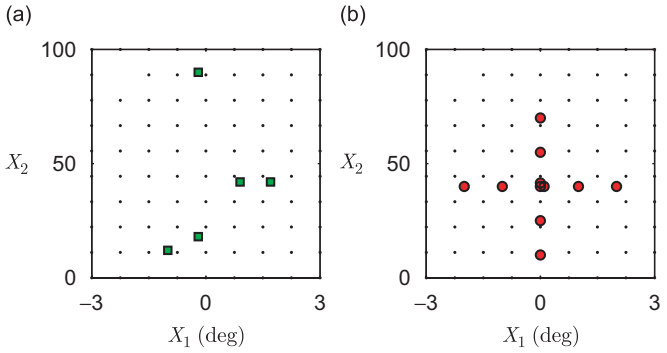

Fig. 17. Locations of $m = 5$ validation data (a) and $n = 11$ calibration data (b) for application #3.

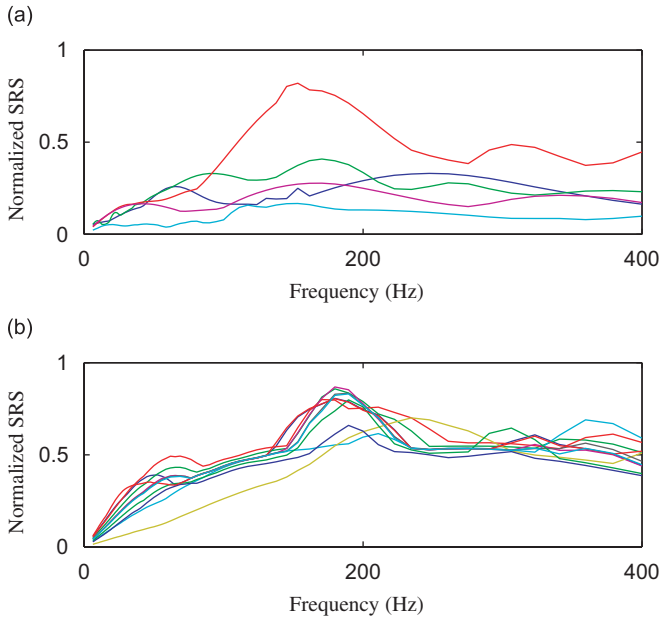

Fig. 18. Normalized shock response spectra (SRS) of internal component response for application #3: (a) $m = 5$ experimental measurements, and (b) $n = 11$ model predictions.

for the purpose of accurately estimating p_F, the probability of system failure defined by Eq. (39). The functional form for each surrogate is defined by Eq. (7), and the parameters for each surrogate are estimated from the calibration data. We note that because $n<15$, there is insufficient calibration data to estimate the coefficients of a fourth-order polynomial and we therefore do not include surrogate g_3 in the collection defined by Eq. (40). Contours of surrogates g_1, g_2, g_4, and g_6 are illustrated by Fig. 19. As for previous applications, the surrogate models are very different, but each is consistent with the available information.

4.3.2. *Optimal model*

The surrogate model probabilities defined by Eq. (3) are listed in the first column of Table 1 assuming $\lambda = 1/2$. By the classical method for model selection, surrogate g_2 is optimal because it has the greatest probability. However, the model probabilities are nearly identical so that there is no clear winner among the collection.

We next apply the proposed decision-theoretic method for model selection so as to explicitly include the purpose of the model, i.e., accurate estimates of system failure, into the selection process; this is not possible

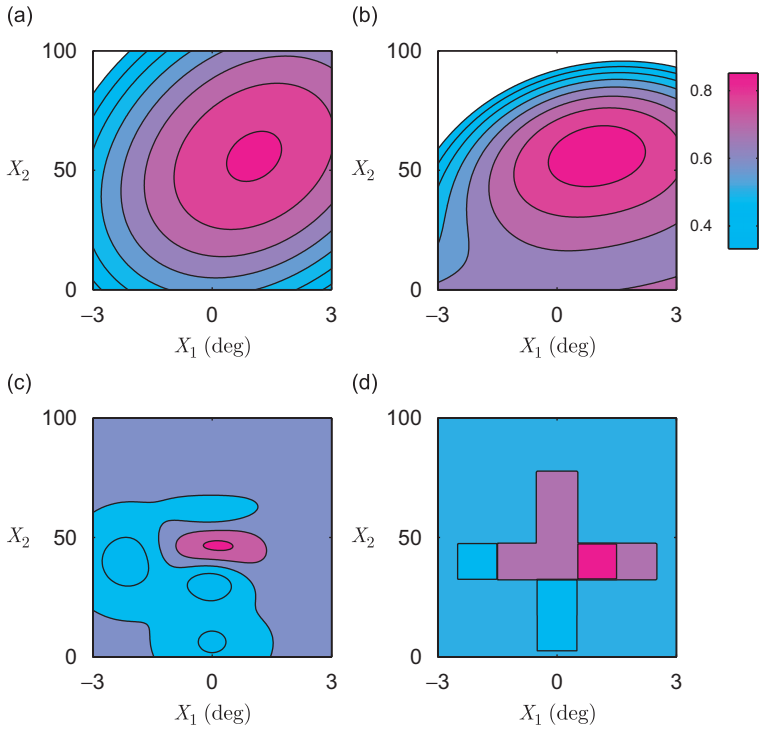

Fig. 19. Contours of surrogate models for application #3: (a) g_1, (b) g_2, (c) g_4, and (d) g_6.

Table 1
Probabilities, p_i, expected utilities, u_i, and probability of failure estimates, $p_{F,i}$, for each surrogate model used for application #3

Model	Classial method	Decision-theoretic method			
	p_i	$p_{F,i}$	u_i ($\bar{\beta}=10$)	u_i ($\bar{\beta}=5$)	u_i ($\bar{\beta}=2$)
g_1	0.200	0	0.0183	0.0091	0.0037
g_2	**0.232**	0	0.0183	0.0091	0.0037
g_4	0.185	0.071	0.0032	0.0032	0.0032
g_5	0.187	0.069	**0.0030**	**0.0030**	**0.0030**
g_6	0.196	0.001	0.0180	0.0090	0.0036

The optimal model is in bold.

with the classical method. Let

$$p_{F,i} = P(g_i(X_1, X_2) > 0.9) \tag{41}$$

denote an estimate for p_F using surrogate $g_i \in \mathcal{G}$. The values for $p_{F,i}$ are listed in the second column of Table 1 illustrating that, with surrogate models g_1, g_2, and g_6, we get estimates of p_F near zero. Models g_4 and g_5 deliver nearly identical estimates for the probability of system failure. Further, these estimates are conservative with respect to the others.

We desire a utility formulation that favors conservative estimates for p_F. An appropriate choice is given by Eq. (22) used for application #1, with ξ_i everywhere replaced by $p_{F,i}$ defined by Eq. (41), since non-conservative estimates for p_F will be assigned a large utility. The expected utilities of each surrogate model are listed in columns 4–6 of Table 1. Three cases are considered, which correspond to three different values for parameter $\bar{\beta} = \beta_2/\beta_1$, where β_1 and β_2 are defined by Eq. (22). For $\bar{\beta} \gg 1$, non-conservative predictions of system failure are highly penalized with respect to conservative predictions; as $\bar{\beta} \to 1$, the utility for

conservative and non-conservative models is identical. As indicated by Table 1, surrogate g_5 is optimal under the decision-theoretic method because it provides conservative estimates for p_F and, consistent with Eq. (6), it has the least expected utility. The optimal model does not change over a wide range of values for $\bar{\beta}$, meaning that precise values for utility parameters β_1 and β_2 are not essential for a rational decision.

5. Conclusions

Methods were developed and applied to select the optimal member from a collection of candidate surrogate models, where each is an approximation for a single comprehensive model. Each model in the collection was consistent with limited information provided by the comprehensive model. Classical methods select the surrogate model that best fits the data provided by the comprehensive model; it was shown that this technique is independent of the model use and, therefore, was inappropriate for some applications. The proposed approach applied techniques from decision theory, where postulated utility functions were used to quantify the model use. A criticism of the method is the possible subjectivity of the utility function since the construction of such a function required a fair understanding of the consequences of unsatisfactory system behavior. We therefore assumed these consequences to be well-specified for all examples considered. Three applications were presented to illustrate the methods. These included surrogate model selection for the purpose of: (1) estimating the minimum of a deterministic function, (2) the design under uncertainty of a simple oscillator, and (3) reliability assessment of a complex engineering system subject to a severe shock and vibration environment. Further application of these methods include constitutive models for material properties of encapsulating foams to mitigate the transmission of mechanical shock loads [10], aerodynamics models for turbulent pressure fluctuations during atmospheric re-entry of a spacecraft [13], and the monitoring of vehicles in the vicinity of critical national assets [14].

Acknowledgments

The author would like to acknowledge M. Grigoriu from Cornell University, and L. Swiler and T. Trucano from Sandia National Laboratories, for their technical comments and review of the manuscript. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

References

[1] W. Gautschi, *Numerical Analysis: An Introduction*, Birkhäuser, Boston, MA, 1997.
[2] M. Buhmann, *Radial Basis Functions: Theory and Implementation*, Cambridge University Press, Cambridge, UK, 2003.
[3] W. Li, S. Padula, Approximation methods for conceptual design of complex systems, *Proceedings of Eleventh International Conference on Approximation Theory*, Gatlinburg, TN, May, 2004, pp. 1–40.
[4] N.A.C. Cressie, *Statistics for Spatial Data*, Wiley, New York, NY, 1993.
[5] A. A. Giunta, L. T. Watson, A comparison of approximation modeling techniques: polynomial versus interpolating models, *Proceedings of 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, St. Louis, MO, September, 1998, pp. 392–404.
[6] J.H. Friedman, Multivariate adaptive regression splines, *Annals of Statistics* 19 (1) (1991) 1–67.
[7] R.G. Ghanem, P.D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Revised ed., Dover Publications Inc., New York, NY, 2003.
[8] G. Kallianpur, *Stochastic Filtering Theory*, Springer, New York, NY, 1980.
[9] R. Jin, W. Chen, T. Simpson, Comparative studies of metamodeling techniques under multiple modeling criteria, *Proceedings of 8th AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, Long Beach CA, September, 2000.
[10] R.V. Field Jr., M. Grigoriu, Model selection in applied science and engineering: a decision-theoretic approach, *Journal of Engineering Mechanics* 133 (7) (2007) 780–791.
[11] M. Grigoriu, D. Veneziano, C.A. Cornell, Probabilistic modeling as decision making, *Journal of Engineering Mechanics* 105 (4) (1979) 585–596.
[12] H. Chernoff, L.E. Moses, *Elementary Decision Theory*, Dover Publications Inc., New York, NY, 1959.
[13] R.V. Field Jr., M. Grigoriu, Optimal stochastic models for spacecraft atmospheric re-entry, *Journal of Sound and Vibration* 290 (3–5) (2006) 991–1014.

[14] R.V. Field Jr., M. Grigoriu, Optimal design of sensor networks for vehicle detection, classification, and monitoring, *Probabilistic Engineering Mechanics* 21 (4) (2006) 305–316.
[15] T. Goel, R.T. Haftka, W. Shyy, N.V. Queipo, Ensemble of surrogates, *Structural and Multidisciplinary Optimization* 33 (3) (2007) 199–216.
[16] C.P. Robert, *The Bayesian Choice*, Second ed., Springer, New York, 2001.
[17] M. Grigoriu, A Decision Theoretic Approach to Model Selection for Structural Reliability, PhD Thesis, Massachusetts Institute of Technology, 1976.
[18] W. Rudin, *Principles of Mathematical Analysis*, Third ed., McGraw-Hill Inc., New York, 1976.
[19] J. D. Griffin, M. S. Eldred, M. L. Martinez-Canales, J.-P. Watson, T. G. Kolda, B. M. Adams, L. P. Swiler, P. J. Williams, P. D. Hough, D. M. Gay, D. M. Dunlavy, J. P. Eddy, W. E. Hart, A. A. Giunta, S. L. Brown, DAKOTA: a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis, Technical Report SAND2006-4055, Sandia National Laboratories, Albuquerque, NM, October 2006. doi:10.2172/895073.
[20] P.E. Gill, W. Murray, M.H. Wright, *Practical Optimization*, Academic Press, San Diego CA, 1981.
[21] A. A. Giunta, M. S. Eldred, J. P. Castro, Uncertainty quantification using response surface approximations, *Proceedings of 9th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability*, Albuquerque NM, July, 2004.
[22] R.Y. Rubinstein, *Simulation and the Monte Carlo Method*, Wiley, New York, NY, 1981.
[23] T.T. Soong, M. Grigoriu, *Random Vibration of Mechanical and Structural Systems*, P T R Prentice-Hall, Englewood Cliffs, NJ, 1993.
[24] V. Picheny, N. Kim, R. T. Haftka, Conservative estimation of cumulative distribution function for probability of failure calculation, *Proceedings of 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Portsmouth, VA, September, 2006.
[25] M.J. Forrestal, V.K. Luk, Penetration into soil targets, *International Journal of Impact Engineering* 12 (3) (1992) 427–444.
[26] B.R. Sorensen, K.D. Kimsey, G.F. Silsby, D.R. Scheffler, T.M. Sherrick, W.S. de Rosset, High velocity penetration of steel targets, *International Journal of Impact Engineering* 11 (1) (1991) 107–119.
[27] R. V. Field, Jr., J. R. Red-Horse, T. L. Paez, Reliability analysis of penetration systems using nondeterministic methods, *Proceedings of 18th International Modal Analysis Conference* (*IMAC*), Society of Experimental Mechanics (SEM), San Antonio, TX, 2000, pp. 363–368.
[28] S. W. Attaway, K. H. Brown, F. J. Mello, M. W. Heinstein, J. W. Swegle, J. A. Ratner, R. I. Zakoks, PRONTO3D Users instructions: a transient dynamic code for nonlinear structural analysis, Technical Report SAND98-1361, Sandia National Laboratories, Albuquerque, NM, January 2000. doi:10.2172/291042.
[29] M. Forrestal, D.Y. Tzou, A spherical cavity-expansion penetration model for concrete targets, *International Journal of Solids and Structures* 34 (31–32) (1997) 4127–4146.
[30] G.E.P. Box, D.W. Behnken, Some new three level designs for the study of quantitative variables, *Technometrics* 2 (4) (1960) 455–475.