ABSTRACT

Response surface methods were originally investigated to provide a mechanism for guiding experimentation in search of optimal settings for design parameters or optimal values of an unknown response. Recently, many additional applications have sparked a resurgence of interest in response surface (or meta-model) methods in the statistical and engineering literatures. Recent literature has addressed more flexible functional forms for modeling the response, new methods of response surface construction, alternate approaches to updating the surface estimate, new sampling methods upon which the response surfaces are based and extensions that can be used to quantify uncertainty in the response and/or in performance characteristics based on the response. The objective of this paper is to provide an overview of response surface methods covering these new developments and reviewing recent applications.

Keywords: Meta-model, Response surface, Surrogate models, Experimental design

1. Introduction

The terminology “response surface” and “response surface modeling” have historically taken a number of meanings in different disciplines, and still have different interpretations in different bodies of literature. The review papers Hill and Hunter (1966), Myers et al. (1989) and Myers (1999) discuss the statistical developments of response surface methodology primarily as an experimental design approach for selecting design parameters for experiments with the objective of optimizing some function of a response. Recently, many additional applications (largely a consequence of the increased use of computational analyses) have broadened the range of application of response surface methods in the statistical and engineering literature. Recent literature has addressed more flexible functional forms for modeling the response, new methods of response surface construction, alternate approaches to updating the surface estimate, new sampling methods upon which the response surfaces are based and extensions that can be used to quantify uncertainty in the response and/or in performance characteristics based on the response. The objective of this paper is to provide an overview of response surface methods covering these new developments and reviewing recent applications.
In many fields of engineering, the term “response surface” is used synonymously with “meta-model” or “surrogate model”, which refer to any relatively simple mathematical relationship between parameters and a response, often based on limited data. We use this later interpretation of a response surface in this paper where we discuss both simple and fairly involved mathematical parameter-response relationships. The terms response surface and meta-model are used interchangeably. While meta-models are still applied frequently to optimization problems, a number of additional applications are quite common. These include: reliability analysis; uncertainty quantification; and model-based prediction (both interpolation and extrapolation). In these application areas, the response surface may be, and often is, utilized to facilitate the selection of additional experiments or simulations. It is also used, however, in other stages of an analysis. The role of the response surface is broadening with new developments in systems design and analysis.

Figure 1 gives a simple illustration of a basic response surface in two dimensional “input” space. The horizontal axes are parameters of the analysis while the vertical axis gives the response at any “point” or “location”. The stars indicate response values $y_i(x_i)$ or just $y_i$, $i = 1,...,n$ at the parameter locations $x_i$ where experiments were performed. We use the term “experiments” to refer to either physical experiments or simulations. Furthermore, we refer to the planning for experimentation as the "experimental design" – whether it is the initial plan or one for supplement experimentation. The parameters of Figure 1 can be either (1) “design parameters” presumably at the control of the experimenter and (2) “environmental variables” that are often assumed to have an associated probability distribution (as indicated in Figure 1). These are parameters that are (typically) controlled for the experiment but are considered variable in an application. Optimization problems usually involve design parameters while reliability or uncertainty quantification applications often involve environmental variables. Some applications involve both types and some parameters don’t fit neatly into either of these categories. We refer to either type as inputs unless further clarity is required.

Often, there is some “performance function” of the response that is of ultimate interest. The performance function can be an important element of the response surface application. Some example applications are listed in Figure 1; but details are deferred to Section 4.

Formally, one can represent a response surface $r$ as a function of the parameters or variables $x$ of the design space $X$ (generally some subset of Euclidian $k$-space, $\mathbb{R}^k$) and of parameters, $\theta$, associated with the family of functions used to construct the surface. For example, a quadratic polynomial response surface in two dimensions will have the form:

$$r(x, \beta) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2. \quad (1)$$

In this case, $k = 2$ and $\theta = \beta$, the vector of surface parameters is six dimensional. The parameters $\beta$ are typically estimated from the experimental data for this model. As a second example, consider a kriging model:
\[ r(x, \beta) = P(x, \beta) + \epsilon(x, C(\xi)) \]

where \( P(x, \beta) \) is a polynomial (perhaps as in eqn. (1)) and \( \epsilon(x, C(\xi)) \) is a zero-mean Gaussian Process (described in Section 2) with spatial covariance function \( C(\xi) \). In this case \( \theta = \{ \beta, \xi \} \). The additional parameters required for full specification of \( C(\xi) \) may be difficult to estimate directly from limited data. Some of the more complicated families of response functions have parametric forms requiring substantial data or strong prior assumptions for their estimation.

The remainder of this paper is composed of three sections covering issues that have received significant attention in the response surface literature recently. Section 2 covers the functional form of the response surface. Issues contrasting the different types of surfaces are discussed. Section 3 addresses response surface construction. Experimental design alternatives are discussed both from the perspective of an initial design and refinements to an existing response surface. Section 4 is focused on common applications. Examples are given for the applications that were listed in Figure 1.

2. Response Surface Families

Issues

The “family” of surfaces selected to model a response can have a substantial impact on the results of an analysis. Figure 2 illustrates a “true” response and three meta-model estimates that are based on the same 15 experimental response values. In selecting a family of response surfaces, one must consider the flexibility of the surfaces. The family of surfaces should be capable of attaining surfaces that meet specific smoothness requirements of an application. Another consideration is the “noise” or variability anticipated (or seen) in the response measurements. Response noise, considered together with the quantity of data available, can impact the surface selection. There is often a tradeoff between assumptions and data requirements that plays a role in the selection. A third consideration in selecting the surface family is the objectives of the analysis. Applications that require estimates of response uncertainty may benefit from using one family or another.

![Figure 2](image)

Figure 2. Examples of different meta-models based on the same response data. The subfigures are: (a) “true” surface; (b) nearest neighbor estimate; (c) kernel regression estimate; and (d) kriging estimate.
In the next subsection, we list and describe briefly several methods for estimating the response. Five families of surface are described. These alternatives are by no means comprehensive; others are available. The final subsection we reflect back on the issues discussed above with reference to the specific surface families.

Surfaces

Non-parametric Regression

Non-parametric regression methods provide some of the simplest approaches to constructing a response surface. Two examples follow. The \( K \)-nearest-neighbor methods uses as the response estimate \( r(x) \), the average value of the \( K \) responses at experimental points closest to \( x \). Figure 2b shows the two-nearest-neighbor response surface. Kernel regression is another method that constructs the response estimate \( r(x) \) as an average of neighboring values; in this case a weighted average. Weights are determined by a kernel function that, generally, has parameters associated with it. The parameters are adjusted (based on the response data) to achieve a balance between variability and bias of the estimated surface. Figure 2c is the kernel regression response estimate based on a triangular kernel function.

Radial Basis Functions

Radial basis functions, see Orr (1999), are radially-symmetric functions that are used to construct the meta-model as:

\[
r(x) = \sum_{i=1}^{n} a_i \phi(\|x-x_i\|),
\]

where the weights \( a_i \) are computed, using the \( n \) response data in the above equation and solving the system of equations, \( \phi \) is a univariate function defined on \([0,\infty)\) and \( \|\cdot\| \) is the Euclidian norm on \( \mathbb{R}^k \). A common choice for the basis functions are the Gaussian functions:

\[
\phi(\|x-x_i\|) = e^{-\|x-x_i\|^2/\sigma^2}.
\]

Polynomial Regression

Polynomial regression methods are popular methods for creating a response surface because the calculations are simple and the resulting function is a closed-form algebraic expression. For example, a quadratic polynomial has the form:

\[
r(x) = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{j=1}^{k} \sum_{j=1}^{k} \beta_{ij} x_i x_j , \quad (2)
\]

where \( \beta_0, \beta_i \) and \( \beta_{ij} \) are estimated coefficients. The number of sampling points \( n \) must be greater than or equal to the number of terms in the polynomial. When there are more sampling points, the system of equations is over determined and a regression procedure is invoked to solve for the coefficients. In this case, the surface does not, in general, match the response values at the sample points. The method of least squares is usually used in the coefficient estimation process. Later in this paper we will refer to the right-hand side of eqn. (2) as a second-order model and to the first two terms of the right-hand side as a first-order model.
Multivariate Adaptive Regression Splines

The multivariate adaptive regression splines (MARS) function approximation method (Friedman, 1991) is based on a recursive partitioning algorithm involving truncated power spline basis functions. The form of the MARS response model is:

\[ r(x) = a_0 + \sum_{m=1}^{M_1} a_m B_m(x) + \sum_{m=M_1+1}^{M_2+M_3} a_m B_m(x_i, x_j) + \ldots, \]

where the \( B_m \) terms are the basis functions, the \( a_m \) terms are the coefficients of the basis functions, \( M_1 \) is the number of one-parameter basis functions, and \( M_2 \) is the number of two-parameter basis functions. Often, either linear or cubic spline basis functions are used. The regression aspect of the MARS algorithm involves a forward/backward stepping process to adaptively add/remove spline basis functions from the model. It is this regression process that generates the \( a_o \) and \( a_m \) terms. The resulting MARS model is a \( C^2 \)-continuous function that will not, in general, match the response data exactly. Like polynomial regression, MARS has the ability to create smooth approximations to noisy data.

Gaussian Processes and Kriging

One family of surfaces that is often used to model a response is the Gaussian Process model \( \eta(x) \) satisfying the property that for any integer \( l \), the joint distribution of \( \eta(x_1), \eta(x_2), \ldots, \eta(x_l) \) is multivariate normal for any locations \( x_1, \ldots, x_l \in \mathbb{X} \), the design space. The general form of the model is parameterized through its mean function \( m(x) \) and spatial covariance function \( \sigma^2 C(x, x') \) for input vectors \( x \) and \( x' \). The mean function is often assumed to be constant but has the general form of a linear model:

\[ m(x) = \sum_{i=1}^{p} \beta_i f_i(x), \]

where the \( f_i \) are arbitrary functions of \( x \) and \( \beta_i \) are their coefficients. The spatial correlation function \( C \) can take a number of parametric forms; see Cressie, 1991. One fairly general correlation function is:

\[ C(x, x') = \prod_{i=1}^{k} C_i \left( \left| x_i - x_i' \right|^q \right) = \prod_{i=1}^{k} e^{-\phi_i \left| x_i - x_i' \right|^q}, \]

for any \( x, x' \in \mathbb{X} \) and parameters \( \phi_i \) and \( q \) and where the \( C_i \) are spatial correlation functions in one dimension. These parameters may be estimated from the data.

One surface estimate based on the Gaussian Process is the kriging estimate. Kriging interpolation techniques were originally developed in the geostatistics and spatial statistics communities; see Matheron 1963. The basic notion that underpins kriging is that the sample response values exhibit spatial correlation (i.e., samples taken close together are likely to have highly correlated response values, whereas samples taken far apart are not). Kriging methods have found wide utility due to their ability to accommodate irregularly spaced data, their ability to model general surfaces that have many peaks and valleys, and their exact interpolation of the given sample response values. The response estimate:

\[ r(x) = \sum_{i=1}^{n} \lambda_i(x)y_i, \]
for any \( x \) and response data \( y_i, i = 1, \ldots, n \) is established by setting up and solving the “kriging equations” for the \( \lambda_i(x) \). This process is also described in Cressie (1991).

### Comparing Surfaces

Reflecting back to the issues set forth at the beginning of this section, it is possible to make some general comments about the major uses of some of the response surface families. “Smoothing” surfaces like the polynomial regression are used most frequently where the response is thought to be adequately modeled using a low-order polynomial and where the data are assumed to contain substantial noise and/or where too few response data are available to capture more detailed failures of the true response. A (possible) advantage of the regression surfaces is that they maintain their “structure” in areas where there are no (or are only sparse) data. By contrast, the non-parametric regression surfaces are determined entirely by the values of neighboring responses. If one is reluctant to assume some overall structure to the region, or if the true surface is thought to have detailed features and resources are available for experimentation to identify these features, these methods may be preferable.

Most other surface families attempt to capture the advantages of both of these approaches. The structure of some of the other families listed above can incorporate a structural component based on the regression model. Alternatively, structure can be provided through the basis functions.

A number of more detailed studies have been published. This task is complicated by the issues listed at the beginning of this section, as well as by the methods used in selecting the experiments (i.e., the experimental design strategy). Different surface families may be preferred for different applications. In spite of these complications, empirical studies can provide useful guidance to surface selection for specific problems. Two are listed below that investigate the meta-model accuracy for a number of experimental design strategies and sample sizes.

**Simpson, et al. (2002)** compare regression models, kriging surfaces, radial basis functions and MARS in terms of their ability to estimate surfaces for two engineering examples of substantially different dimension. Four metrics were evaluated, but because some appeared redundant, only maximum error and root mean square error (RMSE) were reported. Sample sizes were also varied. Their conclusions were that kriging and radial basis functions tend to offer more accurate approximations across, sampling designs and for a wide range of sample sizes. Large sample sizes were required for accuracy with the MARS surface estimate.

**Swiler et al. (2006)** compare kriging, polynomial regression, and MARS surfaces for a number of experimental designs, according to RSME, mean absolute deviation and a mean absolute error criteria. Analysis of variance is used to assess performance difference between the surface families. This study is not yet complete.

### 3. Response Surface Construction and Updating

There are a number of different approaches to experimental design supporting an analysis involving meta-modeling. An appropriate design may depend on the amount of noise in the response, the type of response surface and on the intended use of the surface. In this section, we discuss three approaches to experimental design that we call “space-filling” designs, “classical designs”, and “performance-directed” designs.

#### Sampling and Space Filling Designs

The following sampling designs are described assuming that some probability distribution is available for the input parameters and variables. In situations where this is not the case, a uniform spacing of points can be achieved by assuming a uniform distribution over the ranges of each of the parameters. Assume we are looking for a sample of size \( n \).
Monte Carlo Sampling

Monte Carlo sampling is the simplest sampling approach to determining points for experimentation. It consists of choosing each point \( x \) independently from the joint parameter distribution \( F(x_1, \ldots, x_n) \). For cases where some or all of the inputs are design parameters, a uniform distribution of points can be obtained by assigning them a uniform probability distribution over the appropriate subset of \( X \). The appeal of Monte Carlo sampling is its simplicity. However, for most cases, more efficient methods can be employed. It is often used for problems where the experimentation is inexpensive.

Latin Hypercube Sampling and Orthogonal Arrays

Latin hypercube sampling (LHS), McKay et al. (1979) is a sampling method developed to address the need for more efficient uncertainty assessment. LHS partitions the parameter space into bins of equal probability, with the goal of attaining a more even distribution of sample points in the parameter space than typically occurs with pure Monte Carlo sampling.

Latin hypercube sampling is a stratified sampling method that selects \( n \) different values from each of \( k \) variables \( x_1, \ldots, x_k \) in the following manner. The range of each variable is divided into \( n \) non-overlapping intervals on the basis of equal probability. One value from each interval is selected at random with respect to the probability density in the interval. The \( n \) values thus obtained for \( x_1 \) are paired in a random manner (equally likely combinations) with the \( n \) values of \( x_2 \). These \( n \) pairs are combined in a random manner with the \( n \) values of \( x_3 \) to form \( n \) triplets, and so on, until \( n \) \( k \)-tuples are formed. This is the Latin hypercube sample. The pairing can be accomplished through random ordering (as suggested above) or directed to induce a correlation structure into the sample. One method for inducing correlations among the input variables is given in Iman and Conover (1982).

Latin hypercube sampling is an example of a more general class of space-filling designs that can be constructed from an orthogonal array. See Owen (1992) or Tang (1993) for details on orthogonal array designs.

Halton Sampling

Halton sampling, see Halton, 1960, Halton and Smith, 1964; and Kocis and Whiten, 1997 is one example of a sampling approach known as quasi-Monte Carlo sampling. The goal of quasi-Monte Carlo methods is to produce sequences which have low discrepancy, where discrepancy refers to the non-uniformity of the sample points. Quasi-Monte Carlo methods produce low discrepancy sequences, especially if one is interested in the uniformity of projections of the point sets onto lower dimensional faces of the hypercube (usually 1-D: how well do the marginal distributions approximate a uniform distribution?). The quasi-Monte Carlo Halton sequence is a deterministic sequence determined by a set of prime bases. The Halton sequence in base 2 starts with points 0.5, 0.25, 0.75, 0.125, 0.625, etc. The first few points in a Halton base 3 sequence are 0.33333, 0.66667, 0.11111, 0.44444, 0.77777, etc. Notice that the Halton sequence tends to alternate back and forth, generating a point closer to zero then a point closer to one.

Classical Experimental Design

Classical experimental designs (see, for example John (1971)) can contribute in at least two ways to methodology for meta-model based analysis. First, screening methods often contribute the initial points of experimentation to obtain information on which parameters and/or variables are contributing substantially to differences in response. Screening methods are used to reduce the dimension of the design space where applicable. Screening Designs include supersaturated designs, Plackett Burman designs, and highly fractionated factorial designs. A second set of classical experimental designs that are useful in response surface modeling are second order response surface designs like Box-Behnken or central composite designs developed to efficiently construct quadratic polynomial response surfaces.
Performance-Directed Designs

A third category for response surface designs is what we call "performance-directed" designs. The objective of these designs is to focus the experimentation on the region of the design space that most critically affects performance. In the introduction, we listed several possible performance functions associated with an analysis that might utilize a meta-model. Two of those listed, optimization and reliability analyses, would not, in general, benefit most from a uniform space-filling design or from a classical design constructed to give precise prediction capability across the design space. In most of these types of applications, there are specific regions of the design space where accuracy and precision are critical and other regions where response precision is of little interest.

In order to accomplish this focus for experimentation, response surface methods have been developed to quantify how "fit" a surface is for addressing a specific objective. One major use of fitness criteria is to evaluate the merits of additional experimentation. The terminology "expected improvement" is used in comparing "candidate" sample locations (before the experimentation is performed) to see which is likely to provide the best fit. The candidate location(s) that yields the highest expected improvement (in fitness) are then selected and the experimentation is then performed.

Evaluating fitness involves assessing meta-model uncertainty as is illustrated in Figure 3. Point-wise uncertainty (see Figure 3a) is used for this assessment in some applications. Gaussian process models are frequently used in this capacity. Examples are given in first three applications of Section 4. The probability measure approach "response modeling" is described in Rutherford (2006a) as an alternative that can be used for more general fitness criteria (see Figure 3b).

The response model is a set of discrete response surface “realizations” or “elements” that are used to construct an atomic probability measure over the response. Probabilities associated with specific events based on the response model are approximated by the fraction of realizations in the ensemble that satisfy the event. For an optimization problem, for example, we estimate the probability of the optimal location being contained in a specific region of the input space as the fraction of the realizations with minimum values in this region. The objectives in building the response model are somewhat different than those involved in trying to construct an average or expected response. Differences between the realizations attempt to capture the uncertainty in the true response surface. Rutherford (2006a) provides a detailed description of the process for constructing a response model. Results using this characterization of uncertainty are shown in the last two examples in Section 4.

Figure 3. Two different ways to characterize response uncertainty: (a) point-wise uncertainty – the response surface estimate, at any point $\mathbf{x}$, is accompanied by an estimate of the uncertainty in the surface at $\mathbf{x}$; (b) atomic response measure – probabilities associated with the response are represented through the discrete response elements.
4. Applications

In this final section, we discuss several possible applications for meta-model analysis. These include prediction, uncertainty quantification, optimization and reliability analysis. For these applications, we discuss possible considerations guiding the selection of the response surface family, experimental design considerations and we provide additional references.

Prediction

Prediction (point-wise, for a specific value of \( x \)) is the most straightforward application of response surface methods. The idea is to use the surface value \( r(x_p) \) as a prediction for the response at point \( x_p \) where, in general, there has been no experimentation. Prediction may involve interpolation or extrapolation, and this intended use could impact the choice of surface family. Interpolation implies that \( x_p \) is in the neighborhood of some of the experimental points \( x_1, \ldots, x_n \). Technically, interpolation points are interior to the convex hull of the design points and extrapolation involves points outside this region.

The choice of a response surface family is guided by the considerations above (in Section 2). The additional concern with prediction at points involving extrapolation is that the surface will often be more accurate if it contains some “structural component”. Polynomial regression, for example, will provide predictions outside of \( X \) that are based on the surface structure while the kriging estimate with a constant mean will tend toward that mean as \( x_p \) moves further away from the conditioning data. The nearest neighbor methods will use the response values corresponding to the closest sample points as the basis for estimating \( x_p \).

A number of fitness criteria are available for evaluating the effectiveness of a meta-model used in prediction problems. Sacks et al. 1989 use a Gaussian process model to investigate and provide optimal designs for computer experiments using the fitness criteria: integrated mean squared error, maximum mean squared error and minimum expected posterior entropy. A more general discussion of “criterion-based” designs is given in Santner et al. (2003).

Uncertainty Quantification

Uncertainty quantification refers to the process of propagating input uncertainty through a computational model in order to estimate and characterize the response uncertainty. In the general case, this is accomplished separately for any design parameter value \( x_d \) of interest where the uncertainty is evaluated across environmental variables. Often the computation is made over the range of the design parameters and the environmental variable uncertainty is assumed to be independent of the design parameter location. Technically, for the former case above, the distribution function for design parameter \( x_d \), is computed as:

\[
F(y, x_d) = \int_{X^E} I(r(x) < y) dF(x_e),
\]

where \( I \) is an indicator variable with value 1 when \( r(x) < y \), \( X^E \) represents the space of environmental variables and \( x_e \) specifies the environmental variables.

The role of the response surface is that of providing the estimate \( r(x) \) in the equation above. One can see that the experimental design for uncertainty quantification will benefit from a strategy that reduces the response uncertainty most at points that are more likely to occur in application. Any of the sampling or space filling designs might be appropriate. To use one of the performance directed designs for supplemental experimentation, the integrated mean squared error fitness criterion are provided in Sacks et al. 1989 can be used.
Optimization

For optimization problems, the response surface is used to locate and/or indicate the design parameters associated with an optimal or near-optimal response. In applications where there are both design parameters and environmental variables, the objective is often to find values of the design parameters that optimize the expected response, or some other criteria taken with respect to the environmental variables. The challenging part of these applications is to find the optimal values with a minimal experimental effort.

Figure 4 indicates two possible approaches to optimization using response surface methods. The surface in Figure 1 is used as an example. In the analysis corresponding to Figure 4a, we consider a maximization problem where the analyst may be willing to assume a single maximum response exists in the design region. We assume further that there is substantial noise associated with the response measurements. By contrast, in the analysis corresponding to Figure 4b we consider a minimization problem where the response measurements are fairly precise (or exact). In the discussions below for these two different optimization scenarios, we assume further that sequential or multi-staged sampling is possible so that we can use an "adaptive" experimental design utilizing information from previous experiments.

![Figure 4](image)

Figure 4. Illustration of two different optimization problems. (a) illustrates a possible set of experiments using a crude version of classical RSM for a maximization problem. (b) shows the actual results for a minimization problem using response modeling. The global minimum is indicated by the red square.

Optimization Using Smooth Response Surfaces

Figure 4a illustrates one possible sequence of experiments that might be the result of a classical statistical approach to optimization through response surface modeling (RSM). In that figure, the different-colored circles indicate different stages of experimentation starting near the middle of the design space and terminating in the upper right. Each circle might (depending on the noise level) indicate multiple experiments. The groups of four circles indicate the design and subsequent analyses for a first-order surface, while the single 5-point design (black) indicates the fitting of a second-order response surface. A crude outline of a simple version of this process is given next.

a) Choose a starting point in the design space and perform an initial first-order design and analysis.

b) Based on the current results, perform experiments along the direction of maximum ascent (for a maximization problem) until a decrease in response is detected. Move back to the preceding point.
c) Perform another first-order experiment about the current point. Test for a significant gradient. If found, return to step 2. Otherwise, supplement the design to get second-order information, and fit a second order polynomial response surface.

This crude process can be (and has been) improved in many ways, Khuri and Cornell (1987), Chapter 5, for example, gives a detailed outline for an RSM approach.

*Global Optimization Using Response Surface Methods*

The minimization problem indicated in Figure 4b has 3 local minima at points (0, 0) value 8, (0,1) value 9.42, and (1,.252) value 7.68. The sequential design algorithms based on meta-models choose the next experimental point based on expected improvement of a fitness function. For this application, improvement can be defined according to one of a number of different criteria. One would be -- choose the point \( \mathbf{x} \) that maximizes:

\[
\int (r(\mathbf{x}) < m) dG(r(\mathbf{x}))
\]

(for a minimization problem) where \( m \) is the current minimum and \( G(r(\mathbf{x})) \) is the uncertainty distribution approximation of the response surface at the point \( \mathbf{x} \). This distribution \( G \) is often approximated using a Gaussian Process model. This is roughly the approach taken in Jones, et al. (1998), and later in Huang, et al. (2005). Rutherford (2006b) used a different fitness criteria based on reducing the design space uncertainty:

\[
\sum_{i=1}^{m} \| \mathbf{x}_i^* - \bar{x}^* \|,
\]

where \( \mathbf{x}_i^* \) is the minimum value on the \( i \)th element of the response measure (see Figure 3b and discussion), \( \bar{x}^* \) is the spatial average of these minima and \( \| \| \) is the Euclidian norm on \( \mathbb{R}^k \). The result of using this approach is shown in Figure 4b, showing 26 points of experimentation. Fifteen were used to calculate the initial response model and the remainder were chosen to search for the minimum point.

*Reliability*

Consider the problem where the environmental variables \( x_1 \) and \( x_2 \) have the joint probability distribution \( F(x_1, x_2) \) and the response \( y \) is considered safe when \( y(x_1, x_2) < T \) and failed otherwise. The brute-force approach to estimating reliability for this problem is to sample the pairs \( x_1, \ldots, x_n \) from the distribution \( F \) and estimate reliability as:

\[
\frac{1}{n} \sum_{i=1}^{n} I(y(x_i) < T),
\]

where \( I \) is an indicator function with value 1 when \( y(x) < T \). The engineering literature has addressed a number of approaches to reducing the experimental requirements for this problem; the most popular being reliability methods or limit-state methods (see Hasofer and Lind (1974)). These methods were established to focus the experimentation about the region of highest probability, where the system would fail (the "design point"). A number of related approaches have made direct use of a polynomial regression response surface based on data taken in this same region (see Kim and Na (1997) and Zheng and Das (2000)).
In general, response surface methods applied to this problem use as a reliability estimate:

\[
\int_X I(r(x) < T) dF(x),
\]  

(4)

where \( I \) is again an indicator function. Since \( r(x) \) is simple to evaluate, the above integration is easily performed (perhaps numerically). The challenge in applying response surface methods is in choosing an experimental design that will provide a precise estimate for reliability using a minimum number of sample points. The precision in \( r(x) \) at points where \( r(x) \gg T \) or \( r(x) \ll T \) is not of interest. Experimentation is important in identifying the points where \( r(x) = T \).

To illustrate how the performance directed experimental design approaches might work to make better use of the meta-model for this application, consider Figure 5 below. In this example, Rutherford (2005) used the probability measure approach to select an optimal (3-point) supplemental design for a reliability problem illustrated using the response in Figure 1. The results in Figure 5 show the experimental design points selected for a base-case reliability problem (eqn. (4)) with threshold \( T = 10 \) and uniform distributions for the inputs \( x_1 \) and \( x_2 \). Fitness of the response surface was considered to be the variance in the reliability estimate given in eqn. (4). The 3 base-case design points are indicated by the 3 fuchsia colored x’s. Note how these experiments are focused on the contour \( T = 10 \).

Figure 5. Differences in experimental design to accommodate analysis objectives. The squares (blue), circles (green), triangles (red), and x’s (fuchsia) represent optimal 3-point designs for four different problems. The point above (2) has both a square and triangle. The dots (black) are the initial data upon which the response model is based.

Several optimal designs for related problems are also shown in Figure 5. The assumptions for the other designs differed from this standard reliability problem in several ways. The point was to illustrate how the change in analysis objectives affects the selection of experimental points. Specifically, a loss function \( (y(x) - T)^2 \) was applied to responses where \( y(x) > T \) so that estimating the quantity

\[
\int_X I(y(x) > T)(y(x) - T)^2 dF(x),
\]  

(5)

was the objective of the analysis. Fitness was considered to be the variance in this estimate. The following assumptions were made for these optimal designs shown in Figure 5.
- Blue squares – eqn. (5) was used (note that the optimal design is changed to a region of the higher values that have more impact).
- Green circles -- eqn. (5) was used and the uniform distribution on $x_2$ was changed to a normal distribution with mean .5, and standard deviation .25 (note that the optimal design is in the region of high probability).
- Red triangles -- eqn. (5) was used with $T$ was changed to 12 (note that the optimal design is moved to a region of even higher values).

In summary, recent engineering applications have made substantial use of response surface methods. A number of different surface families have been used for different applications. In many of these applications fitness criteria associated with the surface and the application are used to select supplemental experimental design points so that the surface can be used more efficiently.

References


Myers, R. H., “Response Surface Methodology--Current Status and Future Directions” (with

Adaptive and Neural Computation, Edinburgh University, Scotland, 1999.


Rutherford, B., “A Response-Modeling Alternative to Surrogate Models for Support in Computational
Analyses”, to be published in Reliability *Engineering and System Safety*, 2006a.

Rutherford, B., “A Response Modeling Approach to Global Optimization and OUU”, to be published
in *Optimization and Engineering*, 2006b.


Santrner T. J., Williams, B. J. and Notz, W. I., *The Design and Analysis of Computer Experiments*,


Response Surface Approximations”, submitted to AIAA Nondeterministic Approaches Conference,


Zheng, Y. and Das, P. K., “Improved Response Surface Method and its Application to Stiffened Plate

**Acknowledgement**

Sandia is a multi-program 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.