

A RESPONSE SURFACE TEST BED

WILLIAM R. MCDANIEL AND BRUCE E. ANKENMAN*

Department of Industrial Engineering and Management Sciences, Northwestern University, 2145 Sheridan Road, Tech C210, Evanston, IL 60208-3119, USA

SUMMARY

A method is presented for creating randomly generated polynomial functions to be used as a test bed of simulated response surfaces. The need for the test bed to perform empirical comparisons of experimental design strategies is discussed and the methods used to create the surfaces are explained. An important feature of the test bed is that the user can control some of the characteristics of the surfaces without directly controlling the surface functions. This allows the user to choose the types of surfaces on which a simulation study is run while preserving the random nature of the surfaces needed for a valid simulation study. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: experimental design; polynomial models; effect heredity; hierarchical model; random function; simulation

1. INTRODUCTION

The experimental study of a response surface for finding optimal or at least desirable settings for the factors is known as Response Surface Methodology (RSM) (see Myers and Montgomery [1]). Many classes of experimental designs have been developed for RSM, such as factorials, fractional factorials, Box–Behnken designs, and central composite designs. A natural question is *How well do these designs work?* That is, how well is the true response surface function modeled by the results of an experimental design? While this may seem like a simple question, in practice it is very difficult to answer because the true response surface function is almost never known. One way that researchers have tried to compensate for this problem is to assume a form for the response surface function and then create experimental designs that will optimally fit a function with that form. Many different criteria (usually involving minimization of the variance of some estimators) have been proposed for design optimization [2].

Optimal experimental design is difficult to apply when the form of the response surface is unknown or when the criteria of interest depend on the true response surface function. The concept of optimal design also encounters difficulty when, instead of a single experimental design, a design strategy

composed of several stages of experimentation is used, as in traditional RSM (see Box and Wilson [3]). We are interested in using a simulation to determine how an experimental design or a design strategy performs when presented with different types of response surface functions. To this end, we have developed a test bed that will randomly generate polynomial functions to represent response surfaces in a simulation study. The characteristics of the surface functions created by the test bed are controllable, allowing researchers to specify the types of surface functions on which design strategies will be tested. When an experimental design is applied to a response surface created by the test bed, random errors from a user-specified error distribution are added to the response to simulate observations with noise. In a simulation study, this allows us to compare the fitted surface from an experimental design with the true response surface function to determine the performance of that design.

An important contribution of this research is the ability to control the types of surface functions created by a simulation. Our methods do not precisely control the surface functions, but instead probabilistically control aspects such as the response range and the maximum number of stationary points. Through this control, we are able to create a wide variety of surface types for testing experimental designs. The test bed can then be used to determine how well a particular experimental design performs on specific functional forms or to obtain performance results

*Correspondence to: B. E. Ankenman, Department of Industrial Engineering and Management Sciences, Northwestern University, 2145 Sheridan Road, Tech C210, Evanston, IL 60208-3119, USA.

that demonstrate the experimental design's robustness across a wide variety of response surface functions.

Work similar to ours was carried out by Smith [4,5], who studied optimization techniques in a simulation environment using empirical studies, and Barton [6], who compared several minimization algorithms on a set of test problems in the presence of random noise.

The work of Hamada and Wu [7] also gives some insight on the expected form of response surface functions. They examined Plackett–Burman designs to obtain information on interactions from initial screening designs, and base their work on the principles of *effect sparsity* and *effect heredity*. Effect sparsity refers to the conjecture that only a few of the many proposed factors are very important in determining the outcome of an experiment. Effect heredity refers to the relationship between the main effects and any two-factor interactions. Hamada and Wu [7] assumed that if a two-factor interaction is significant, then at least one of the main effects involved in the interaction is also significant. They go on to introduce ways to obtain information on such interactions from screening designs, including using several real-life examples that back up these principles. Subsequently, Chipman *et al* [8] extended this idea to higher-order interactions.

The next section presents a discussion of the surface characteristics that the user can control when using the test bed. Section 3 contains a detailed description of our methodology, and Section 4 provides some examples of how the user input affects the characteristics of the simulated surfaces.

2. SURFACE CHARACTERISTICS AND CONTROL

One of the main issues addressed in this paper is how to create a randomly generated response surface function and still maintain some control over the characteristics of the surface. We have chosen to restrict the test bed to polynomial functions because polynomials can be used to represent a wide variety of surfaces. Also, since polynomial models are most often used to model response surfaces, it follows that many practitioners believe that polynomial functions capture the most significant features of the true response surfaces in their applications. We create the polynomial surface functions in two steps. The first step creates the form of the function; that is, all terms that are in the polynomial are selected. The result is a polynomial function where all non-zero terms are defined, but the coefficients are symbols, not numbers. This function is what we call the *surface form*. The

second step assigns numerical values to the symbolic coefficients. Once the numerical values are assigned to the coefficients, a response surface is completely defined and we will continue to call this the *response surface function*.

When using the test bed to create random response surface functions, the user is given partial control of the six following key surface characteristics:

- (1) effect sparsity;
- (2) bumpiness;
- (3) response range;
- (4) flatness;
- (5) effect heredity;
- (6) random error.

Each of these characteristics is now described in the context of the test bed.

The first characteristic to be controlled is effect sparsity. In the test bed, effect sparsity is the proportion of the factors under consideration that are *active*. An active factor is defined as a factor that appears in at least one term of the response surface function.

Another important characteristic to control is the *bumpiness* of the surface. We use the term *bumpiness* loosely to refer to the prevalence of maximum, minimum, and inflection points on the surface. We feel that neither uni-modal surface functions nor multi-dimensional cosine functions have stationary points that are placed randomly enough to reflect real-world surfaces. However, some control on the maximum number of stationary points is necessary.

A third feature of the surfaces that we want to control is the range of the response. To accomplish this we first define the *region of operability* as the range of factor values over which we will define values of the response. Restricting the region of operability gives us the opportunity to control the range of the response over a finite set. Even with this restriction, controlling the range of the response values is a difficult problem, especially when the polynomial function has many higher-order terms. Our methods allow the user to specify a target range for the response, denoted \mathbf{R}_y . We do not guarantee that the response will remain in \mathbf{R}_y over the region of operability, but instead ensure that most of the surface is in the target range.

We next define a concept called *flatness*, which is the extent to which local deviations in the surface are small with respect to the target response range, \mathbf{R}_y , in the region of operability. For the test bed, flatness is a number, denoted r , that is specified by the user. As r is increased, the height of the local maxima and depth of the local minima on the surface are both decreased.

Effect heredity is the relationship between higher and lower order effects. If we think of a two-factor interaction as the effect that changing the level of one factor has on the main effect of another factor, then it seems unlikely that an interaction term will appear when both main effect terms are null. The test bed provides the opportunity to explore the potential impact that this assumption may have on various design and analysis strategies. Control of effect heredity is carried out through user-specified conditional probabilities.

The last characteristic that we are interested in controlling is the *random error* of the observed response. When experiments are run on physical systems, the observations of the response are subject to error. We want to include an error distribution with the surface functions that we create in order to make the surfaces more realistic and to help determine which experimental designs or design strategies perform the best in the presence of noise. However, it is important that this error distribution be controllable relative to the underlying surface. Thus, the error distribution is directly specified by the user and it is up to the user to set this distribution in proper relationship to the target response range, \mathbf{R}_y .

3. DESCRIPTION OF THE TEST BED METHODOLOGY

The test bed consists of a computer program that produces a 'true' response surface function whose characteristics are probabilistically controlled by user inputs. The following notation is introduced to describe the test bed.

Let $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \dots, x_{i,k})'$ represent a vector of values for the k independent variables in a single run of an experiment and let the $(n \times k)$ matrix \mathbf{X} represent an experiment with n runs, where the i th row of \mathbf{X} corresponds to \mathbf{X}_i' . Without loss of generality, we assume that all variables are continuous. Let $\mathbf{e} = (e_1, e_2, \dots, e_n)'$ be a vector of error terms with joint density function $f_e(\cdot|\xi)$, where the vector ξ represents the parameters of the distribution. Now let $\mathbf{y} = (y_1, y_2, \dots, y_n)'$ be a vector of values of the response of interest where y_i is the response corresponding to the values of the independent variables in vector \mathbf{x}_i . Then we define the true response surface at any point \mathbf{x}_i to be $g(\mathbf{x}_i)$. The function $g(\mathbf{x}_i)$ is defined over the region \mathbf{C} in \mathcal{R}^k , where \mathbf{C} is the region of operability. The function $g(\mathbf{x}_i)$ is a full polynomial function in all k factors which includes all terms up to order d ($d = 6$ in the test bed). The coefficients of g , represented by the w -length vector of parameters γ

in \mathcal{R}^w , are determined by user input, where w can be calculated as

$$w = \sum_{i=2}^d \binom{k}{i} + kd$$

The function $g(\mathbf{x}_i)$ is then written as $g(\mathbf{x}_i, \gamma)$. The response is determined by

$$y_i = g(\mathbf{x}_i, \gamma) + e_i$$

where e_i is the i th element of $e \sim f_e(\cdot|\xi)$.

Figure 1 is a block diagram of the test bed, which shows how user input and random number generation is combined to produce a vector, \mathbf{y} , of observed values for an experimental design, \mathbf{X} .

Using the notation above, we now describe in detail how the user input and random number streams are used to create a response surface function with the desired effect sparsity, effect heredity, response range, bumpiness, flatness, and error distribution. Since g is always a full sixth-order polynomial in k factors, the only control that the user has over the response surface function is through the assignment of the parameter values in γ . As discussed previously, these values are assigned in two steps. The first step is the surface form selection, which determines the coefficients in γ that will be assigned a value of zero, effectively removing certain terms from the response surface function. The result of this step is the surface form, which is a polynomial function with symbolic coefficients on each non-zero term. The second step is the response surface generation, which assigns real numbers to the non-zero coefficients in the surface form. The result of this step is the response surface function, which is a specific instance of the surface form.

3.1. Surface form selection

Figure 2 shows the five sub-steps (in the numbered boxes) of surface form selection. In sub-step 1 of Figure 2, the user begins characterizing a surface form by stating the number of factors, k , that are of interest and the region of operability of each factor. The maximum number of factors in the test bed is 15, which is also the default number of factors.

Let $\mathbf{l} = (l_1, l_2, \dots, l_k)$ be a vector whose elements represent the lower bounds on the region of operability for each of the k factors, with a default value of $\mathbf{l} = (0, 0, \dots, 0)$. Also, let $\mathbf{u} = (u_1, u_2, \dots, u_k)$ be a vector whose elements represent the upper bounds on the region of operability with a default value of $\mathbf{u} = (100, 100, \dots, 100)$. Then the region of operability is

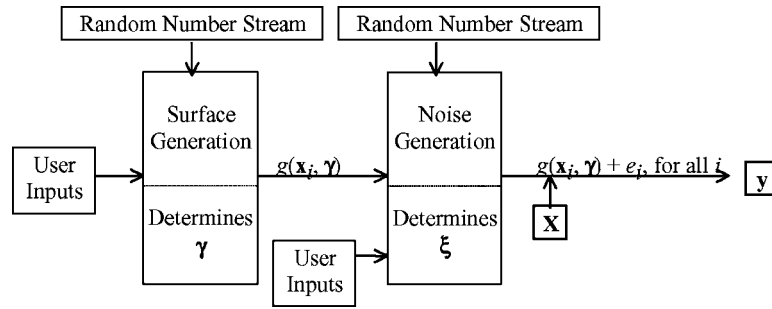


Figure 1. The basic structure and construction of the response surface simulator

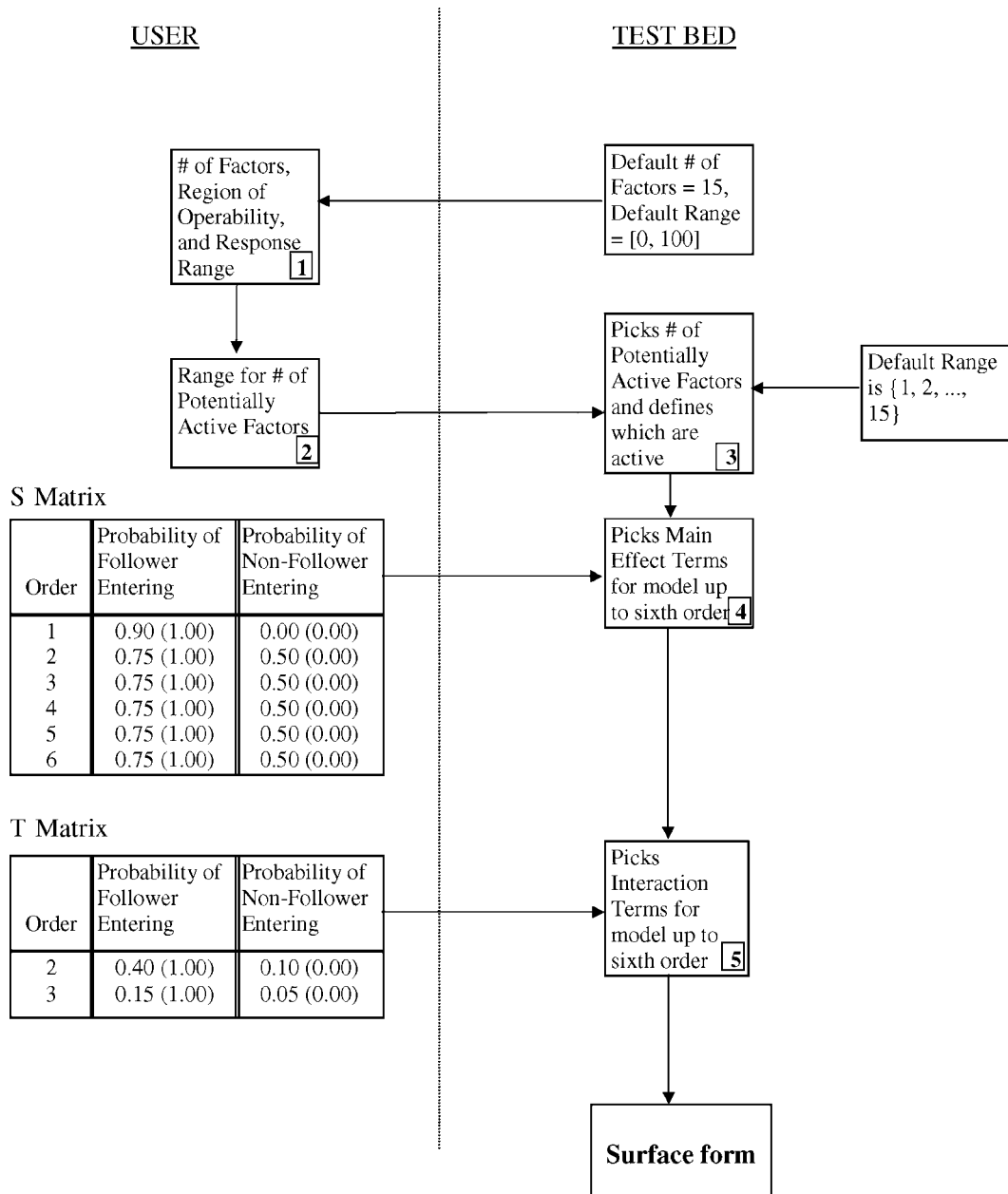


Figure 2. Flowchart of the process of converting user input into a surface form

defined as

$$\mathbf{C} = \{\mathbf{x} = (x_1, x_2, \dots, x_k) \mid l_1 \leq x_1 \leq u_1, \\ l_2 \leq x_2 \leq u_2, \dots, l_k \leq x_k \leq u_k\}$$

Now let l_y be the target lower bound on the response range and u_y be the target upper bound on the response range. Then $\mathbf{R}_y = \{y \mid l_y \leq y \leq u_y\}$ is the target response range. This range will be used to scale the response surface such that a substantial portion of $g(\mathbf{x}_i, \gamma)$ is within \mathbf{R}_y in the region of operability. The default for the response range is [0, 100].

As an example, assume that $k = 3$. Then after sub-step 1, $g(\mathbf{x}_i, \gamma)$ would have 84 terms and can be written as

$$y = \gamma_0 + \gamma_1^{(1)}x_1 + \gamma_2^{(1)}x_2 + \gamma_3^{(1)}x_3 \\ + \gamma_{12}x_1x_2 + \dots + \gamma_1^{(2)}x_1^2 + \dots + \gamma_3^{(6)}x_3^6$$

where $\gamma_j^{(i)}$ refers to coefficient of the i th order main effect term for factor j and γ_{ijk} is the coefficient for the interaction between factors i , j , and k . In the test bed's current implementation, interaction terms are limited to second and third order. The main effect terms up through to sixth order are retained at this stage.

In sub-step 2, the user provides inputs to control the number of the k available factors that are active. A *potentially active factor* is defined as a factor that has a non-zero probability of appearing in the surface form in sub-steps 4 and 5 of the surface form selection procedure (see below). If, during sub-steps 4 and 5, a term containing one of the potentially active factors is selected to have a non-zero value, that factor will appear in the response surface function and thus becomes an active factor. Factors not chosen to be potentially active cannot appear in the response surface function and thus are *inactive factors*.

The number of potentially active factors, f , is randomly determined. The user, however, provides a range of integer values and a probability for each of these integers to become f . The default for the range of integers is $\{1, 2, \dots, k\}$ and, by default, the distribution is discrete uniform, where any integer in the range is equally likely to become f . The range of integers may be changed to any range of integers between 1 and k . Once f is chosen, the active factors are picked randomly from the set of all factors such that each of the k factors has an equal chance to become a potentially active factor, as shown in sub-step 3 of Figure 2. If, for example, $k = 3$, $f = 2$, and the first two factors were chosen to be in the model,

the surface form after sub-step 3 would be

$$y = \gamma_0 + \gamma_1^{(1)}x_1 + \gamma_2^{(1)}x_2 + \gamma_{12}x_1x_2 + \gamma_1^{(2)}x_1^2 \\ + \gamma_2^{(2)}x_2^2 + \gamma_{112}x_1^2x_2 + \gamma_{122}x_1x_2^2 + \gamma_1^{(3)}x_1^3 \\ + \gamma_2^{(3)}x_2^3 + \gamma_1^{(4)}x_1^4 + \gamma_2^{(4)}x_2^4 + \gamma_1^{(5)}x_1^5 + \gamma_2^{(5)}x_2^5 \\ + \gamma_1^{(6)}x_1^6 + \gamma_2^{(6)}x_2^6$$

Next, as shown in sub-step 4 of Figure 2, the test bed determines which of the main effect terms in the surface form will have a value of zero and which will have a non-zero value. A surface form may contain any main effect term up to sixth order, and the user must provide a set of probabilities that help to determine the non-zero terms. Effect heredity and the number of possible stationary points in the surface form are affected by this stage of the surface form selection procedure.

To specify the main effects that are to appear in the surface form, the user first supplies the probability, $s_{1,1}$, that a potentially active factor will have a first-order (linear) term. Notice that this implies that a factor does not have to have a first-order main effect to be considered active. For terms of higher (second to sixth) order, the user must supply two conditional probabilities. The first is the conditional probability that a higher-order term, such as x_2^4 , follows a main effect term in the same active factor of one order lower, like x_2^3 . Such terms are known as *followers* because they follow from the previous terms in the model. We will denote these probabilities $s_{i,1}$ for $i = 2, 3, \dots, 6$. More formally,

$$s_{i,1} = \Pr(\gamma_j^{(i)} \neq 0 \mid \gamma_j^{(i-1)} \neq 0) \\ \text{for } 1 \leq j \leq f \text{ and } 2 \leq i \leq 6$$

For convenience of notation, we have reordered the factors such that the first f factors are the potentially active factors.

The second probability is the conditional probability that a higher-order term of a potentially active factor appears when that factor does not have a term of one order lower. These probabilities are denoted $s_{i,2}$ for $i = 2, 3, \dots, 6$. Such terms are known as *non-followers*. We define these probabilities as

$$s_{i,2} = \Pr(\gamma_j^{(i)} \neq 0 \mid \gamma_j^{(i-1)} = 0) \\ \text{for } 1 \leq j \leq f \text{ and } 2 \leq i \leq 6$$

The default value of the first conditional probability is one, and the default for the second conditional probability is zero, giving a full sixth-order model in the main effect terms, but each of these may be changed to any value on the range [0, 1].

These conditional probabilities are split to give users the flexibility to control how strongly effect heredity applies to the surface form. The conditional probabilities are placed in a (6 × 2) matrix **S**, with elements $s_{i,j}$, where $s_{1,2} = 0$ and all other elements are defined as above. In the example input shown in Figure 2, the matrix **S** is

$$\mathbf{S} = \begin{bmatrix} 0.90(1.00) & 0.00(0.00) \\ 0.75(1.00) & 0.50(0.00) \\ 0.75(1.00) & 0.50(0.00) \\ 0.75(1.00) & 0.50(0.00) \\ 0.75(1.00) & 0.50(0.00) \\ 0.75(1.00) & 0.50(0.00) \end{bmatrix}$$

where the values in parentheses in the figure are the default values for each of the elements. Returning to our example, if this **S** matrix were applied, the surface form after sub-step 4 may be

$$y = \gamma_0 + \gamma_1^{(1)}x_1 + \gamma_2^{(1)}x_2 + \gamma_{12}x_1x_2 + \gamma_{112}x_1^2x_2 + \gamma_{122}x_1x_2^2 + \gamma_2^{(2)}x_2^2 + \gamma_1^{(3)}x_1^3 + \gamma_2^{(3)}x_2^3 + \gamma_1^{(4)}x_1^4 + \gamma_2^{(4)}x_2^4 + \gamma_2^{(5)}x_2^5 + \gamma_2^{(6)}x_2^6$$

reflecting that many of the higher-order terms were randomly chosen to have a non-zero value. Some, however, were eliminated from the surface form by setting their coefficient equal to zero.

After the main effect terms are selected, the test bed selects the interaction terms in the surface form, as shown in sub-step 5 of Figure 2. Since we currently limit interactions in the test bed to order three, all coefficients for interactions above order three are set to zero. The first probability for the second-order interactions is the conditional probability that a higher-order term follows from a first-order main effect that has already been selected to appear in the model. *Following* in this context means that a second-order interaction can be formed by multiplying any single potentially active factor by a main effect that is already in the model. The other probability for the second-order interactions is the conditional probability that a second-order term appears that is made up of two potentially active factors that do not have main effect terms in the model. These terms are known as non-followers. These probabilities are designated $t_{1,1}$ and $t_{1,2}$ and are defined as

$$t_{1,1} = \Pr(\gamma_{ij} \neq 0 \mid \gamma_i^{(1)} \neq 0 \cup \gamma_j^{(1)} \neq 0) \text{ for } 1 \leq i < j \leq f$$

and

$$t_{1,2} = \Pr(\gamma_{ij} \neq 0 \mid \gamma_i^{(1)} = 0 \cap \gamma_j^{(1)} = 0) \text{ for } 1 \leq i < j \leq f$$

Similarly, the first conditional probability specified by the user for the third-order interactions is the conditional probability that an interaction appears that contains a second-order term that is present in the surface form. It is important to remember that these third-order interaction terms can follow from either second-order interaction terms or second-order main effect terms in the surface form. The other conditional probability for third-order interactions is the conditional probability that an interaction appears that does not contain a second-order term that is present in the model. These non-following third-order interactions may contain active factors that have first-order main effects, but no factor combinations that are found in any second-order terms in the model. These two conditional probabilities are designated $t_{2,1}$ and $t_{2,2}$ and are defined as

$$t_{2,1} = \Pr(\gamma_{ijk} \neq 0 \mid \gamma_{ij} \neq 0 \cup \gamma_{ik} \neq 0 \cup \gamma_{jk} \neq 0) \text{ where } i \leq j \leq k \text{ and } i \neq k$$

and

$$t_{2,2} = \Pr(\gamma_{ijk} \neq 0 \mid \gamma_{ij} = 0 \cap \gamma_{ik} = 0 \cap \gamma_{jk} = 0) \text{ where } i \leq j \leq k \text{ and } i \neq k$$

In the above definitions, assume that if $i = j$, then $\gamma_{ij} = \gamma_i^{(2)}$.

Once again, the default values for the first conditional probability is one for each order, and the default for the second conditional probability is zero for each order, but these can be changed to any valid probability. These conditional probabilities form a (2 × 2) matrix **T**, with elements $t_{i,j}$ as defined above. In the example input shown in Figure 2, the matrix **T** takes on the value

$$\mathbf{T} = \begin{bmatrix} 0.40(1.00) & 0.10(0.00) \\ 0.15(1.00) & 0.50(0.00) \end{bmatrix}$$

where the values in parentheses in the figure are the default values for each of the elements. After this **T** matrix is applied to our example, the final version of the surface form may look like

$$y = \gamma_0 + \gamma_1^{(1)}x_1 + \gamma_2^{(1)}x_2 + \gamma_2^{(2)}x_2^2 + \gamma_1^{(3)}x_1^3 + \gamma_2^{(3)}x_2^3 + \gamma_1^{(4)}x_1^4 + \gamma_2^{(4)}x_2^4 + \gamma_2^{(5)}x_2^5 + \gamma_2^{(6)}x_2^6 \tag{1}$$

demonstrating that the second-order interaction term had a 60% chance of having its coefficient set equal to zero, on the condition that it would follow the main effect terms of x_1 and x_2 .

The surface form is now fully defined and will be denoted $g^*(\mathbf{x}_i, \gamma^*)$ to indicate that all terms that

have been assigned coefficients of zero in the model selection procedure have been removed from the function and the parameter vector. Consequently, γ^* is a much smaller vector than γ .

3.2. Response surface function generation

The creation of the surface form addresses the effect sparsity, the effect heredity, and the bumpiness desired by the user. However, the range of the response and the flatness of the surface will be primarily affected by the generation of the response surface function. This is the function that results when values are assigned to the coefficients, γ^* in the surface form.

We have chosen a regression technique for assigning values to the coefficients, γ^* , in the surface forms. The method, roughly stated, chooses many points at random in and around the region of operability. Each of these points is randomly assigned a response value in the user-specified range. The surface form, $g^*(\mathbf{x}_i, \gamma^*)$, is then fitted to these points using a least-squares fit for γ^* . Since the response values are contained in the user-specified region, the fitted surface is also approximately contained in the region.

The method begins by defining the *surface creation region* as a region larger than the region of operability by 30% in each direction for each potentially active factor. Let $\mathbf{v}' = (v_1, v_2, \dots, v_f)$ where $v_i = (u_i - l_i) * 0.30$ for $1 \leq i \leq f$. The surface creation region then would be

$$\mathbf{D} = \{\mathbf{x} = (x_1, x_2, \dots, x_f) \mid (l_1 - v_1) \leq x_1 \leq (u_1 + v_1), (l_2 - v_2) \leq x_2 \leq (u_2 + v_2), \dots, (l_f - v_f) \leq x_f \leq (u_f + v_f)\}$$

Thus, if the default region of operability for a variable, [0, 100], is used, the surface creation region is [-30, 130]. The surface creation region is larger than the region of operability to reduce the effect of steep slopes in the surface near the edges of the region of operability. The region was expanded by 30% in each direction based on our experience with the edge behavior of fitted polynomial surfaces. A smaller expansion would increase the risk of steep slopes near the boundaries and a larger expansion might require more points (adding computation time and data storage needs) to adequately cover the surface creation region.

Once the surface creation region is defined, we select points at random in this region and assign these points random response values that are uniformly distributed in \mathbf{R}_y . Recall that f is the number of

potentially active factors from the model selection procedure and thus represents the dimension of the subspace in which the surface resides. Let p be the number of non-zero terms in the surface form, i.e. the length of the vector γ^* , and let r be the flatness index chosen by the user (flatness has a default value of two). To give the surface appropriate flatness in the space of active factors, the number of random points for fitting the regression is set at p^*r^*f , where this term is always rounded up to an integer value.

Notice that there are $p^*(r^*f - 1)$ error degrees of freedom in the regression, so the value of r must be greater than $1/f$ so that the error degrees of freedom for the regression is greater than zero. In the context of fitting a response surface function to a surface form, this is similar to allowing for an average of r^*f degrees of freedom per parameter. While this allows us to relate the number of points in the regression to the number of parameters estimated, determining how the points are spaced over the surface creation region requires a different point of view. In a sense, we can describe how densely the surface creation region is populated by sample points by considering the distance between any point and its nearest neighbor. Since all points are chosen randomly and uniformly, this distance increases as the dimension of the factor space increases [9]. So, the nearest-neighbor distance is approximately proportional to r^*p , allowing control of the density of the points in the factor space by changing r .

The random points chosen are stored in a matrix \mathbf{X}_r with dimension $(p^*r^*f) \times (f)$. Let the i th row of \mathbf{X}_r be a row vector, \mathbf{x}'_{ri} , selected randomly from a multivariate uniform distribution spanning the surface creation region, \mathbf{D} . Since each row is an f -length row vector and there are p^*r^*f rows, then $(p^*r^*f^2)$ random numbers are generated to create \mathbf{X}_r . In addition, a vector of responses \mathbf{y}_r of length p^*r^*f is created where each element y_{ri} is chosen randomly and uniformly from \mathbf{R}_y .

Let $\mathbf{h}(\mathbf{x}_i)$ be a vector-valued function such that $g^*(\mathbf{x}_i, \gamma^*) = \mathbf{h}'(\mathbf{x}_i)\gamma^*$. For example, if we were fitting a response surface function to the surface form represented in equation (1), $\mathbf{h}(\mathbf{x}_i)$ would be a transformation such that:

$$\mathbf{h}'(x_1, x_2) = (1, x_1, x_2, x_2^2, x_1^3, x_2^3, x_1^4, x_2^4, x_2^5, x_2^6)$$

To assign values to the p elements of γ^* , we create a $((p^*r^*f) \times (p))$ matrix, \mathbf{H} , such that the i th row of \mathbf{H} is a p -length vector $\mathbf{h}'(\mathbf{x}_{ri})$. The fitted coefficients for the response surface function are then $\gamma_r^* = (\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'\mathbf{y}_r$. The response surface function

Table 1. Levels of the treatments used in the bumpiness experiment

Treatment	Levels						
	0	1	2	3	4	5	6
S matrix (S)	Just linear	All linear, some quadratic	All linear, all quadratic	All linear, some quadratic or cubic	Low probability of all up to sixth order	Medium probability of all up to sixth order	All up to sixth order
T matrix (T)	None	All two-factor and three-factor interactions	N/A	N/A	N/A	N/A	N/A

N/A = not applicable.

is $g^*(\mathbf{x}_i, \gamma_r^*) = \mathbf{h}'(\mathbf{x}_i)\gamma_r^*$. Clearly many different response surface functions can be generated from the same surface form by choosing a new set of points to create a new \mathbf{H} and \mathbf{y}_r , just as many surface forms can be created from the same \mathbf{S} and \mathbf{T} matrices.

Once the response surface function is defined, responses can be observed with or without adding an error term. The error term distribution is completely determined by user inputs. Typical examples may be normal or uniform. The default form of the error distribution is $N(0, 1)$ and the user has the option to change the mean or variance or to make the error model a different distribution. Other options for controlling the error distribution may be pursued in future research. S-Plus[®] (see [10]) code of the implementation of this test bed may be found at <http://www.iems.nwu.edu/bea/testbed>.

4. EXAMPLES AND DISCUSSION

In this section, we provide some examples of the use of the input controls to change the characteristics of the surfaces. The \mathbf{S} and \mathbf{T} matrices control the bumpiness under the conjecture that increasing the number of higher-order terms in the underlying surfaces created will increase the bumpiness of these surfaces. We created surfaces by systematically varying these matrices to increase the number of higher-order terms. Seven different \mathbf{S} matrices were used, ranging from one that only allowed linear terms in the model to one that included all higher-order terms up to sixth-order in the model. Two different \mathbf{T} matrices were used; one allowed no interactions and the other included all two- and three-factor interactions. If we think of these as two treatments with the levels shown in Table 1, all combinations of levels of the two treatments were used and thus 14 types of surfaces were created.

To demonstrate how the \mathbf{S} and \mathbf{T} matrices affect the bumpiness of the surface, a sample of each of

the 14 surfaces was selected and a three-dimensional surface plot was drawn for one pair of factors per surface. To allow for plotting, all other factors are held constant at the center point values (50) and the response surface is shown over the region of operability of the selected factors. These surfaces are shown in Figure 3 without any error component so that the variations in the true response surface functions can be seen.

In Figure 3, the first column shows the surfaces for each level of the \mathbf{S} -matrix factor with the \mathbf{T} -matrix factor set at its low level (no interactions). The surface at the lowest level of the \mathbf{S} -matrix factor is simply a plane cutting through space, so there is no bumpiness, while the surface at the highest level of this factor was very bumpy with local minima and maxima. The second column in Figure 3 shows surfaces for each level of the \mathbf{S} -matrix factor at the high level of the \mathbf{T} -matrix factor (all two- and three-factor interactions). Although the difference is subtle, the two columns do demonstrate how adding interactions to the surfaces changes them. In the second column, surfaces twist as interaction effects become important.

In similar ways, the other test bed controls can be shown to produce surfaces that have characteristics that conform to the users desires (see [11]). This is an important capability, because the test bed's intended use is to compare the effectiveness of experimental design strategies through simulation studies. The strategies often rely on assumptions on the underlying nature of the surfaces, such as the presence of strong linear effects, local minimum and maximum points, and sparse interactions. Creating surface functions that to varying degrees reflect or violate these assumptions allows the user to determine how effective the strategy is when the assumptions hold and how robust the strategies are when they do not.

In McDaniel and Ankenman's paper [12], the test bed described in this paper is used to study a problem

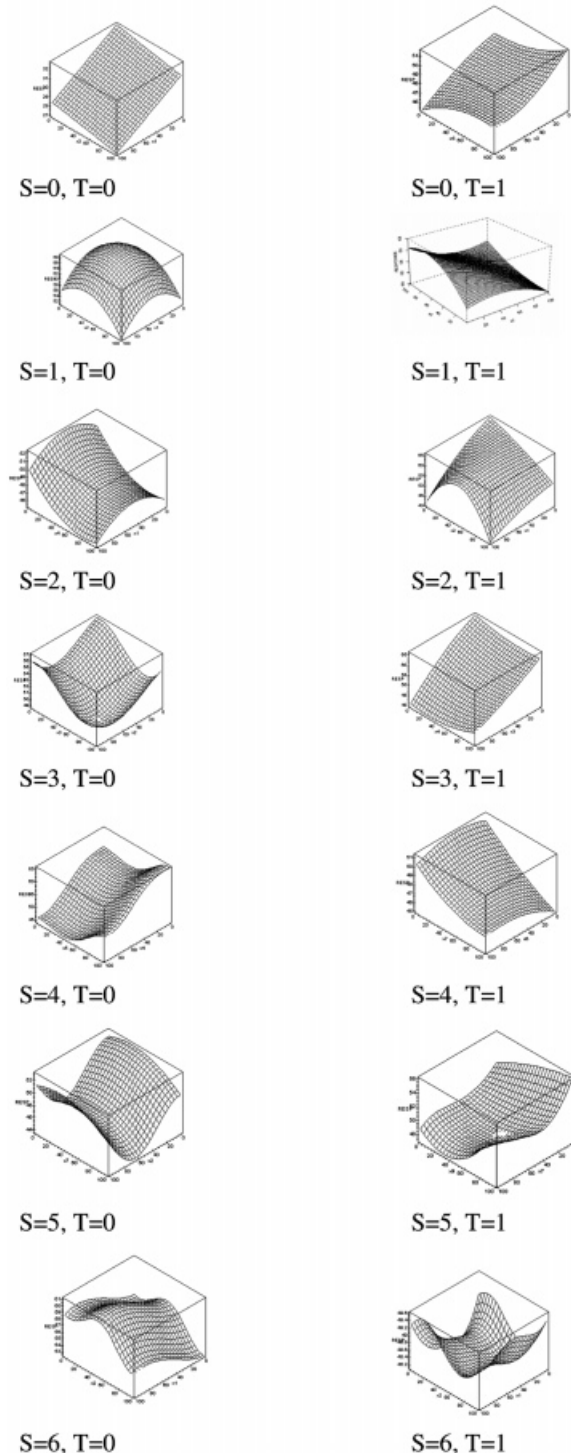


Figure 3. Comparison of the surface bumpiness for the different levels of the **S** matrix and the **T** matrix

we call the *small factor change* problem. This is a problem that is often encountered when the response of interest must be changed by a certain amount, but because of other considerations the change in the factor settings must be minimized. The question that we are interested in is *What is an appropriate experimental design strategy to use when faced with this problem?* A large scale simulation was run which included seven different design strategies and 1225 different response surface functions created by the test bed. A version of the traditional response surface methodology was shown to be the best performer for solving the small factor change problem.

ACKNOWLEDGEMENTS

The authors would like to thank Dan Meyer, Liz Schiferl, Barry Nelson, Ajit Tamhane, Kwok Tsui, and Ted Allen for helpful discussions during this research. The research was supported in part by the Morris E. Fine Junior Professorship in Materials and Manufacturing and the Ellis R. Ott Scholarship.

REFERENCES

1. Myers RH, Montgomery DC. *Response Surface Methodology*. John Wiley: New York, 1995.
2. Pukelsheim F. *Optimal Design of Experiments*. John Wiley: New York, 1993.
3. Box GEP, Wilson KB. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society, Ser. B* 1951; **13**(1):1–45.
4. Smith DE. Studies with a prototype 'OPTIMIZER' for use in computer simulation. *17th Conference on the Design of Experiments in Army Research, Development and Testing*, 1971.
5. Smith DE. An empirical investigation of optimum-seeking in the computer simulation situation. *Operations Research* 1972; **13**:166–173.
6. Barton RR. Minimization algorithms for functions with random noise. *American Journal of Mathematical and Management Sciences* 1984; **4**:109–138.
7. Hamada M, Wu CFJ. Analysis of designed experiments with complex aliasing. *Journal of Quality Technology* 1991; **24**(3):130–137.
8. Chipman H, Hamada M, Wu CFJ. A Bayesian variable selection approach for analyzing designed experiments with complex aliasing. *Technometrics* 1997; **39**:372–381.
9. Weber R, Blott S. An approximation-based data structure for similarity search. *Technical Report 24*, ESPRIT project HERMES 9141 1997.
10. Statistical Sciences, Inc. *S-PLUS User's Manual, Version 3.3 for Windows*. Statistical Sciences, Inc.: Seattle, 1995.
11. McDaniel WR. Using simulation to compare experimental design strategies for quality improvement with minimal changes to factor levels. *PhD Thesis*, Department of Industrial Engineering and Management Sciences, Northwestern University 1999 (unpublished).
12. McDaniel WR, Ankenman BE. Comparing experimental design strategies for quality improvement with minimal changes to factor levels. *Quality and Reliability Engineering International* 2000; **16**:355–362.

Authors' biographies:

William McDaniel received his PhD in Industrial Engineering/Management Science from Northwestern University in 1999, where his research interests included response surface methodology, design of experiments, and simulation. He is currently working as an Operations Research Consultant for ZS Associates in Evanston, IL.

Bruce Ankenman received his PhD in Industrial Engineering from University of Wisconsin-Madison in 1995. He is now an Assistant Professor in the Department of Industrial Engineering and Management Sciences at Northwestern University. His research interests include design of experiments, robust design, and experiments involving variance components and dispersion effects.