STRATEGIES FOR OPTIMIZATION OF MULTIPLE-RESPONSE SIMULATION MODELS

William E. Biles
University of Notre Dame
James J. Swain
Air Products and Chemicals, Inc.

ABSTRACT

This paper examines several procedures for optimizing simulation models having controllable input variables $x_i, i = 1,\ldots,n$ and yielding responses $\eta_j, j = 1,\ldots,m$. This problem is often formulated as a constrained optimization problem, or it can be formulated in one of several multiple-objective formats, including goal programming. Whatever the mode of problem formulation, the optimization of multiple-response simulations can be approached through direct search methods, a sequence of first-order response-surface experiments, or by applying mathematical programming techniques to a set of second-order response surfaces.

INTRODUCTION

A computer simulation can be regarded as a "black-box" into which values for n controllable input variables $x_i, i = 1,\ldots,n$ are combined in some manner to produce values for m responses η_j, $j = 1,\ldots,m$. These responses are usually measures of effectiveness for the system under study.

A given simulation trial at a specific set of values for the input variables X^k yields an observation y_j^k for each system response η_j. The responses are also affected by a set of p uncontrollable factors $z_\ell, \ell = 1,\ldots,p$, causing the true response η_j at X^k to suffer a disturbance ε_j, so that

$$y_j^k = \eta_j(X^k) + \varepsilon_j, \quad j = 1,\ldots,m \tag{1}$$

In many instances, the purpose of computer simulation is to evaluate the qualitative effects of various policies for operating a system. That is, each of the input variables $x_i, i = 1,\ldots,n$ has a set of distinct, qualitative levels. To find the best operating policy for the system under study, it is necessary only to perform simulation trials of sufficient duration at each possible combination of values for the controllable input variables, selecting as a solution the combination yielding the most favorable results. Often the objective of computer simulation is to find optimum values for controllable input variables $x_i, i = 1,\ldots,n$ which can assume either a wide range of discrete values or an infinite number of real values. The latter problem would fall into the domain of classical optimization except that (a) the functional forms of the response relationships

$$\eta_j = g_j(x_1,\ldots,x_n), \quad j = 1,\ldots,m \tag{2}$$

are usually unknown and (b) an observation y_j^k of the system response η_j at X^k contains random error ε_j due to the unpredictable effects of the uncontrollable factors $z_\ell, \ell = 1,\ldots,p$.

Most of the response surface methods described in this paper require the assumption that the error component ε_j be a normally distributed random variable with

$$E(\varepsilon_j) = 0, \qquad j = 1,\ldots,m \tag{3}$$

$$\mathrm{Var}(\varepsilon_j) = \sigma_j^2, \qquad j = 1,\ldots,m \tag{4}$$

This assumes that the actual response surfaces are the expected values of the observed responses. Since an observation y_j^k from a simulation trial is usually the sample mean of s realizations of the response variable, the variance σ_j^2 is actually $1/s$ as large as the variance of the population of the response variable. Indeed, the minimization of σ_j^2 through so-called "variance reduction" techniques is an essential feature of the simulation-optimization strategies discussed here, although variance reduction will not specifically be addressed here.

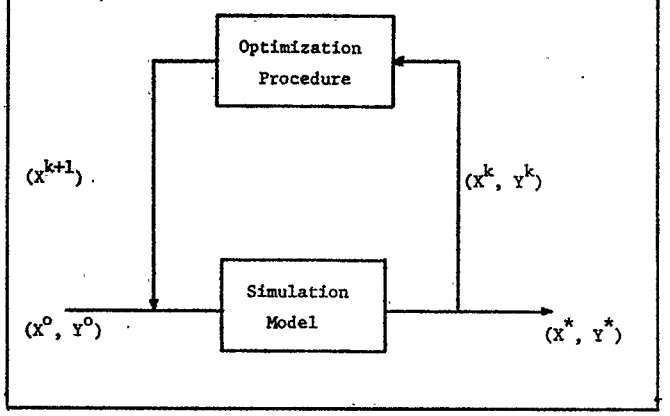
Note that we have m separate $(n+1)$-dimensional surfaces, each having its own characteristic random variation as expressed in (3) and (4). The problem at hand is to have the computer simulation model which produces these m responses either automatically or interactively controlled by an optimization procedure. The interface between the optimization method and the simulation model is illustrated in a general sense in Illustration 1. The extent of the interaction between the simulationist and the optimization procedure can be almost none, as in the case of a fully algorithmic technique, to almost completely sequential decision-making. This paper describes several such techniques.

METHODS OF PROBLEM FORMULATION

We shall consider two basic approaches to formulating the problem of optimizing a multiple-variable, multiple-response simulation model. One approach is the familiar constrained optimization formulation

ILLUSTRATION 1

The Interface Between Optimization and Simulation



in which one of the system responses, say η_1, is to be maximized or minimized, subject to maintaining the remaining m-1 response within prescribed bounds. The second approach is the multiple-objective formulation, in which the m responses are either weighted to form a single objective or treated in manner akin to goal programming. Each of these two approaches is described in the following sections.

Constrained Optimization

Under the constrained optimization approach, the problem is stated as

$$\text{Max(or Min)} \quad \eta_1 = g_1(x_1,\ldots,x_n) \tag{5}$$

subject to the constraints

$$a_i \leq x_i \leq c_i, \quad i = 1,\ldots,n \tag{6}$$

$$\eta_j = g_j(x_1,\ldots,x_n) \left\{ \begin{matrix} \geq \\ = \\ \leq \end{matrix} \right\} d_j, \quad j = 2,\ldots,m \tag{7}$$

The constraints expressed in (6) are bounds on the controllable input variables $x_1,\ldots,x_n$ and are typically known a priori. Thus the bounds (6) generally form the known experimental region prior to conducting simulation trials. In contrast to that, the response functions $g_j(x_1,\ldots,x_n)$ in (7) are not usually known a priori and the responses η_j must be estimated experimentally via simulation. Hence simulation trials performed at points satisfying (6) may yield responses violating (7).

To complicate matters even more, the random error ε_j can lead to erroneous decisions relative to the constraints in (7), leading the simulationist to believe that a given response is feasible when it is not, and vice versa. The same is true relative to the objective function in (5). That is, one simulation trial can appear to represent an improvement over another when the true response at this particular set of values $x_1,\ldots,x_n$ does not. These difficulties can be countered through variance reduction techniques.

Multiple Objective Optimization

One approach to a multiple objective formulation is to assign weights $w_j, j = 1,\ldots,m$ to the m responses and form a single objective function

$$\text{Max(or Min)} \; W = \sum_{j=1}^{m} w_j g_j(x_1,\ldots,x_n) \tag{8}$$

The bounds (6) still apply, so that the problem remains one of constrained optimization, but one in which the entire feasible region is known a priori. The weights $w_j, j = 1,\ldots,m$ are typically assigned through the subjective judgment of the decision-maker in the system being simulated. These weights are usually normalized, so that

$$\sum_{j=1}^{m} w_j = 1 \tag{9}$$

One frequently encounters the situation in which certain of the responses η_j are to be maximized and others minimized. This case is handled by maximizing the negative of those functions which are to be minimized, so that the objective function in (8) is rearranged to the form

$$\text{Max } W = \sum_{j=1}^{s} w_j g_j(x_1,\ldots,x_n) - \sum_{j=s+1}^{m} w_j g_j(x_1,\ldots,x_n) \tag{10}$$

where s functions are maximized and m-s functions are minimized.

A second approach to the multiple objective formulation is one which casts the problem in the format

$$\text{Max} \quad U[g_1(x_1,\ldots,x_n),\ldots, \ldots, g_m(x_1,\ldots,x_n)] \tag{11}$$

subject to the bounds in (6). The formulation in (10) is a special case of that in (11), in which $U[g_j(x_1,\ldots,x_n)]$ is a linear additive function. Montgomery and Bettencourt [15] discuss various formulations of the multiple objective optimization problem, as well as several approaches to its solution, and demonstrate its application to multiple-response simulation.

Another multiple objective optimization formulation is that called goal programming. This procedure is initiated by establishing a set of goals in terms of the m system responses. These goals are expressed as

$$G_j = g_j(x_1,\ldots,x_n), \qquad j = 1,\ldots,m \tag{12}$$

Each goal must have an associated right-side value

d_j; that is,

$$G_j = g_j(x_1,\ldots,x_n) \left\{ \begin{matrix} \leq \\ = \\ \geq \end{matrix} \right\} d_j, j = 1,\ldots,m \qquad (13)$$

Hence each goal can be expressed as an equality

$$g_j(x_1,\ldots,x_n) + n_j - p_j = d_j, j = 1,\ldots,m \qquad (14)$$

where n_j is a negative deviation from d_j, and p_j is a positive deviation. Either n_j or p_j must be zero in any given solution, and both could be zero. Next, each of the m goals G_j is assigned to a priority level $P_k, k = 1,\ldots,\ell$, where P_1 represents the highest priority and P_ℓ the lowest. For any goal falling within a given priority level P_k, the decision-maker is unable to say that one goal is preferred to another. The final step in problem formulation is to combine these several levels of goals into an achievement function, which has the form

$$A = \{P_1(\overline{n}_1,\overline{p}_1),\ P_2(\overline{n}_2,\overline{p}_2),\ldots,P_\ell(\overline{n}_\ell,\overline{p}_\ell)\} \qquad (15)$$

This achievement function is simply an ordered ℓ-vector. Its structure is predicated on one of the following procedures for achieving the j-th goal:
(a) To equal or exceed d_j, minimize n_j
(b) To equal or be less than d_j, minimize p_j
(c) To equal d_j, minimize $(n_j + p_j)$
A solution $(x^*_1,\ldots,x^*_n)$ is considered optimal if, for this solution the corresponding value A^* is the same as or preferred to any other value A.

Thus, the general goal programming problem is to find $x_1,\ldots,x_n$ so as to minimize the ordered vector (15) such that the goals (14) are satisfied and

$$\begin{aligned} x_i &\geq 0, \quad i = 1,\ldots,n \\ n_j &\geq 0, \quad j = 1,\ldots,m \\ p_j &\geq 0, \quad j = 1,\ldots,m \end{aligned} \qquad (16)$$

The functions $g_j(x_1,\ldots,x_n)$ in (14) are generally unknown, but are usually assumed to be nonlinear. Any technique proposed for solving this problem in the simulation domain must provide experimental estimates of these unknown functions, as well as a mathematical procedure for optimization. Moreover, the experimental observations are produced via simulation - each simulation trial at a point $x_1,\ldots,x_n$ produces m responses $y_j = 1,\ldots,m$. Biles [3] has described the application of nonlinear goal programming to the multiple-response simulation problem.

The following sections describe several optimization procedures which can be applied to one or more of the above formulations of the multiple-response simulation problem.

OPTIMIZATION TECHNIQUES

Various procedures have been applied in combining optimization and simulation to seek the "optimum" solution to systems possessing a single response η. The multiple-response problem described here is complicated by the necessity to observe several responses at once, and to incorporate these values into the optimization technique. But many of the same techniques that have been applied successfully to the single-response problem can, with appropriate modifications, be extended to accommodate multiple responses. Moreover, these modified procedures are often applicable to more than one of the aforementioned formulations of the multiple-response problem.

The optimization procedures described below fall into three categories: (1) direct search techniques, (2) first-order response surface methods, and (3) second-order response surface procedures. Although numerous techniques will be cited, only a few broadly stated procedures will be outlined here due to space limitations. It should be remembered that, although we may refer to "optimization" techniques, the classical notion of an "optimum" solution is inapplicable due to the presence of the sampling error ε_j associated with each response variable η_j. Rather we shall seek a solution which hopefully lies close to the true solution. In a more formal sense, we might state that we are to some degree, say 90%, confident that the true solution lies within some interval about our estimated solution.

Direct Search Methods

Direct search methods are those which, applied in a purely computational manner, do not require the use of derivatives. These methods progress through a sequence of points according to some algorithm. Typical of this class of optimization techniques are the pattern search algorithm by Hooke and Jeeves [11], sequential simplex search by Spendley, Hext and Himsworth [19], and the so-called "complex" search method by M. J. Box [6]. In general, these direct search procedures make rapid early progress toward an "optimum", but iterate laboriously as a solution is neared. This is particularly true in the presence of random error, as encountered in simulation.

Among the direct search techniques, Box's "complex" method [6] is most easily adapted to a multiple-response environment. It also performs better than any of the other direct search techniques in the face on random error and constraints. In fact, "complex" search is not at all complex, but derives its name from a contraction of the words "constrained simplex": it evolved from the sequential simplex method [19] and the necessity to deal with constraints. Noh [17] has suggested a mofication of Box's method [6] which makes it especially suitable for the multiple-response simulation problem. The following procedure describes a generalized "complex" procedure as it might be applied to the multiple-response simulation problem:

1. Randomly generate a set of $N \geq n+2$ search points satisfying the known bounds (6).
2. Perform a simulation trial at each of these N search points and record the mN estimated responses $y^\ell_j, j = 1,\ldots,m;\ \ell = 1,\ldots,N$.
3. Where a given search point X^k is observed to violate one or more constraints (7), if such constraints apply with the particular problem formulation being employed, generate a replacement search point $X^{k'}$, perform a simulation

trial at $X^{k'}$, and record the m estimated responses at $X^{k'}$.

4. After N feasible search points have been established, evaluate the objective function for each of these N points. This "objective function" might be η_1 in (5), W in (10), U in (11), or A in (15). Among these N search points, find the worst point X^w; that is, the search point giving the least desirable value of the objective function. Define X^c as the centroid of the N-1 remaining points. Project from X^w through X^c to the image point $X^{w'}$. If the known bounds (6) are violated by this move, shorten the step to $X^{w'}$ until no violation occurs. Perform a simulation trial at $X^{w'}$.
5. Repeat steps 3 and 4 until a solution (X^*,Y^*) is obtained which represents the best solution that can be achieved within the available computer time.

A significant advantage of complex search is that, once N feasible simulation trials (and perhaps several infeasible trials) have been performed, trials are conducted one at a time thereafter. The search can be continued as long as improved solutions are obtained. If several successive simulations are performed at scattered points around the known experimental region without achieving an improved solution, however, the search can be terminated and the best available solution adopted.

First-Order Response Surface Methods

First-order response surface methods attempt to accomplish experimentally what the "method of steepest ascent" accomplishes computationally. From a current point X^k, a designed experiment is conducted (with a simulation trial at each design point) to estimate the gradient direction $\nabla g(X^k)$. Simulation trials are then conducted at points along this direction to a new point X^{k+1} which represents the best solution obtained along $\nabla g(X^k)$. This process is an experimental approximation of

$$X^{k+1} = X^k + \lambda^k[\nabla g(X^k)] \tag{17}$$

The step length λ^k can be estimated by a line search or by a regression procedure as described by Biles [1,2].

The gradient direction $\nabla g(X^k)$ is estimated by placing an appropriate first-order experimental design, such as a 2^n factorial, 2^{n-p} fractional factorial, or n-dimensional simplex design (see Myers [16]) around the current point X^k. Brooks and Mickey [7] have shown that the n-dimensional simplex design, which employs n+1 design points at the vertices of a regular simplex, gives the greatest efficiency in terms of information per design point. That is, the n+1 observations $y^\ell, \ell = 1,\ldots,n+1$ are the minimum number from which the multiple linear regerssion model

$$\hat{y} = b_o + \sum_{i=1}^{n} b_i x_i \tag{18}$$

can be estimated. Since the gradient direction $\nabla g(X^k)$ is mathematically defined as the n-vector of first partial derivatives of $g(X)$ evaluated at X^k, it is clear that $\nabla g(X^k)$ is simply the n-vector of regression coefficients, or

$$\nabla g(X^k) = (b_1,\ldots,b_n)' \tag{19}$$

In the multiple-response simulation problem, a simulation trial is conducted at each design point in the n-simplex design and the m observations y_j^ℓ, $j = 1,\ldots,m$ are recorded at each design point. Multiple linear regression (see Draper and Smith [8]) is applied separately to each set of observations (assuming independence among the m responses), producing the m models

$$y_j = b_{j,0} + \sum_{i=1}^{n} b_{j,i}x_i, \quad j = 1,\ldots,m \tag{20}$$

and thus the m gradient vectors

$$\nabla g_j(X^k) = (b_{j,1},\ldots,b_{j,n})', \quad j = 1,\ldots,m \tag{21}$$

These estimates can then be employed in any one of several optimization schemes to produce an improved solution X^{k+1}. A generalized procedure for accomplishing this improved solution, and an estimated "optimum", will be described later. But first it is necessary to give attention to the experimental design employed to estimate the gradient vectors $\nabla g_j(X^k)$, $j = 1,\ldots,m$.

As stated earlier, the n-simplex design is the most economical design for estimating the first-order model (18). It is usually desirable to minimize the variances of the regression coefficients b_i, $i = 1,\ldots,n$. To accomplish this the first-order experimental design should be orthogonal. That is, the placement of the N experimental points (in our case, simulation trials) is described by the N by n design matrix D, where

$$D = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ x_{1N} & x_{2N} & \cdots & x_{nN} \end{bmatrix} \tag{22}$$

Then an N by n+1 matrix X is constructed by placing a unit vector to the left of D. Thus,

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{n1} \\ 1 & x_{12} & x_{22} & \cdots & x_{n2} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 1 & x_{1N} & x_{2N} & \cdots & x_{nN} \end{bmatrix} \tag{23}$$

It is usually convenient to code the design levels, so that

$$\sum_{u=1}^{N} x_{iu}^2 = N$$

$$i = 1,\ldots,n \qquad (24)$$

$$\sum_{u=1}^{N} x_{iu} = 0$$

If the actual value of the u-th level of the i-th variable is ε_{iu}, then the corresponding coded value is

$$x_{iu} = \frac{\varepsilon_{iu} - \bar{\varepsilon}_{iu}}{S_i} \qquad (25)$$

$$\text{where } \bar{\varepsilon}_i = \left(\sum_{u=1}^{N} \varepsilon_{iu}\right)/N \qquad (26)$$

$$\text{and } S_i = \sqrt{\sum_{u=1}^{N} (\varepsilon_{iu} - \bar{\varepsilon}_i)^2/N} \qquad (27)$$

Then

$$X'X = \begin{bmatrix} N & 0 & 0 & \ldots & 0 \\ 0 & N & 0 & \ldots & 0 \\ . & & & & . \\ . & & & & . \\ 0 & 0 & 0 & \ldots & N \end{bmatrix} \qquad (28)$$

Since the n+1 vector of regression coefficients is estimated by the least-squares relation

$$\bar{b} = (X'X)^{-1} X' \bar{y} \qquad (29)$$

the variance of b_i, $i = 1,\ldots,n$ is

$$\text{Var } (b_i) = \sigma^2/N, \quad i = 1,\ldots,n \qquad (30)$$

where σ^2 is the variance of the error ε. Since we are interested in m different responses y_j, $j=1,\ldots,m$, equations (29) and (30) can be generalized to

$$\bar{b}_j = (X'X)^{-1} X' \bar{y}_j \;, \; j = 1,\ldots,m \qquad (29a)$$

$$\text{Var } (b_j,i) = \sigma_j^2/N, \quad \begin{matrix} i = 1,\ldots,n \\ j = 1,\ldots,m \end{matrix} \qquad (30a)$$

For any orthogonal first-order design, the results in (28)-(30) are true, giving a "minimum variance" design. The 2^n factorial and 2^{n-p} fractional factorial designs are orthogonal and minimum variance. Orthogonal n-simplex designs can be easily constructed by starting with any orthogonal N by N matrix O and proceeding as illustrated in the following 2-dimensional example:

$$O = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 0 & 2 \end{bmatrix}$$
$$\sqrt{3}/\sqrt{3} \quad \sqrt{3}/\sqrt{2} \quad \sqrt{3}/\sqrt{6}$$

Multiple each element in the j-th column of matrix O by the term shown at the bottom of the column, which is

$$\sqrt{N} \Big/ \sqrt{\sum_{i=1}^{N} O_{j,i}^2},$$

to give the matrix X. In this example, we get

$$X = \begin{bmatrix} 1 & \sqrt{3/2} & -1/\sqrt{2} \\ 1 & -\sqrt{3/2} & -1/\sqrt{2} \\ 1 & 0 & 2/\sqrt{2} \end{bmatrix}$$

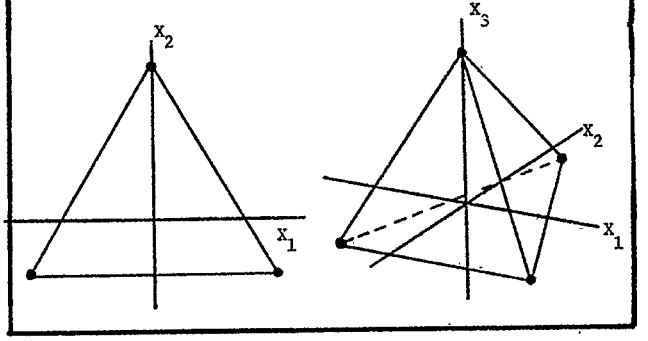
From this it is seen that the design matrix is

$$D = \begin{bmatrix} \sqrt{3/2} & -1/\sqrt{2} \\ -\sqrt{3/2} & -1/\sqrt{2} \\ 0 & 2/\sqrt{2} \end{bmatrix}$$

which describes the vertices of an equilateral triangle in 2-dimensional space, as shown in Illustration 2. Other choices of X are possible simply by choosing different starting orthogonal matrices O. These other choices of X simply represent different orientations of the n-simplex. The X'X matrix is diagonal for each, and thus they are all orthogonal. See Myers [16] for a complete treatment of simplex designs.

ILLUSTRATION 2

First-Order n-Simplex Response Surface Designs



One additional requirement for the n-dimensional simplex to place two or more design points at the centroid of the simplex, which corresponds to X^k, the current point in the search. These points allow degrees of freedom for testing lack of fit, and they provide a starting point for the subsequent line search along $\nabla g(X^k)$.

Biles [1] has described a first-order procedure for approaching the constrained formulation of the multiple-response simulation problem. This procedure consists of a gradient search along $\nabla g_1(X^k)$ as long as the search is interior to the constraints (6) and (7), but follows the gradient projection direction (see Rosen [18]) once one or more constraints are encountered. Swain [20] has compared several first-order and second-order techniques, including those of Rosen [18], Klingman and

Himmelblau [14], and Zoutendijk [21]. Swain reported very little difference among these techniques in terms of experimental requirements, and hence simulation computer time, but saw somewhat greater variability in computational requirements for the optimization algorithms. The Zoutendijk technique required significantly greater computation than the other methods.

Biles [3] demonstrated both first-order and second-order approaches to a nonlinear goal programming formulation of the multiple-response simulation problem. These approaches are based on Ignizio's adaptation [12] of the method of Griffith and Stewart [10] to goal programming and, like the constrained procedures mentioned previously, combine simulation, experimental design and mathematical programming.

The following generalized procedure is followed in employing a first-order response surface approach to the multiple-response simulation problem. The particular problem formulation and optimization procedure will govern the precise sequence of steps in implementing this procedure.

1. Identify the known experimental region $a_i \leq x_i \leq c_i$, $i = 1,\ldots,n$. Select a starting point X^0 within this region. With X^0 as its center, array an orthogonal first-order n-simplex response surface design within a selected design radius. Place $n_c = n/2 \geq 2$ points at the design center X^0 (coded as the $\underline{0}$ - vector).
2. Perform simulation trials at each of the $N = n_c + n+1$ experimental design points and record the response y_j^{ℓ}, $j = 1,\ldots,m$; $\ell = 1,\ldots,N$. Using multiple linear regression, fit linear models of the form (20).
3. Apply the appropriate optimization technique to locate the next center point in the search.
4. Repeat steps 1-3 until an "optimum" solution is located. It may be appropriate to add design points to complete a second-order response surface design to test this optimum solution. The procedure for accomplishing this is described in the next section.

Second-Order Response Surface Methods

The first-order response surface techniques described previously involve an adaptation of mathematical programming in which experimentation forms an essential component in the sequence of search steps. This section described procedures in which experimentation and computation are performed separately, applying nonlinear mathematical programming techniques to functions that are estimated through designed experiments. The following discussion treats these two aspects separately.

Experimental designs for second-order response surfaces involve at least three levels of each controllable input variable x_i, $i = 1,\ldots,n$. The design that is automatically suggested by this requirement is the 3^n factorial design. The number of design points needed for this is prohibitive, however. Box and Wilson [4] have devised a workable alternative to the 3^n factorial system through the development of central composite designs. These designs involved adding 2n radial points plus center points to a 2^n factorial or 2^{n-p} fractional factorial structure. Given that p is zero for a full factorial design, the number of design points needed for the central composite design is $2^{n-p} + 2n + n_c$, where n_c points are placed at the design center, coded as described previously.

The n-simplex design can be extended to allow estimation of the second-order response surface model

$$y = b_0 + \sum_{i=1}^{n} b_i x_i + \sum_{i=1}^{n} b_{ii} x_i^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} b_{ij} x_i x_j \qquad (31)$$

Note that this model has $(n+1)(n+2)/2$ regression coefficients. It is a convenient coincidence that by adding a design point for each edge of the n-simplex, exactly $(n+1)(n+2)/2$ design points are obtained. Of course, n_c design points are placed at the design center to enable testing for lack of fit.

The placement of the design points is as follows:

1. $n+1$ design points are placed at the vertices of a regular n-simplex as described earlier. $n_c = n/2 \geq 2$ design points are placed at the centroid of the n-simplex.
2. $(n+2)(n+3)/2$ design points are placed at locations

$$X^{ji} = a(X^i + X^j)/2 \qquad (32)$$

where a is chosen to make the design either equiradial or biradial. For an equiradial design, the value of a should be computed from the relations

$$a = \sqrt{\frac{(p-c)^2 + (n-1)(q-c)^2}{0.5(p+q-2c)^2 + (n-2)(q-c)^2}} \qquad (33)$$

$$p = \frac{\sqrt{n+1} + n - 1}{n\sqrt{2}} \qquad (34)$$

$$q = \frac{\sqrt{n+1} - 1}{n\sqrt{2}} \qquad (35)$$

$$c = \frac{p + (n-1)q}{(n+1)} \qquad (36)$$

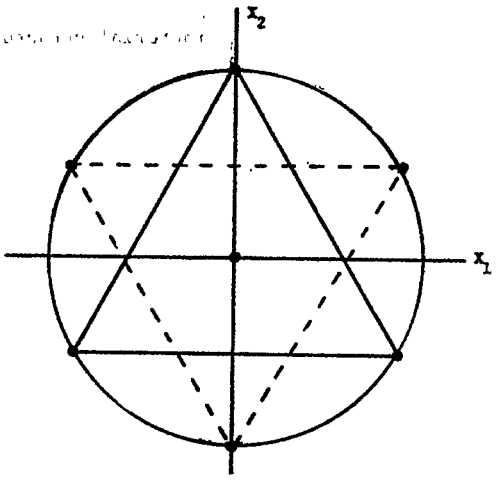
A suitable biradial design is obtained with $a = 1$, although any value of a other than that from (33) will produce a biradial design. Illustration 3 shows both an equiradial and a biradial design.

It is important to note that, despite starting with an orthogonal n-simplex first-order design, the second-order designs constructed in the manner described above are not generally orthogonal. But other desirable statistical properties are often sought with second-order designs, including rotatability and uniform precision (see Myers [16]). The equiradial design constructed from a regular n-simplex according to relations (32)-(36) is both rotatable and uniform precision.

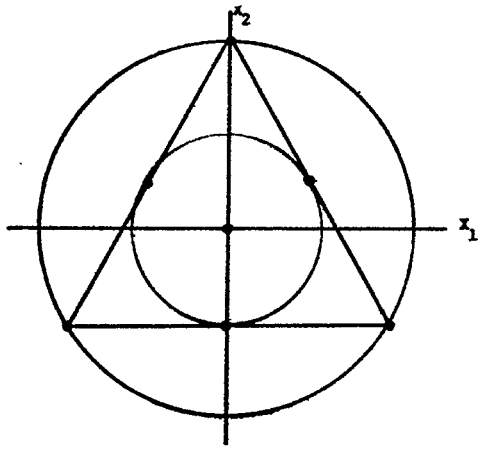
After the experimental design for the second-order problem has been constructed, simulation trials are performed at each design point and the m responses from each trial are recorded. Multiple linear regression techniques are employed to fit m models

ILLUSTRATION 3

Second-Order n-Simplex Response Surface Designs



(a) Equiradial Design



(b) Biradial Design

of the form stated in (31); that is,

$$y_k = b_{k,0} + \sum_{i=1}^{n} b_{k,i}x_i + \sum_{i=1}^{n} b_{k,ii}x_i^2 + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} b_{k,ij}x_i x_j, \quad k=1,\ldots,m \qquad (37)$$

These m models are then treated as known functions in formulating a mathematical programming problem in either the constrained format represented by equations (5)-(7), the weighted objective function format expressed by (10) and (6), the multiple-objective format stated in (11) and (6), or the nonlinear goal programming scheme shown in (14)-(16). It is usually worthwhile to perform local simulation experiments around the predicted "optimum" to test the validity of this solution.

SUMMARY

This paper has proposed several strategies for formulating and solving the multiple-response simulation problem. The great scarcity of literature reporting efforts to apply optimization to multiple-response computer simulations attests either to the need to have available workable methodology or to the utter lack of such a need. Only future efforts and publications will tell the truth of that situation.

ACKNOWLEDGMENTS

This research was initiated under grant NSF ENG74-19313 sponsored by the Engineering Division of the National Science Foundation. It is continuing under contract N00014-76-C-1021 sponsored by the Office of Naval Research.

BIBLIOGRAPHY

1. Biles, W.E., "A Gradient-Regression Search Procedure for Simulation Experimentation," Proceedings of the 1974 Winter Simulation Conference, Washington, D.C., 1974, p. 491-497.

2. Biles, W.E., "Response Surface Methods for Optimization of Constrained Multivariable Processes," Industrial and Engineering Chemistry: Process Design and Development, Vol. 14, No. 2, 1975, p. 152-158.

3. Biles, W.E., "Optimization of Multiple-Objective Computer Simulations: A Nonlinear Goal Programming Approach," Proceedings of the 1977 AIDS Conference, Chicago, 1977.

4. Box, G.E.P., and K. B. Wilson, "On the Experimental Attainment of Optimum Conditions," Journal of the Royal Statistical Society, B, Vol. 13, 1951, p. 1-45.

5. Box, G.E.P., and J. S. Hunter, "Multifactor Experimental Designs for Exploring Response Surfaces," Annals of Mathematical Statistics, Vol. 28, 1957, p. 195-242.

6. Box, M.J., "A New Method of Constrained Optimization and a Comparison with Other Methods," Computer Journal, Vol. 8, 1965.

7. Brooks, S.H., and M.R. Mickey, "Optimum Estimation of Gradient Direction in Steepest Ascent Experiments," Biometrics, Vol. 17, No. 1, 1961.

8. Draper, N.R., and H. Smith, Applied Regression Analysis, John Wiley and Sons, New York, 1966.

9. Dyer, J.S., "Interactive Goal Programming," Management Science, Vol. 19, No. 1, 1973.

10. Griffith, R.E., and R.A. Stewart, "A Nonlinear Programming Technique for the Optimization of Continuous Processing Systems, Management Science, Vol. 7, 1961.

11. Hooke, R. and T.A. Jeeves, "Direct Search Solution of Nu merical and Statistical Problems," Journal of the Association of Computing Machinery, Vol. 8,1961.

12. Ignizio, J.P., Goal Programming and Extensions, Lexington Books, D.C., Heath and Company, Lexington, Mass., 1976.

13. Kleijnen, J.P.C., Statistical Techniques in Simulation, Part I, Marcel Dekker, New York, 1975.

14. Klingman, W.R., and D.M. Himmelblau, "Nonlinear Programming with the Aid of the Multiple Gradient Summation Technique," Journal of the Association of Computing Machinery, Vol. 11, 1964.

15. Montgomery, D.C. and V.M. Bettencourt, "A Review of Multiple Response Surface Methods in Computer Simulation," 1976 Fall ORSA/TIMS Meeting, Miami Beach, 1976.

16. Myers, R.H., Response Surface Methodology, Allyn and Bacon, Boston, Mass, 1971.

17. Noh, J.C., "A Two Phase Complex Method for Nonlinear Process Optimization," 45th National ORSA/TIMS Meeting, Boston, Mass, 1974.

18. Rosen, J.B., "The Gradient Projection Method for Nonlinear Programming: Part II - Nonlinear Constraints," Journal of the Society of Nonlinear and Applied Mathematics, Vol. 8, 1961.

19. Spendley, W., and G.R. Hext, and R.F., Himsworth, "Sequential Application of Simplex Designs in Optimization and Evolutionary Operations," Technometrics, Vol. 4, 1962.

20. Swain, J.J., "Efficient Strategies for Constrained Experimental Optimization," unpublished Master's thesis, University of Notre Dame, 1977.

21. Zoutendijk, G., Methods of Feasible Directions, Elsevier, New York, 1960.