Proceedings of the 1992 Winter Simulation Conference
ed. J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson

SEQUENTIAL EXPERIMENTAL DESIGNS FOR SIMULATION METAMODELING

Joan M. Donohue
College of Business Administration
University of South Carolina
Columbia, South Carolina 29208

Ernest C. Houck
R. B. Pamplin College of Business
Virginia Polytechnic Institute & State University
Blacksburg, Virginia 24061

Raymond H. Myers
Department of Statistics
Virginia Polytechnic Institute & State University
Blacksburg, Virginia 24061

ABSTRACT

A procedure is developed for the construction of sequential simulation designs for the estimation of first- and second-order response surface metamodels. The first stage of experimentation involves the use of a fractional two-level factorial design augmented with replicated center points. Information obtained from this experimental design is used to estimate the "optimal" location of the *factorial* design points for the second stage of experimentation. Two types of performance criteria are considered in the specification of the factor settings: (1) integrated mean squared error of the predicted response variable, and (2) integrated mean squared error of the response function slopes. Additional data is collected in the second stage using a different fraction of the two-level factorial design. If quadratic curvature is indicated, a third stage of experimentation is performed to collect data for the axial portion of a central composite design. Two performance criteria are considered in the specification of the optimal *axial* levels: (1) integrated variance error of the predicted response variable, and (2) integrated variance error of the response function slopes.

The selection of factor levels in the second and third stages also depends on the strategy used in assigning random number streams to the stochastic components of the simulation model. We investigate three assignment methods (independent streams, common streams, and the assignment rule blocking strategy), and we develop sequential design plans for each strategy.

1 INTRODUCTION

Attention in this paper is focused on the estimation of a response surface metamodel for the approximation of an unknown functional relationship that exists between a controllable set of input variables and a simulated response variable. The final objective of the simulation study may be optimization, prediction, sensitivity analysis, or other considerations, and therefore we develop the sequential design plans using two different types of experimental design criteria. The first criterion involves minimizing the errors associated with predicting the *value* of the response variable, whereas the second criterion involves minimizing the errors associated with estimating *changes* in the value of the response variable. The first criterion is useful when the objective of the simulation metamodeling process is prediction of system response. The second criterion, on the other hand, is applicable when the experimenter's goal is to study the effect on system response of changes in the values of the input variables (sensitivity analysis). In the context of response surface methodology (RSM), this second criterion is useful in the early stages of an RSM study—when the experimenter has not yet located the region of the factor space that contains the optimal response. The first criterion is applicable in the final stages of an RSM study—when prediction of the optimum response and estimation of the optimum operating conditions are frequent objectives of experimenters. (See the texts by Box and Draper 1987, Khuri and Cornell 1987, and Myers 1976 for additional information on RSM.)

Regardless of the purpose and goals of the study, the simulation metamodeler must choose an appropriate region of the factor space for experimentation. The sequential design procedure assumes that a first- or second-order metamodel will be adequate for describing the response variable as a function of the input factors. Therefore, in defining the *region of interest*, the experimenter needs to choose the ranges of the input factors such that the response surface, within that region, is approximately linear or quadratic in nature. We assume that the region of interest (after coding the input variables) is cuboidal in shape and specified by the experimenter's choice of the high and low levels of each input factor. Note that the high and low factor levels associated with the region of interest are different from the maximum and minimum factor levels which are typically associated with a larger *operability region*.

In addition to selecting the region of interest, the method of assigning random number streams to the experimental design points needs to be chosen by the experimenter. We have developed sequential design plans for the following three assignment methods:

1. independent random number streams (**IR**): the use of independent stream sets for each stochastic model component on each simulation run,
2. common random number streams (**CR**): the use of independent stream sets for the stochastic components of the simulation model (within a run), and the use of a common set of streams for each nonreplicated simulation run (between runs), and
3. assignment rule blocking strategy (**AR**): the use of the common streams strategy within one block of an orthogonally blockable experimental design and the use of an antithetic set of streams in the second orthogonal block (Schruben and Margolin, 1978).

We assume that the experimenter is collecting the simulated response data in a manner that would provide independent response observations if the **IR** strategy were used. For example, the response of interest could be the average value of some performance measure during a terminating simulation or, it could be the average value of a steady state performance measure obtained from a simulation in which the initial transient period was deleted. When the response data is collected in this manner, and the **IR** strategy is used to assign random number streams to the stochastic model components, the simulated response observations are independent and the simulation output can be analyzed using standard statistical techniques. However, when the **CR** and **AR** strategies are used, the common and antithetic stream sets induce correlation among the simulated responses and, therefore, more sophisticated statistical analysis techniques are required. The potential benefits of these strategies (e.g., improved prediction of the responses and reduced variances of some metamodel parameters) may offset the increased computational efforts. Implementation of either of these correlation induction strategies requires careful synchronization of the random number streams in order to maximize the potential benefits of the strategy. Additionally, a pilot study must be performed to ascertain whether or not correlation is actually being induced among the responses, to perform appropriate statistical tests for the assumed correlation structure, and to estimate the magnitudes of the induced correlations.

There are many other important issues that simulation metamodelers must address. For example, the input distributions of the stochastic model components, the starting conditions of the simulation, the length of the simulation run (and possibly the transient period), the input factors needed to model the system response, etc., must be determined by the experimenter. (See Law and Kelton 1991 for a thorough discussion of these and other related issues.) In this research, we assume that the experimenter has developed a valid simulation model of a system that can be used to generate independent response observations. This simulation model will generate the data needed to develop a response surface metamodel of the system and, in turn, the metamodel will be used to *efficiently* study and/or optimize the system response variable of interest.

2 PREVIOUS RESEARCH

Many studies have examined the performance of the three random number assignment strategies noted earlier. Schruben and Margolin (1978) developed the **AR** strategy and investigated its performance relative to the **IR** and **CR** strategies. In the context of fitting a first-order metamodel, using a single replication of an experimental design that partitions into two orthogonal blocks, these authors found that the **AR** strategy was the preferred procedure when the performance criterion was the minimization of either the trace or the determinant of the estimator covariance matrix. Kiefer (1978) extended these results to additional variance-oriented criteria, such as the minimization of the maximum eigenvalue of the estimator covariance matrix and the minimization of the maximum diagonal element of the estimator covariance matrix. Three additional variance-oriented criteria (prediction variance, integrated variance, and variance of slopes) were examined by Hussey, Myers and Houck (1987a, 1987b) in the context of fitting first- and second-order metamodels using single replications of experimental designs that partition into an even number of orthogonal (or nearly orthogonal) blocks. Similar to the earlier research of Schruben and Margolin, their results indicated a preference for the **AR** strategy, and this preference existed regardless of the number of blocks and whether the design blocking was exactly orthogonal.

The **IR**, **CR**, and **AR** strategies were further studied by Donohue, Houck and Myers (1992a, 1992b) using two integrated mean squared error performance criteria (MSE of the predicted response and MSE of the response function slopes). These authors considered both first- and second-order experimental designs and assumed that the predicted response could be biased due to misspecification of the metamodel. Additionally, these authors extended the original formulations of the **CR** and **AR** strategies by accommodating both replicated center runs and three-block designs through the use of independent stream sets. Results of their research indicated a preference for the **AR** strategy even when independent stream sets were used to replicate center runs and/or generate design points in a third block.

Research has also been conducted to address the statistical analysis issues associated with the **CR** and **AR** strategies. Due to the correlation that is induced between responses generated using these strategies, standard statistical procedures that assume independent responses cannot be used. Joshi and Tew (1992), as well as Kleijnen (1992), developed statistical analysis and validation procedures for the **CR** strategy. Nozari, Arnold and Pegden (1987) discussed the validity of the assumptions used in the theoretical development of the **AR** strategy and presented inferential procedures for the analysis of response data simulated using the assignment rule. Additional research on the validation of assumptions underlying the **AR** strategy was performed by Tew and Wilson (1992a, 1992b). These authors presented a statistical test for lack-of-fit to the postulated metamodel and provided formal inferential procedures for examining a number of the assumptions associated with the **AR** strategy.

3 PRELIMINARY INFORMATION

The problem under consideration is the development of a simulation metamodel for the estimation of a response variable, Y, as a function of a set of k controllable factors, ξ_i $(i = 1, \ldots, k)$, within a specified region of the factor space. An experimental design plan consists of N design points which specify the levels of the factors on each simulation run. The $N \times k$ design matrix $(N > k)$ can be written as

$$\boldsymbol{D} = \begin{bmatrix} \xi_{11} & \xi_{21} & \cdots & \xi_{k1} \\ \xi_{12} & \xi_{22} & \cdots & \xi_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ \xi_{1N} & \xi_{2N} & \cdots & \xi_{kN} \end{bmatrix} \tag{1}$$

where ξ_{iu} $(i = 1, \ldots, k;\ u = 1, \ldots, N)$ denotes the level of the ith factor on the uth simulation run. The ξ_{iu} represent coded values of continuous, quantitative input variables, such that +1, 0, and −1, respectively, correspond to the high, middle, and low levels of each factor for the specified region of interest. We denote each row of $\boldsymbol{D}$, the uth design point, as a k-dimensional row vector $\boldsymbol{\xi}'_u$ $(u = 1, \ldots, N)$. The random number streams associated with $\boldsymbol{\xi}'_u$ are denoted

$$\boldsymbol{R}_u = \boldsymbol{R}_{1u}, \ldots, \boldsymbol{R}_{mu} \tag{2}$$

where the column vector $\boldsymbol{R}_{ju}$ $(j = 1, \ldots, m;\ u = 1, \ldots, N)$ is the jth stream of random numbers used on the uth simulation run, and m is the number of stochastic components in the simulation model. Using the design matrix in (1) and the random number streams in (2), a set of N simulated responses, Y_u $(u = 1, \ldots, N)$ are obtained. The values of Y_u would typically be the average response during the simulation run, possibly after the deletion of a transient period. The simulation metamodel representing the relationship between the response variable and the k controllable input factors can be written as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{3}$$

where $\boldsymbol{Y}$ and $\boldsymbol{\epsilon}$ are N-dimensional column vectors of the response observations and random errors, respectively, $\boldsymbol{\beta}$ is a p_i-dimensional column vector of model coefficients, $\boldsymbol{X}$ is an $N \times p_i$ matrix of the regressor terms in the metamodel, and p_i $(i = 1, 2)$ is the number of parameters in a response surface metamodel of order i. In the case of a first-order metamodel, which includes an intercept term and k linear terms, we have $p_1 = 1 + k$. A second-order metamodel additionally includes k quadratic and $\binom{k}{2}$ second-order interaction terms, yielding $p_2 = 1 + 2k + \binom{k}{2}$. We assume that the expected value and dispersion matrix of the vector of random errors are $\mathrm{E}(\boldsymbol{\epsilon}) = \boldsymbol{0}$ and $\mathrm{cov}(\boldsymbol{\epsilon}) = \boldsymbol{\Sigma}$, with $\boldsymbol{\Sigma}$ positive definite, $\mathrm{var}(\epsilon_u) = \sigma^2$ and $\mathrm{cov}(\epsilon_u, \epsilon_v) = \rho_{uv}\sigma^2$, where ρ_{uv} is the correlation between ϵ_u and ϵ_v, $(u, v = 1, \ldots, N)$. Under these assumptions, the covariance matrix of the response vector becomes $\mathrm{cov}(\boldsymbol{Y}) = \boldsymbol{\Sigma}$. Additionally, the hypothesis tests for lack-of-fit to the assumed metamodel require the usual assumption of normally distributed ϵ_u. The homogeneity of variance assumption is consistent with the previous research discussed in §2, however, in situations where the assumption is not valid, we suggest that an appropriate variance stabilizing transformation be applied to the response data (see Box and Draper 1987, pp. 283-291). For further discussion on the homogeneity of variance assumption, see Donohue, Houck and Myers (1992b), Nozari, Arnold and Pegden (1987), Schruben and Margolin (1978), and Tew and Wilson (1992a).

The metamodel parameters in (3) can be estimated using ordinary least squares (OLS) or generalized least squares (GLS), assuming the $\boldsymbol{X}$ matrix is nonsingular. These estimators are given by

$$\hat{\boldsymbol{\beta}} = \begin{cases} (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} & \text{for OLS} \\ (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} & \text{for GLS.} \end{cases} \tag{4}$$

The GLS estimator has the disadvantage of being dependent on the covariance structure of the response observations, and the estimation of $\boldsymbol{\Sigma}$ requires replication of the design points. However, the dispersion matrices of both the OLS and GLS estimators in (4),

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = \begin{cases} (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} & \text{for OLS} \\ (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1} & \text{for GLS,} \end{cases}$$

depend on $\boldsymbol{\Sigma}$. Since the "optimal" sequential designs that we develop for the **CR** and **AR** strategies depend on the $\mathrm{cov}(\hat{\boldsymbol{\beta}})$, a pilot study is needed to estimate the magnitudes of the induced correlations in $\boldsymbol{\Sigma}$. Even for the **IR** strategy, with $\boldsymbol{\Sigma}_{\mathrm{IR}} = \sigma^2 \boldsymbol{I}_N$, we recommend that a pilot study be performed in order to check for violations of the normality and homogeneity of variance assumptions.

The following two sections present the experimental design criteria used to determine the optimal sequential design plans and the correlation induction strategies used to assign random number streams to the stochastic model components.

3.1 Design Criteria

Response surface models are generally first- or second-order regression models that are intended to provide the experimenter with some knowledge about the nature of the true, but unknown and frequently complicated, relationship between the response and the input variables. Box and Draper (1959) presented a design criterion that protects against model misspecification by incorporating both variance and bias errors of the predicted response. Since the assumed metamodel in (3) can only approximate the true functional relationship, Box and Draper proposed that the experimental design plan be developed so as to protect against biases in the predicted response due to model terms of order one degree higher than the fitted metamodel. In the case of a first-order model, protection against bias due to unfitted quadratic and two-way interaction terms would be desired and, in the case of a second-order model, the experimental design plan should protect against bias due to unfitted third-degree terms.

Box and Draper's design criterion is the minimization of the integrated mean squared error of the predicted response variable, normalized with respect to the number of experimental design points and the experimental error variance. Letting Ξ denote the region of interest, $\boldsymbol{\xi}$ denote a point within Ξ, $\hat{Y}_{(\boldsymbol{\xi})}$ denote the fitted response at the point $\boldsymbol{\xi} \in \Xi$, and $Y_{(\boldsymbol{\xi})}$ denote the "true" response (assuming the true metamodel is of order one degree higher than the fitted model), the integrated *MSE of response* criterion becomes the minimization of

$$\begin{aligned} \mathsf{J} &= \frac{N\Omega_\Xi}{\sigma^2}\int_\Xi \mathrm{E}\left\{\left[\hat{Y}_{(\boldsymbol{\xi})} - \mathrm{E}(Y_{(\boldsymbol{\xi})})\right]^2\right\} d\boldsymbol{\xi} \\ &= \mathsf{V} + \mathsf{B} \end{aligned} \tag{5}$$

where $\Omega^{-1} = \int_\Xi d\boldsymbol{\xi}$ is the volume of the region of interest and V and B are the integrated variance and bias errors of the predicted response, defined as

$$\begin{aligned} \mathsf{V} &= \frac{N\Omega_\Xi}{\sigma^2}\int_\Xi \mathrm{var}\left(\hat{Y}_{(\boldsymbol{\xi})}\right) d\boldsymbol{\xi} \\ \mathsf{B} &= \frac{N\Omega_\Xi}{\sigma^2}\int_\Xi \mathrm{bias}^2\left(\hat{Y}_{(\boldsymbol{\xi})}\right) d\boldsymbol{\xi}. \end{aligned} \tag{6}$$

An extension of the integrated MSE of response criterion, involving the partial derivatives, or slopes, of the response function was developed by Myers and Lahoda (1975). Similar to Box and Draper, these authors assume that the experimental design plan should provide protection against bias due to unfitted model terms of order one degree higher than those in the fitted metamodel. The integrated *MSE of slopes* criterion, however, considers only the variance and bias errors that affect the gradient information obtained from the metamodel coefficients. Therefore, it is a useful criterion when estimating the rate of change in the response variable is more important than estimating the value of the response variable. This criterion, as presented by Myers and Lahoda, calls for the minimization of

$$\begin{aligned} \mathsf{J}^* &= \frac{N\Omega_\Xi}{\sigma^2}\int_\Xi \mathrm{E}\left\{\left[\hat{\gamma}_{(\boldsymbol{\xi})} - \gamma_{(\boldsymbol{\xi})}\right]'\left[\hat{\gamma}_{(\boldsymbol{\xi})} - \gamma_{(\boldsymbol{\xi})}\right]\right\} d\boldsymbol{\xi} \\ &= \mathsf{V}^* + \mathsf{B}^* \end{aligned} \tag{7}$$

where $\hat{\gamma}_{(\boldsymbol{\xi})}$ and $\gamma_{(\boldsymbol{\xi})}$ are k-dimensional column vectors of the partial derivatives of the fitted and true metamodels (e.g., the third row of $\gamma_{(\boldsymbol{\xi})}$ is partial derivative of $Y_{(\boldsymbol{\xi})}$ with respect to the third regressor variable, ξ_3), and V^* and B^* are the integrated variance and bias errors of the response function slopes, defined as

$$\begin{aligned} \mathsf{V}^* &= \frac{N\Omega_\Xi}{\sigma^2}\int_\Xi \mathrm{var}\left(\hat{\gamma}_{(\boldsymbol{\xi})}\right) d\boldsymbol{\xi} \\ \mathsf{B}^* &= \frac{N\Omega_\Xi}{\sigma^2}\int_\Xi \mathrm{bias}^2\left(\hat{\gamma}_{(\boldsymbol{\xi})}\right) d\boldsymbol{\xi}. \end{aligned} \tag{8}$$

The first stage of the sequential design procedure presented in §4 is performed in order to estimate the bias error resulting from unfitted second-order terms. In the second stage, the design is augmented with additional points such that the combined design plan minimizes **J** or **J*** for a fitted first-order metamodel. If there is an indication of quadratic curvature in the response surface, then a third stage of the sequential design process is performed. The combined, three-part design plan is a central composite design that minimizes **V** or **V*** for a fitted second-order metamodel. We call these sequential designs plans "Min-**V**/Min-**J**" and "Min-**V***/Min-**J***" designs.

3.2 Assignment Strategies

In addition to the selection of a design criterion, the experimenter must choose a method of assigning random number streams to design points. Similar to the earlier research noted in §2, we consider three assignment strategies: **IR**, **CR**, and **AR**. In this section, we present the structure of $\boldsymbol{\Sigma}$, the covariance matrix of the response observations for each strategy. The basic form of this matrix,

$$\boldsymbol{\Sigma} = \mathrm{cov}(\boldsymbol{Y}) = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1N} \\ \rho_{21} & 1 & & \vdots \\ \vdots & & \ddots & \rho_{N-1,N} \\ \rho_{N1} & \rho_{N2} & \cdots & 1 \end{bmatrix} \sigma^2, \tag{9}$$

is a result of the homogeneity of variance assumption discussed in §2.

For the **IR** strategy, a different random number stream is assigned to the m stochastic model components on the N simulation runs, thereby requiring mN stream sets. This strategy generates independent response observations ($\rho_{uv} = 0$, for all $u, v = 1, \ldots, N$), and the covariance matrix in (9) becomes $\boldsymbol{\Sigma}_{\mathrm{IR}} = \sigma^2 \boldsymbol{I}_N$.

For the **CR** strategy (as applied in this research), a common set of m streams is used for all non-center points, and independent stream sets are used for center runs. Letting N_{C} denote the number of center runs, the **CR** strategy requires $m(N_{\mathrm{C}} + 1)$ stream sets for the experimental design. Following the assumptions of Schruben and Margolin (1978), we assume that pairs of responses generated with common streams are positively correlated with a constant magnitude of ρ_+ ($0 \leq \rho_+ < 1$); that is, $\mathrm{cov}(Y_u, Y_v) = \sigma^2 \rho_+$ if Y_u and Y_v are generated with a common set of random number streams, say $\boldsymbol{R}_1$. If the response vector is partitioned as $\boldsymbol{Y}' = (\boldsymbol{Y}'_{N_{\mathrm{NC}}} | \boldsymbol{Y}'_{N_{\mathrm{C}}})$, where $\boldsymbol{Y}_{N_{\mathrm{NC}}}$ consists of N_{NC} non-center points and $\boldsymbol{Y}_{N_{\mathrm{C}}}$ consists of N_{C} center points, then the partitioned covariance matrix for the **CR** strategy becomes

$$\boldsymbol{\Sigma}_{\mathrm{CR}} = \left[\begin{array}{cccc|c} 1 & \rho_+ & \cdots & \rho_+ & \\ \rho_+ & 1 & & \vdots & \boldsymbol{0}_{N_{\mathrm{NC}}}\boldsymbol{0}'_{N_{\mathrm{C}}} \\ \vdots & & \ddots & \rho_+ & \\ \rho_+ & \rho_+ & \cdots & 1 & \\ \hline \multicolumn{4}{c|}{\boldsymbol{0}_{N_{\mathrm{C}}}\boldsymbol{0}'_{N_{\mathrm{NC}}}} & \boldsymbol{I}_{N_{\mathrm{C}}} \end{array}\right] \sigma^2 \tag{10}$$

where $\boldsymbol{0}_i$ is an i-dimensional column vector of zeros.

The **AR** strategy requires that the experimental design partition into orthogonal blocks. In this research, a common set of streams, say $\boldsymbol{R}_1$, is used in the first block, and the antithetic set of streams, $\overline{\boldsymbol{R}}_1 = \boldsymbol{1} - \boldsymbol{R}_1$, is used in the second block (where $\boldsymbol{1}$ is an appropriately dimensioned matrix of ones). The first and second blocks are different fractions of a two-level factorial design. Center runs and any axial points for a central composite design are placed in a third block and use independent stream sets. Again, we follow the assumptions of Schruben and Margolin (1978) and assume that pairs of responses generated with common streams (within either of the first two blocks) are positively correlated with a constant magnitude of ρ_+, and pairs of responses generated with antithetic stream sets (one from each of the first two blocks) are negatively correlated with a constant magnitude of $-\rho_-$ ($0 \leq \rho_- < 1$); that is, $\mathrm{cov}(Y_u, Y_v) = -\sigma^2 \rho_-$ if Y_u and Y_v are generated with an antithetic set of random number streams, say $\boldsymbol{R}_1$ and $\overline{\boldsymbol{R}}_1$. If the response vector is partitioned as $\boldsymbol{Y}' = (\boldsymbol{Y}'_{N_1} | \boldsymbol{Y}'_{N_2} | \boldsymbol{Y}'_{N_3})$, where $\boldsymbol{Y}'_{N_i}$ consists of the N_i points in the ith block, then the partitioned covariance matrix for the **AR** strategy becomes

$$\boldsymbol{\Sigma}_{\mathrm{AR}} = \left[\begin{array}{c|c|c} (1-\rho_+)\boldsymbol{I}_{N_1} + \boldsymbol{1}_{N_1}\boldsymbol{1}'_{N_1}\rho_+ & -\boldsymbol{1}_{N_1}\boldsymbol{1}'_{N_2}\rho_- & \boldsymbol{0}_{N_1}\boldsymbol{0}'_{N_3} \\ \hline -\boldsymbol{1}_{N_2}\boldsymbol{1}'_{N_1}\rho_- & (1-\rho_+)\boldsymbol{I}_{N_2} + \boldsymbol{1}_{N_2}\boldsymbol{1}'_{N_2}\rho_+ & \boldsymbol{0}_{N_2}\boldsymbol{0}'_{N_3} \\ \hline \boldsymbol{0}_{N_3}\boldsymbol{0}'_{N_1} & \boldsymbol{0}_{N_3}\boldsymbol{0}'_{N_2} & \boldsymbol{I}_{N_3} \end{array}\right] \sigma^2 \tag{11}$$

where $\boldsymbol{1}_i$ is an i-dimensional column vector of ones.

4 SEQUENTIAL DESIGN PROCEDURE

The sequential design procedure for the estimation of a first- or second-order metamodel consists of two primary stages (stages one and two), an optional third stage, and a pilot study that is required when the **CR** or **AR** strategy is used. In this section, we present the sequential procedures of developing both Min-**V**/Min-**J** and Min-**V***/Min-**J*** experimental design plans.

As an example of the three stages of the design procedure, consider the three factor experimental design in Table 1. The right-hand side of the table specifies

Table 1: A $k = 3$ Sequential Design

	ξ_1	ξ_2	ξ_3	IR	CR	AR
$D_1 =$	$+1$	$+1$	$+1$	R_1	R_1	R_1
	$+1$	-1	-1	R_2	R_1	R_1
	-1	$+1$	-1	R_3	R_1	R_1
	-1	-1	$+1$	R_4	R_1	R_1
	0	0	0	R_5	R_2	R_2
	0	0	0	R_6	R_3	R_3
	0	0	0	R_7	R_4	R_4
$D_2 =$	$+g$	$+g$	$-g$	R_8	R_1	$\overline{R}_1$
	$+g$	$-g$	$+g$	R_9	R_1	$\overline{R}_1$
	$-g$	$+g$	$+g$	R_{10}	R_1	$\overline{R}_1$
	$-g$	$-g$	$-g$	R_{11}	R_1	$\overline{R}_1$
$D_3 =$	$+\alpha$	0	0	R_{12}	R_1	R_6
	$-\alpha$	0	0	R_{13}	R_1	R_7
	0	$+\alpha$	0	R_{14}	R_1	R_8
	0	$-\alpha$	0	R_{15}	R_1	R_9
	0	0	$+\alpha$	R_{16}	R_1	R_{10}
	0	0	$-\alpha$	R_{17}	R_1	R_{11}

the random number streams assigned to each design point for the **IR**, **CR**, and **AR** strategies. The design matrix for the first stage, D_1, is a fractional factorial design (here, a one-half fraction) with the levels of the coded input variables, ξ_{iu}, at the ± 1 extremes of the cuboidal region of interest. The design is also augmented with replicated center runs ($\xi_{iu} = 0$, for $i = 1, \cdots, k$) in order to estimate the quadratic bias parameter needed for minimization of **J** in (5), or **J*** in (7), in the second stage. The design matrix for the second stage, D_2, is a different fraction of the factorial design, and the levels of the input variables are located at $\pm g$. The combined design plan of the first two stages is a first-order experimental design that minimizes the estimated value of **J** or **J*** through an appropriate choice of the factor level g. Utilizing information from the fitted first-order metamodel, a statistical test for quadratic lack-of-fit is performed. If a second-order metamodel is needed to adequately describe the response surface, then a third stage is performed in which the factorial design is augmented with D_3, the axial portion of a central composite design. The value of α is selected so as to minimize **V** in (6), or **V*** in (8), of a second-order metamodel. The combined design plan, utilizing all three stages, is a Min-**V**/Min-**J** or Min-**V***/Min-**J*** experimental design.

4.1 Pilot Study

For all three assignment strategies, a pilot study should be performed to check for violations of the normality and homogeneity of variance assumptions. We recommend that the pilot study consist of the n fractional factorial design points used in the first stage of the sequential design procedure. (For example, $n = 4$ for the design in Table 1.) Tew and Wilson (1992a) recommend that the number of independent replications for the pilot study, r, be within the range $\max\{2n, 16\} \leq r \leq 32$. These authors develop a test for multivariate normality based on the Shapiro-Wilk statistic and they recommend that the response data at each of the n design points in the pilot study be checked for univariate normality using normal probability plots. If the multivariate or univariate normality assumptions do not hold, then the experimenter should apply an appropriate "normalizing" transformation to the response data (see Box and Draper 1987, p. 281).

The homogeneity of variance assumption can be checked using Bartlett's test (1937) when the response data is normally distributed or, if slight deviations from normality exist, any of the robust tests investigated by Conover, Johnson and Johnson (1981) can be used. If the homogeneity of variance assumption is violated, then the experimenter should use an appropriate "variance-stabilizing" transformation (see Box and Draper 1987, p. 283).

For the **CR** and **AR** strategies, the pilot study is also used to check for violations of the assumed covariance structure in (10) or (11). Morrison (1990, p. 294) provides a likelihood-ratio test for the assumed structure of Σ under the **CR** strategy and Tew and Wilson (1992a) present a similar likelihood-ratio test for the **AR** strategy. If the assumed covariance structure appears to be incorrect, then the experimenter should carefully check the simulation program for proper synchronization of the random number streams (see Law and Kelton 1991, p. 619). When the experimenter has taken the necessary steps to achieve proper synchronization, yet the assumed covariance structure under the **CR** or **AR** strategy appears to be incorrect, we recommend that the **IR** strategy be used.

For experimenters using the **CR** or **AR** strategy, estimates of the induced correlation magnitudes are needed for computations performed in stages one and two. Silver and Dunlap (1987) note that when averaging a set of correlation coefficients, the average Pearson correlation coefficient is biased negatively, whereas the back transform of the average Fisher's z-transformation ($z = \tanh^{-1}\rho$) tends to be biased positively, but the magnitude of the latter bias is smaller. These authors, as well as Rao (1973), recommend the use of Fisher's z-transformation for estimating an average correlation coefficient. Using the z-transformation, the estimate of ρ_+ for the **CR** strategy becomes $\hat{\rho}_+ = (e^{2\bar{z}} - 1)/(e^{2\bar{z}} + 1)$, where $\bar{z}$ is the average of the $n(n-1)/2$ z-transformations of the pairwise correlations in a pilot study of n design points. The estimates, $\hat{\rho}_+$ and $\hat{\rho}_-$, for the **AR** strategy are similarly computed, using $n(n-2)/4$ positive correlations and $n^2/4$ negative correlations. Rao (1973, p. 434) also provides a useful statistical test for the homogeneity of a set of correlation coefficients. We recommend that this test be performed on the set of correlations used to compute $\hat{\rho}_+$ and $\hat{\rho}_-$.

Equation (4) indicates that when the metamodel parameters are estimated using GLS, an estimate of $\boldsymbol{\Sigma}$ is also required. Since the pilot study only replicates design points in the first stage of the three-stage experimental design, we cannot estimate $\boldsymbol{\Sigma}$ for the entire experimental design unless we assume that the structure of $\boldsymbol{\Sigma}$ will be the same in each stage. Empirical results of the research studies noted in §2 indicate that this assumption appears to be reasonable. Therefore, we suggest that the experimenter compute $\widehat{\boldsymbol{\Sigma}}_{\mathrm{CR}}$ and $\widehat{\boldsymbol{\Sigma}}_{\mathrm{AR}}$ using equations (10) and (11), replacing ρ_+, ρ_-, and σ^2 with their estimators $\hat{\rho}_+$, $\hat{\rho}_-$, and $\hat{\sigma}^2$. In following section, we provide an appropriate estimator of σ^2 using response data collected from the center points in the first stage of the design.

4.2 First Stage

The purpose of the first stage of the sequential design procedure is to estimate the quadratic bias parameter needed for minimization of **J** in (5) or **J*** in (7). A fractional two-level factorial design, augmented with replicated center runs, is used in this stage (see the design matrix $\boldsymbol{D}_1$ in Table 1). The center runs use independent random number streams for all three strategies, but the factorial points use a common set of streams for the **CR** and **AR** strategies. For the **AR** strategy, this fractional factorial design is one of two orthogonal blocks (the second block is the design used in stage two).

Box and Draper (1987, pp. 72, 189), Khuri and Cornell (1987, p. 167), and Myers (1976, p. 116) indicate that independently replicated center runs provide estimates of both σ^2 and the quadratic bias parameter $\theta\phi = N(\sum_{i=1}^{k}\beta_{ii})^2/\sigma^2$. The pure experimental error variance is computed from the independently replicated center runs as

$$\hat{\sigma}^2 = \frac{\sum_{u=1}^{N_{\mathrm{C}}}(Y_u - \overline{Y}_{\mathrm{C}})^2}{N_{\mathrm{C}} - 1} \quad \text{for } u \in \{\text{center runs}\}$$

where $\overline{Y}_{\mathrm{C}}$ is the average response at the N_{C} center points. If we assume that the true response surface is second-order, then an unbiased estimator of the sum of the pure quadratic coefficients, $\sum_{i=1}^{k}\beta_{ii}$, is the difference $\overline{Y}_{\mathrm{NC}} - \overline{Y}_{\mathrm{C}}$, where $\overline{Y}_{\mathrm{NC}}$ is the average response at the $N - N_{\mathrm{C}}$ non-center (factorial) points. A reasonable estimator of the quadratic bias parameter is then

$$\widehat{\theta\phi} = \frac{N}{\hat{\sigma}^2}(\overline{Y}_{\mathrm{NC}} - \overline{Y}_{\mathrm{C}})^2.$$

The optimal levels for the design points in stage two of the sequential design procedure, g, can then be determined from the Min-**J** or Min-**J*** values of the pure second-order design moment, defined as $\lambda = \sum_{u=1}^{N}\xi_{iu}^2/N$, for $i = 1, \cdots, k$. For the combined design of stages one and two, this design moment, denoted λ_2, is computed as

$$\lambda_2 = \frac{N_{\mathrm{F}}}{2N_2}\left(1 + g^2\right) \tag{12}$$

where N_{F} is the total number of factorial design points and N_2 is the total number of design points in the combined design for stages one and two. (For example, $N_{\mathrm{F}} = 8$ and $N_2 = 11$ for the design in Table 1.) The optimal value of λ_2, denoted $\lambda_2^\dagger$, is "as large as possible" for a Min-**J*** design. For a cuboidal region in the coded design variables, we specify

$$\lambda_2^\dagger = 1 \qquad \text{for a Min-}\mathbf{J}^*\text{ design.} \tag{13}$$

The optimal value of $\lambda\,(0 < \lambda \leq 1)$ for a Min-**J** design, obtained by setting the derivative $d\mathbf{J}/d\lambda_2$ equal to zero, tends to be smaller than the Min-**J*** value in (13). For the **CR** and **AR** strategies, this optimal design moment depends on the induced correlation magnitudes ($\hat{\rho}_+$ and $\hat{\rho}_-$) and the parameter estimation technique (OLS or GLS). Due to space limitations, we refer the reader to Donohue, Houck and Myers (1992a) for the equations needed to compute the Min-**J** value of $\lambda_2^\dagger$. Having determined the value of $\lambda_2^\dagger$, we solve equation (12) for the optimal level of the factorial design points in the second stage,

$$g^\dagger = \sqrt{\frac{2N_2\lambda_2^\dagger}{N_{\mathrm{F}}} - 1}. \tag{14}$$

In some instances, additional center runs are needed in stage two in order to achieve the optimal value of $g^\dagger$ in (14).

4.3 Second Stage

The purpose of the second stage is to estimate a first-order metamodel using a design that, when combined with stage one, minimizes the estimated mean squared error (**J** or **J***). The optimal levels of the factorial design points for this second stage are $g^{\dagger}$ (see the design matrix $\boldsymbol{D}_2$ in Table 1). Independent random number streams are used for the **IR** strategy in stage two but the **CR** strategy uses the common set of streams from stage one again in stage two. For the **AR** strategy, the second stage represents the second orthogonal block (see Myers 1976, p. 180, for the orthogonal blocking requirements in first-order designs) and the stream set that is antithetic to the common set used in stage one is used in stage two.

An additional objective of the second stage is to perform a hypothesis test for quadratic lack-of-fit. A test of the null hypothesis $\mathrm{H}_0\colon \sum_{i=1}^{k} \beta_{ii} = 0$ versus $\mathrm{H}_a\colon \sum_{i=1}^{k} \beta_{ii} \neq 0$ is provided by Khuri and Cornell (1987, p. 167) and Myers (1976, p. 116). The test statistic, $F = (N_C N_F(\overline{Y}_{NC} - \overline{Y}_C)^2/(\hat{\sigma}^2(N_C + N_F))$ follows an F_{1,N_C-1} distribution. If the null hypothesis cannot be rejected, then the third stage of the sequential design procedure need not be performed since a first-order metamodel appears to be adequate. However, the experimenter should test for lack-of-fit due to unfitted interaction terms (see Box and Draper 1987, p. 73 and Khuri and Cornell 1987, p. 155) and include these terms in the fitted first-order metamodel if necessary.

When the null hypothesis of no quadratic lack-of-fit is rejected, indicating the need for quadratic terms in the metamodel, a third stage should be performed in order to fit a second-order metamodel. By augmenting the design of stages one and two with the axial points of a central composite design, a second-order metamodel can be estimated. Since the factorial design does not provide the estimates of cubic bias that would be needed to minimize **J** or **J*** of a second order design, we instead minimize **V** in (6), or **V*** in (8), in this third stage through an appropriate choice of the axial levels, α. Similar to approach taken in stage one, we determine the value of the pure second-order design moment that results in a minimum value of **V** or **V***, for $0 < \lambda \leq 1$. For the combined three-design, this design moment, denoted λ_3, is computed as

$$\lambda_3 = \frac{N_F}{2N_3}\left(1 + g^2\right) + \frac{2\alpha^2}{N_3} \tag{15}$$

where N_3 is the total number of design points in the combined design of stages one, two and three. Donohue, Houck and Myers (1992c) present scalar equations for **V** and **V*** that can be used to iteratively determine the optimal value of λ_3, denoted $\lambda_3^{\dagger}$. These equations are developed for each of the following situations: **IR** strategy; **CR** using OLS or GLS estimation; and **AR** using OLS or GLS. Empirical results show that $\lambda_3^{\dagger}$ tends to be larger for the Min-**V*** designs than for the Min-**V** designs, and it tends to be larger for OLS estimation than for GLS estimation. Inserting the optimal value of λ_3 into equation (15) and solving for α, we find that the optimal level of the axial design points for third stage is

$$\alpha^{\dagger} = \sqrt{\frac{N_3\lambda_3^{\dagger}}{2} - \frac{N_F}{4}(1 + g^2)}.$$

In some instances additional center runs are needed to achieve the optimal value of $\alpha^{\dagger}$ in stage three.

4.4 Optional Third Stage

The third stage of the sequential design procedure augments the factorial design of stages one and two with the axial portion of a central composite design. The optimal levels of the axial design points in the third stage, $\alpha^{\dagger}$, are selected for minimization of the integrated variance error (**V** or **V***) of a second-order design (see the design matrix $\boldsymbol{D}_3$ in Table 1). The **IR** and **AR** strategies use independent random number streams for the axial points, but the **CR** strategy uses the common stream set used in stages one and two. We recommend that this stage be performed when there is evidence of quadratic lack-of-fit in the first-order metamodel. The Min-**V**/Min-**J** or Min-**V***/Min-**J*** central composite design provides the response data needed to fit a second-order metamodel with linear, quadratic and two-way interaction terms.

5 SUMMARY

We have presented a method of estimating first- and second-order simulation metamodels using a sequential design procedure. The Min-**V**/Min-**J** and Min-**V***/Min-**J*** that we develop can be used with one of three random number assignment strategies (**IR**, **CR**, or **AR**) and with either of two parameter estimation techniques (OLS or GLS). Empirical evidence generated by the authors indicates superior performance of the **AR** strategy, particularly when GLS estimation is used. Since analytical comparisons of the performance of the designs cannot be made, we suggest that further empirical research be conducted using a variety of true response functions.

REFERENCES

Bartlett, M. S. 1937. Properties of sufficiency and statistical tests. *Proceedings of the Royal Society A* 160:268–282.

Box, G. E. P., and N. R. Draper. 1959. A basis for the selection of a response surface design. *Journal of the American Statistical Association* 54:622–654.

Box, G. E. P., and N. R. Draper. 1987. *Empirical model-building and response surfaces.* New York: John Wiley & Sons.

Conover, W. J., M. E. Johnson, M. E., and M. M. Johnson. 1981. A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics* 21:351–361.

Donohue, J. M., E. C. Houck, and R. H. Myers. 1992a. Simulation designs and correlation induction for reducing second-order bias in first-order response surfaces. Technical Report, Department of Statistics, Virginia Tech, Blacksburg, Virginia.

Donohue, J. M., E. C. Houck, and R. H. Myers. 1992b. Simulation designs for quadratic response surface models in the presence of model misspecification. To appear in *Management Science.*

Donohue, J. M., E. C. Houck, and R. H. Myers. 1992c. Sequential experimental designs for the estimation of first- and second-order simulation metamodels. Working paper, Department of Management Science, University of South Carolina, Columbia, South Carolina.

Hussey, J. R., R. H. Myers, and E. C. Houck. 1987a. Correlated simulation experiments in first-order response surface designs. *Operations Research* 35:744–758.

Hussey, J. R., R. H. Myers, and E. C. Houck. 1987b. Pseudorandom number assignments in quadratic response surface designs. *IEE Transactions* 19:395–403.

Joshi, S., and J. D. Tew. 1992. Statistical analysis and validation procedures under the common random number correlation-induction strategy. To appear in *European Journal of Operational Research.*

Khuri, A. I., and J. A. Cornell. 1987. *Response surfaces: Designs and analyses.* New York: Marcel Dekker.

Kiefer, J. 1978. Comment: Pseudorandom number assignment in statistically designed simulation and distribution sampling experiments. *Journal of the American Statistical Association* 73:523–524.

Kleijnen, J. P. C. 1992. Regression metamodels for simulation with common random numbers: Comparison of validation tests and confidence intervals. To appear in *Management Science.*

Law, A. M., and W. D. Kelton. 1991. *Simulation modeling and analysis.* 2d ed. New York: McGraw-Hill.

Morrison, D. F. (1990). *Multivariate statistical methods.* 3d ed. New York: McGraw-Hill.

Myers, R. H. 1976. *Response surface methodology.* Ann Arbor, Michigan: Edward Brothers.

Myers, R. H., and S. J. Lahoda. 1975. A generalization of the response surface mean square error criterion with a specific application to the slope. *Technometrics* 17:481–486.

Nozari, A. S., S. F. Arnold, and C. D. Pegden. 1987. Statistical analysis for use with the Schruben and Margolin correlation induction strategy. *Operations Research* 35:127–139.

Rao, C. R. 1973. *Linear statistical inference and its applications.* 2d ed. New York: John Wiley & Sons.

Schruben, L. W., and B. H. Margolin. 1978. Pseudorandom number assignment in statistically designed simulation and distribution sampling experiments. *Journal of the American Statistical Association* 73:504–520.

Silver, N. C., and W. P. Dunlap. 1987. Averaging correlation coefficients: Should Fisher's Z transformation be used? *Journal of Applied Psychology* 72:146–148.

Tew, J. D., and J. R. Wilson. 1992a. Validation of simulation analysis methods for the Schruben-Margolin correlation-induction strategy. *Operations Research* 40:87–103.

Tew, J. D., and J. R. Wilson. 1992b. Estimating simulation metamodels using combined correlation-based variance reduction techniques. To appear in *IEE Transactions.*

AUTHOR BIOGRAPHIES

JOAN M. DONOHUE is an Assistant Professor of Management Science at the University of South Carolina. Her research interests include the design of simulation experiments, response surface metamodeling, and variance reduction techniques.

ERNEST C. HOUCK is a Professor of Management Science at Virginia Polytechnic Institute & State University. His research interests include experimental design for simulation metamodeling, correlation induction strategies, and neural networks.

RAYMOND H. MYERS is a Professor of Statistics at Virginia Polytechnic Institute & State University. His research interests include response surface methodology and regression analysis.