INFORMS Journal on Computing
Vol. 12, No. 4, Fall 2000

0899-1499/100/1204-0272 $05.00

A Revised Simplex Search Procedure for Stochastic Simulation Response Surface Optimization

DAVID G. HUMPHREY / *Nortel Networks, Operations Research Department, Research Triangle Park, NC 27713, Email: humphre@nortelnetworks.com*

JAMES R. WILSON / *Department of Industrial Engineering, North Carolina State University, Raleigh, NC 27695, Email: jwilson@eos.ncsu.edu, Web: http://www.ie.ncsu.edu/jwilson*

(Received: September 1997; revised: April 2000; accepted: May 2000)

We develop a variant of the Nelder-Mead (NM) simplex search procedure for stochastic simulation optimization that is designed to avoid many of the weaknesses encumbering similar direct-search methods—in particular, excessive sensitivity to starting values, premature termination at a local optimum, lack of robustness against noisy responses, and computational inefficiency. The Revised Simplex Search (RSS) procedure consists of a three-phase application of the NM method in which: (a) the ending values for one phase become the starting values for the next phase; (b) the step size for the initial simplex (respectively, the shrink coefficient) decreases geometrically (respectively, increases linearly) over successive phases; and (c) the final estimated optimum is the best of the ending values for the three phases. To compare RSS versus NM and procedure RS+S9 due to Barton and Ivey, we summarize a simulation study based on four selected performance measures computed for six test problems that include additive white-noise error, with three levels of problem dimensionality and noise variability used in each problem. In the selected test problems, RSS yielded significantly more accurate estimates of the optimum than NM or RS+S9, and both RSS and RS+S9 required roughly four times as many function evaluations as NM.

Stochastic simulation optimization can be viewed as finding a combination of (deterministic) input parameters (factor levels or design variables) that yields the optimal expected value of a user-specified (random) output response generated by the simulation model. Let the d-dimensional design point $\mathbf{x} \equiv [x_1, \ldots, x_d]$ represent the factor-level combination specifying the policy governing operation of the simulation model under the associated scenario (or alternative system configuration). Thus the components of $\mathbf{x}$ together with a (possibly infinite) stream of random numbers constitute the full set of inputs to the simulation model; and we let $\mathbf{Y}(\mathbf{x}) \equiv [Y_1(\mathbf{x}), \ldots, Y_p(\mathbf{x})]$ denote the p-dimensional random vector of output responses generated on one run of the simulation at design point $\mathbf{x}$. With respect to optimization of system performance, we assume that one of the components of $\mathbf{Y}(\mathbf{x})$, say the initial element $Y_1(\mathbf{x})$, is the primary response of interest; and we let $\theta(\mathbf{x}) = \mathrm{E}[Y_1(\mathbf{x})]$ denote the response surface function to be optimized. We define the *region of interest* for the optimization procedure,

$$\Xi \equiv \{\mathbf{x} \in R^d: \mathbf{x} \text{ defines feasible system operating conditions}\}, \quad (1)$$

where R^d denotes d-dimensional Euclidean space. Assuming that the primary performance measure is expected total cost and thus should be minimized, we seek to determine

$$\theta^* \equiv \min_{\mathbf{x}\in\Xi} \theta(\mathbf{x}) \text{ and } \mathbf{x}^* \equiv \arg\min_{\mathbf{x}\in\Xi} \theta(\mathbf{x}),$$

the minimum cost and the optimal design point defining the minimum-cost system configuration.

In this article we formulate, implement, and evaluate a stochastic simulation optimization procedure that incorporates many desirable properties of the well-known Nelder-Mead (NM) simplex search procedure (Nelder and Mead 1965) while avoiding some of the critical weaknesses of this procedure—in particular, excessive sensitivity to starting values, premature termination at a local optimum, lack of robustness against noisy responses, and computational inefficiency (Parkinson and Hutchinson 1972, Barton and Ivey 1996). In Section 2 we give a formal algorithmic statement of the Revised Simplex Search (RSS) procedure. In Section 3 we summarize the main figures of merit that we used to evaluate and compare procedures for stochastic simulation optimization, and we analyze the significant factors that affect the performance of simplex-search-type procedures. Section 4 contains a summary of a comprehensive Monte Carlo comparison of procedure RSS versus the classical procedure NM as well as procedure RS+S9, a variant of NM that was developed by Barton and Ivey (1996). Finally, in Section 4 we recapitulate the main findings of this work, and we present recommendations for future research. Although this paper is based on Humphrey (1997), some of our results were also presented in Humphrey and Wilson (1998).

1. Revised Simplex Search (RSS) Procedure

In this section we describe the operation of the RSS procedure, and we introduce the symbolism required to specify precisely the steps of the procedure. RSS operates in three

Subject classifications: Simulation, design of experiments: direct-search optimization techniques
Other key words: Simplex search procedures.

phases indexed by the phase counter φ, and within each phase, each additional stage q involves generating a new simplex from the current simplex via the operations of reflection, expansion, contraction, or shrinkage (as described below) until the termination criterion for the current phase (also described below) is satisfied. In the current phase φ and stage q of procedure RSS, we let $\mathbf{x}_i \equiv [x_{i,1}, \ldots, x_{i,d}]$ denote the ith vertex of the latest simplex generated by RSS for $i = 1, \ldots, d + 1$, $q = 0, 1, \ldots$, and $\varphi = 1, 2, 3$. (Although $\mathbf{x}_i^{(q,\varphi)}$ might be a more complete notation for the ith vertex of the qth simplex generated in phase φ, we suppress the exponent $^{(q,\varphi)}$ for simplicity since no confusion can result from this usage.) In the initial phase of operation of RSS, the user provides the initial vertex $\mathbf{x}_1 \equiv [x_{1,1}, \ldots, x_{1,d}]$ that defines the starting point for the overall search procedure. In terms of the step size parameter τ, the initial step size ν_1 for the first phase is $\nu_1 = \max\{1, \tau|x_{1,j}| : j = 1, \ldots, d\}$. The remaining vertices of the initial simplex are given by $\mathbf{x}_{i+1} = \mathbf{x}_1 + \nu_1\mathbf{e}_i$ for $i = 1, \ldots, d$, where $\mathbf{e}_i$ is the d-dimensional unit vector with one in the ith component and zeros elsewhere. In the initial phase of operation of RSS, the coefficient for the shrinkage operation has the value $\delta_1 = 0.5$ originally recommended by Nelder and Mead (1965).

With respect to the current (latest) simplex generated in phase φ of the operation of procedure RSS, we let $\hat{\theta}(\mathbf{x}_i)$ denote the simulation-based estimate of the objective function value $\theta(\mathbf{x}_i)$ at vertex $\mathbf{x}_i$ for $i = 1, \ldots, d + 1$, and we let $\mathbf{x}_{\max}$ denote the vertex of the current simplex yielding

$$\hat{\theta}_{\max} = \hat{\theta}(\mathbf{x}_{\max}) \equiv \max\{\hat{\theta}(\mathbf{x}_i) : 1 \leqslant i \leqslant d + 1\}. \tag{2}$$

In similar fashion, we define $\mathbf{x}_{\min}$ and $\hat{\theta}_{\min} = \hat{\theta}_{\min}(\mathbf{x}_{\min})$, and we let $\mathbf{x}_{\text{ntw}}$ denote the vertex of the current simplex yielding $\hat{\theta}_{\text{ntw}} = \hat{\theta}(\mathbf{x}_{\text{ntw}})$, the next-to-worst (second largest) of the response surface estimates observed at the vertices of the current simplex. When it is not important to emphasize the vertex upon which quantities like $\hat{\theta}(\mathbf{x}_{\max})$ depend, we will use the alternative notation $\hat{\theta}_{\max}$ for simplicity. For $q = 0, 1, \ldots$, the qth stage within phase φ of procedure RSS begins by computing the centroid of all the vertices in the current simplex except $\mathbf{x}_{\max}$,

$$\mathbf{x}_{\text{cen}} = \frac{1}{d}\left\{\left[\sum_{i=1}^{d+1} \mathbf{x}_i\right] - \mathbf{x}_{\max}\right\}. \tag{3}$$

Phase φ of procedure RSS ends when RSS generates a new simplex that is sufficiently "small" to satisfy the termination criterion. Then the phase counter φ is incremented by one and procedure RSS is restarted, provided $\varphi \leqslant 3$.

At the beginning of phase φ of procedure RSS for $\varphi = 2$ and 3, we take the initial step size to be $\nu_\varphi = \frac{1}{2}\nu_{\varphi-1}$ so that the initial step size decreases geometrically over successive phases. Similarly, we take $\delta_\varphi = \delta_{\varphi-1} + 0.2$ for $\varphi = 2$ and 3 so that the shrink coefficient increases linearly over successive phases until it reaches the value 0.9 recommended by Barton and Ivey (1996) for optimization of noisy functions. For $\varphi = 1, 2, 3$, we let $\hat{\mathbf{x}}^*(\varphi)$ denote the final estimate of the optimal solution delivered in phase φ. Then we take as the final estimated optimum the best of the ending values for all three phases. A flow chart of RSS is depicted in Figure 1, and a formal statement of the algorithm is given below. The basis for the design of procedure RSS is detailed in Section 2 below.

Steps of Procedure RSS

0. ***Set Up Phase 1.*** Initialize the following: the phase counter $\varphi \leftarrow 1$; the iteration (stage, simplex) counter, $q \leftarrow 0$; the shrink coefficient used in phase 1, $\delta_1 \leftarrow 0.5$; the initial step size used in phase 1,

$$\nu_1 \leftarrow \begin{cases} \max\{\tau|x_{1,j}| : j = 1, \ldots, d\}, & \text{if } \mathbf{x}_1 \neq \mathbf{0}_d, \\ 1, & \text{otherwise;} \end{cases} \tag{4}$$

and the other vertices of the initial simplex in phase 1,

$$\mathbf{x}_{i+1} \leftarrow \mathbf{x}_1 + \nu_1\mathbf{e}_i \text{ for } i = 1, \ldots, d. \tag{5}$$

Go to step 1.

1. ***Attempt Reflection.*** Form a new simplex by reflecting $\mathbf{x}_{\max}$ through the centroid $\mathbf{x}_{\text{cen}}$ of the remaining vertices of the current simplex to obtain the reflected point

$$\mathbf{x}_{\text{refl}} \leftarrow \mathbf{x}_{\text{cen}} + \alpha(\mathbf{x}_{\text{cen}} - \mathbf{x}_{\max}),$$

where $\alpha = 1.0$ is the reflection coefficient of Nelder and Mead (1965). If

$$\hat{\theta}_{\min} \leqslant \hat{\theta}_{\text{refl}} \leqslant \hat{\theta}_{\text{ntw}}, \tag{6}$$

that is, if the reflected point $\mathbf{x}_{\text{refl}}$ yields a response no worse (no larger) than the next-to-worst vertex $\mathbf{x}_{\text{ntw}}$ in the current simplex but does not yield a better (smaller) response than the best vertex $\mathbf{x}_{\min}$, then replace the worst vertex $\mathbf{x}_{\max}$ in the current simplex by the reflected point $\mathbf{x}_{\text{refl}}$,

$$\mathbf{x}_{\max} \leftarrow \mathbf{x}_{\text{refl}}; \tag{7}$$

and go to step 6 to test the termination criterion. If the condition (6) for accepting the reflection is not satisfied, then go to step 2.

2. ***Attempt Expansion.*** If

$$\hat{\theta}_{\text{refl}} < \hat{\theta}_{\min} \tag{8}$$

so that the reflected point $\mathbf{x}_{\text{refl}}$ is better than the best vertex $\mathbf{x}_{\min}$ in the current simplex, then extend the search in the direction $\mathbf{x}_{\text{refl}} - \mathbf{x}_{\text{cen}}$ to yield the expansion point

$$\mathbf{x}_{\exp} \leftarrow \mathbf{x}_{\text{cen}} + \gamma(\mathbf{x}_{\text{refl}} - \mathbf{x}_{\text{cen}}),$$

where $\gamma = 2.0$ is the expansion coefficient of Nelder and Mead (1965). If $\hat{\theta}_{\exp} < \hat{\theta}_{\min}$, then accept the expansion and replace $\mathbf{x}_{\max}$ by $\mathbf{x}_{\exp}$ in the current simplex,

$$\mathbf{x}_{\max} \leftarrow \mathbf{x}_{\exp};$$

and go to step 6. If $\hat{\theta}_{\exp} \geqslant \hat{\theta}_{\min}$, then reject the attempted expansion and replace $\mathbf{x}_{\max}$ by $\mathbf{x}_{\text{refl}}$ in the current simplex,

$$\mathbf{x}_{\max} \leftarrow \mathbf{x}_{\text{refl}};$$

and go to step 6. Finally if the condition (8) for attempting expansion is not satisfied, then go to step 3.

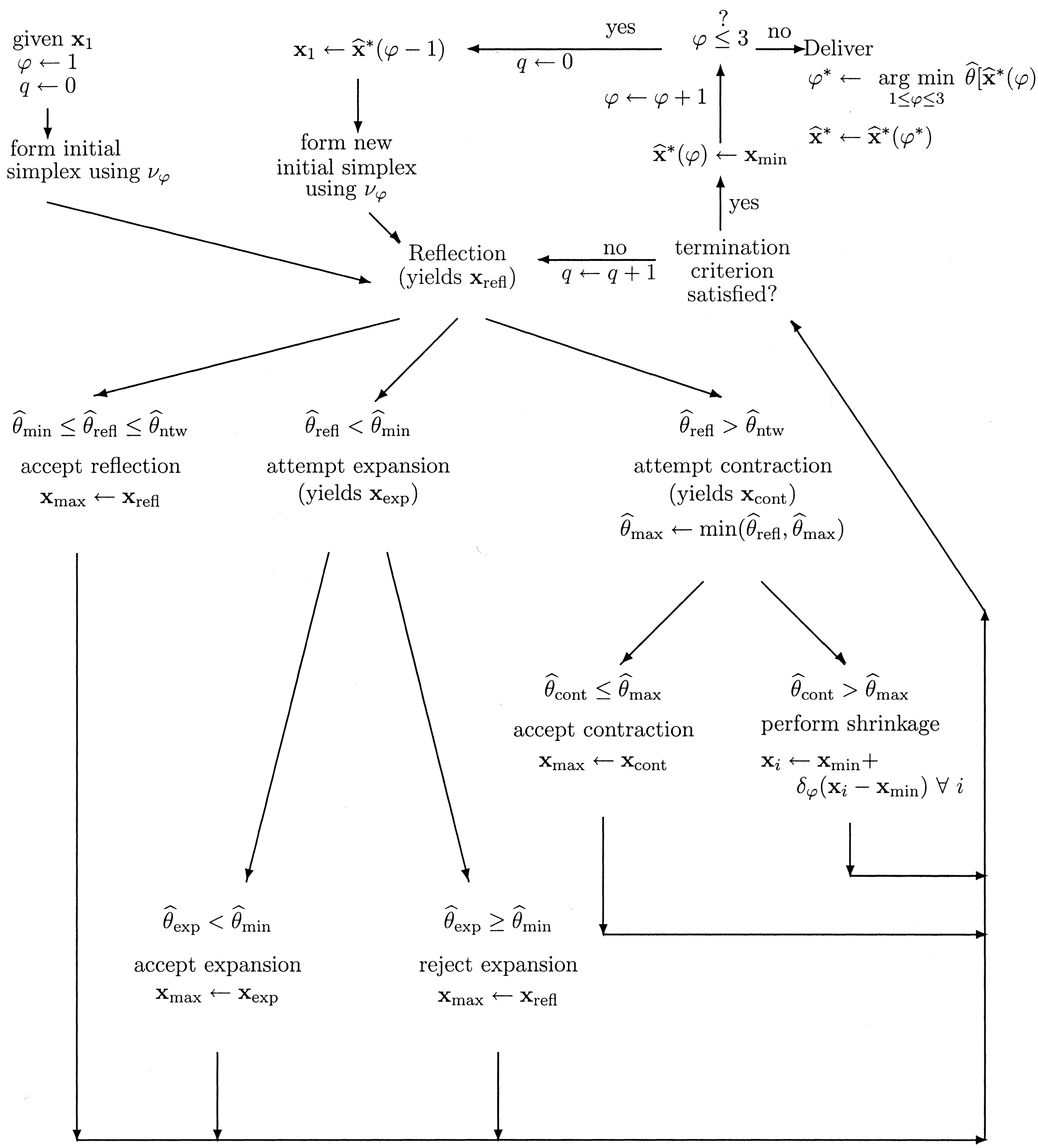

Figure 1. Flow chart of procedure RSS.

3. ***Set Up Attempted Contraction.*** If

$$\hat{\theta}_{\text{refl}} > \hat{\theta}_{\text{ntw}}$$

so that the reflected point $\mathbf{x}_{\text{refl}}$ yields a worse (larger) response than the next-to-worst vertex $\mathbf{x}_{\text{ntw}}$ of the current simplex, then reduce the size of the current simplex—either by a contraction or a more drastic shrinkage. To set up this reduction in the size of the simplex, update the worst vertex in the current simplex as follows:

$$\text{if } \hat{\theta}_{\text{refl}} \leqslant \hat{\theta}_{\max}, \text{ then} \left\{ \begin{array}{l} \mathbf{x}_{\max} \leftarrow \mathbf{x}_{\text{refl}} \\ \hat{\theta}_{\max} \leftarrow \hat{\theta}_{\text{refl}} \end{array} \right\}.$$

Compute the contraction point

$$\mathbf{x}_{\text{cont}} \leftarrow \mathbf{x}_{\text{cen}} + \beta(\mathbf{x}_{\max} - \mathbf{x}_{\text{cen}}),$$

where $\beta = 0.5$ is the contraction coefficient of Nelder and Mead (1965).

4. ***Contract Simplex in One Direction.*** If

$$\hat{\theta}_{\text{cont}} \leqslant \hat{\theta}_{\max},$$

so that the contracted point $\mathbf{x}_{\text{cont}}$ yields a response no worse (no larger) than the worst vertex $\mathbf{x}_{\max}$ of the current simplex, then replace $\mathbf{x}_{\max}$ in the current simplex by $\mathbf{x}_{\text{cont}}$,

$$\mathbf{x}_{\max} \leftarrow \mathbf{x}_{\text{cont}};$$

and go to step 6; otherwise go to step 5.

5. ***Shrink Entire Simplex.*** If the contracted point $\mathbf{x}_{\text{cont}}$ yields a worse (larger) response than every vertex in the current simplex including $\mathbf{x}_{\max}$ so that the shrinkage condition

$$\hat{\theta}_{\text{cont}} > \hat{\theta}_{\max}$$

is satisfied, then reduce the lengths of all edges of the current simplex with common endpoint $\mathbf{x}_{\min}$ by the shrinkage factor δ_φ, yielding a new simplex with vertices

$$\mathbf{x}_i \leftarrow \mathbf{x}_{\min} + \delta_\varphi(\mathbf{x}_i - \mathbf{x}_{\min}) \text{ for } i = 1, \ldots, d+1; \quad (9)$$

and go to step 6.

6. ***Test Termination Criterion for Current Phase.*** After each reflection, expansion, contraction, or shrinkage, apply the termination criterion

$$\max_{1 \le i \le d+1} \|\mathbf{x}_i - \mathbf{x}_{\min}\| \le \begin{cases} \eta_1\|\mathbf{x}_{\min}\|, & \text{if } \|\mathbf{x}_{\min}\| \ne 0, \\ \eta_2, & \text{otherwise,} \end{cases} \quad (10)$$

where η_1 and η_2 are user-specified tolerances and the maximum is taken over all vertices in the current simplex. If the termination condition (10) is not satisfied, then increment the iteration counter $q \leftarrow q + 1$ and go to step 1. If the termination condition (10) is satisfied, then go to step 7.

7. ***Terminate Current Phase.*** Record the termination point of the current phase

$$\hat{\mathbf{x}}^*(\varphi) \leftarrow \mathbf{x}_{\min}, \quad (11)$$

increment the phase counter,

$$\varphi \leftarrow \varphi + 1,$$

and go to step 8.

8. ***Test Final Termination Criterion.*** If $\varphi > 3$, then compute the final estimate $\hat{\mathbf{x}}^*$ of the global optimum according to

$$\varphi^* \leftarrow \arg\min\{\hat{\theta}[\hat{\mathbf{x}}^*(\varphi)] : \varphi = 1, 2, 3\} \text{ and } \hat{\mathbf{x}}^* \leftarrow \hat{\mathbf{x}}^*(\varphi^*);$$

finally deliver $\hat{\mathbf{x}}^*$ and $\hat{\theta}(\hat{\mathbf{x}}^*)$ and stop. If $\varphi \le 3$, then go to step 9.

9. ***Set Up Next Phase.*** Initialize the following: the iteration counter $q \leftarrow 0$; the first vertex of the initial simplex in the current phase,

$$\mathbf{x}_1 \leftarrow \hat{\mathbf{x}}^*(\varphi - 1); \quad (12)$$

the initial step size for the current phase,

$$\nu_\varphi \leftarrow \tfrac{1}{2}\nu_{\varphi-1}; \quad (13)$$

the shrink coefficient for the current phase,

$$\delta_\varphi \leftarrow \delta_{\varphi-1} + 0.2; \quad (14)$$

and the other vertices of the initial simplex in the current phase,

$$\mathbf{x}_{i+1} \leftarrow \mathbf{x}_1 + \nu_\varphi \mathbf{e}_i \text{ for } i = 1, \ldots, d. \quad (15)$$

Go to step 1.

2. Development of Procedure RSS

In this section we formulate the principal performance measures that we used to evaluate and compare procedures for stochastic simulation optimization, and we summarize our preliminary analysis of the significant factors affecting the performance of simplex-search-type procedures. This analysis formed the basis for the design of procedure RSS.

2.1 Formulation of Performance Measures for Simulation Optimization

We used four figures of merit to evaluate stochastic simulation optimization procedures: (a) logarithm of the number of function evaluations; (b) absolute percentage deviation of the estimated optimal function value from the true optimal function value; (c) maximum over all coordinates of the absolute percentage deviation of the estimated optimum from the true optimum taken with respect to each coordinate separately; and (d) average over all coordinates of the absolute percentage deviation of the estimated optimum from the true optimum taken with respect to each coordinate separately.

2.1.1 Logarithm of Number of Function Evaluations

To measure the computational work performed by a simulation optimization procedure, we compute the (natural) logarithm of the total number of function evaluations required by the procedure before it terminates and delivers the final estimates $\hat{\theta}(\hat{\mathbf{x}}^*)$ and $\hat{\mathbf{x}}^*$:

$$L \equiv \ln(\text{total number of function evaluations required}). \quad (16)$$

The logarithmic transformation in (16) is used to obtain approximately normal observations with a common variance to which we can apply standard statistical techniques such as analysis of variance and multiple comparisons procedures; see Anderson and McLean (1974). Although L is widely used in experimental comparisons of simulation optimization procedures (Barton and Ivey 1996), it should be recognized that in the optimization of a large-scale stochastic simulation model, each function evaluation represents a separate simulation run, and different runs may require substantially different amounts of execution time to deliver the corresponding function values. Thus in general L provides at best a rough indication of the total computational work required by a simulation optimization procedure.

2.1.2 Final Function Value

Provided that the optimal function value $\theta^* \ne 0$, we use the absolute percentage deviation

$$D = \left| \frac{\hat{\theta}^*(\hat{\mathbf{x}}^*) - \theta^*}{\theta^*} \right| \quad (17)$$

as a measure of the accuracy of the final result delivered by a simulation optimization procedure. When averaged over independent replications of each procedure applied to a given test problem, the quantity (17) provides a dimensionless figure of merit that allows us to compare the performance of simulation optimization procedures across different test problems. All the test problems used in this work were specifically constructed to have nonzero optimal function values.

2.1.3 Coordinatewise Maximum Absolute Percentage Deviation from Global Optimum

The third performance measure is the maximum over all j (for $1 \leqslant j \leqslant d$) of the absolute percentage deviation of $\hat{x}_j^*$ (the jth coordinate of the estimated optimum $\hat{\mathbf{x}}^*$) from x_j^* (the jth coordinate of the true optimum $\mathbf{x}^*$), provided that each $x_j^* \neq 0$:

$$B = \max_{1 \leqslant j \leqslant d} \left| \frac{\hat{x}_j^* - x_j^*}{x_j^*} \right|. \tag{18}$$

When there are multiple optima, we evaluate the right-hand side of (18) for each optimum, and we take the smallest of these quantities as the final value of B. When averaged over independent replications of each procedure applied to a given test problem, the quantity (18) provides another dimensionless figure of merit that allows us to compare the performance of simulation optimization procedures across different test problems. All of the test problems used in this work were specifically constructed to have optima with all coordinates having nonzero values.

2.1.4 Coordinatewise Average Absolute Percentage Deviation from Global Optimum

The final performance measure is the average computed over all j (for $1 \leqslant j \leqslant d$) of the absolute percentage deviation of $\hat{x}_j^*$ (the jth coordinate of the estimated optimum $\hat{\mathbf{x}}^*$) from x_j^* (the jth coordinate of the true optimum $\mathbf{x}^*$), provided that each $x_j^* \neq 0$:

$$A = \frac{1}{d} \sum_{j=1}^{d} \left| \frac{\hat{x}_j^* - x_j^*}{x_j^*} \right|. \tag{19}$$

When there are multiple optima, we evaluate the right-hand side of (19) for each optimum and take the smallest of these quantities as the final value of A. We believe that A provides the best overall characterization of the accuracy with which a simulation optimization procedure estimates the true optimum.

No single performance measure can tell the entire story about the performance of a particular search procedure, but we believe that (16)–(19) provide meaningful information that can be aggregated over different test problems to yield a comprehensive basis for comparison of selected simulation optimization procedures.

2.2 Significant Factors Affecting Performance of Simplex-Search-Type Procedures

In seeking to formulate a simplex-search-type procedure that avoids some of the drawbacks of the Nelder-Mead procedure when it is applied to optimization of noisy response functions, we identified three significant factors affecting the performance of such procedures: sizing the initial simplex, restarting the search, and adjusting the shrink coefficient δ. Each of these factors will be discussed briefly; a more detailed analysis of these factors is given in Humphrey (1997) and in Humphrey and Wilson (2000).

2.2.1 Sizing the Initial Simplex

Our preliminary experimentation showed that starting a simplex-search-type procedure with a larger initial simplex generally improved the performance of the procedure. The idea behind starting with a larger initial simplex is straightforward. A smaller simplex starting far from the true optimum will have to iterate many times (mostly through reflections and expansions) in order to move into a neighborhood of the optimum in which the response surface is well behaved. Along the way to such a neighborhood, any errant contractions or shrinkages of the current simplex will significantly slow the procedure's progress toward the optimum. By comparison, a larger simplex starting far from the optimum can make much faster progress toward the optimum by "covering more ground" with each reflection or expansion; and in this situation any errant contractions or shrinkages will have a less severe effect on the procedure's progress toward the optimum. Although Parkinson and Hutchinson (1972) observed similar effects in their extensive numerical evaluation of the performance of simplex-search-type procedures for optimization of deterministic response functions, the situation is much less clear-cut when such procedures are applied to noisy responses. Based on a preliminary simulation study similar to that described in Section 3 below, we found that taking the initial step size parameter $\tau = 4.0$ in (4) appeared to yield the best overall performance for procedure RSS.

2.2.2 Restarting the Search

Our preliminary experimentation also revealed that to guard against premature termination at a false optimum, the most effective action was to step away from the current termination point, restart the search procedure with a new initial simplex, and compare the resulting alternative termination points. Parkinson and Hutchinson (1972) observed similar effects with deterministic response functions. Based on a preliminary simulation study similar to that described in Section 3 below, we found that significantly improved performance of RSS was obtained by restarting the procedure twice—thus in effect we designed RSS to operate in three phases and finally deliver the best solution taken over all three phases. Displays (12)–(15) specify the restart step for phase φ of procedure RSS as it depends on the results of phase $\varphi - 1$ for $\varphi = 2$ and 3. Notice that on each successive phase of operation of procedure RSS, the initial step size is reduced by 50% compared to the initial step size used in the previous phase.

2.2.3 Adjusting the Shrink Coefficient

Another change incorporated into procedure RSS involves the shrink coefficient δ. Every time a shrinkage is performed, each edge of the simplex is rescaled by the factor δ; and since $0 < \delta < 1$, the overall size of the simplex is substantially reduced. Based on a preliminary simulation study similar to that described in Section 3 below, we obtained better performance in the first phase of operation of procedure RSS by using the shrink coefficient value $\delta_1 = 0.5$ that was originally recommended by Nelder and Mead (1965); but in the later phases of operation of RSS, we obtained greater protection against premature termination with the larger shrink coefficient values $\delta_2 = 0.7$ and $\delta_3 = 0.9$. In procedure RS+S9, Barton and Ivey (1996) fixed the shrink coefficient at the value 0.9 to reduce the likelihood of premature termination; moreover, after each shrinkage operation (9) is performed, procedure RS+S9 requires resampling the response at the anchor point $\mathbf{x}_{\min}$ and then reranking and relabeling the vertices of the new simplex (that is, $\mathbf{x}_{\max}$, $\mathbf{x}_{\text{ntw}}$, $\mathbf{x}_{\min}$, etc.) before attempting the next reflection operation.

The motivation for the shrink-coefficient assignments $\delta_1 = 0.5$, $\delta_2 = 0.7$, $\delta_3 = 0.9$ used in procedure RSS is that during the earlier, "hill-climbing" phases of the operation of RSS, the current simplex is usually far from the optimum, and the likelihood of an errant shrinkage should be relatively low since the topology of the response surface often has a larger effect on the behavior of the search procedure than the noise in the sampled responses. If during its first phase of operation procedure RSS detects what appears to be nonconvex behavior in the responses observed at the vertices of the current simplex so that a shrinkage operation should be performed, then the smaller value of δ_1 should enable the shrink operation to be more effective in positioning the simplex in a locally convex neighborhood of the optimum. If several errant shrinkages are performed during the earlier phases of the search and the simplex becomes too small to make effective progress toward the optimum, then procedure RSS attempts to compensate for this through the formation of a new initial simplex at the start of each of the later phases of the search. Moreover in the later phases of the search, the simplex is usually in a subregion of the region of interest (1) where the response surface is relatively flat so that the noise in the sampled responses typically has a larger effect on the behavior of the search procedure; consequently the likelihood of performing an errant shrinkage should be higher than it was in the earlier phases of the search. Because shrinkages drastically reduce the size of the simplex, we attempt to protect against errant shrinkages (and consequently premature termination) in the later phases of the search by increasing the value of the shrink coefficient δ_φ for successive values of the phase counter φ.

3. Experimental Performance Evaluation

In this section we describe the problems used in testing procedure RSS and comparing its performance with that of procedures NM and RS+S9. We also provide a summary and analysis of the experimental results.

3.1 Description of Test Problems

We selected six problems to serve as a test-bed for comparing the performance of procedure RSS with that of procedures NM and RS+S9. Similar problems were used in the experimental performance evaluation of Parkinson and Hutchinson (1972) for optimization of deterministic response functions and in the study of Baron and Ivey (1996) for optimization of noisy response functions. To mimic the behavior of responses generated by a stochastic simulation model, we took each sampled response to have the form $\hat{\theta}(\mathbf{x}) = \theta(\mathbf{x}) + \zeta$, where $\theta(\cdot)$ is one of the test functions described below and the additive white-noise error term ζ is randomly sampled from a normal distribution with a mean of zero and a standard deviation that is systematically varied to examine the effect of increasing levels of noise variability on the selected simplex-search-type procedures. For all three procedures, we used the common termination criterion (10) with $\eta_1 = 5.0 \times 10^{-6}$ and $\eta_2 = 10^{-20}$ to provide an equitable basis for comparing the performance of these procedures. For each test problem described below, we specify the function to be minimized, the starting point used by each search procedure, the optimal function value, and the point(s) corresponding to the optimal function value. A complete description of all test problems is given in Humphrey (1997).

3.1.1 Test Problem 1: Variably Dimensioned Function

The variably dimensioned function is defined as

$$\theta(\mathbf{x}) = \sum_{i=1}^{d+2} [f_i(\mathbf{x})]^2 + 1,$$

where

$$f_i(\mathbf{x}) = x_i - 1 \text{ for } i = 1, \ldots, d,\; f_{d+1}(\mathbf{x}) = \sum_{j=1}^{d} j(x_j - 1),$$

$$\text{and } f_{d+2}(\mathbf{x}) = \left[\sum_{j=1}^{d} j(x_j - 1)\right]^2.$$

The initial point is given by $\mathbf{x}_1 \equiv [x_{1,1}, x_{1,2}, \ldots, x_{1,d}]$, where $x_{1,j} = 1 - (j/d)$, $j = 1, \ldots, d$. The optimal function value of $\theta^* = 1$ is achieved at the point $\mathbf{x}^* = [1, \ldots, 1]$. Figure 2 depicts the variably dimensioned function for the case $d = 2$.

3.1.2 Test Problem 2: Trigonometric Function

The trigonometric function is defined as

$$\theta(\mathbf{x}) = \sum_{i=1}^{d} [f_i(\mathbf{x})]^2 + 1,$$

where

$$f_i(\mathbf{x}) = d - \sum_{j=1}^{d} \cos(x_j - 1) + i[1 - \cos(x_i - 1)] - \sin(x_i - 1),\; i = 1, \ldots, d.$$

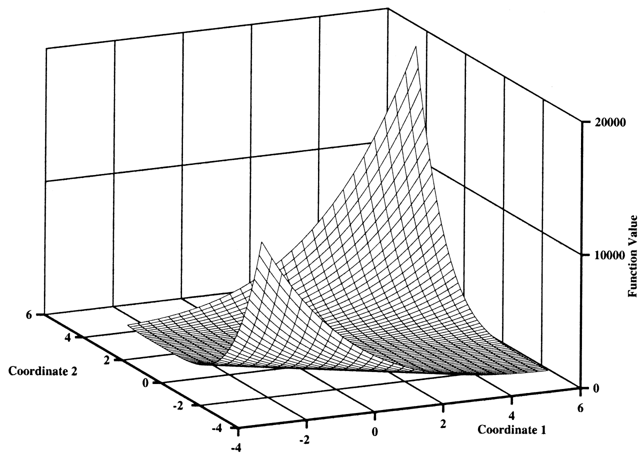

Figure 2. Test problem 1: variably dimensioned function for $d = 2$.

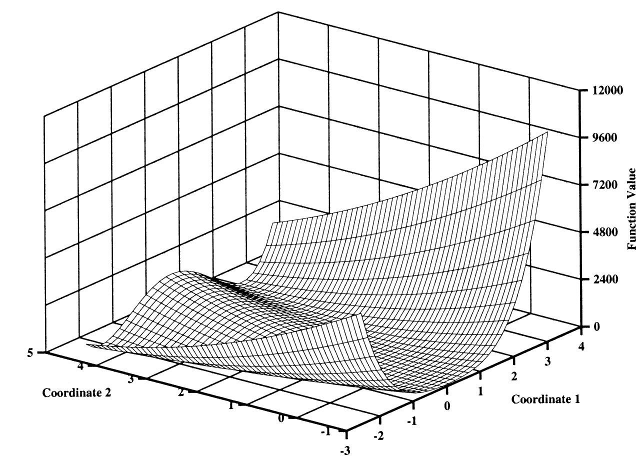

Figure 4. Test problem 3: extended Rosenbrock function for $d = 2$.

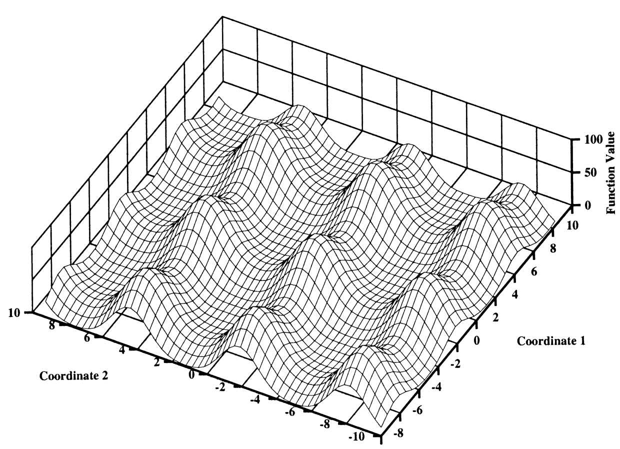

Figure 3. Trigonometric function (test problem 2) for $d = 2$.

We used the starting point $\mathbf{x}_1 \equiv [1/d, \ldots, 1/d]$. The optimal value of $\theta^* = 1$ is achieved at every point in the lattice of points given by

$$\mathbf{x}^*_{k_1 k_2 \cdots k_d} = [1 + 2\pi k_1, \ldots, 1 + 2\pi k_d],$$

$$\text{where } k_j = 0, \pm 1, \pm 2, \ldots, \text{ for } j = 1, \ldots, d. \quad (20)$$

Figure 3 depicts the trigonometric function for $d = 2$. When a given search procedure terminates on this test problem, we determine which of the optimal points specified by (20) is closest in Euclidean distance to the final estimate $\hat{\mathbf{x}}^*$, and we use that optimum for calculating performance measures A and B as specified in displays (19) and (18), respectively.

3.1.3 Test Problem 3: Extended Rosenbrock Function

The extended Rosenbrock function is defined as

$$\theta(\mathbf{x}) = \sum_{i=1}^{d} [f_i(\mathbf{x})]^2 + 1,$$

where

$$\left.\begin{aligned} f_{2i-1}(\mathbf{x}) &= 10(x_{2i} - x_{2i-1}^2) \\ f_{2i}(\mathbf{x}) &= (1 - x_{2i-1}) \end{aligned}\right\} \text{ for } i = 1, \ldots, d/2.$$

The initial point is given by $\mathbf{x}_1 \equiv [-1.2, 1, \ldots, -1.2, 1]$, and the optimal value of $\theta^* = 1$ occurs at $\mathbf{x}^* = [1, \ldots, 1]$. Figure 4 depicts the extended Rosenbrock function for the case $d = 2$.

3.1.4 Test Problem 4: Extended Powell Singular Function

The extended Powell singular function is defined as

$$\theta(\mathbf{x}) = \sum_{i=1}^{d} [f_i(\mathbf{x})]^2 + 1,$$

where

$$\left.\begin{aligned} f_{4i-3}(\mathbf{x}) &= x_{4i-3} + 10x_{4i-2} - 11 \\ f_{4i-2}(\mathbf{x}) &= \sqrt{5}(x_{4i-1} - x_{4i}) \\ f_{4i-1}(\mathbf{x}) &= (x_{4i-2} - 2x_{4i-1} + 1)^2 \\ f_{4i}(\mathbf{x}) &= \sqrt{10}(x_{4i-3} - x_{4i})^2 \end{aligned}\right\} \text{ for } i = 1, \ldots, d/4,$$

so that the dimensionality d of the input vector $\mathbf{x}$ must be a multiple of 4. The starting point is $\mathbf{x}_1 \equiv [3, -1, 0, 1, \ldots, 3, -1, 0, 1]$, and the optimal value of $\theta^* = 1$ is achieved at the point $\mathbf{x}^* = [1, \ldots, 1]$.

3.1.5 Test Problem 5: Brown's Almost-Linear Function

Brown's almost-linear function is defined as

$$\theta(\mathbf{x}) = \sum_{i=1}^{d} [f_i(\mathbf{x})]^2 + 1,$$

Table I. Optimal Points of Brown's Almost-Linear Function (Test Problem 5)

d	Optimal Points
2	$x_1^* = [1, 1]$, $x_2^* = [1/2, 2]$
10	$x_1^* = [1, \ldots, 1]$, $x_2^* = [0.9794304, \ldots, 0.9794304, 1.2056959]$
18	$x_1^* = [1, \ldots, 1]$, $x_2^* = [0.9937218, \ldots, 0.9937218, 1.1130085]$

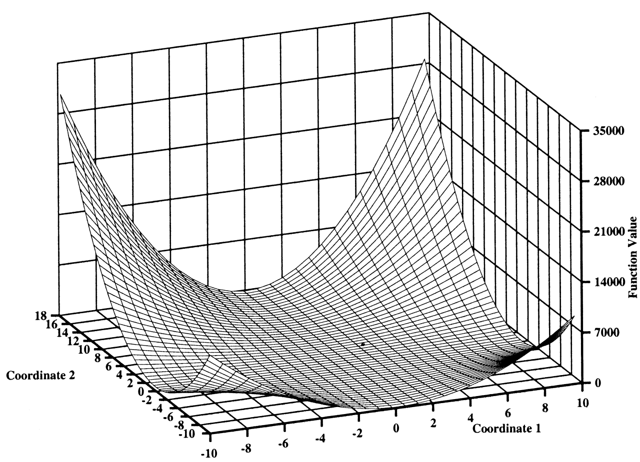

Figure 5. Test problem 5: Brown's almost-linear function for $d = 2$.

where

$$f_i(\mathbf{x}) = x_i + \sum_{j=1}^{d} x_j - (d+1) \text{ for } i = 1, \ldots, d-1,$$

$$\text{and } f_d(\mathbf{x}) = \left(\prod_{j=1}^{d} x_j \right) - 1.$$

The initial point is given by $\mathbf{x}_1 \equiv [1/2, \ldots, 1/2]$, and the optimal function value is $\theta^* = 1$, which occurs at two different points for all values of d. Moré et al. (1981) show that $\theta(\mathbf{x}^*) = 1$ at $\mathbf{x}^* = (\alpha, \ldots, \alpha, \alpha^{1-d})$, where α satisfies

$$d\alpha^d - (d+1)\alpha^{d-1} + 1 = 0. \qquad (21)$$

Notice that $\alpha = 1$ is a solution of (21) for every value of d. We computed the following additional real solutions of (21): $\alpha = \frac{1}{2}$ for $d = 2$; $\alpha = 0.9794304$ for $d = 10$; and $\alpha = 0.9937218$ for $d = 18$. Table I specifies the optimal points $\mathbf{x}_1^*$, $\mathbf{x}_2^*$ for the values of d used in our experimental performance evaluation. Figure 5 depicts Brown's almost linear function for the case $d = 2$. When a given search procedure terminates on this test problem, we determine which of the two optimal points is closer in Euclidean distance and we use that optimal point for calculating performance measures A and B.

Table II. Coefficients $\{c_i: 1 \leqslant i \leqslant d\}$ of Corana Function (Test Problem 6)

i	c_i	i	c_i	i	c_i	i	c_i
1	1	6	10	11	100	16	1000
2	1000	7	100	12	1000	17	1
3	10	8	1000	13	1	18	10
4	100	9	1	14	10	19	100
5	1	10	10	15	100	20	1000

3.1.6 Test Problem 6: Corana Function

In some respects, the Corana function (Corana et al. 1987) represents the most difficult test problem used in our experimental performance evaluation. We take

$$\Xi \equiv \{\mathbf{x} \in R^d: -a_i \leqslant x_i \leqslant a_i \text{ for } i = 1, \ldots, d\},$$

where $a_i = 10^4$ for $i = 1, \ldots, d$, and we define a set of "pockets" within Ξ as follows:

$$\Upsilon_{k_1,\ldots,k_d} \equiv \{\mathbf{x} \in \Xi: k_i s_i - t_i < x_i - 1 < k_i s_i + t_i \quad \text{for } i = 1, \ldots, d\},$$

where $k_1, \ldots, k_d$ are integers, the vectors $\mathbf{t} = (t_1, \ldots, t_d)$ and $\mathbf{s} = (s_1, \ldots, s_d)$ are composed of positive real numbers, and $t_i < s_i/2$ for $i = 1, \ldots, d$. From this we define Υ to be the family of open, disjoint, rectangular subdomains of R^d within Ξ defined as follows:

$$\Upsilon \equiv \bigcup_{k_1=-\infty}^{+\infty} \cdots \bigcup_{k_d=-\infty}^{+\infty} \Upsilon_{k_1,\ldots,k_d} - \Upsilon_{0,\ldots,0}.$$

The Corana function is defined by

$$\theta(\mathbf{x}) = \begin{cases} 1 + \sum_{i=1}^{d} c_i (x_i - 1)^2, & \text{for } \mathbf{x} \in \Xi - \Upsilon, \\ 1 + \psi \sum_{i=1}^{d} c_i (z_i - 1)^2, & \text{for } \mathbf{x} \in \Upsilon, \end{cases}$$

where

$$z_i = \left\{ \begin{array}{ll} k_i s_i + t_i, & \text{if } k_i < 0, \\ 0, & \text{if } k_i = 0, \\ k_i s_i - t_i, & \text{if } k_i > 0, \end{array} \right\} \text{ for } i = 1, \ldots, d.$$

The initial point is given by $\mathbf{x}_1 \equiv [2, \ldots, 2]$, and the optimal function value $\theta^* = 1$ is achieved at the point $\mathbf{x}^* = [1, \ldots, 1]$. We used $s_i = 0.2$, $t_i = 0.05$ for $i = 1, \ldots, d$, and $\psi = 0.15$ as is usually done in applications of the Corana function. Table II specifies the coefficients $\{c_i: 1 \leqslant i \leqslant d\}$ used in our experimentation with the Corana function. Figure 6 depicts the Corana function for the case that $d = 2$ and $c_1 = c_2 = 1$. Note that a value of $c_2 = 1000$ (as is used in our analysis, but not shown in Figure 6) causes the response surface to be extremely steep in the second coordinate direction.

3.2 Summary of Experimental Results

In the experimental performance evaluation, we sought to include "low," "medium," and "high" levels of dimensionality and noise variability. For the "low" level of dimensionality, $d = 2$ is the natural choice. Since the literature indicates that simplex-search type procedures tend to perform well

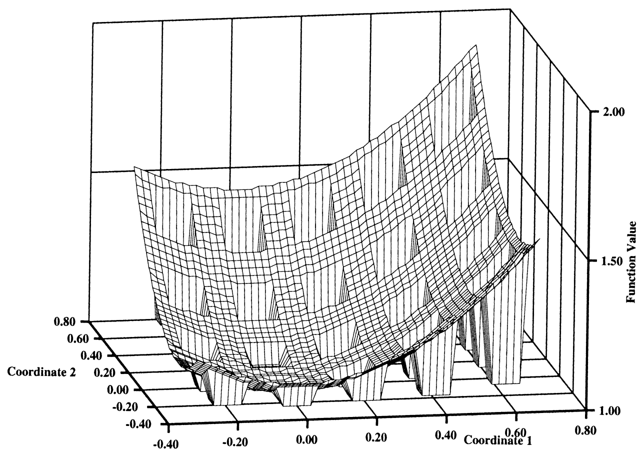

Figure 6. Plot of Corana function (test problem 6) for $d = 2$.

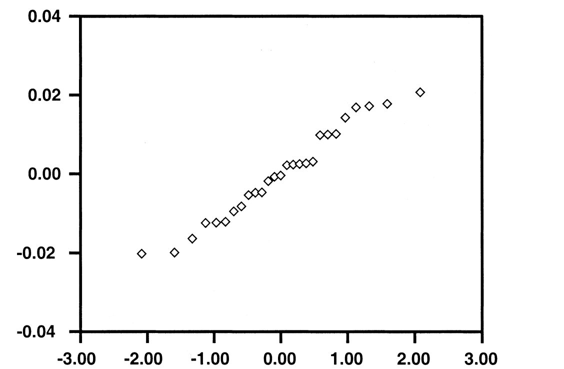

Figure 7. Normal probability plot of estimated residuals in the ANOVA model (22) for performance measure L of test problem 1.

for $d \leqslant 10$ (Nelder and Mead 1965, Barton and Ivey 1996), we took $d = 10$ as the "medium" level of dimensionality for all test problems except problem 4, where we took $d = 8$ to satisfy the requirements of the extended Powell singular function. Based on our previous computational experience with the Nelder-Mead procedure in a wide variety of statistical-estimation problems involving minimization of functions of up to 20 independent variables (Wagner and Wilson 1996, Kuhl and Wilson 2000), we took $d = 18$ as the "high" level of dimensionality for all test problems except problem 4, where we took $d = 16$.

To gauge the effect of increasing levels of noise variability on the selected simplex-search-type procedures, we took each sampled response to have the form $\hat{\theta}(\mathbf{x}) = \theta(\mathbf{x}) + \zeta$, where $\theta(\cdot)$ is one of the selected test functions described in Section 3.1 and ζ is randomly sampled from a normal distribution with a mean of zero and a standard deviation of 0.75, 1.0, or 1.25 times the magnitude of the optimal response $|\theta^*|$. This arrangement provided "low," "medium," and "high" levels of variation around the true underlying response surface relative to the optimal function value $\theta^* = 1$ that was common to all six test problems.

Our study of the ith problem ($1 \leqslant i \leqslant 6$) constituted a complete factorial experiment in which there were three factors each at three levels as defined below:

$P_j \equiv j$th level of optimization procedure

$$= \begin{cases} \text{NM} & \text{for } j = 0, \\ \text{RSS} & \text{for } j = 1, \\ \text{RS+S9} & \text{for } j = 2; \end{cases}$$

$Q_k \equiv k$th level of problem dimensionality

$$= \begin{cases} 2 \text{ (4 in problem 4)} & \text{for } k = 1, \\ 10 \text{ (8 in problem 4)} & \text{for } k = 2, \\ 18 \text{ (16 in problem 4)} & \text{for } k = 3; \end{cases}$$

and

$N_l \equiv l$th level of noise standard deviation

$$= \begin{cases} 0.75|\theta^*| & \text{for } l = 1, \\ 1.00|\theta^*| & \text{for } l = 2, \\ 1.25|\theta^*| & \text{for } l = 3. \end{cases}$$

Within the ith experiment and for each of the selected performance measures that were observed on the mth replication of the treatment combination (P_j, Q_k, N_l), we postulated a statistical model of the form

$$Z_{ijklm} = \beta_0 + \beta_P W_{P_j} + \beta_Q W_{Q_k} + \beta_N W_{N_l} + \beta_{PQ} W_{P_j} W_{Q_k} + \beta_{PN} W_{P_j} W_{N_l} + \beta_{QN} W_{Q_k} W_{N_l} + \varepsilon_{ijklm}, \qquad (22)$$

where $1 \leqslant i \leqslant 6$, $1 \leqslant m \leqslant 9$, and the "coded" independent variables W_{P_j}, W_{Q_k}, and W_{N_l} are defined as follows:

$$W_{P_j} = \begin{cases} -1, & \text{for } j = 0, \\ 0, & \text{for } j = 1, \\ +1, & \text{for } j = 2; \end{cases}$$

$$W_{Q_k} = \begin{cases} -1, & \text{for } k = 1, \\ 0, & \text{for } k = 2, \\ +1, & \text{for } k = 3; \end{cases} \quad \text{and } W_{N_l} = \begin{cases} -1, & \text{for } l = 1, \\ 0, & \text{for } l = 2, \\ +1, & \text{for } l = 3. \end{cases}$$

We used the statistical model (22) to perform analysis of variance (ANOVA) and appropriate follow-up multiple comparisons procedures for each of the performance measures L, D, B, and A; and for these performance measures, the dependent variable Z_{ijklm} is given by L_{ijklm}, D_{ijklm}, B_{ijklm}, or A_{ijklm}, respectively.

To assess the validity of the statistical model (22) on which our experimental performance evaluation is based, we examined the estimated residuals for this model using normal probability plots and the Shapiro-Wilk test for normality (Shapiro and Wilk 1965). Figures 7 to 10, respectively, display normal probability plots for the residuals corresponding to the performance measures L, D, B, and A in test problem 1. The P-values for the Shapiro-Wilk test statistics corresponding to Figures 7 through 10 are 0.99, 0.98, 0.99,

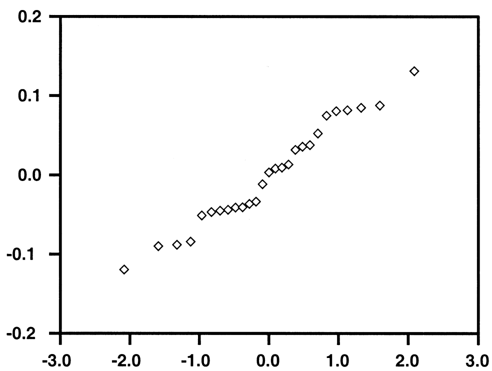

Figure 8. Normal probability plot of estimated residuals in the ANOVA model (22) for performance measure D of test problem 1.

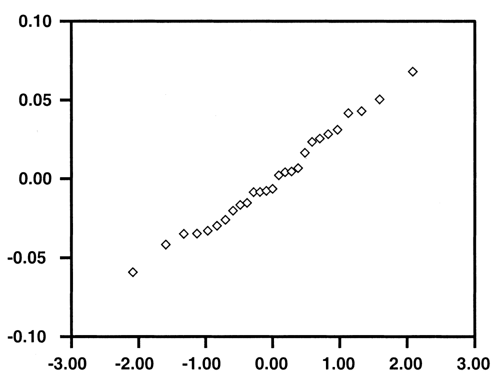

Figure 9. Normal probability plot of estimated residuals in the ANOVA model (22) for performance measure B in test problem 1.

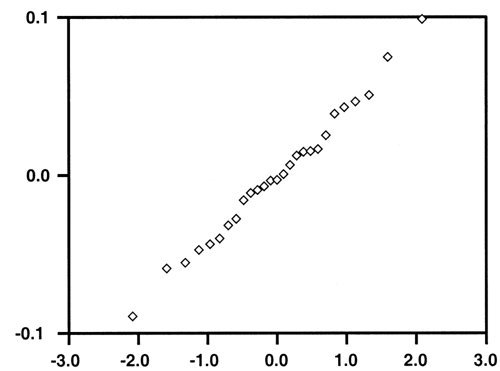

Figure 10. Normal probability plot of estimated residuals for ANOVA model (22) of performance measure A in test problem 1.

and 0.98, respectively. These results are representative of the other 20 cases discussed in Humphrey (1997). We believe these results provide substantial visual and statistical evidence that the residuals associated with the ANOVA model (22) are approximately normally distributed with a constant variance across all levels of the independent variables P_j, Q_k, and N_l. For a more detailed discussion of the validation of (22) including formal statistical tests for homogeneity of the response variances, see Humphrey and Wilson (2000).

We computed average performance measures for each problem i ($1 \leqslant i \leqslant 6$) and optimization procedure j ($0 \leqslant j \leqslant 2$) as follows:

$$\bar{L}_{ij} \equiv \frac{1}{81} \sum_{k=1}^{3} \sum_{l=1}^{3} \sum_{m=1}^{9} L_{ijklm};$$

and $\bar{D}_{ij}$, $\bar{B}_{ij}$, and $\bar{A}_{ij}$ are defined similarly. For problem i separately ($1 \leqslant i \leqslant 6$), we used the ANOVA procedure of SAS (SAS Institute 1989) to compare $\bar{L}_{i0}$, $\bar{L}_{i1}$, and $\bar{L}_{i2}$ via a Ryan-Einot-Gabriel-Welsch multiple comparisons F-test (Einot and Gabriel 1975) with level of significance 0.05. This test looks for significant differences among the means, and groups the means accordingly (where means not significantly different from each other are placed within the same group). This type of test was performed 24 times so that each of the three optimization procedures was compared against the other two procedures with respect to each of the four performance measures on all six test problems. It should be recognized that the overall level of significance 0.05 for the multiple-comparisons tests discussed in the next section applies to each combination of test problem and performance measure separately. Because the performance of the selected optimization procedures differed so drastically between test problems, it was necessary to analyze the results for each test problem as a separate experiment. Table III summarizes the results of these F-tests, and Section 3.3 below contains an analysis of these results.

3.3 Analysis of Experimental Results

Most of the analysis of this section is based directly on the information presented in Table III. We consider each of the performance measures separately over the six problems studied. Humphrey (1997) provides complete details on the analysis of the experimental results.

3.3.1 ANOVA Results

The ANOVA results provide additional evidence of the adequacy of the statistical model (22) for the purposes of this study. All R^2 values are above 0.93, and most are above 0.99. For each test problem, the ANOVA reveals two significant main effects—problem dimensionality and optimization procedure. As the dimensionality of each test problem increased, optimization of the associated function became more difficult; and this phenomenon resulted in large F-values for the dimensionality factor (Q_k).

For each test problem, the corresponding ANOVA for (22) also reveals that optimization procedure (P_j) is a highly significant main effect. As suggested by Table III and elaborated in the next four subsections, the principal source of this significant effect was the generally superior performance of procedure RSS versus procedures NM and RS+S9 with respect to the performance measures A, B, and D.

Table III. Results of Multiple Comparisons Tests on Procedures NM, RS+S9, and RSS for Level of Significance 0.05

Prob	Performance Measure											
	$\bar{L}_{ij}$			$\bar{D}_{ij}$			$\bar{B}_{ij}$			$\bar{A}_{ij}$		
i	Proc j	Value	Gr*	Proc j	Value	Gr*	Proc j	Value	Gr*	Proc j	Value	Gr*
1	RS+S9	6.86	1	NM	5.10	1	NM	1.28	1	NM	0.41	1
	RSS	6.68	2	RS+S9	4.94	2	RS+S9	1.28	1	RS+S9	0.40	1
	NM	5.57	3	RSS	0.48	3	RSS	0.38	2	RSS	0.19	2
2	RS+S9	7.06	1	NM	0.22	1	NM	0.47	1	NM	0.29	1
	RSS	6.83	2	RSS	0.12	2	RS+S9	0.39	2	RS+S9	0.24	2
	NM	5.56	3	RS+S9	0.10	2	RSS	0.35	3	RSS	0.20	3
3	RS+S9	6.69	1	NM	20.2	1	NM	2.01	1	RS+S9	1.04	1
	RSS	6.68	1	RS+S9	20.0	1	RS+S9	2.01	1	NM	1.04	1
	NM	5.50	2	RSS	18.2	2	RSS	1.74	2	RSS	0.96	2
4	RSS	7.25	1	NM	11.0	1	NM	1.66	1	NM	0.84	1
	RS+S9	7.02	2	RS+S9	10.1	2	RS+S9	1.59	2	RS+S9	0.79	2
	NM	5.78	3	RSS	3.76	3	RSS	0.95	3	RSS	0.41	3
5	RS+S9	6.89	1	NM	1.85	1	RS+S9	0.78	1	RS+S9	0.32	1
	RSS	6.68	2	RS+S9	1.56	1	NM	0.78	1	NM	0.30	1
	NM	5.57	3	RSS	0.53	2	RSS	0.55	2	RSS	0.29	1
6	RS+S9	6.98	1	RS+S9	245.2	1	RSS	1.75	1	NM	0.99	1
	RSS	6.91	2	NM	238.3	2	NM	1.53	2 1	RS+S9	0.92	1
	NM	5.60	3	RSS	229.8	3	RS+S9	1.40	2	RSS	0.92	1

*Grouping of procedures with nonsignificant differences in performance based on Ryan-Einot-Gabriel-Welsch multiple comparison procedure.

Table IV. Relative Computational Effort of Procedures

Procedure	Problem 1	2	3	4	5	6	Avg
NM	1	1	1	1	1	1	1
RSS	3.2	3.6	3.1	4.5	3.1	3.9	3.6
RS+S9	3.6	4.7	3.3	3.5	3.8	4.3	3.8

Of the two-factor interactions represented in (22), the only significant effect is the interaction of problem dimensionality with search procedure. Unfortunately, we have been unable to draw any general conclusions about the relative advantages or disadvantages of the three search procedures with increasing dimensionality. We believe that this issue should be the subject of future investigation.

3.3.2 Number of Function Evaluations

In terms of the logarithm of the number of function evaluations performed, Table III shows that procedures RSS and RS+S9 were roughly comparable, while procedure NM required significantly less work than RSS or RS+S9 to deliver a final answer. If we take the number of function evaluations for procedure NM as a baseline, then from Table IV we see that procedures RSS and RS+S9 generally required about four times as many function evaluations as procedure NM.

3.3.3 Final Function Value at Estimated Optimum

The results presented in Table III for the performance measure D warrant further discussion. In every test problem except problem 2, procedure RSS yielded an average value of D that is significantly smaller than the average D-values produced by either NM or RS+S9. In problem 2 (that is, the trigonometric function with the lattice (20) of optimal points), procedures RSS and RS+S9 yielded results that are not statistically distinguishable from each other but are significantly better than the results of procedure NM. Moreover, notice that with respect to the performance measure D, procedure RSS performed *much* better than either RS+S9 or NM on two of the six problems (namely, problems 1 and 4).

3.3.4 Maximum Relative Component Deviation from Global Optimum

With respect to the performance measure B, Table III shows that procedure RSS significantly outperformed both procedures NM and RS+S9. In problem 1, procedure RSS has a $\bar{B}$-value of about 0.38 while the corresponding $\bar{B}$-values for procedures RS+S9 and NM are each about 1.28. The results are less dramatic for problems 2–5, but they still clearly favor

RSS. In problem 6, however, there is no clear-cut distinction between the performances of the three procedures.

3.3.5 Average Relative Component Deviation from Global Optimum

With respect to the performance measure *A*, Table III shows that procedure RSS significantly outperformed procedures NM and RS+S9 in the first four problems. In problems 5 and 6 there are no significant differences in the performances of the three procedures.

4. Conclusions and Recommendations for Future Research

4.1 Conclusions

The results of our experimental performance evaluation of procedures NM, RS+S9, and RSS show that in the six test problems, procedure RSS required roughly as much work as procedure RS+S9 and about four times as much work as procedure NM. However, in four of the six test problems, procedure RSS significantly outperformed RS+S9 and NM with respect to all measures of convergence to the optimum; and in the other two test problems, RSS consistently delivered results at least as good as the results for procedures NM and RS+S9. Although such experimental results are extremely difficult to generalize, they do suggest that significant improvements in the performance of simplex-search-type procedures can be achieved by exploiting the principal features of procedure RSS—namely, a multiphase approach in which (a) the search is restarted in the second and subsequent phases; (b) in successive phases the size of the initial simplex is progressively reduced while the shrink coefficient is progressively increased to provide adequate protection against premature termination; and (c) in the end the best solution is taken over all phases of the search procedure.

4.2 Recommendations for Future Research

The analysis in Section 4 raises questions and issues that merit consideration for future work. The suite of six test problems should be enlarged to provide for analysis on a collection of test problems that encompasses an even broader range of the following factors: degree of difficulty, dimensionality, and response surface geometry. The experimental performance evaluation should also be expanded to include other variants of procedure NM.

Another promising area for future research is a more detailed study of the effects of dimensionality on the performance of procedure RSS. While our study looked at dimensionalities $d = 2, 10, 18$ ($d = 4, 8, 16$ for test problem 4), a more detailed examination of dimensionalities within and above this range could probably provide additional insight into how procedure RSS performs as the number of design variables changes.

Finally, an effort should be made to formulate some rules of thumb for the use of procedure RSS in general applications. Particular issues of interest are how to set the starting point $\mathbf{x}_1$ and the initial step size parameter τ for the search procedure. We believe that future progress in the development of effective and efficient simplex-search-type procedures will depend critically on the development of generally applicable, robust techniques for adjusting these quantities to the problem at hand.

References

Anderson, V.L., R.A. McLean. 1974. *Design of Experiments: A Realistic Approach*. Marcel Dekker, Inc., New York.

Barton, R.R., J.S. Ivey, Jr. 1996. Nelder-Mead simplex modifications for simplex optimization. *Management Science* **42** 954–973.

Corana, A., M. Marchesi, C. Martini, S. Ridella. 1987. Minimizing multimodal functions of continuous variables with the "simulated annealing" algorithm. *ACM Transactions on Mathematical Software* **13** 262–280.

Einot, I., K.R. Gabriel. 1975. A study of the powers of several methods of multiple comparisons. *Journal of the American Statistical Association* **70** 574–583.

Humphrey, D.G. 1997. A revised simplex search procedure for stochastic simulation response-surface optimization. Ph.D. Dissertation, Department of Industrial Engineering, North Carolina State University, Raleigh, NC.

Humphrey, D.G., J.R. Wilson. 1998. A revised simplex search procedure for stochastic simulation response-surface optimization. D.J. Medeiros, E.F. Watson, J.S. Carson, M.S. Manivannan, eds. *Proceedings of the 1998 Winter Simulation Conference*. Institute of Electrical and Electronics Engineers, Piscataway, NJ. 751–759.

Humphrey, D.G., J.R. Wilson. 2000. A revised simplex search procedure for stochastic simulation response-surface optimization. Technical Report, Department of Industrial Engineering, North Carolina State University, Raleigh, NC. ftp://ftp.ncsu.edu/pub/eos/pub/jwilson/rssv4c.pdf [accessed April 5, 2000].

Kuhl, M.E., J.R. Wilson. 2000. Least squares estimation of nonhomogeneous Poisson processes. *Journal of Statistical Computation and Simulation* **67** 75–108. ftp://ftp.ncsu.edu/pub/eos/pub/jwilson/jscs30.pdf [accessed April 5, 2000].

Moré, J.J., B.S. Garbow, K.E. Hillstrom. 1981. Testing unconstrained optimization software. *ACM Transactions on Mathematical Software* **7** 17–41.

Nelder, J.A., R. Mead. 1965. A simplex method for function minimization. *Computer Journal* **7** 308–313.

J.M. Parkinson, D. Hutchinson. 1972. An investigation into the efficiency of variants on the simplex method. F.A. Lootsma, ed. *Numerical Methods for Non-linear Optimization*. Academic Press, London. 115–135.

SAS Institute, Inc., 1989. *SAS/STAT User's Guide, Version 6, Fourth Edition*. SAS Institute Inc., Cary, NC.

Shapiro, S.S., M.B. Wilk. 1965. An analysis of variance test for normality. *Biometrika* **52** 591–611.

Wagner, M.A.F., J.R. Wilson. 1996. Using univariate Bézier distributions to model simulation input processes. *IIE Transactions* **28** 699–711.