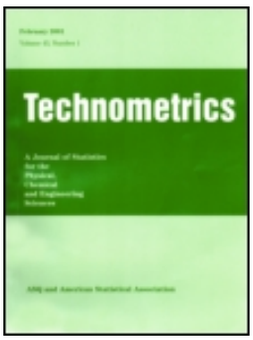

Technometrics

ISSN: 0040-1706 (Print) 1537-2723 (Online) Journal homepage: https://www.tandfonline.com/loi/utch20

Generalized Simulated Annealing for Function Optimization

Ihor O. Bohachevsky , Mark E. Johnson & Myron L. Stein

To cite this article: Ihor O. Bohachevsky , Mark E. Johnson & Myron L. Stein (1986) Generalized Simulated Annealing for Function Optimization, Technometrics, 28:3, 209-217, DOI: 10.1080/00401706.1986.10488128

To link to this article: https://doi.org/10.1080/00401706.1986.10488128

TECHNOMETRICS, AUGUST 1986, VOL. 28, NO. 3

This article is dedicated to the memory of Mark Kac.

Generalized Simulated Annealing for Function Optimization

Ihor O. Bohachevsky, Mark E. Johnson, and Myron L. Stein

Statistics and Operations Research
Los Alamos National Laboratory
Los Alamos, NM 87545

A generalized simulated annealing method has been developed and applied to the optimization of functions (possibly constrained) having many local extrema. The method is illustrated in some difficult pedagogical examples and used to solve a problem analyzed by Bates (*Technometrics*, 25, pp. 373–376, 1983), for which we identify an improved optimum. The sensitivity of the solution to changes in the constraints and in other specifications of the problem is analyzed and discussed.

KEY WORDS: Sensitivity analysis; Optimal design; Multiple extrema; Boltzmann's distribution; Biased random walk; Metropolis algorithm.

1. INTRODUCTION

Optimization problems surface regularly in statistical applications. Some recent examples from this journal include the construction of optimal designs (Johnson and Nachtsheim 1983) and parameter estimation in reliability models (Lawless 1983), in kinetic models (Ziegel and Gorman 1980), and in nonlinear problems (Schwetlick and Tiller 1985). Occasionally, direct focus has been on the optimization methods themselves, as in Olsson and Nelson (1975), which showcased Nelder and Mead's simplex scheme.

There has been considerable recent interest in global function optimization. Levy and Montalvo (1985) proposed a deterministic scheme that has achieved good empirical performance. Their approach requires calculation of the second derivatives of the objective function. Researchers at Erasmus University (Boender and Rinnooy Kan 1983; Boender 1984; Rinnooy Kan and Timmer 1984) examined stochastic approaches, usually of the multiple-starting-point–local-improvement variety. Conlon (1985) programmed Price's (1977) algorithm and reported the capability of solving multiple extrema problems. The preceding methods do not enjoy the same qualitative features as the generalized simulated annealing method. A thorough comparison of recently published algorithms on test problems may be in order.

In this article we present a generalized simulated annealing method for function optimization that has considerable potential in statistical applications. The advantages of the method are the ability to migrate through a sequence of local extrema in search of the global solution and to recognize when the global extremum has been located. The presentation contains descriptions of the ideas that underlie the generalized simulated annealing method and of some successful applications. We are currently investigating the general characteristics of the problems to which the method is applicable, the approaches that may lead to the a priori estimates of the parameters that control the algorithm's performance for a given objective function, and ways to improve the efficiency of the method.

The method derives its name from the annealing process of metals in which many final crystalline configurations (corresponding to different energy states) are possible depending on the rate of the cooling process. The method of simulated annealing can be traced to Metropolis [see Kirkpatrick, Gelatt, and Vechhi (1983) for a detailed historical discussion], who attempted to simulate the behavior of an ensemble of atoms in equilibrium at a given temperature.

Our article begins with a brief summary of the simulated annealing method and presents its generalization in Section 2 together with an illustration on a function with many local optima. In Section 3 a problem from the optimal design literature is examined (Bates 1983) and a substantially better optimum is determined. In Section 4 a sensitivity analysis for this problem is reported and it is shown that the constraints imposed by the experimenters greatly in-

fluence the character of the resulting optimal designs. We conclude in Section 5 with a discussion of the potential of the generalized simulated annealing method and an indication of other applications.

2. THE GENERALIZED SIMULATED ANNEALING METHOD

The development of the standard simulated annealing method was motivated by the behavior of mechanical systems with a very large number of degrees of freedom. In this context very large means so large as to preclude any exhaustive analysis of the possible states of the system and admitting only the possibility of a statistical analysis. According to the general principles of physics, any such system will, given the necessary freedom, tend toward the state of minimum energy. Therefore a mathematical model of the behavior of such a system will contain a method for minimizing a certain function—namely, the total energy of the system.

For example, atoms of a molten metal when cooled to a freezing temperature will tend to assume relative positions in a lattice in such a way as to minimize the potential energy of their mutual forces. Because of a very large number of atoms and possible arrangements, the final state will most likely correspond to only a local energy minimum and not a global one. The solidified metal may be reheated and cooled slowly with the hope that it will then migrate to a lower energy state. In metallurgy that process is called annealing; therefore, the method that mathematically models it is called simulated annealing.

Simulated annealing, then, is a convenient way to find a global extremum of a function that has many local extrema and may not be smooth (the method does not require calculation of derivatives). The method may be used to find either minima or maxima. It is convenient to explain it, however, with the assumption that the extremum is a minimum and that the objective function is positive. A way of removing these assumptions is indicated at the end of this section.

The method is a biased random walk that samples the objective function in the space of the independent variables; it is executed in the following manner. Starting at an initial point (chosen randomly or specified, depending on available information), the corresponding initial value of the objective function, ϕ_0, is calculated. Next, a random point is chosen on the surface of the unit n-dimensional hypersphere, where n is the dimensionality of the problem; this specifies a random direction in which a step of size Δr is taken. The magnitude of Δr depends on the properties of the objective function and on the desired accuracy and resolution; its determination may require some experimentation.

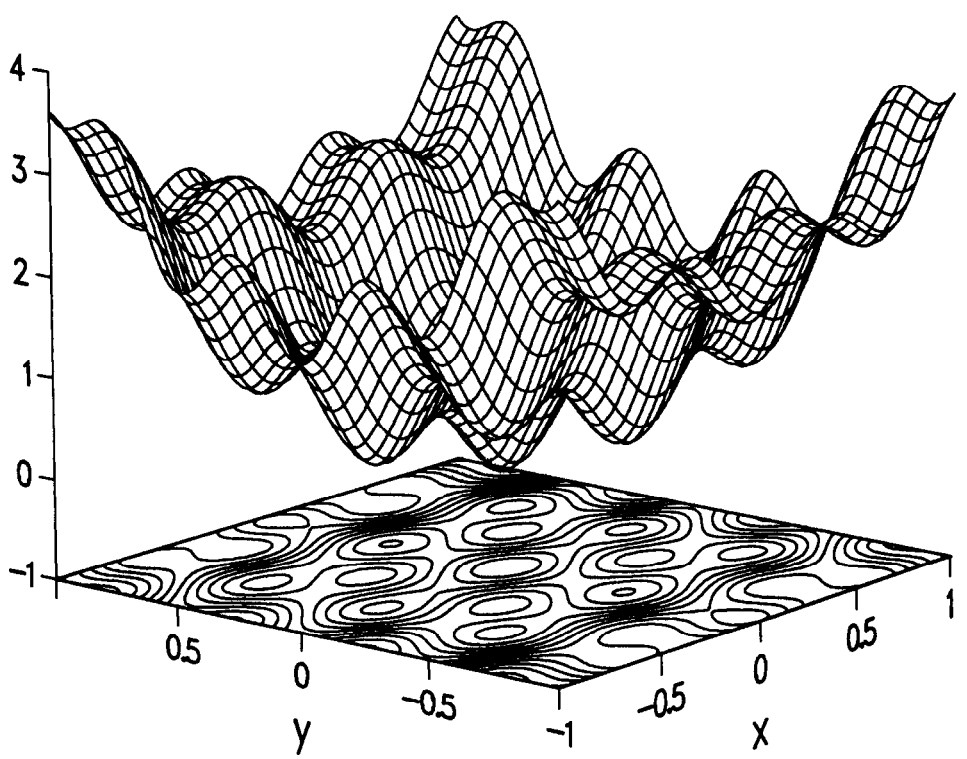

Figure 1. Objective Function With Many Local Minima and a Global Minimum at (0, 0). The function is $\phi_1(x, y) = ax^2 + by^2 - c\cos(\alpha x) - d\cos(\gamma y) + c + d$, *where* $a = 1, b = 2, c = .3, d = .4, \alpha = 3\pi, \gamma = 4\pi$.

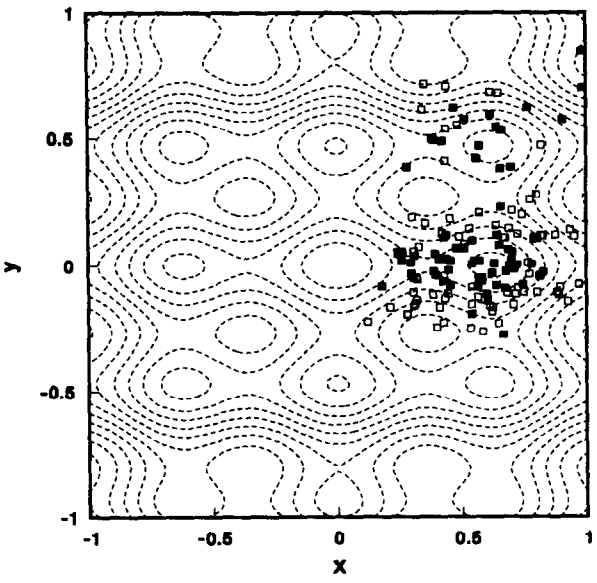

Figure 2. Standard Simulated Annealing Method Example With $\Delta r = .15$, $\beta = 4.5$, and $N = 150$. Solid squares are accepted steps, open squares are rejected steps. The function is $\phi_1(x, y) = ax^2 + by^2 - c\cos(\alpha x) - d\cos(\gamma y) + c + d$, where $a = 1$, $b = 2, c = .3, d = .4, \alpha = 3\pi, \gamma = 4\pi$. The starting point is (1, 1).

An alternate procedure for obtaining a random direction is to choose n random numbers, $\{\theta_i\}_{i=1}^n$, from the uniform distribution on $[-1, 1]$ and to convert them into direction cosines, $\theta_i(\sum_{i=1}^n \theta_i^2)^{-1/2}$. We used both procedures without significant differences in the results.

At the endpoint of the step the new value of the objective function, ϕ_1, is calculated and the step is accepted with probability p given by:

$$p = 1 \qquad \text{if } \Delta\phi = \phi_1 - \phi_0 \leq 0$$
$$= \exp(-\beta\Delta\phi) \quad \text{if } \Delta\phi > 0,$$

where β is a positive parameter; a way to determine its value is indicated later. This means that beneficial steps ($\phi_1 \leq \phi_0$) are accepted unconditionally but the detrimental steps ($\phi_1 > \phi_0$) are accepted according to an auxiliary experiment. A random number ρ is generated from the uniform distribution on (0, 1) and is compared to the value $\exp(-\beta\Delta\phi)$. If $p < \exp(-\beta\Delta\phi)$, then the step is accepted; otherwise it is rejected and a step in a new random direction is attempted. Thus the probability of accepting an increment of ϕ depends on the size of the increment: the larger the increment, the smaller the probability of its acceptance. The probability of accepting a detrimental increment is always greater than zero, however, so the path may walk out of local and global minima. The number of steps required for that depends on the magnitude of β that must be adjusted for each particular objective function. Analytical considerations suggest that a reasonable value of β should be such that the inequalities $.5 < \exp(-\beta\Delta\phi) < .9$ hold because probabilities too close to unity lead to inefficient searches (virtually every step is accepted) whereas probabilities much below .5 require too many function evaluations to escape from a local minimum. A reasonable value of the step size, Δr, should be such as to allow escape from a local minimum in a few (2–3) steps.

The expression for the conditional acceptance probability p is motivated by the physical process modeled by the simulated annealing method. In statistical mechanics the probability that the system will transit from the state with energy E_1 to the state with energy E_2 where $E_2 - E_1 = \Delta E > 0$ is $\exp(-\Delta E/kT)$, where k is the Boltzmann constant and T is the absolute temperature. Thus $\beta = 1/kT$, and the lower the temperature the smaller the probability of transition to a higher energy state.

The performance of the standard simulated annealing method is illustrated in Figures 1 and 2. Figure 1 shows an objective function of two variables, $\phi(x, y)$, that has many local minima and one global minimum at $x = y = 0$ [$\phi(x, y) = ax^2 + by^2 - c\cos(\alpha x) - d\cos(\gamma y) + c + d$]. Figure 2 shows the steps of the random walk with $\beta = 4.5$ and the step size $\Delta r = .15$. The solid squares indicate the accepted steps, and the open squares indicate the rejected ones; this result indicates that 150 steps are not sufficient to locate the global minimum unambiguously. The random path wanders in and out of various extrema and does not provide a logical termination criterion. A subsequent analysis of the clusters of attempted steps is needed to identify the global extremum. Because of this property, the method may require an excessively large number of steps to find a

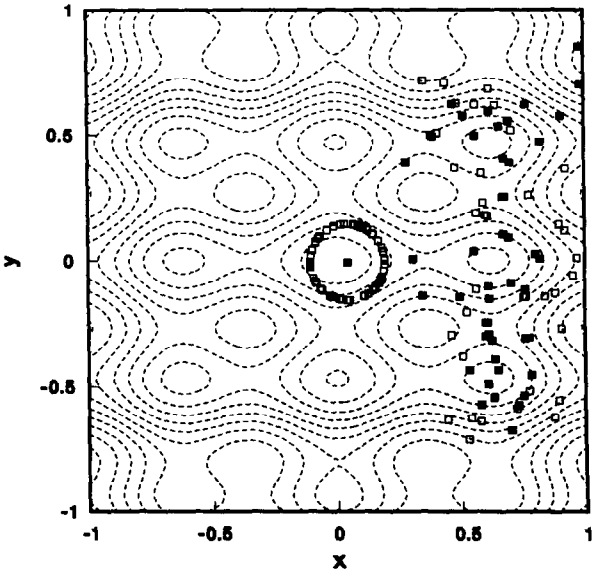

Figure 3. Generalized Simulated Annealing Method Example With $g = -1$, $\Delta r = .15$, $\beta = 3.5$, and $N = 150$ (same function ϕ_1 as in Fig. 2). The starting point is (1, 1).

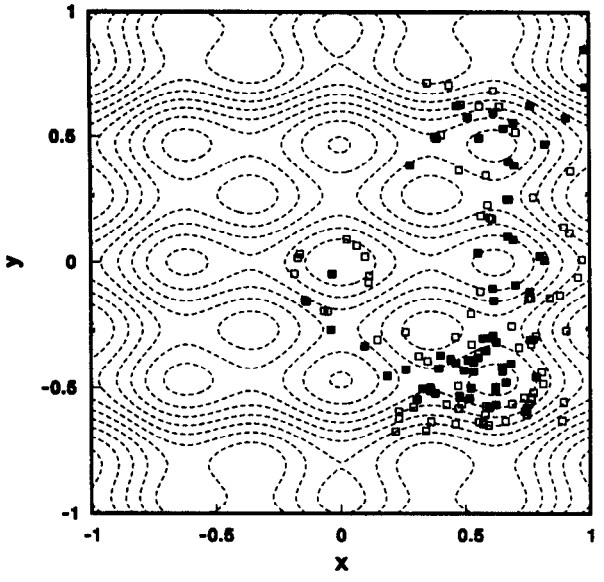

Figure 4. Generalized Simulated Annealing Method Example With $g = -1$, $\Delta r = .15$, $\beta = 4.0$, and $N = 150$ (same function ϕ_1 as in Fig. 2). The starting point is (1, 1).

satisfactory solution. In contrast, an equal number of steps with the generalized simulated annealing method, introduced in the next paragraph, leads to an unambiguous solution, as illustrated in Figure 3.

To overcome the shortcomings of the standard simulated annealing method, we generalized the method by making the probability of accepting a detrimental step tend to zero as the random walk approaches the global extremum. This scheme leads to a path that terminates at the global extremum and thus requires fewer steps than the standard annealing method. Assuming that the objective function equals zero at the global minimum, the required behavior of the probability is easily accomplished by setting $p = \exp(-\beta\phi^g\Delta\phi)$, for $\Delta\phi > 0$, where g is an arbitrary negative number; for $g = 0$, we recover the standard simulated annealing method. We indicate later how to remove the restriction on the value of the objective function at the global extremum. To summarize the generalized simulated annealing method, the following algorithm is offered.

Assume $\phi(\mathbf{x})$ is the function to be minimized with $\mathbf{x}$ restricted to Ω, a subset of R^n. Let ϕ_m be the value of ϕ at the global optimum.

1. Let $\mathbf{x}_0$ be an arbitrary starting point (either specified or selected randomly from Ω).
2. Set $\phi_0 = \phi(\mathbf{x}_0)$. If $|\phi_0 - \phi_m| < \varepsilon$, stop.
3. Random direction. Generate n independent standard normal variates $Y_1, \ldots, Y_n$ and compute the components of $\mathbf{U}$: $U_i = Y_i/(Y_1^2 + \cdots + Y_n^2)^{1/2}$, $i = 1, \ldots, n$.
4. Set $\mathbf{x}^* = \mathbf{x}_0 + (\Delta r)\mathbf{U}$.
5. If $\mathbf{x}^* \notin \Omega$, return to step 3. Otherwise, set $\phi_1 = \phi(\mathbf{x}^*)$ and $\Delta\phi = \phi_1 - \phi_0$.
6. If $\phi_1 \le \phi_0$, set $\mathbf{x}_0 = \mathbf{x}^*$ and $\phi_0 = \phi_1$. If $|\phi_0 - \phi_m| < \varepsilon$, stop. Otherwise, go to step 3.
7. If $\phi_1 > \phi_0$, set $p = \exp(-\beta\phi_0^g\Delta\phi)$. (a) Generate a uniform 0–1 variate V. (b) If $V \ge p$, go to step 3. (c) If $V < p$, set $\mathbf{x}_0 = \mathbf{x}^*$, $\phi_0 = \phi_1$, and go to step 3.

The performance of the generalized simulated annealing algorithm is shown in Figure 3 for $g = -1$, $\Delta r = .15$, $\beta = 3.5$; a total of 150 steps were taken. The global extremum is clearly indicated by the circular pattern of steps pertaining to unsuccessful attempts to wander out of the global minimum. Figure 4 illustrates how a larger value of β (=4) inhibits but does not eliminate the ability of the algorithm to climb out of a local minimum. In applications the random walk is terminated when an arbitrarily prescribed number of successive trials (say 20–50) fails to produce a single acceptable increment. The global minimum is the final accepted point (i.e., the center of the circle on whose circumference are the points of this sequence of randomly distributed rejected steps).

The performance of the algorithm for different contour shapes is illustrated in Figures 5 and 6 to demonstrate that it does not depend on any special property of the objective function shown in Figure 1. From these exercises we learned that the algorithm performs well (i.e., it requires a small number of trials to locate the global extremum) when the value of β is selected so that 50%–90% of the detrimental steps are accepted. This observation is consistent

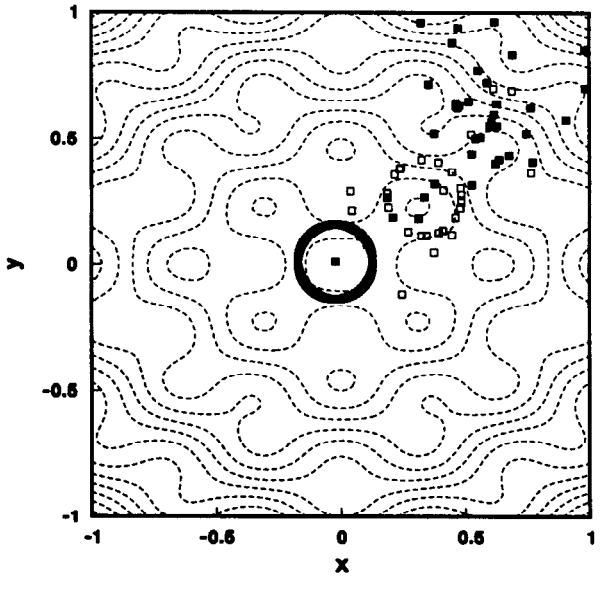

Figure 5. Generalized Simulated Annealing Method Example With $g = -1$, $\Delta r = .15$, $\beta = 3.0$, and $N = 400$. The function $\phi_2(x, y) = ax^2 + by^2 - c[\cos(\alpha x)\cos(\gamma y)] + c$, with the coefficients as in Figure 2. The starting point is (1, 1).

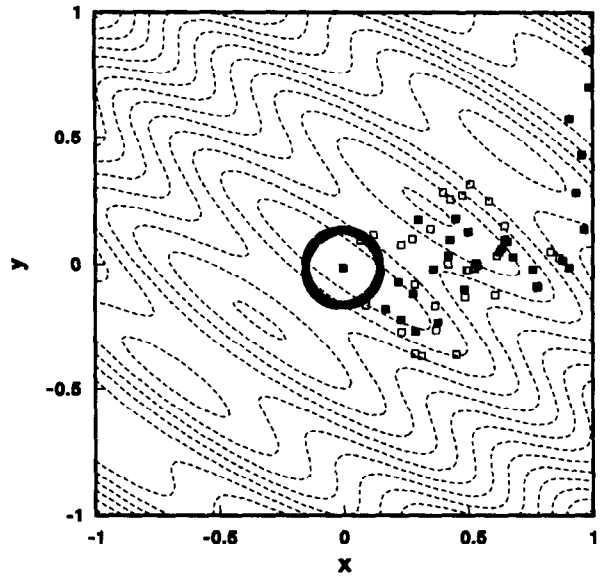

Figure 6. Generalized Simulated Annealing Method Example With $g = -1$, $\Delta r = .15$, $\beta = 3.0$, and $N = 400$. The function ϕ_3 $(x, y) = ax^2 + by^2 - c\cos(\alpha x + \gamma y) + c$, with the coefficients as in Figure 2. The starting point is (1, 1).

with the analytical estimates deduced previously, $.5 < \exp(-\beta\Delta\phi/\phi) < .9$. Figure 7 illustrates, for a representative value of β, the number of function evaluations that were required to locate the global minimum of a function of three variables with numerous local minima (an extension of the function shown in Fig. 1 to three dimensions). These results suggest that the number of trials required to find the solution does not increase very rapidly with the dimensionality of the problem. The number of steps required to obtain typical 2-dimensional results shown in Figures 5 and 6 ranged between 300 and 500 whereas the number of steps required to solve a corresponding 3-dimensional problem clustered between 600 and 800.

Figure 7 also indicates the number of times each problem was solved with different initial points and different parameter values to arrive at some of the general observations presented in this article.

The restrictions imposed on the behavior of the objective function in the preceding presentation of the generalized simulated annealing method are easily removed by observing that the maximization of any objective function ϕ is trivially equivalent to the minimization of $\phi_{max} - \phi$, where ϕ_{max} is the value of ϕ at the global maximum. This substitution also ensures the vanishing of the new objective function at the global extremum; to ensure vanishing in the case of minimization, we work with $\phi - \phi_{min}$, where ϕ_{min} is the value of ϕ at the global minimum. The primary difficulty is in the fact that, in general, the value of ϕ_{min} or ϕ_{max} is not known a priori. (There are many problems, however, in which that value is known and only its location requires determination). When that is the case, we begin in the minimization problem with some tentative estimate of ϕ_{min} and continue the random walk until $\phi - \phi_{min}$ becomes negative; at that time we decrease ϕ_{min}, continue the search, and repeat the process when necessary. Of course, if the initial estimate of ϕ_{min} is too small, it is increased as

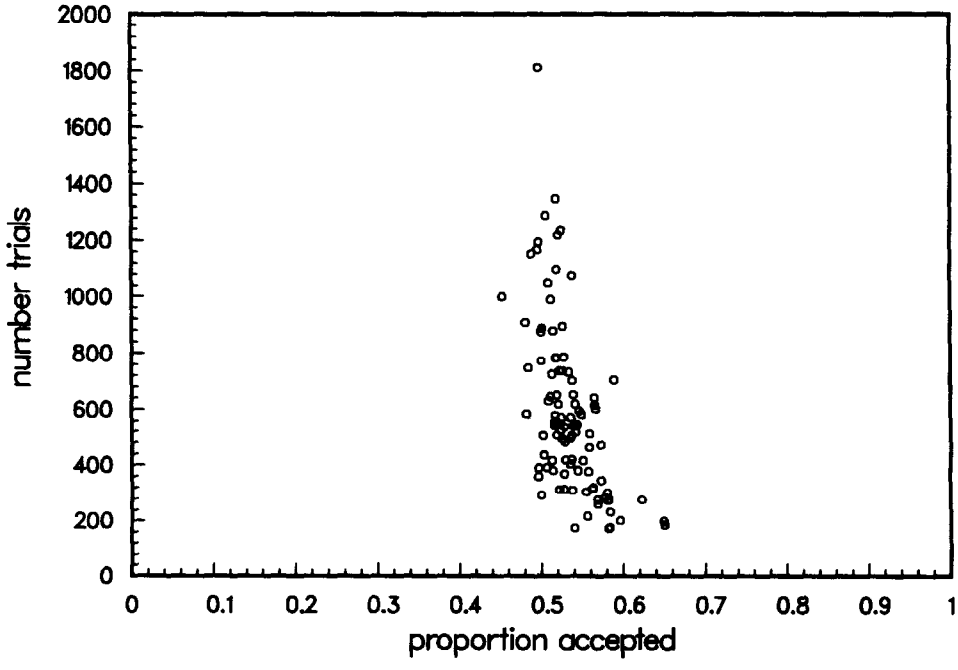

Figure 7. Variation in Proportion of Detrimental Steps Accepted for a 3-Dimensional Version of the ϕ_1 Function Using the Generalized Annealing Method With $\beta = 3.5$ and $\Delta r = .15$.

required to ensure the vanishing of the objective function. The magnitudes of the required adjustments must be estimated for each particular problem.

3. AN OPTIMAL DESIGN PROBLEM

Bates (1983) described an experiment that requires determination of the concentration of a neurotransmitter in solution as a function of immersion time of a brain tissue slice. The functional form considered was

$$c_i = \theta_1[\exp(-\theta_3 t_{i-1}) - \exp(-\theta_3 t_1)] + \theta_2(t_i - t_{i-1}), \tag{3.1}$$

where c_i is the concentration and t_i is the time (measured from the first immersion) when the tissue is moved from vial i to vial $i+1$. The object of the design problem is to select the set of times t_i that maximizes the precision in estimating θ_3. Because the parameters θ_1 and θ_2 appear linearly in (3.1) for fixed θ_3, their unknown values do not affect the optimal design. The unknown value of θ_3 does matter, but the experimenters have evidence to believe that θ_3 is restricted to the rather narrow range ($.2 \le \theta_3 \le .3$). The resulting design based on assuming $\theta_3 = .25$ is, of course, only locally optimal in the sense of Chernoff (1962).

The approach used to select the times of immersion involves the D-optimal criterion adapted to nonlinear functions (Box and Lucas 1959). For (3.1), we form the X matrix whose ith row is

$$[\exp(-\theta_3 t_{i-1}) - \exp(-\theta_3 t_i),\ t_i - t_{i-1},\ -\theta_1 t_{i-1}\exp(-\theta_3 t_{i-1}) + \theta_1 t_i \exp(-\theta_3 t_i)].$$

The D-optimal design is obtained by selecting the t_i so that the determinant of $X'X$ is maximized. Additional constraints imposed by the experimenters were: 11 vials were to be used (so that X is an 11×3 matrix), the tissue should remain in each vial at least 1 minute (so that $t_i - t_{i-1} \ge 1$), and the duration of the experiment would be at most 30 minutes ($t_{11} \le 30$).

Bates's best local optimum was found to be at the set of times (2.7, 3.7, 4.7, 5.7, 12.9, 13.9, 14.9, 15.9, 16.9, 17.9, 30) with corresponding $|X'X| = 71$. The generalized simulated annealing algorithm of Section 2 located the vector of times (3.2, 11.2, 12.2, 13.2, 14.2, 15.2, 16.2, 17.2, 18.2, 19.2, 30) for which $|X'X| = 105.3$. To ensure confidence that the global optimum was located, we reproduced the calculation several tens of times with different initial points. In these calculations the value of the step size, Δr, ranged between .005 and .01 and the value of β ranged between 10 and 100, clustering near 75; we found that the calculations are not sensitive to the exact value of β. The maximum value $|X'X|_m$, required to make the objective function $\phi = |X'X|_m - |X'X|$ vanish at the global optimum, was increased from 71 to 105.3 in approximately 1% increments and g was set equal to -1.

The relative efficiency computed as $(71/105)^{1/3} = .88$ is useful to compare our result to that of Bates. This value suggests that the volume of our confidence ellipsoids for the estimates of $(\theta_1, \theta_2, \theta_3)$ will be 88% the size of Bates's ellipsoids.

The number of function evaluations required to locate the global extremum with the generalized simulated annealing method depends on the accuracy of the initial estimate of the optimal value of the objective function and on the appropriate value of β. Therefore it is difficult to obtain a general and absolute measure of the algorithm's performance. Most of the solutions presented in this section required several thousand trials; this is a very small number relative to the total volume of the design space that must be searched, which is 30^{11} (11 variables ranging from 1 to 30). In spite of this large volume, simulated annealing methods (standard and general) work because they benefit from the constraints that reduce the volume of the space where the optimal solution must lie.

In the present problem the constraints $t_{i+1} - t_i \ge 1$ for all i imply that the volume, V, in which the optimal solution is to be determined is $V = \prod_{i=1}^{11} \tau_i$, where τ_i is the length of the interval to which t_i is restricted to satisfy the constraints. By using the fact that the geometric mean is always less than or equal to the arithmetic mean and that $\sum_{i=1}^{11} \tau_i \le 30$, we obtained the estimate

$$V \le (30/11)^{11} \simeq 62{,}090.$$

In view of this estimate, it is not surprising that a biased sample of a few thousand trials suffices to locate the optimal solution accurately. The trials that violate the constraints are discarded prior to the evaluation of the objective function and do not contribute significantly to the cost of the computation.

Table 1. Solutions for Various θ_3

θ_3	$\{t_i\}_{opt}$	$\lvert X'X\rvert_m$
.20	3.9, 12.0, 13.0, 14.0, 15.0, 16.0, 17.0, 18.0, 19.0, 20.0, 30.0	90.6
.25	3.2, 11.2, 12.2, 13.2, 14.2, 15.2, 16.2, 17.2, 18.2, 19.2, 30.0	105.3
.30	2.9, 10.2, 11.2, 12.2, 13.2, 14.2, 15.2, 16.2, 17.2, 18.2, 30.0	107.4

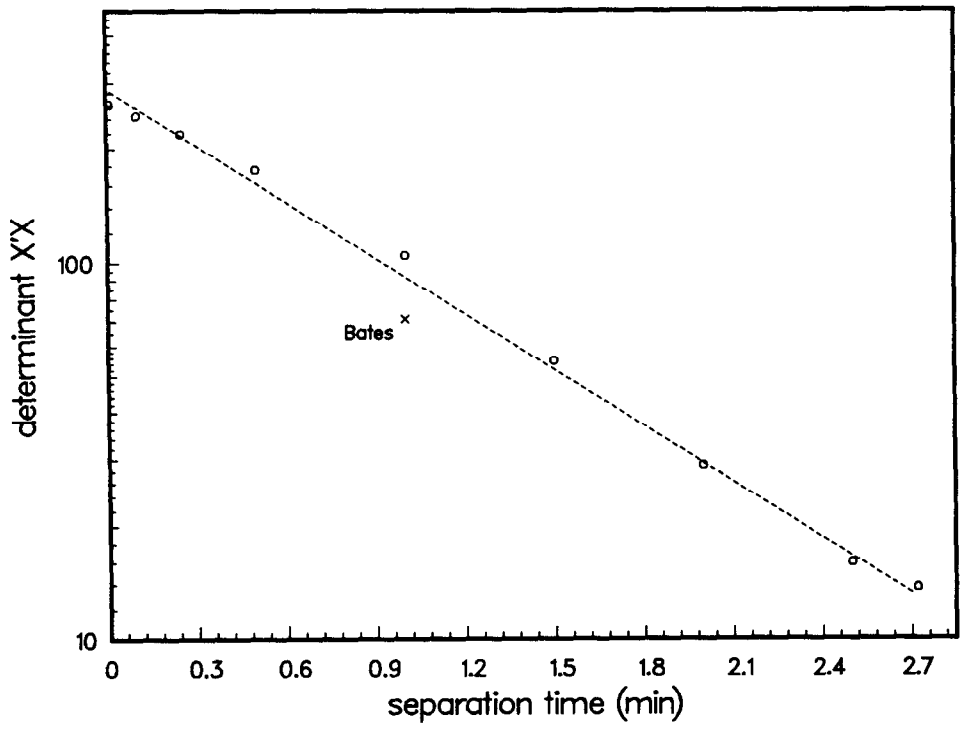

Figure 8. Log-Linear Plot of Optimal Objective Function Values as a Function of the Minimal Immersion Time Constraint. The dashed line is an eyeball trend line.

4. SENSITIVITY ANALYSES

A report of research results should contain (in addition to the presentation of the solution) a discussion of the solution's behavior, dependence on constraints, sensitivity to changes in parameters that specify the problem, and other information that may enhance the understanding of the results. This is es-

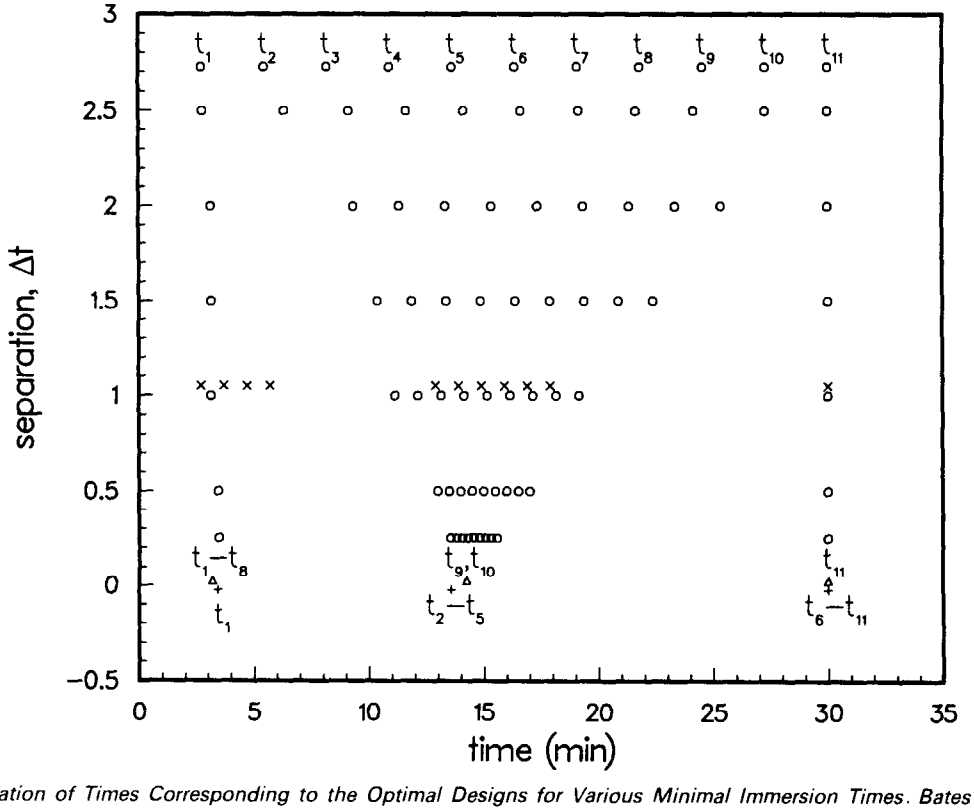

Figure 9. Location of Times Corresponding to the Optimal Designs for Various Minimal Immersion Times. Bates's local optimal solution is indicated with crosses. For Δt near zero, multiple global solutions can occur; two are indicated here (with + and with Δ).

Table 2. Solutions for Various Numbers of Vials

No. of Vials	$\{t_i\}_{opt}$	$\lvert X'X\rvert_m$
10	3.3, 11.7, 12.7, 13.7, 14.7, 15.7, 16.7, 17.7, 18.7, 30.0	121.9
11	3.2, 11.2, 12.2, 13.2, 14.2, 15.2, 16.2, 17.2, 18.2, 19.2, 30.0	105.3
12	3.2, 10.8, 11.8, 12.8, 13.8, 14.8, 15.8, 16.8, 17.8, 18.8, 19.8, 30.0	89.9

pecially true for the present problem because of the obvious dependence of the solution on the constraints. In this section we present the results of such sensitivity analyses. The purpose is to determine and to quantify the parametric dependencies and the trends and thus point out to the experimenters the potential gains that may be realized from modifications of the experimental procedures.

Bates noted that his design did not change appreciably when θ_3 was .2 or .3 instead of .25. To assess this statement at the global solutions, we used the generalized simulated annealing method and obtained designs for values of θ_3 as presented in Table 1. These results confirm Bates's observation.

All optimal solutions are at the boundaries determined by the separation constraints. Therefore it is of considerable interest to know how those constraints affect the value of the optimal solution. Figure 8 shows the dependence of the maximum value of $|X'X|$ on the minimum separation interval $\Delta t = t_{i+1} - t_i$. We see that $|X'X|$ increases exponentially as the separation constraint is relaxed. Therefore it behooves the experimenters to examine this trade-off carefully. Figure 9 shows the distribution of optimal sampling times for various Δt. Bates's solution is indicated by crosses ($\Delta t = 1$) and is clearly different from the global optimum.

An interesting feature of this model is the occurrence of several distinct sampling time distributions that yield the same value of $|X'X|_m = 262$ when the required separation distance is reduced to a very small value ($t_{i+1} - t_i \geq 10^{-3}$); two of those distributions are indicated in Figure 9.

To estimate the effect of the number of vials on the optimal experimental design, we obtained the results shown in Table 2 (assuming $\Delta t \geq 1$, $\theta_3 = .25$). These results indicate that the potential advantages of experiments with fewer vials should be examined.

Finally, we consider the impact of adjusting the duration of the experiment from the nominal 30 minutes. Clearly, if comparable results could be obtained in a 20- or 25-minute experiment, a reduced time would be preferable. Similarly, if the gains for running 35 or 40 minutes were substantial, a change in procedure could be warranted. We obtained the results shown in Table 3 ($\theta_3 = .25$, 11 vials). We conclude that a small change in the duration of the experiment leads to a very large change in performance.

Box, Hunter, and Hunter (1978) criticized optimal design construction in the following way: "In recent years the study of 'optimal' design has become separated from real experimentation with the predictable consequence that its limitations have not been stressed or, often, even realized" (p. 472).

The experience with the generalized simulated annealing method reported in this article provides information about the effects of the somewhat arbitrary constraints imposed by the experimenter. The results not only indicate the directions in which the experiment should be modified to realize maximum benefits, but also provide quantitative estimates of these benefits. Such insights should be of interest to experimenters.

5. SUMMARY OBSERVATIONS

Nonlinearities of objective functions and multiplicity of local extrema have been stumbling blocks in the search for methods capable of locating global extrema. The generalized simulated annealing method presented in this article appears to overcome these difficulties for a realistic class of problems—in particular, the determination of D-optimal experimental designs with numerous constraints. We have also demonstrated the necessity of sensitivity analyses to assess the impact of experimenter-imposed constraints.

The generalized simulated annealing algorithm de-

Table 3. Solutions for Various Stopping Times

Time	$\{t_i\}_{opt}$	$\lvert X'X\rvert_m$
25	3.0, 9.6, 10.6, 11.6, 12.6, 13.6, 14.6, 15.6, 16.6, 17.6, 25.0	35.3
30	3.2, 11.2, 12.2, 13.2, 14.2, 15.2, 16.2, 17.2, 18.2, 19.2, 30.0	105.3
35	3.6, 13.8, 14.8, 15.8, 16.8, 17.8, 18.8, 19.8, 20.8, 21.8, 35.0	226.4

tailed in Section 2 is easy to implement. A list of the computer code that produced Figure 3 is available from Mark E. Johnson upon request.

Examples of other applications of simulated annealing are the optimization of integrated circuit designs (Kirkpatrick et al. 1983), solutions of the traveling salesman problem (Bonomi and Lutton 1984), the construction of evolutionary trees (Lundy 1985), determination of optimal designs for polynomial regression (Haines 1986) and the attempt of Bohachevsky, Viswanathan, and Woodfin (1984) to develop an "intelligent" optical design program.

ACKNOWLEDGMENTS

We gratefully acknowledge helpful discussions with R. L. Bivins, D. G. Brooks, R. D. Cook, and T. P. Cotter and the computing assistance of D. O. Dickman and E. L. Woods in the early phases of this work.

[*Received January 1985. Revised February 1986.*]

REFERENCES

Bates, D. (1983), "The Derivative of $X'X$ and Its Uses," *Technometrics*, 25, 373–376.

Boender, C. G. E. (1984), "The Generalized Multinomial Distribution: A Bayesian Analysis and Applications," unpublished Ph.D. dissertation, Erasmus University, Dept. of Mathematics.

Boender, C. G. E., and Rinnooy Kan, A. H. G. (1983), "Bayesian Stopping Rules for a Class of Stochastic Global Optimization Methods," Report 8319/0, Erasmus University, Dept. of Mathematics.

Bohachevsky, I. O., Viswanathan, V. K., and Woodfin, G. (1984), "An 'Intelligent' Optical Design Program," in *Proceedings of the International Society for Optical Engineering*, pp. 104–112.

Bonomi, E., and Lutton, J.-L. (1984), "The N-City Travelling Salesman Problem: Statistical Mechanics and the Metropolis Algorithm," *SIAM Review*, 26, 551–568.

Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978), *Statistics for Experimenters*, New York: John Wiley.

Box, G. E. P., and Lucas, H. L. (1959), "Design of Experiments in Non-Linear Situations," *Biometrika*, 46, 77–90.

Chernoff, H. (1962), "Optimal Accelerated Life Designs for Estimation," *Technometrics*, 4, 381–408.

Haines, L. M. (1986), "The Application of the Annealing Algorithm to the Construction of Exact Optimal Designs in Polynomial Regression Models," unpublished manuscript.

Johnson, M. E., and Nachtsheim, C. J. (1983), "Some Guidelines in Constructing Exact D-Optimal Designs on Convex Design Spaces," *Technometrics*, 25, 271–277.

Kirkpatrick, S., Gelatt, C. D., Jr., and Vechhi, M. P. (1983), "Optimization by Simulated Annealing," *Science*, 220, 671–680.

Lawless, J. F. (1983), "Statistical Methods in Reliability," *Technometrics*, 25, 305–335.

Levy, A. U., and Montalvo, A. (1985), "The Tunneling Algorithm for the Global Minimization of Functions," *SIAM Journal of Scientific and Statistical Computing*, 6, 15–29.

Lundy, M. (1985), "Applications of the Annealing Algorithm to Combinatorial Problems in Statistics," *Biometrika*, 72, 191–198.

Olsson, D. M., and Nelson, L. S. (1975), "The Nelder–Mead Simplex Procedure for Function Minimization," *Technometrics*, 17, 45–51.

Price, W. L. (1977), "Optimization," *The Computer Journal*, 20, 367–370.

Rinnooy Kan, A. H. G., and Timmer, G. T. (1984), "A Stochastic Approach to Global Optimization," Report 8419/0, Erasmus University, Dept. of Mathematics.

Schwetlick, H., and Tiller, V. (1985), "Numerical Methods for Estimating Parameters in Nonlinear Models With Errors in the Variables," *Technometrics*, 27, 17–24.

Ziegel, E. R., and Gorman, J. W. (1980), "Kinetic Modelling With Multiresponse Data," *Technometrics*, 22, 139–151.