School of Operations Research
and Industrial Engineering
College of Engineering
Cornell University
Ithaca, N.Y. 14853

Technical Report No. 715
August 1986
(Revised September 1987)

A Review of Techniques for Simulation Optimization

by
S.H. Jacobson
L.W. Schruben

This research was supported by AT&T Information Systems and Air Force System Command, Rome Air Development Center, Griffiss Air Force Base, NY 13441-5700 by subcontract from Syracuse University, Syracuse, NY.

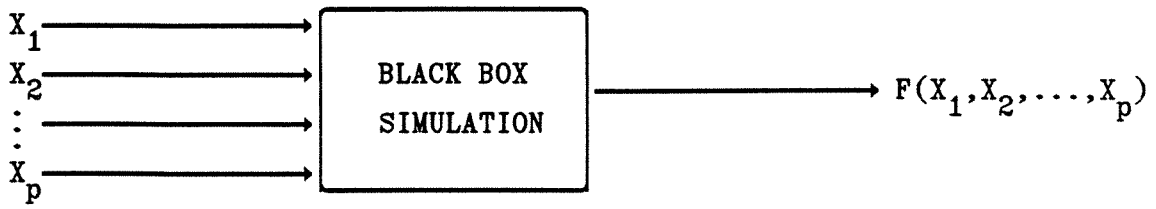
ABSTRACT

Techniques for discrete event simulation optimization are classified into three groups: path search methods, pattern search methods and random methods. Each class is described and a survey of recent literature is presented in this article.

INTRODUCTION

This paper is a summary of research and applications that have been done in simulation optimization over the past 15 years. Its purpose is to give a quick tour of the literature over this time span. Although no general method has been devised which works well on an arbitrary simulation, there has been a considerable amount of work done in this area. Over the years a number of other review papers have been written on this subject ([26],[27],[34],[49],[54],[66],[75]) as well as a comparison of optimization algorithms on stochastic functions [5].

The problem is to optimize a black box simulation expected response $E(F(X))$, over a region S, with respect to a set of p input factors, $X \varepsilon S \subset R^p$. The following diagram depicts a black box simulation:



where $X=(X_1,X_2,\ldots,X_p)$. Our goal is to optimize (maximize or minimize) the expected simulation output response $E(F(X))$ over all possible

feasible input factor settings X. The difficulty with this problem is that the output F(X) is a random variable equal to the sum of a constant and a random variable which represents the noise of the system. (i.e. $F(X)=E(F(X))+\varepsilon$ where ε is a random variable with $E(\varepsilon)=0$ and $V(\varepsilon)<+\infty$). Therefore, although we want to optimize E(F(X)), only F(X) can be observed. This puts the simulation optimization problem in the class of stochastic optimization problems, which are known to be difficult to solve. In practice, many simulations and real-life systems of interest have noisy objective functions. Therefore the solution to the simulation optimization problem must not only present an efficient and effective technique of optimizing a system's expected response, which of itself may be difficult, but it must also eliminate or sufficiently reduce the effect of noise such that the optimization procedure can locate the optimal factor setting of the expected simulation output response. It is this point which makes the problem difficult to solve.

The discussion presented in this paper focuses on discrete event simulation, although some papers on continuous event simulation are mentioned ([11],[66]). All methods involve searching for local optima. Global simulation optimization is not considered. However, some of the work in this area, such as confidence interval estimation of global optima ([12],[36],[51]), could possibly be adapted for use in stochastic systems. This may prove to be a rich area of future research. Here we classify the different approaches of searching for a local optima into 3 basic categories: path search methods, pattern search methods, and random methods. Path search methods have been the most widely studied approach for solving the simulation optimization problem, hence dominate the discussion and references in this paper. Pattern search methods and random methods are mentioned, with emphasis

on the types of methods which have been presented in the literature.

Path Search Methods

Path search methods involve estimating a direction to move from a current factor setting to an improved point in the feasible factor set. Typically only local information is used. Once a direction of movement is found, a distance to move in that direction is determined. The most common direction of movement is the gradient.

A widely studied approach in this class of techniques is Response Surface Methodology (RSM) (see [15], [39] or [56] for general reviews of RSM). RSM uses a set of sample observations around a particular observation, and attempts to fit a polynomial response surface to these points. There are many types of experimental designs specifically tailored to choosing these points [20]. One hopes that the shape of the response function is captured by the polynomial fit to the data. The polynomial is usually of first or second order.

Conventional RSM involves fitting a linear regression model around a current factor setting, and using this linear model to estimate an improving direction from the current point. Suppose k factor settings are needed for the experimental design used to fit the model. Let $Y=X\beta+\varepsilon$, where X is a kxp+1 matrix whose rows are the factor settings around the current point, β is a p+1x1 vector of unknown coefficients for the linear model, Y is a kx1 vector of response measurements corresponding to the k factor settings in X, and ε is the random noise of the system, which we will assume has mean zero. The ordinary least square estimate of β is $\hat{\beta}=(X^TX)^{-1}X^TY$. Therefore $\hat{Y}=X\hat{\beta}$ defines a

linear response surface, from which an optimal direction (e.g. the gradient) is easily obtained. One then moves in this direction a distance which optimizes the response function (line search). This is called Phase I of RSM. This procedure is repeated until the linear model stops providing a sufficiently good fit (e.g. the gradient of the linear response surface stops improving the response function value) [72]. Then Phase II is implemented. Here a quadratic model is fit to the response. The quadratic model is as follows:

$$Y_h = \beta_0 + \sum_{i=1}^{p} X_{hi}\beta_i + \sum_{\substack{i=1 \\ i\le j}}^{p}\sum_{j=1}^{p} X_{hi}X_{hj}\beta_{ij} + \varepsilon_h \qquad \text{where } h=1,2,\ldots k.$$

One then uses this fit to obtain a direction, the gradient, which will improve the response. The same procedure, as described in Phase I, is done in Phase II, such that an improved factor setting is obtained. When the length of the direction chosen to improve the factor setting gets sufficiently close to zero (e.g. $||d|| \le \delta$ for some $\delta \ge 0$, where d is the gradient of the response surface obtained in Phase II), the procedure stops.

An advantage of this method is that it is founded on a statistical theory that is easy to understand. The method is also easy to implement [72]. The major disadvantage of RSM is that a large number of simulation runs may be needed. Choosing the optimal set of runs to make involves deciding among many types of experimental designs for the data.

A significant amount of research has been done on RSM. Smith has used ideas involving first order polynomial fits (Phase I) and second order polynomial fits (Phase II), as desribed above ([47],[72],[73],[74],[75],[76],[77]). He has also done work in trying to

screen out unimportant factors so that the number of variables in the polynomial would be small [71]. Daugherty and Turnquist have done RSM, subject to constraints based on the cost to run the simulation [18]. They have also used spline functions [35] to create artificial experimental points [19]; this enabled them to obtain a reasonably good fit of the model using fewer simulation runs, hence make more efficient use of the data. Spline functions, rather than ordinary polynomial functions, were used, since spline functions give a smoother fit. The response surface was also fit over the whole feasible region, as opposed to a localized part of it.

Heller and Staats [38] used RSM, in conjunction with optimization techniques which are gradient based, to solve problems subject to costs and constraints. They called the direction obtained a direction of "Cheapest Ascent". They modified Zoutendijk's method of Feasible Directions [85] for this approach. Zoutendijk's method is a gradient-based optimization algorithm for constrained optimization problems. It involves creating a linear program, with the feasible region defined by first derivative linear approximations of all the binding constraints. One then solves the LP to obtain a direction of ascent, which is used to move to an improved factor setting. Therefore, suppose one wants to solve the program:

$$\text{Maximize } f(x)$$

$$\text{Subject to} \quad g_i(x) \geq 0 \quad i=\{1,2,\ldots,m\}$$

If the current factor setting is x^k, then we can obtain a feasible direction of ascent, d, by solving the linear program:

Maximize α

$$\text{Subject to} \quad \nabla f(x^k)d-\alpha\geq 0$$
$$\nabla g_i(x^k)d-\alpha\geq 0$$
$$|d_j|\leq 1 \text{ where } j\varepsilon\{i\varepsilon\{1,2,\ldots,m\}:g_i(x^k)=0\}$$
$$\text{and } d=(d_1,d_2,\ldots,d_m)^T$$

Zoutendijk's method is a relaxation method. Convergence may be slow, depending on the feasible region's steepness ([46],[84]).

Mihram [50] used RSM, with simplex experimental designs [13]. A simplex experimental design requires p+1 points where $X\varepsilon R^p$ (hence the name simplex). These p+1 points are all the same distance from the current point. Moreover, the distance between any two of these p+1 points is the same. Biles did multiple response surface fitting ([7],[9]) and multiple objective optimization ([6],[8],[10]). He also did some work in first and second order response surface fitting, with direct search along the direction of the gradient, and gradient projection methodology [46]. Gradient projection methods project the gradient of the objective function onto the boundary of the feasible region, hence yielding a feasible improving direction, which is not necessarily optimal, but is easy to obtain. Cooley and Houck [17] looked at the use of common and antithetic pseudorandom number streams as a means of reducing the variance of an estimated response surface. They demonstrated RSM, with the variance reduction modification, by optimizing an inventory system model. Safizadeh and Thornton [67] extended this work by considering alternative ways of using the pseudorandom number streams.

Eldridge [23] used a 5-phase procedure. His phases were: 1) screening (determining which variables are significant to the model), 2) grouping the variables into subspaces, which can be evaluated individually, 3) optimizing

over each of these subspaces, 4) fitting an approximating function over each subspace, 5) identifying the global solution over the entire space. A novel aspect of Eldridge's paper (the techniques he used were all standard), was the use of a random factorial design to better identify the optima of a multimodal response surface. This design is a combination of a complete factorial design and a random balance design [68]. This design was used, in Phase II, to break the region into a set of unimodal surfaces, hence simplifying the problem and enabling one to solve each unimodal surface individually. Using the theory from analysis of variance, a standard F-test, with one degree of freedom, was implemented to identify subspaces where a quadratic model could be fit. The F-test identified whether there were cubic or higher order effects present in a region, by measuring if such terms were significant. If the test showed there were such terms (i.e. a subspace could be adequately fit only by a cubic or higher order polynomial), it was broken down further until all subspaces could be fit to quadratic models.

Montgomery et al. looked at RSM, by considering different experimental designs suitable for fitting second order models to simulation responses [55]. In another paper they also considered screening methods which reduced the number of factors [52]. In addition, they looked at multivariate RSM using the Geoffrion-Dyer Interactive Vector Maximal Algorithm [53]. This algorithm is a multicriteria optimization procedure, which they adapted for use in multiple response surface optimization. It involves a concave utility function U, which gives a measure of the relative weighting between each of the response surfaces. Using gradient information from the utility function and the response surfaces, an approximation to the tradeoff weights between the response surfaces is determined. The algorithm then uses these weights to move toward an improved solution.

All of the above references involve modifying RSM to obtain better fits to the response function of the simulation so as to obtain gradient estimates that will most likely point in a response improving direction. They either use different experimental designs or screen out insignificant factors. Work has also been done on experimental designs to obtain better fits, as described, for example, in ([16],[52],[55]). Other reviews ([26],[27],[54]) contrast RSM with other methods. Smith did a comparison study [75] between seven different techniques. It was found that for a unimodal response function, RSM with 2^k factorial design and Random Search yielded the largest relative gain of the response function towards its optima. Single Factor or Univariate Search yielded the smallest relative gain.

Stochastic Quasi-Gradient methods use finite-difference Monte Carlo estimates of the gradients, applied to standard gradient-based optimization algorithms. Ermoliev [24] discussed these methods and their applications.

Stochastic Approximation (SA) ([43],[64]) is another path search method. Using the notation defined previously, let X_n be the factor setting for the n^{th} iteration of the algorithm, where the optimal factor setting is X_*. If $g(X_n)$ denotes an estimate of the gradient of $E(F(.))$ at X_n, then SA chooses the next factor setting $X_{n+1}=X_n+\lambda_n g_n$ where λ_n is a scalar parameter which satisfies the convergence criteria given by Dvoretsky [22]. Dvoretsky generalized the criteria given by Kiefer and Wolfowitz, and Robbins and Monro. The Kiefer-Wolfowitz (K-W) algorithm defines g to be a 2-sided gradient estimate of F. We assume the variance of F(X) is finite for all $X \varepsilon S$, and E(F(X)) is unimodal. The convergence criteria for the K-W algorithm is: Given two sequences, α_n and γ_n, satisfying the convergence criteria $(\Sigma(\alpha_n)=+\infty,\ \mathrm{Lim}(\gamma_n)=0,\ \Sigma((\alpha_n/\gamma_n)^2)<+\infty)$, let $X_{n+1}=X_n+(\alpha_n/2\gamma_n)(F(X_n+\gamma_n)-F(X_n-\gamma_n))$, then X_n converges in mean square and

with probability 1 to a local maxima of F. Even if the convergence criteria are satisfied, experience has shown that convergence may be very slow and require a large number of simulation runs (a general summary of such results, including convergence rates, can be found in [1],[81]).

Azadivar used one-sided [2] and two-sided [3] gradient estimates in applying SA (in [3], only the sign of the gradient estimate was used). Kushner and Gavin [44] developed a generalized SA method by searching for, as opposed to using predetermined, step lengths to move in the direction of the gradient estimate. Pflug [63] used two-sided finite difference gradient approximations with SA; he called his approach the "Stochastic Quasigradient method". Ruppert et al. [65] used SA applied to a Monte Carlo simulation of a fish harvesting model. Glynn and Sanders [33] proposed a gradient estimation technique, using likelihood ratios and Monte Carlo estimation procedures. Fox [30] estimated gradients for transient markov chains. He also gave a complexity analysis of the estimation procedure [29] and compared it to the approaches presented in [30].

Much recent work has been done to obtain an estimate of the gradient, by Ho, Suri, Zazanis et al. ([40],[41],[79],[80],[82]). They have developed a technique called Perturbation Analysis (PA). PA uses a basic idea from calculus, the chain rule, to obtain a gradient estimate using only one run of a simulation, for any $p \geq 1$. If w is the simulation response (e.g. the average waiting time in a queue, for an M/M/1 queue), s is the measure of interest (e.g. the service time of a customer), and X is the factor of interest (e.g. the service rate), then we have $(\partial w/\partial X)=(\partial w/\partial s)(\partial s/\partial X)$. They assume that $\partial s/\partial X$ is known and can be evaluated for a given s and X. They make the basic assumption that the sequence of events remains the same for both X and $X+\Delta X$, where ΔX is an infinitesimal perturbation. If this assumption holds, the

effect of ΔX can be measured as it propogates through the simulation. This effect is kept track of in accumulators (e.g. $(\partial s/\partial X)$ evaluated at s and X, is added to the accumulator whenever a customer is served in the current busy period), one for each of the indices of the factor X. If the effect of the perturbation dies out in the simulation run, (e.g. a busy period ends), the accumulator values are recorded and reset to zero. The total sum of all the effects of a perturbation will then yield the gradient of a response of the simulation. It is important to note that one must be able to calculate the derivative of the measure of interest, so that the accumulator can correctly sum the effect of the perturbation on the given response. Although this method has many desirable properties, such as its simplicity and that it takes very little work to implement the algorithm within a simulation (one need add only a few lines to a computer program for each factor; one to update the accumulator so the perturbations effects are being measured, and one to reset the accumulator at points where the perturbation's effects have died out), it has so far only been successfully applied to certain queueing models.

An important analysis of PA was done by Heidelberger [37], where necessary and sufficient conditions for PA to produce strongly consistent estimates were given for regenerative processes. He also identified and quantified a major weakness of PA: the assumption that the order of events does not change when a factor is perturbed by an infinitesimal amount. This assumption is typically not true. Heidelberger's work explains why PA does not give good estimates in certain simulation models.

As described above, gradient estimation has been the major focus of path search methods. In general, estimating gradient directions can be both difficult and expensive, using the techniques presented. This is why very

little work has gone into estimating hessians or developing the quasi-newton method to solve the simulation optimization problem. Zazanis and Suri [82] used PA to estimate the hessian of a G/G/1 queue from a single sample path. Such hessian estimates could be used in a quasi-newton algorithm. The difficulty with using this method in a stochastic environment is that a product of two matrix estimates (the inverse of the hessian and the gradient) must be taken. The inverse of the hessian estimate may be very poor, even if the estimate of the hessian is reasonably good, especially when the condition number of the hessian matrix is large (i.e. the ratio of the largest and the smallest singular values of the hessian matrix is large). If an efficient technique of reducing the noise effect on the hessian matrix estimate, and its inverse, is developed, more work will be done with quasi-newton methods for solving the simulation optimization problem.

Pattern Search Methods

Pattern search methods are those which require no gradient estimates, nor randomization procedures, but rather, use some characteristic or pattern of the observations, to obtain an improved point

The most common method in this class is the Hooke and Jeeves method (HJ) [42]. HJ is based on the idea that if a direction has produced a favorable change in the optimal value, then one should continue to move in this direction. It uses the pattern from which previous improving changes have been made, to obtain better factor settings, and eventually, the optimal setting. One initially selects a set of incremental values for each factor. Starting at an initial point, one checks, univariately, if positive or negative incremental value changes in the factor settings produce improved response values. This yields a new setting for each factor. One then moves directly from the initial point in the direction towards, and through, this new point. This procedure is continued until optimal changes cannot be made with the given incremental values. Then the incremental values are decreased and the procedure is repeated from the beginning. When the incremental values reach a prespecified tolerance, the procedure is terminated and the current factor settings are reported as the solution. Nozari and Morris ([58],[59]) have applied this method in conjunction with the Dudewicz-Dalal Selection and Ranking method [21]. The Dudewicz-Dalal method is a two-staged sampling procedure, used to identify the best factor setting between a set of k factor settings. Let $\mu_{[k-1]}$ be the mean of the second best system, $\mu_{[k]}$ is the mean of the best system and δ is a prespecified indifference measure between the best and second best systems. The procedure results in a probability of correctly selecting the optimum that is greater than P^*, provided

$\mu_{[k-1]}-\mu_{[k]} \geq \delta > 0$ (if we are searching for a minima), where P^* is a prespecified probability. The first stage involves running the simulation for each of the factor settings to be compared. The observations must be independent to apply the Dudewicz-Dalal method. Therefore, the simulation runs must be long enough so that the batch sizes are sufficiently large to ensure independence. The second stage uses the Dudewicz-Dalal method, together with the information obtained in the first stage, to determine a search direction which is used in the HJ algorithm. Therefore, the HJ algorithm is modified by having it obtain its search direction from the two stage procedure described above. Pegden and Gately also applied the Hooke and Jeeves method, to problems written in SLAM [61] and GASP IV [62].

Another pattern search method which has been studied is the Simplex method ([57],[60],[78]). The Simplex method starts with a set of p+1 factor settings in R^p (hence the name simplex). Then, by comparing their response values, it eliminates the factor setting with the smallest function value (in the case of maximization) and replaces it by a new factor setting, determined by the centroid of the p remaining factor settings and the eliminated factor setting. The resulting simplex either grows or shrinks, depending on the value of the new factor setting. The procedure is repeated until no more improvements can be made by eliminating the smallest valued point and the resulting final simplex is small. Meier [48] looked at applying the simplex method to simulation studies. A Univariate Search method, applied to simulation programs written in GPSS, by Lefkowitz and Schriber [45] has also been looked into. Recently, Schruben [70] applied the Frequency Domain method [69] of simulation output analysis to optimize a simulation output response. Schruben showed that a simulation output response could be optimized by using information obtained from its power spectrum. More work is being done, using

the frequency domain approach, including the estimation of gradients and higher order sensitivities.

Random Methods

Random methods are those which use a random approach to select factor settings, with the hope of obtaining an improving, and eventually, optimal response. The major problem with these methods is that they are slow to converge (if they converge at all) to an optima. Previous information is typically not used at each iteration. Therefore, the method requires a large number of simulation runs. The reliability of the solution can be expressed only in terms of probabilities. These probability statements are in turn based on rather restrictive assumptions.

Very little has been done with this method in trying to solve the simulation optimization problem. Garcia-Diaz et al. [32] used the Out-of-Kilter Algorithm [28] with Monte-Carlo sampling to optimize a production transportation problem which was transformed into a network flow problem. The Out-of-Kilter Algorithm determines the minimum cost flow of a source to sink network. It does this by sequentially obtaining solutions to the primal and dual linear programs for the network. Using the complementary slackness property, the algorithm terminates with the optimal solution. The Monte-Carlo sampling procedure is what made their approach random. Smith looked at random search as a possible method ([72],[75],[77]), and Farrell mentioned random methods in his review and comparison papers ([26],[27]). Fox [31] presented ideas using quasirandom numbers as a possible alternative to pseudo-random number streams. Quasirandom numbers try to minimize the discrepancy of a

sample, which is a measure of how good the spacing is. More formally, given a set of points $x^1,x^2,\ldots,x^N \varepsilon I^s$ and a subset $G \subset I^s$, define the counting function $S_N(G)$ as the number of points $x^i \varepsilon G$. For each $x=(x_1,x_2,\ldots,x_s)\varepsilon I^s$, let G_x be the rectangular s-dimensional region $G_x=[0,x_1)x[0,x_2)x\ldots x[0,x_s)$ with volume $x_1x_2\ldots x_s$. Then the discrepancy of the points $x^1,x^2,\ldots,x^N$ is

$$D^*(x^1,x^2,\ldots,x^N)= \operatorname{Sup}_{x\varepsilon I^s} |S_N(G_x)-Nx_1x_2\ldots x_s|$$

It is believed that the smallest discrepancy possible is $O(\text{Log}_p N)$, where N is the number of points generated, and p is the dimension of each point. Bratley and Fox [14] compared different quasirandom generators which give number streams with the smallest known discrepancy. This idea holds promise in solving the simulation optimization problem, in conjunction with Random methods, although at present it has not been fully developed or tested in a stochastic optimization environment.

Glynn [34] gives a survey of Monte Carlo algorithms for Stochastic Optimization, with particular reference to the analysis and comparison of the convergence rates between the different approaches. Zheng [83] outlines an integral approach to solve the global optimization problem. Most optimization algorithms try to find an improving direction. Zheng tries to minimize the volume of space between the function he is trying to maximize and a plane whose level is set by the current largest value the function is known to take. The implementation of his algorithm uses the Monte Carlo technique to estimate this volume and the largest value of the function. Evtushenko [25] gave an algorithm for lipshitz continuous functions (i.e. there exists a $k\geq 0$, the lipshitz constant, such that $||f(x)-f(y)||\leq k||x-y||$ for all x,y over the set for which f is defined), which introduced the integral approach to optimization. His algorithm is space-covering based (i.e. one tries to

eliminate regions where the optima cannot occur. When all regions where an improved point might exist have been eliminated, the algorithm stops. Regions are eliminated based on the number of points one has evaluated the function at and the best value of the function one has obtained so far). It has not been successfully implemented since obtaining the lipshitz constant can be quite costly. Zheng's modifications get around this problem, at the expense of efficiency and speed of convergence.

CONCLUSIONS

The simulation response optimization problem has been considered and classified into three basic categories: path search methods, pattern search methods and random methods. Although there have been a significant amount of research in the area, no general approach has been developed into an efficient and practical algorithm. Barton [4] gives testing strategies for comparing simulation optimization techniques, with emphasis on the choice of test functions and the random variability of this function. The literature clearly points out that particular simulations, due to their structure and form, seem to yield more consistent results for certain methods. For example, simulations which are designed using a queueing model may yield the best result using Perturbation Analysis, while a simulation which appears to have interaction between the factors or a more clearly defined polynomial relation between the factor settings and the response may yield the best results using RSM. Efficient allocation of simulation runs is the main consideration in choosing a method; such runs can be quite costly. Each method has their advantage and disadvantage. Most methods consider very basic local

optimization ideas (e.g. gradient estimates). In a stochastic environment, it is difficult to use higher level optimization techniques. Until new ideas are brought into this area, these methods will continue to be applied. However, with a growing interest in this field, it seems likely that there will be new approaches discovered and developed which will take advantage of more sophisticated mathematical and optimization ideas.

BIBLIOGRAPHY

[1] Albert, A.E., Gardner, L.A., Jr., 'Stochastic Approximation and Nonlinear Regression', M.I.T. Press, Cambridge, Massachusetts, Research Monograph No. 42, 1967.

[2] Azadivar, F., "A Simulation Optimization Approach to Optimum Storage and Retrieval Policies in an Automated Warehousing System", Proceedings of the 1984 Winter Simulation Conference, P.207-214.

[3] Azadivar, F., Talavage, J.J., "Optimization of Stochastic Simulation Models ", Mathematics and Computers in Simulation", Vol. 22, No. 3, North Holland Publishing Co., September 1980, P.231-241.

[4] Barton, R.,"Testing Strategies for Simulation Optimization", Technical Report, S.O.R.I.E., Cornell University, 1987.

[5] Barton, R., "Minimization Algorithms for Functions with Random Noise", American Journal of Mathematical and Management Sciences, Vol. 4, No. 1&2, P.109-138, 1984.

[6] Biles, W.E., Swain, J.J., "Mathematical Programming and the Optimization of Computer Simulations", Math Programming Studies, Vol. 11, P.189-207, 1979.

[7] Biles, W.E., "Optimization of Multiple Response Simulation Models", South Bend, Indiana: University of Notre Dame, 1978.

[8] Biles, W.E., "Optimization of Multiple-Objective Computer Simulations: a Non-Linear Goal Programming Approach", Proceedings of the 1977 AIDS Conference, Chicago, 1977.

[9] Biles, W.E., "Strategies for Optimization of Multiple Response Simulation Models", Proceedings of the 1977 Winter Simulation Conference, P.134-142.

[10] Biles, W.E., "A Gradient Regression Search Procedure for Simulation Experimentation" Proceedings of the 1974 Winter Simulation Conference, P.491-497.

[11] Birta, L.G., "A Parameter Optimization Module for CSSL-Based Simulation Software", Simulation, Vol. 28, No. 4, P.113-123, 1977.

[12] Boender, C.G.E., Rinnooy Kan, A.H.G., Stougie, L., Timmer, G.T., "A Stochastic Method for Global Optimization", Mathematical Programming, Vol. 22, P.125-140, 1982

[13] Box, G.E.P., Behnken, D.W., "Simplex-Sum Designs: A Class of Second Order Rotatable Designs Derivable from those of First Order", Annals of Mathematical Statistics, Vol. 31, P.838-864, 1960.

[14] Bratley, P., Fox, B.L., "Implementing Sobol's Quasirandom Sequence Generator", Technical Report, University of Montreal, Montreal, Quebec, Canada, September, 1986.

[15] Brightman, H.J., "Optimization Through Experimentation: Applying Response Surface Methods", Decision Sciences, Vol. 9, P.481-495, 1978.

[16] Burdick, D.S., Naylor, T.H., "Design of Computer Simulation Experiments for Industrial Systems", CACM, Vol. 9, No. 5, P.329-338, 1966.

[17] Cooley, B.J., Houck, E.C., "A Variance-Reduction Strategy for RSM Simulation Studies", Decision Sciences, Vol. 13, P.303-321, 1982.

[18] Daugherty, A.F., Turnquist, M.A., "Budget Constrained Optimization of Simulation Models Via Estimation of Their Response Surfaces", Operations Research, Vol. 29, No. 3, P.485-500, 1981.

[19] Daugherty, A.F., Turnquist, M.A., "Simulation Optimization Using Response Surfaces based on Spline Approximations", Proceedings of the 1978 Winter Simulation Conference, Vol. 1, P.182-193.

[20] Davies, O.L. (Editor), 'The Design and Analysis of Industrial Experiments', Longman Group Limited, London and New York, 1978.

[21] Dudewicz, E.J., Dalal, S.R., "Allocation of Measurements in Ranking and Selection with Unequal Variances", Sankhya, B37, P.28-78, 1975.

[22] Dvoretsky, A., "On Stochastic Approximation", Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, P.39-59, 1956.

[23] Eldridge, D.L., "Experimental Optimization of Statistical Simulation", Proceedings of the 1974 Winter Simulation Conference, P.503-510.

[24] Ermoliev, Y., "Stochastic Quasi-Gradient Methods and their Application to System Optimization", Stochastics, Vol. 9, P. 1-36.

[25] Evtushenko, Y.P., "Numerical Methods for Finding Global Extrema of a Non-Uniform Mesh", USSR Computing Machines and Mathematical Physics, Vol. 11, P.1390-1403, 1971.

[26] Farrell, W., "Literature Review And Bibliography of Simulation Optimization", Proceedings of the 1977 Winter Simulation Conference, P.117-124.

[27] Farrell, W. et al., "Optimization Techniques for Computerized Simulation Models", CACI Inc., California, 1975.

[28] Ford, L.R., Jr., Fulkerson, D.R., 'Flows in Networks', Princeton Press, 1962.

[29] Fox, B.L., "Complexity of Gradient Estimation for Transient Markov Chains", Technical Report #753, S.O.R.I.E., Cornell University, July, 1987.

[30] Fox, B.L., "Gradient Computation for Transient Markov Chains", Technical Report, S.O.R.I.E., Cornell University, 1987.

[31] Fox, B.L., "Implementation and Relative Efficiency of Quasirandom Sequence Generators", Working Paper, University of Montreal, August, 1984.

[32] Garcia-Diaz, A., Hogg, G.L., Phillips, D.T., Worrell, E.J., "Combined Simulation and Network Analysis of a Production Distribution System", Simulation, Vol. 40, No. 2, P.59-66, 1983.

[33] Glynn, P.W., Sanders, J.L., "Monte Carlo Optimization of Stochastic Systems: Two New Approaches", Proceedings ASME Computers in Engineering Conference, 1986, To Appear.

[34] Glynn, P.W., "Optimization of Stochastic Systems", Proceedings of the 1986 Winter Simulation Conference, Washington, D.C..

[35] Greville, T.N.E., 'Theory and Application of Spline Functions', Academic Press, New York, 1969.

[36] Haan, L. de, "Estimation of the Minimum of a Function Using Order Statistics", Journal of the American Statistical Association, Vol.76, No.373, P.467-469, 1981.

[37] Heidelberger, P., "Limitations of Infinitesimal Perturbation Analysis", IBM Research Report RC 11891, IBM Watson Research Labs, Yorktown Heights, N.Y., 10598, 1986.

[38] Heller, N.B., Staats, G.E., "Response Surface Optimization when Experimental Factors are Subject to Costs and Constraints", Technometrics, Vol. 15, No. 1, P.113-123, 1973.

[39] Hill, W.J., Hunter, W.G., "A Review of Response Surface Methodology: A Literature Survey", Technometrics, Vol. 8, No. 4, P.571-590, 1966.

[40] Ho., Y.C., "Perturbation Analysis Methods for Discrete Event Dynamical Systems and Simulations", Proceedings of the 1984 Winter Simulation Conference, P.171-173.

[41] Ho, Y.C., Cao, X., "Perturbation Analysis And Optimization of Queueing Networks", Journal of Optimization Theory and Application, Vol. 40, No. 4, P.559-582, 1983.

[42] Hooke, R., Jeeves, T.A., "A Direct Search Solution of Numerical and Statistical Problems", JACM, Vol. 8, No. 2, P.212-229, 1961.

[43] Kiefer, J., Wolfowitz, J., "Stochastic Approximation of the Maximum of a Regression Function", Annals of Mathematical Statistics, Vol. 23, P.462-466, 1952.

[44] Kushner, H.J., Gavin, T., "A Versatile Method for the Monte Carlo Optimization of Stochastic Systems", Int. Journal Of Control, Vol. 18, No. 5, P.963-975, 1982.

[45] Lefkowitz, R.M., Schriber, T.J., "Use of an External Optimization Algorithm with a GPSS Model", Proceedings of the 1971 Winter Simulation Conference, P.162-171.

[46] Luenberger, D.G., 'Introduction to Linear and Nonlinear Programming', Addison-Wesley, Reading, Massachusetts, 1973.

[47] Mandelbaum, J., Jorgenson, F.L., Smith, D.E., Storck, C.E., "Man-Machine Role Identification in Seeking Improved Solutions to Large-Scale Computer Simulation Problems", Proceedings of 31st Military Operations Research Symposium, 1973.

[48] Meier, R.C., "The Application of Optimum-Seeking Techniques to Simulation Studies: A Preliminary Evaluation", Journal of Financial and Quantitative Analysis, Vol. 2, P.31-51, 1967.

[49] Meketon, M., "Optimization Methods in Simulation: A Survey of Recent Results", Presented at the 1983 Winter Simulation Conference.

[50] Mihram, G.A., "An Efficient Procedure for Locating the Optimal Simulation Response", Proceedings of the 4th Conference on the Application of Simulation, P.154-161, 1970.

[51] Monroe, H.M., Sielken, R.L., "Confidence Limits for Global Optima Based on Heuristic Solutions to Difficult Optimization Problems", American Journal of Mathematical and Management Sciences, Vol. 4, No.1&2, P.139-168, 1984.

[52] Montgomery, D.C., "Methods for Factor Screening in Computer Simulation Experiments", Technical Report, Georgia Institute of Technology, 1979.

[53] Montgomery, D.C., Bettencourt, V.M., Jr., "Multiple Response Surface Methods in Computer Simulation", Simulation, Vol. 29, No. 4, P.113-121, 1977.

[54] Montgomery, D.C., Bettencourt, V.M., Jr., "A Review of Multiple Response Surface Methods in Computer Simulation", 1976 ORSA\TIMS Meeting.

[55] Montgomery, D.C., Evans, D.M., Jr., "Second Order Response Surface Designs in Computer Simulation", Simulation, Vol. 25, No. 6, P.169-178, 1975.

[56] Myers, R.H., 'Response Surface Methodology', Allyn-Bacon, Boston, 1971.

[57] Nelder, J.A., Mead, R., "A Simplex Method for Function Minimization", Computer Journal, Vol. 7, P.308-313, 1965.

[58] Nozari, A., Morris, J.S., "Application of an Optimization Procedure to Steady-State Simulation", Proceedings of the 1984 Winter Simulation Conference, P.217-219.

[59] Nozari, A., Morris, J.S., "An Optimization Procedure for Simulation Models", Working Paper, College of Business and Economics, University of Idaho, Moscow, Idaho, 1983.

[60] Ollson, D.M., Nelson, L.S., "The Nelder-Mead Simplex Procedure for Function Minimization", Technometrics, Vol. 17, No. 1, P.45-51, 1975.

[61] Pegden, C.D., Gately, M.P., "A Decision-Optimization Module for SLAM", Simulation, Vol. 34, No. 1, P.18-25, 1980.

[62] Pegden, C.D., Gately, M.P., "Decision Optimization For GASP IV Simulation Models", Proceedings of the 1977 Winter Simulation Conference, P.125-133.

[63] Pflug, G. Ch., "Optimizing Simulated Systems", Simuletters, Vol. 15, No. 4, P.6-9, 1984.

[64] Robbins, H., Monro, S., "A Stochastic Approximation Method", Annals of Mathematical Statistics, Vol. 22, P.400-407, 1951.

[65] Ruppert, D., Reish, R.L., Deriso, R.B., Carroll, R.J., "Optimization Using Stochastic Approximation and Monte Carlo Simulation (with Application to Harvesting of Atlantic Menhaden)", Biometrics, Vol. 40, P.535-545, 1984.

[66] Rustagi, J.S., "Optimizing Methods in Simulation", Technical Report, Department of Statistics, Ohio State University, July 1981.

[67] Safizadeh, M.H., Thornton, B.M., "An Alternative Variance-Reduction Strategy for RSM Simulation Studies", Decision Sciences, Vol. 13, P.322-330, 1982.

[68] Satterthwaite, F.E., "Random Balance Experimentation", Technometrics, Vol. 1, No. 2, 1959.

[69] Schruben, L.W., Cogliano, V., " An Experimental Procedure for Simulation Response Surface Model Identification ", CACM, to appear, 1987.

[70] Schruben, L.W., " Simulation Optimization Using Frequency Domain Methods ", Proceedings of the 1986 Winter Simulation Conference, Washington, D.C..

[71] Smith, D.E., Mauro, C.A., "Factor Screening in Computer Simulation", Simulation, Vol. 38, No. 2, P.49-54, 1982.

[72] Smith, D.E., "Optimization of a Computer Simulation Response", Technical Report 106-3, Presented at ORSA\TIMS Joint National Meeting, November 1976.

[73] Smith, D.E., "Automatic Optimum-Seeking Program for Digital Simulation", Simulation, Vol. 27, No. 1, P.27-32, 1976.

[74] Smith, D.E., "Requirements of an Optimizer for Computer Simulation", Naval Research Logistics Quarterly, Vol. 20, No. 1, P.161-179, 1973.

[75] Smith, D.E., "An Empirical Investigation of Optimum Seeking in the Computer Simulation", Operations Research, Vol. 21, P. 475-497, 1973.

[76] Smith, D.E., Storck, C.E., "Research on an Optimizer Computer Program for use in Simulation Studies", Technical Report, AD 766-089/7, 1973.

[77] Smith, D.E., "Studies with a Prototype Optimizer for use in Computer Simulation", Proceedings of the 17th Conference on the Design of Experiments in Army Research, Development and Testing, 1972.

[78] Spendley, W., Hext, G.R., Himsworth, F.R., "Sequential Application of Simplex Designs in Optimisation and Evolutionary Operations", Technometrics, Vol. 4, No. 4, P.441-461, 1962.

[79] Suri, R., "Infinitesimal Perturbation Analysis for General Discrete Events System", Presented at IEEE decision and Control Conference, 1983.

[80] Suri, R., Zazanis, M.A., "Perturbation Analysis Gives Strongly Consistent Estimates for the M/G/1 Queue", Proceedings of the 1984 IEEE Decision and Control Conference.

[81] Wilde, D.J., 'Optimum Seeking Methods', Prentice-Hall, Englewood Cliffs, N.J., 1964.

[82] Zazanis, M.A., Suri, R., "Estimating Second Derivatives of Performance Measures for G/G/1 Queues from a Single Sample Path", Working Paper, Harvard University, Division of Applied Sciences, 1985.

[83] Zheng, Q., "Theory and Methods for Global Optimization --- An Integral Approach", Presented at Optimization Days, Montreal, Canada, 1986.

[84] Zlobec, S., Mathematical Programming Lecture Notes, Personal Communications, McGill University, Department of Mathematics and Statistics, 1978.

[85] Zoutendijk, G., 'Methods of Feasible Direction', Amsterdam, Elsevier, 1960.