Proceedings of the 1987 Winter Simulation Conference
A. Thesen, H. Grant, W. David Kelton (eds.)

FUTURE DIRECTIONS IN RESPONSE SURFACE METHODOLOGY FOR SIMULATION

James R. Wilson
School of Industrial Engineering
Purdue University
West Lafayette, IN 47907, U.S.A.

ABSTRACT

As a tool for gradient estimation and sensitivity analysis in discrete simulation, response surface methodology possesses noteworthy advantages in comparison to some of the more recently developed techniques. This paper surveys future directions for research, development, and application of response surface methodology in discrete simulation.

1. INTRODUCTION

In the past few years there have been significant advances in the areas of gradient estimation and sensitivity analysis for discrete simulation. Beyond the traditional approach of response surface methodology (RSM) (Box and Wilson 1951), some radically different techniques have emerged: (a) frequency-domain methods (Cogliano 1982, Schruben 1986, Schruben and Cogliano 1987); (b) perturbation analysis (Ho, Eyler, and Chien 1979; Suri and Zazanis 1987); and (c) likelihood-ratio methods (Glynn 1986, Reiman and Weiss 1986). Along with these dramatic new developments, there has been a resurgence of interest in RSM itself. In this brief survey of future directions for research, development, and application of RSM in discrete simulation, we begin by examining some of the reasons for the renewed interest in RSM.

In comparison to the new techniques for gradient estimation and sensitivity analysis, the mathematical and statistical foundations of RSM are not only more transparent but also more completely developed. Frequency-domain methods depend heavily on the theory of Fourier transforms and generalized functions. Moreover, some of the basic theory supporting frequency-domain methods is incomplete—for example, the problem of controlling gain has not yet been rigorously resolved. As it was originally formulated, infinitesimal perturbation analysis was severely limited in the class of queueing systems to which it could be validly applied (Heidelberger 1987). Although recent extensions of perturbation analysis appear to have eliminated these limitations (Ho, Suri, Cao, Cassandras, and Zazanis 1987), there is still no clear-cut characterization of the class of discrete systems to which perturbation analysis can be validly applied. The development of such a characterization seems to be a formidable theoretical problem. Likelihood-ratio methods depend heavily on measure-theoretic probability, and currently they have only been adapted to the regenerative method of simulation analysis. Moreover, likelihood-ratio methods are presently limited to estimation of parameter sensitivities for Markov chains and Poisson arrival processes. In contrast to the more advanced theoretical machinery required by the newer methods for gradient estimation and sensitivity analysis, RSM is simply based on well-known principles of regression analysis and analysis of variance.

The renewed interest in RSM is also partly due to the comparative ease with which RSM can be applied in practice. Within each run, frequency-domain methods require careful indexing of simulation-generated observations together with sinusoidal variation of selected input variables according to this "time" index. Such variation is frequently difficult to arrange appropriately, especially for discrete input variables. Another practical difficulty in implementing frequency-domain methods for a large number of input factors is the assignment of driving frequencies to all of these input factors so as to avoid confounding effects of interest. With regard to perturbation analysis and likelihood-ratio methods, it is unclear how either of these techniques can be generally implemented in large-scale discrete simulation experiments. In contrast, RSM can be applied to any discrete simulation since it is not an *intrusive* procedure—that is, RSM does not require manipulation or restructuring of the internal operation of the simulation. Thus from both a theoretical and practical standpoint, RSM possesses distinct advantages for gradient estimation and sensitivity analysis in discrete simulation experiments.

2. FUTURE DIRECTIONS FOR RSM

The main disadvantage of RSM is the cost of making all of the runs required by the usual experimental designs. This cost includes the experimenter's time, which is often the factor of overriding practical importance. Of course the cost of RSM also includes the total computing time for the simulation runs; and this cost can be prohibitive when the response of interest has a large variance so that excessive run lengths are required. The cost problem for simulation-oriented RSM naturally motivates the search for effective variance reduction techniques that can be incorporated into the overall experimental design.

2.1. Induced-Correlation Strategies for RSM

Schruben and Margolin (1978) developed the first widely accepted strategy for inducing correlation among responses in a simulation experiment in order to improve the efficiency of a metamodel estimator. Refinements and extensions of the Schruben-Margolin strategy should constitute a major portion of future research on RSM.

In the development of more efficient induced-correlation strategies for RSM, one of the most important issues is the extension of the Schruben-Margolin strategy to accommodate experimental designs that are not two-level factorial (or fractional factorial) designs. Specifically, we need an extended "assignment rule" for dedicating random number streams to blocks in three-level factorial designs, composite designs, and factor-screening designs. Note that although Cooley and Houck (1982, 1983) advocated using the original Schruben-Margolin assignment rule in central composite designs, this approach will yield *identical* responses for all runs performed at the center point within a block. Now in this situation there is no residual error that is independent of the block effect; thus the basic assumptions underlying the Schruben-Margolin strategy are violated. Clearly we need new induced-correlation strategies that are adapted to the composite designs used so frequently in RSM.

It is also necessary to extend the Schruben-Margolin strategy to handle more general "repeated-measures" type covariance structures for the response taken across the points of the experimental design. Although we frequently encounter unequal response variances across the runs of a large-scale simulation experiment, the more serious problem seems to be marked disparities among the correlations induced between runs by the Schruben-Margolin strategy. Even after a variance-stabilizing transformation has been successfully applied to the simulation-generated responses, greatly differing correlations are frequently observed within blocks (Tew and Wilson 1986). We need analogues of the basic results of Schruben and Margolin (1978) for more flexible forms of the covariance matrix of the response. For example, it would be desirable to develop a variant of the Schruben-Margolin strategy for designs having a first-order autoregressive covariance structure within each block; see Chapter 8 of Johnston (1972). Such a development should include optimal metamodel point estimators for a large class of experimental designs with this covariance structure and for a large class of design optimality criteria. Moreover, an appropriate assignment rule should be developed to induce such a covariance structure in a large class of discrete simulation models.

Along with the development of new induced-correlation strategies for simulation, we require new statistical analysis techniques that exploit the more general repeated-measures covariance structures. Given such an induced covariance matrix, we can construct the required analysis technique if we can identify the multiplicity of each distinct eigenvalue of the covariance matrix and if we can find at least one eigenvector of that matrix which does not depend on the unknown parameters in the matrix. For example, in the standard formulation of the Schruben-Margolin strategy for an m-point experimental design with response variance σ^2 and with induced positive correlation ρ_1 within blocks and induced negative correlation ρ_2 between blocks, we have the following properties for the response covariance matrix: (a) The eigenvector $\mathbf{1}_m$ (consisting of m 1's) corresponds to the eigenvalue $\sigma^2[1+(\frac{1}{2}m-1)\rho_1+\frac{1}{2}m\rho_2]$ with multiplicity 1; (b) The eigenvector $[\mathbf{1}'_{m/2} \;\; -\mathbf{1}'_{m/2}]'$ corresponds to the eigenvalue $\sigma^2[1+(\frac{1}{2}m-1)\rho_1-\frac{1}{2}m\rho_2]$ with multiplicity 1; and (c) The only remaining distinct eigenvalue is $\sigma^2(1-\rho_1)$ with multiplicity $m-2$. This is the essential information from which Nozari, Arnold, and Pegden (1987) developed optimal hypothesis-testing procedures and confidence-region estimation procedures for the postulated metamodel under the Schruben-Margolin strategy. Tew and Wilson (1985, 1986) also used this information to construct their modified likelihood-ratio test for validating the assumed form of the response covariance matrix. Similar results can be developed for more general covariance structures.

2.2. Control-Variates Strategies for RSM

The Monte Carlo method of control variates has the potential for extremely effective use in simulation-based RSM studies, and the computational overhead imposed by this technique is usually negligible relative to the cost of performing the basic simulation experiment. Nozari, Arnold, and Pegden (1984) developed a method for applying control variates to the estimation of a univariate simulation metamodel. It would be desirable to extend this method to accommodate a covariance structure between the response and the controls that is not homogeneous across the points of the experimental design. Where it is possible to do so, we should also exploit controls with a known covariance matrix to obtain more reliable controlled metamodel estimators; see Bauer, Venkatraman, and Wilson (1987).

2.3. Combined Correlation Strategies for RSM

Among future directions for research and development for simulation-oriented RSM, one of the most promising seems to be the integration of induced-correlation methods with the method of control variates. Tew and Wilson (1987) have taken some preliminary steps in this direction, and their results suggest that all of the correlation-based variance reduction techniques can be combined effectively in the analysis of simulation metamodels. However, much fundamental work remains to be done before such combined correlation strategies can be routinely applied to large-scale simulation experiments. In particular, all of the considerations discussed in Subsections 2.1 and 2.2 apply here: (a) We need RSM strategies that can handle more general repeated-measures covariance structures induced between the responses and/or the controls across the points of the design; (b) We need an extended assignment rule for dedicating random number streams to block effects and/or control variates so as to ensure the validity of the postulated covariance structure and to maximize the efficiency gain; (c) We need a suitably generalized procedure for validating all of the assumptions underlying such a combined strategy; and (d) We need corresponding hypothesis-testing and confidence-region estimation techniques for the postulated metamodel. Where possible, we should also have the means for exploiting control variates with a known covariance matrix. All of these issues require careful mathematical and statistical analysis as well as a thorough consideration of the logical structure and operation of large-scale simulation models.

2.4. Extension to Multiresponse Metamodels

All of the foregoing discussion was limited to univariate metamodels. Most real-world simulation models generate several responses of interest, and the analysis of multiresponse metamodels should be another key direction for future research. Currently there are no induced-correlation strategies for multiresponse simulation. Although Porta Nova and Wilson (1986) have developed a control-variates strategy for analysis of multiresponse metamodels, their methods should be generalized to accommodate nonhomogeneous covariances between the responses and the controls across the points of the experimental design. Currently there are no combined correlation strategies for multiresponse simulation experiments.

3. CONCLUSIONS

Response surface methodology offers excellent opportunities for application as well as methodological development in the areas of gradient estimation and sensitivity analysis in discrete simulation. Much research is currently in progress on induced-correlation strategies, control-variates strategies, and combined correlation strategies for RSM. Some work is also underway on multiresponse metamodels. Nevertheless, much remains to be done. There is good reason to believe that these efforts will yield results of substantial practical benefit and that RSM will continue to play a prominent role in the design and analysis of discrete simulation experiments.

REFERENCES

Bauer, K. W., Venkatraman, S., and Wilson, J. R. (1987). Estimation procedures based on control variates with known covariance matrix. In: *Proceedings of the 1987 Winter Simulation Conference* (A. Thesen, F. H. Grant, and W. D. Kelton, eds.). Institute of Electrical and Electronics Engineers, Atlanta, Georgia, to appear.

Box, G. E. P. and Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society* **B13**, 1-38.

Cogliano, V. J. (1982). Sensitivity analysis and model identification in simulation studies. Unpublished Ph.D. Dissertation, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York.

Cooley, B. J. and Houck, E. C. (1982). A variance-reduction strategy for RSM simulation studies. *Decision Sciences* **13**, 303-321.

Cooley, B. J. and Houck, E. C. (1983). On an alternative variance-reduction strategy for RSM simulation studies. *Decision Sciences* **14**, 134-137.

Glynn, P. W. (1986). Stochastic approximation for Monte Carlo optimization. In: *Proceedings of the 1986 Winter Simulation Conference* (J. R. Wilson, J. O. Henriksen, and S. D. Roberts, eds.). Institute of Electrical and Electronics Engineers, Washington, D.C., 356-365.

Heidelberger, P. (1987). Limitations of infinitesimal perturbation analysis. Technical Report RC 11891, IBM Thomas J. Watson Research Center, Yorktown Heights, New York.

Ho, Y. C., Eyler, M. A., and Chien, T. T. (1979). A gradient technique for for general buffer storage design in a serial production line. *International Journal of Production Research* **17**, 557-580.

Ho, Y. C., Suri, R., Cao, X. R., Cassandras, C., and Zazanis, M. A. (1987). Perturbation analysis (PA) of discrete event systems—limited or unlimited? Spring Joint National Meeting of The Institute of Management Sciences and the Operations Research Society of America, New Orleans, Louisiana.

Johnston, J. (1972). *Econometric Methods*, Second Edition. McGraw-Hill, New York.

Nozari, A., Arnold, S. F., and Pegden, C. D. (1984). Control variates for multipopulation simulation experiments. *IIE Transactions* **16**, 159-169.

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

Porta Nova, A. M. and Wilson, J. R. (1986). Using control variates to estimate multiresponse simulation metamodels. In: *Proceedings of the 1986 Winter Simulation Conference* (J. R. Wilson, J. O. Henriksen, and S. D. Roberts, eds.). Institute of Electrical and Electronics Engineers, Washington, D.C., 326-334.

Reiman, M. I. and Weiss, A. (1986). Sensitivity analysis via likelihood ratios. In: *Proceedings of the 1986 Winter Simulation Conference* (J. R. Wilson, J. O. Henriksen, and S. D. Roberts, eds.). Institute of Electrical and Electronics Engineers, Washington, D.C., 285-289.

Schruben, L. W. (1986). Simulation optimization using frequency domain methods. In: *Proceedings of the 1986 Winter Simulation Conference* (J. R. Wilson, J. O. Henriksen, and S. D. Roberts, eds.). Institute of Electrical and Electronics Engineers, Washington, D.C., 366-369.

Schruben, L. W. and Cogliano, V. J. (1987). An experimental procedure for simulation response surface model identification. *Communications of the ACM* **30**, 716-730.

Schruben, L. W. and Margolin, B. H. (1978). Pseudorandom number assignment in statistically designed simulation and distribution sampling experiments. *Journal of the American Statistical Association* **73**, 504-525.

Suri, R. and Zazanis, M. A. (1987). Perturbation analysis gives strongly consistent sensitivity estimates for the M/G/1 queue. Technical Report, Division of Applied Sciences, Harvard University, Boston, Massachusetts.

Tew, J. D. and Wilson, J. R. (1985). Validation of correlation-induction strategies for simulation experiments. In: *Proceedings of the 1985 Winter Simulation Conference* (D. T. Gantz, G. C. Blais, and S. L. Solomon, eds.). Institute of Electrical and Electronics Engineers, San Francisco, California, 190-195.

Tew, J. D. and Wilson, J. R. (1986). Validation of correlation-induction strategies for simulation experiments. Research Memorandum 86-12, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.

Tew, J. D. and Wilson, J. R. (1987). Metamodel estimation using integrated correlation methods. In: *Proceedings of the 1987 Winter Simulation Conference* (A. Thesen, F. H. Grant, and W. D. Kelton, eds.). Institute of Electrical and Electronics Engineers, Atlanta, Georgia, to appear.

AUTHOR'S BIOGRAPHY

JAMES R. WILSON is an Associate Professor in the School of Industrial Engineering at Purdue University. He received a B.A. in mathematics from Rice University in 1970, and M.S. and Ph.D. degrees in industrial engineering from Purdue University in 1977 and 1979 respectively. He has been involved in various simulation studies while working as a research analyst for the Houston Lighting & Power Company (1970—72) and while serving as a U.S. Army officer (1972—75). From 1979 to 1984, he was an Assistant Professor in the Mechanical Engineering Department of The University of Texas at Austin. His current research interests include simulation output analysis, variance reduction techniques, ranking-and-selection procedures, and stopping rules. He is a member of ASA, IIE, ORSA, SCS, and TIMS.

James R. Wilson
School of Industrial Engineering
Purdue University
West Lafayette, IN 47907, U.S.A.
(317) 494-5408
wilsonj@gb.ecn.purdue.edu