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

FACTOR SCREENING OF MULTIPLE RESPONSES

Lori S. Cook
Department of Industrial Engineering
University of Louisville
Louisville, Kentucky 40292

ABSTRACT

The majority of the simulation literature has focused on single response models with a limited number of input variables. However, in most real world simulation applications, there are multiple performance measures of interest and numerous input variables. The purpose of this session is to focus on the application of factor screening designs for multiple response simulation models. The intent is to provide the practicing simulationist with general guidelines to reduce the original number of input factors to the subset of factors which exert the most significant impact on the multiple performance measures. This session will review factor screening analysis, multiple response simulation analysis and the application of factor screening in the presence of multiple responses.

1 INTRODUCTION

Simulation is defined as the representation of the dynamic behavior of a system by moving it from state to state in accordance to well defined operating rules (Pritsker, 1986). Simulation provides a method of analysis for large and complex systems which cannot be evaluated through mathematical modeling. In simplified form, a simulation model can be thought of as a "black box" where controllable and uncontrolled values are inputted and combined to generate a set of response values. The response values correspond to system performance measures which are to be evaluated.

The previous decade proved to be the turning point for the usage of simulation modeling and analysis through the rapid advancement of simulation languages to the market place. It is no longer a requirement for the simulationist to have an advance degree to understand the concepts and application of simulation. With the continual improvement of simulation products from a design and operating platform, simulation can now be used at lower technological levels. One of the most significant contributions to the simulation world has been the introduction of simulation on the personal computing level. Today, you will find a personal computer on almost any engineer's or scientist's desktop. Hence, simulation technology is readily available to all which are interested.

A common managerial perception is that a simulationist is a computer programmer and that anyone who possesses computer skills can perform a simulation analysis. In fact, it has been estimated that only one-third of the time involved in a simulation project is spent on the actual coding of the model. The real effort lies in gathering the input data and analyzing the simulation output. The danger of providing a very technical tool to the non-technical person is the potential misuse of the tool. Often the non-technical person attempting to apply simulation does not have the statistical knowledge necessary to adequately use the tool and the potential misuse of simulation can be a very costly experience.

There is an excessive amount of simulation literature relating to a variety of issues from design and development of models, to applications of models and the statistical implications of simulation models. Regardless of the abundance of literature, there is a vast number of practicing simulationist which possess inadequate statistical skills. Often the simulation output analysis is performed in a "hit or miss" sequence. This implies that there is not a designed experiment to ensure the proper evaluation process. The simulation experiments are randomly selected with little thought concerning the statistical implications of the selected experiments. To ensure a successful simulation projects it is essential for the practicing simulationist to use a properly designed experiment.

2 REVIEW OF FACTOR SCREENING FOR SIMULATION ANALYSIS

Factor screening is the process of determining the subset of factors in a simulation model which exert the greatest impact on the set of response variables. Mauro and Smith (1984) stated the following goals of factor screening: 1) to classify as important as many of the truly important factors as possible, 2) to avoid declaring unimportant factors as important, 3) to accomplish these objectives using the smallest number of simulation runs possible.

The methodology of factor screening was introduced by Watson (1961). He made the following assumptions if f factors where to be tested for their effect on the response:

1. All factors have independently the same prior probability of being effective.
2. Effective factors have the same effect.
3. There are no interactions present.
4. The required design exists.
5. The directions of possible effects are known.
6. The error of all observations are independently normal with a constant known variance.
7. $f=gk$, where $g=$ number of groups and $k=$ number of factors per group.

After Watson introduced the concept of factor screening, a number of researchers notably Kleijnen (1975, 1987), Mauro (1982, 1984, 1986), Smith (1982, 1984) and Montgomery (1979), have challenged the underlying assumptions of factor screening. Their research has indicated that the assumptions are very robust and the violation of these assumptions does not prohibit the use of factor screening techniques. In addition, a number of researchers have focused on issues concerning optimal grouping policies, multi-stage procedures and minimizing the number of runs required. However, the factor screening research has yet to adequately address the situation in which multiple responses occur.

To date, there have been only a limited number of real world applications of factor screening in simulation modeling. Cochran and Chang (1990) used a two-stage group screening experimental design to investigate which subset of variables display the most important impact on the optimum response variable. After employing the two stage group screening process, Response Surface Methodology was used to determine the optimum value of the variables. The paper considered multiple input parameters but did not explicitly consider the situation in which there are multiple response variables.

Rooda and Schilden (1982) applied a two-stage group screening process in conjunction with multiple regression analysis to the simulation of maritime transport and distribution by sea-going barges. The twenty-nine individual factors were aggregated into eight groups. A resolution IV fractional factorial design was used for the eight groups and the results were analyzed through estimated weighted least squares. A metamodel was derived and cross-validated. One-half of the groups were eliminated through significance testing. In the second stage, the remaining groups were disjoined into the original individual factors. The twelve individual factors where analyzed through a resolution III fractional factorial design. Similar to the first stage, a metamodel was derived and cross-validated.

Biles and Hatfield (1991) detailed a factor screening and region reduction approach for multiple response simulation models. The initial simulation runs were executed based on a Resolution IV fractional factorial design. Next, a first-order metamodel was established through the multiple simulation responses. Then the input factors were screened based on significance testing. Finally, a region reduction process was applied to the significant input factors. The technique is likely to be effective in factor screening, but as the number of responses increases it will not be very effective in the region reduction process.

3 REVIEW OF MULTIPLE RESPONSE SIMULATION ANALYSIS

Simulation provides a tool in which the relationship of controllable and uncontrollable input variables to output variables can be evaluated. The vast majority of simulation research has focused on the uni-response simulation model. However, very few models are simply stated as uni-response models. Typically, there are multiple responses which are of interest to the simulationist. This poses a difficult problem to the simulationist trying to evaluate and compare multiple responses. The following three techniques have been suggested for the analysis of the multiple response simulation model.

3.1 Performing Multiple Univariate Analyses on Common Data Set

One approach for analyzing multiple responses is to use univariate analysis on the same data set {(Naylor 1966, 1967) (Hunter and Naylor, 1970) (Shannon, 1975)}. In this approach, the interdependence among the response variables is not considered. For example, if several univariate tests are performed at $\alpha = .10$ significance level, the significance level for the entire study becomes $1 - (1 - \alpha)^p$, where p is the number of measures of univariate tests performed on the data set. If five individual tests ($p=5$) were performed at $\alpha = .10$ significance level, the resulting experimentwise level would be $1 - (1 - .10)^5 = 0.41$. This would undoubtedly be an unacceptable level for any simulationist.

Another methodology for analyzing multiple response models is to use the Bonferroni inequality approach {(Kleijnen, 1980) (Miller, 1981) (Balci and Sargent, 1981)}. This is accomplished by adjusting the individual significance levels to achieve the desired experimentwise error rate. The individual significance levels are set at α_E / p, where α_E represents the acceptable experimentwise error rate. For example, if $p=5$ and the desired experimentwise error rate is equal to .10, the univariate level would be .02. The Bonferroni inequality approach is noted as being a very conservative type of analysis. However, it provides a simple means for analyzing multiple responses and is superior to simply using univariate analysis on the same data set.

Johnson and Wichern (1982) state that the Bonferroni intervals can effectively be used when the number of confidence intervals are small. Charnes (1991) suggests that if the number of confidence intervals are large, the intervals may be very wide and provide minimal accuracy. Charnes recommends the multivariate batch means (MBM) method of constructing a joint confidence region. Charnes and Kelton (1988) provided a simulation application to compare multivariate output analytic methods and demonstrated the advantages of the MBM technique.

3.2 Combining Responses into a Response Function

Another approach used to address multiple response models has been the construction of a criterion or utility function {(Kotler, 1970) (Montgomery and Bettencourt, 1977) (Biles and Swain, 1979) (Clayton, 1983) (Rees, 1985) (Biles, 1987)}. In this approach the multiple responses are combined into a single function using a subjective weighting scheme to eliminate the multiple response issue. However, just as in goal programming types of applications, questions surface concerning the validity of the weighting schemes used and the actual construction of the function. To avoid the subjective weighting schemes, a total cost function can be used as a weighted objective function. The total cost function inherently weights the function and eliminates the subjective weighting.

3.3 Performing Multivariate Statistical Tests

A number of researchers have proposed the use of multivariate statistical methods for the analysis of the multiple response model. The proposed methods consist of Hotelling's T^2 test {(McArdle, 1977) (Balci and Sargent, 1981) (Schruben, 1981) (Seila, 1984)} factor analysis (Clark, 1983) canonical correlation analysis (Friedman, 1987) and multivariate analysis of variance (Friedman, 1984, 1986). These types of analysis have shown promising results, but further research needs to be performed.

4 FACTOR SCREENING METHODOLOGY

The intent of this session is to provide a methodology in which to combine factor screening techniques with multiple response analysis. Factor screening should only be used in certain simulation modeling situations. The appropriate use of factor screening is a function of: the number of factors involved, the number of simulation runs required, the number of simulation runs available for use and the available time for analysis.

The number of input factors dictates if there is a need for factor screening analysis. If the number of input factors is relatively small , factor screening would not be necessary. It would be more appropriate to use a 2^n full factorial or 2^{n-p} fractional factorial design. However, if there are a large number of input factors, it would be more appropriate to use a screening methodology to determine the subset of factors which are most significant. By screening the input parameters, one can also determine the level of accuracy required for the significant input parameters. Hence, for a significant input factor parameter the data estimates should have a higher degree of accuracy for the next phase of the simulation model.

The number of simulation runs required, the number available for use and the time available are all important considerations of factor screening. The factor screening methodology discussed should only be applied to determine the subset of important factors within the simulation model. It is not intended for predicting the actual values of the performance measures. Therefore, the simulationist should use the minimum number of runs necessary for screening to achieve the desired accuracy the simulationist is willing to accept.

4.1 Experimental Design Strategy

The experimental design strategy for factor screening will vary depending on the level of saturation of the design. The three levels of saturations are unsaturated ($k<N$), saturated ($k=N-1$) and supersaturated ($k \geq N$), where k is the number of factors and N is the number of runs. In the unsaturated and saturated situation, the simulationist has more runs available than factors for the screening process. In the supersaturated situation, the

simulationist has more factors to screen than runs available. The following will discuss 2^n full factorial designs, 2^{n-p} fractional factorial, supersaturated design and group screening designs.

A 2^n full factorial design is a formation which contains every possible combination of the **n** factors, each at two levels. Typically, the two levels represent the "high" and "low" values for the input parameter. The design allows all effects and interactions to be analyzed. The two-level factorial designs are very useful in the screening process, but often require an excessive number of runs. For example, if a simulation model has 10 input factors, the 2^{10} factorial design would require 1024 runs before replicating. The full factorial design would be recommended only in situations in which a small number of factors were to be screened (≤ 5 factor).

Due to the excessive number of runs often required by the full factorial design, the 2^{n-p} fractional factorial is often used in the screening process. These designs utilize a 1/p fraction of the 2^n design points in the full factorial design (Box and Hunter 1961). The design assumes that the higher order interactions are negligible and the main effects and low order interactions are obtained by running only a fraction of the full factorial. The degree of acceptable confounding determines the resolution of the design. In resolution III designs, no main effect is confounded with any other main effect, but main effects are confounded with two-factor interactions and two factor interactions with one another. In resolution IV designs, no main effect is confounded with any other main effect or two factor interaction but two factor interactions are confounded with one another. In resolution V designs, no main effect or two factor interaction is confounded with any other main effect or two factor interaction, but two factor interactions are confounded with three factor interactions.

Mauro (1986) provided an overview of supersaturated designs for use in factor screening . He discussed random balance, systematic supersaturated, group screening, modified group screening, T-optimal, R-optimal, and search designs. In the supersaturated situation, he recommends using group screening except when the number of runs is severely limited. For this situation, he recommends a systematic supersaturated designs.

Group Screening is based on aggregating individual factors into groups. The groups containing the individual factors are then treated as single factors. The level of a group factor is achieved by assigning the individual factors inside a group to their "high" or "low" levels together. If a group is deemed not significant, it is concluded that all factors within the group are unimportant. Then individual factors of the significant groups are analyzed to determine the significant individual factors.

4.2 Analysis of Factor Screening Designs for Multiple Response Simulation Model

The techniques recommended for use in screening multiple response simulation scenarios consist of utility functions, modified Bonferroni inequality approach and multivariate statistical analysis. A utility function could be used if the model inherently displays a function which incorporates the multiple responses. For example, in an inventory simulation the total cost function incorporates the multiple responses into a weighted utility function. By developing a meta-model and analyzing the change in the total cost function, the input factors can appropriately be screened. In the modified Bonferroni inequality approach, the individual significance levels would be assigned in inverse proportion to a weighted utility function. Current research is examining the use of the multivariate statistical analysis procedures. The type of analysis technique used will be based on the structure and nature of the simulation model.

4.3 Methodology

If screening is deemed necessary, a sequential approach is recommended. Due to the importance of limiting the number of simulation runs the appropriate experimental design strategy must be initially determined. If the number of input factors is small ($\leq$ 5 factor), a full factorial or resolution IV design would be recommended.

If the number of factors exceeds 5, then the number of simulation runs required would have to be examined. It is recommended that the initial design should be of a least resolution IV because of the confounding pattern of main effects with the two-way interactions in the resolution III design. Therefore, a resolution IV fractional factorial design is recommended for situations which have 6 to 11 input factors.

If the number of input factors exceeds 11, a group screening approach is recommended. If the number of groups is less than or equal to 5, a full factorial design should be used for the initial screening. Subsequent screening designs would be based on the number of individual factors remaining in the significant groups. If the number of groups is between 6 and 11, a fractional factorial design should be used for the initial screening. Subsequent screening designs would be based on the number of individual factors remaining in the significant groups. If the number of groups exceeds 11, a systematic supersaturated design should be used.

After the experimental design strategy has been selected, the type of analysis procedure should be determined. It is recommended that either utility function, modified bonferroni or multivariate analysis is used. The methodology selected will be dependent on the structure and nature of the simulation model.

Additional consideration should be given to the number of replications for each of the experimental design points. In the screening process, it is suggested that at a minimum one simulation replication of each design point in conjunction with the antithetic of each design point should be executed. Under no circumstances should only one replication be performed.

5 CONCLUSIONS

Most real world simulation models have a large number of input variables and will exhibit multiple response measures. The practicing simulationist must have a means to determine the most important input factors and a methodology by which to analyze the multiple output responses. This session investigated the feasibility of using factor screening designs for multiple response simulation models. The intent is to provide general guidelines for the simulationist to use for factor screening. The factor screening methodology discussed should only be used to determine the subset of important factors within the simulation model. It is not intended for predicting actual values of the performance measures.

To date, there have been only a limited number of real world applications of factor screening in simulation modeling. Kleijnen (1987) noted that "in academic studies the system is often so small that screening is not necessary; in practical studies the system is often large but the statistical know-how is missing." Current research is actively pursuing more efficient techniques for the factor screening process in the presence of multiple responses.

REFERENCES

Balci, O. and R. Sargent. 1981. A methodology for cost-risk analysis in the statistical validation of a simulation model. *Communications of the ACM* 24(11):190-197.

Biles, W.E. 1984. Design of simulation experiments. *Proceedings of the 1984 Winter Simulation Conference*. 99-104.

Biles, W.E. and I.T. Hatfield. 1991. Simultaneous factor screening and region reduction in computer simulation experiments. *Computers and Industrial Engineering* 21(1-4).

Biles, W. E. and H. Ozmen. 1987. Optimization of simulation responses in a multicomputing environment. *Proceedings of the 1987 Winter Simulation Conference*. 402-408.

Biles, W. E. and J. J. Swain. 1979. Mathematical programming and the optimization of computer simulations. *Mathematical Programming Studies on Engineering Optimization* (R. S. Dembro and M. Avriel, Eds.) Amsterdam: North Holland Publishing Company. 189-207.

Box, G. E. and J. S. Hunter. 1961. The 2^{k-p} fractional factorial designs, part I. *Technometrics* 3(3):311-351.

Box, G. E. and J. S. Hunter. 1961. The 2^{k-p} fractional factorial designs, part II. *Technometrics* 3(4):449-458.

Charnes, J. M. 1991. Multivariate simulation output analysis. *Proceedings of the 1991 Winter Simulation Conference*. 187-192.

Charnes, J. M. and W. D. Kelton. 1988. A comparison of confidence region estimators for multivariate simulation output. *Proceedings of the 1988 Winter Simulation Conference*. 458-465.

Clark, T. D., M. M. Kabil and M. I. M. Mostafa. 1983. An analysis and simulation of an experimental Suez Canal Traffic Control System. *Proceedings of the 1983 Winter Simulation Conference*. 311-318.

Clayton, E., W. Weber and B. Taylor. 1982. A goal programming approach to the optimization of multiresponse simulation models. *IIE Transactions* 14(4):282-287.

Cochran, J. K. and J. Chang. 1990. Optimization of multivariate simulation output models using a group screening method. *Computers and Industrial Engineering* 18(1):95-103.

Cook, L. S. 1992. Factor Screening Designs for Multiple Response Computer Simulations. Unpublished dissertation: Department of Industrial Engineering, University of Louisville.

Friedman, L. W. 1984. Establishing functional relationships in multiple response simulation: the multivariate general linear metamodel. *Proceedings of the 1984 Winter Simulation Conference*. 285-289.

Friedman, L. W. 1984. Multivariate simulation output analysis: past, present, and future. *Proceedings of the 1984 Winter Simulation Conference*. 277-281.

Friedman, L. W. 1984. Statistical considerations in computer simulation: the state of the art. *Journal of Statistical Computation and Simulation* 1:237-263.

Friedman, L. W. 1987. Systems simulation design and analysis of multivariate response simulations: the state of the art. *Behavioral Science* 32:138-148.

Friedman, L. W. 1986. The analysis of multiple response simulation output data: experiments of comparison. *Computer and Operations Research* 13(6):647-652.

Friedman, L. W. 1989. The multivariate metamodel in queuing system simulation. *Computers and Industrial Engineering* 16(2):329-337.

Hunter, J. S. and T.H. Naylor. 1970. Experimental design for computer simulation experiments. *Management Science* 16(7):422-434.

Johnson, R. A. and D. W. Wichern. 1988. *Applied Multivariate Statistical Analysis*, 2d. ed. Englewood Cliffs, NJ: Prentice Hall.

Kleijnen, J. P. 1980. Analysis of simulation with common random numbers. *Simuletter* 11(2):7-10.

Kleijnen, J. P. 1975. Screening designs for poly-factor experimentation. *Technometrics* 17(4):487-493.

Kleijnen, J. P. 1987. Simulation with too many factors: review of random and group-screening designs. *European Journal of Operational Research* 31:31-36.

Kleijnen, J. P. 1975. *Statistical Techniques in Simulation, Vol. I and II*. New York: Marcel Dekker.

Kleijnen, J. P. 1987. *Statistical Tools for Simulation Practitioners*. Marcel Dekker: New York.

Kotler, P. 1970. A guide to gathering expert estimates. *Business Horizons* 13(5):79-87.

Mauro, C. A. 1986. Efficient identification of important factors in large scale simulations. *Proceedings of the 1986 Winter Simulation Conference*. 296-305.

Mauro, C. A. 1984. On the performance of two-stage group screening. *Technometrics* 26(3):255-264.

Mauro, C. A. and D. E. Smith. 1982. The performance of two-stage group screening in factor screening experiments. *Technometrics* 24(2):325-330.

Mauro, C. A. and D. E. Smith. 1984. Factor screening in simulation: evaluation of two strategies based on random balance sampling. *Management Science* 30(2):209-221.

McArdle, J. J. 1977. The statistical analysis of Monte Carlo simulation data using the techniques of discrete multivariate analysis. *Proceedings, Computer Science and Statistics: The 10th Annual Symposium on the Interface*. 241-246.

Meketon, M. 1987. Optimization in simulation: a survey of recent results. *Proceedings of the 1987 Winter Simulation Conference*. 58-67.

Miller, R. 1981. *Simultaneous Statistical Inference*. New York: Springer-Verlag.

Montgomery, D. C. 1984. *Design and Analysis of Experiments*. New York: John Wiley & Sons.

Montgomery, D. C. 1979. *Methods for Factor Screening*. Technical report: Georgia Institute of Technology.

Montgomery D. C. and V. M. Bettencourt. 1977. Multiple response surface methods in computer simulation. *Simulation* 29(4):113-121.

Naylor, T. H., J. Balintfy, D. S. Burdick and K. Chu. 1966. *Computer Simulation Techniques*. New York: John Wiley & Sons.

Naylor, T. H., D. S. Burdick, and W. Sasser. 1967. Computer simulation experiments with economic systems: the problem of experimental design. *Journal of ASA* 62(320):1315-1337.

Neter, J., W. Wasserman, and M. Kutner. 1990. *Applied Linear Statistical Models*, 3d. ed. Homewood, IL: Richard D. Irwin.

Patel, M. S. 1962. Group screening with more than two stages. *Technometrics* 4:209-217.

Patel, M. S. and J. A. M. Ottieno. 1984. Optimum two-stage group screening designs. *Communications in Statistics Theory & Methods* 12(21):2649-2663.

Pritsker, A. A. B. 1986. *Introduction to Simulation and SLAM II*. West Lafayette, IN: Systems Publishing Corporation.

Rees, L., E. Clayton, and B. Taylor. 1985. Solving multiple response simulation models using a unified response surface methodology with a lexicographic goal programming framework. *IIE Transactions* 17(1):47-57.

Rooda, J. E. and W. Schilden. 1982. Simulation of Maritime transport and distribution by sea-going barges: an application of multiple regression analysis and factor screening. *Bulk Solids Handling* 2(4).

Schruben, L. W. 1981. Control of initialization bias in multivariate simulation responses. *Communications of the ACM* 24(4):246-252.

Seila, A. 1984. Multivariate simulation output analysis. *American Journal of Mathematical and Management Sciences* 4(3&4):313-334.

Shannon, R. E. 1975. Simulation: a survey with research suggestions. *AIIE Transactions* 7(3):289-301.

Smith, D. E. and C. A. Mauro. 1982. Factor screening in computer simulation. *Simulation* 38(2):49-54.

Watson, G. S. 1961. A study of group screening methods. *Technometrics* 3:371-388.

AUTHOR BIOGRAPHY

LORI S. COOK is a Ph.D. candidate in the Department of Industrial Engineering at the University of Louisville. She received her BS and M.Eng. in Industrial Engineering from the University of Louisville. In addition, she has attended the University of Kentucky and the University of Cincinnati. She has worked for Wright-Patterson Air Force Base, Armco Eastern Steel Division and Kentucky Fried Chicken Corporation where she performed simulation analysis of complex systems. Her research interests include the application of optimization and simulation modeling to complex systems. She is a member of Alpha Pi Mu, Tau Beta Pi, IIE, TIMS and ORSA.