Proceedings of the 2000 Winter Simulation Conference
J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, eds.

ROBUST DESIGN: SEEKING THE BEST OF ALL POSSIBLE WORLDS

Susan M. Sanchez

Operations Research Department
Systems Management Department
Naval Postgraduate School
Monterey, CA 92943, U.S.A.

ABSTRACT

We describe a framework for analyzing simulation output in order to find solutions that will work well after implementation. We show how the use of a loss function that incorporates both system mean and system variability can be used to efficiently and effectively carry out system optimization and improvement efforts. For models whose behavior depends on quantitative factors, we illustrate how robust design can be accomplished by using simple experimental designs in conjunction with response-surface metamodels. The results can yield new insights into system behavior, and may lead to recommended system configurations that differ substantially from those selected by analysis solely on the basis of mean response. We assume a knowledge base at the level of Chapter 12 of *Simulation Modeling and Analysis* (Law and Kelton 2000) but will review essential elements and distribute illustrative examples at the session.

1 INTRODUCTION

What is robust design? It is a system optimization and improvement process that springs from the view that a system should not be evaluated on the basis of mean performance alone. In addition to exhibiting an acceptable mean performance, a "good" system must be relatively insensitive to uncontrollable sources of variation present in the system's environment.

Some notation and terminology will be necessary before the robust design process and its benefits can be fully discussed. However, it is important to state up front that the purpose of robust design is to lead to better decisions: better in terms of implementation, better in terms of the level and consistency of performance, better in terms of cost, and better in terms of insights into the drivers of system performance.

The robust design approach originated in quality planning and engineering product design activities (Taguchi and Wu 1980; Taguchi 1986 1987). Taguchi found that it was often more costly to control causes of manufacturing variation than to make a process insensitive to these variations, and through the use of simple experimental designs and loss functions it was often possible to greatly improve product performance by "building in" the quality. Taguchi's philosophy and strategy have been widely praised in both the applied statistics and manufacturing communities, (Pignatiello and Ramberg 1991) but many of the methods and tactics he advocates are often controversial (Box 1988, Ramberg, Pignatiello and Sanchez 1991, Nair et al. 1992). The approach described in this paper (see also Sanchez et al. 1996 and 1998, Ramberg et al. 1991) combines Taguchi's strategy and response-surface metamodeling techniques. The additional insights that can be gained make this approach particularly beneficial when analyzing simulations of complex systems.

In the simulation context, robust design can be viewed from two slightly different perspectives. One view is that simulation is used primarily as a surrogate for a real system, because of the cost and time required to make and observe changes in a real system. From this perspective, the total time required to perform a robust design experiment is greatly reduced, but the designs and analyses used are the same as those that would be applied to a physical system if cost and time permitted. Applications have included the product designers' uses of computer models for experimentation in place of physical prototypes, particularly in the semiconductor industry (Sacks et al. 1989, Welch et al. 1990). These experiments have typically involved Monte Carlo simulation, although clearly robustness can be used as a criteria for evaluating discrete-event simulation systems as well. Those who use simulation to study systems primarily because of the difficulty of experimenting on the real system may realize the benefits of improved performance and decreased cost cited by many manufacturers if they decide to evaluate performance in terms of robustness.

A larger view of the simulation process is also possible. A simulation model is constructed assuming a variety of system inputs (e.g., distributional forms and characteristics, simplifying assumptions, level of detail) which are unlikely to be completely accurate. Model verification and validation are important issues in the field, as is simulation sensitivity analysis. From this perspective, one can view robust design as a process of simulation optimization, where the "best" answer is not overly sensitive to small changes in the system inputs. If robust systems are identified, then the actual results are more likely to conform to the anticipated results after implementation.

The robustness criteria can be applied to rank a discrete number of alternatives, which result from changing the settings of some or all of the inputs to the simulation model (or system). Alternatively, if some or all of the input factors are quantitative, one can construct metamodels of the simulation which describe how the system performance varies as a function of the input factors. There are many approaches to metamodeling, but response surface metamodels work well in the robust design context. Metamodels provide much more information about the underlying system than haphazard investigation of a few alternatives. Thus, if the goal of the analyst is to optimize or improve the model's performance, and flexibility exists in the settings of the decision factor levels (as in prospective studies), then building a metamodel is appropriate. The actual number of configurations studied, and the form (linear, quadratic, etc.) of the resulting metamodel are dependent on the experimental design chosen. Note that first-order models may not suffice for complex discrete-event simulations: performance is often highly nonlinear, even over a relatively restricted range of factor settings.

The construction of metamodels is facilitated by the use of experimental design techniques. Simulation analysts have the luxury of controlling all inputs to the simulation (including random number seeds, etc.): this means they have more flexibility in designing the experiment, and more opportunities for exploiting the additional degree of control, than do those experimenting directly on real systems.

In this tutorial our focus is on the robust design process for discrete-event simulation experiments. We begin with a discussion of the terminology used in the robust design approach. We then discuss tactical issues, such as appropriate experimental designs, metamodel construction, robust design identification, analysis, We conclude with a summary of the benefits of a robust design approach, emphasizing the types insights into system behavior that can be achieved.

2 NOTATION AND TERMINOLOGY

In this section, we present notation and discuss the terminology related to the robust design process.

2.1 Factor Classification

In systems where stochastic variation is present, the response exhibits random fluctuation or variation. In order to achieve systems or products for which the variation around the target value is low, several steps are necessary. First, one must identify factors in the system which are anticipated to affect the system response. Factors are classified as *decision factors*, *noise factors*, or *artificial factors*.

The decision factors are those which are controllable in the real world setting modeled by the simulation. Noise factors are not easily controllable or controllable only at great expense in the real-world setting. Noise factors include sources of variation within the real-world system (i.e., within a manufacturing plant) as well as exogenous factors (such as customer and supplier characteristics). Finally, artificial factors are those simulation-specific variables such as the initial state of the system, the warm-up period (truncation point), termination conditions (run duration), and random number stream(s) (seed, antithetic switch).

The distinction between decision factors and noise factors is often recognized in simulation experiments, but rarely used to develop the experimental design or affect the analysis of the simulation results. However, as we discuss in Section 3.1, the classification is important. It is necessary for determining system robustness, and also presents an opportunity for reducing the number of runs required by concentrating sampling efforts on assessing decision factor effects. This additional layer of control made possible by the artificial factors can also be exploited in the experimental design (Schruben et al. 1992). This is not new—it is the basis of many variance reallocation techniques.

2.2 Performance Evaluation

The analyst begins by specifying some performance characteristic of special interest, and an associated target value τ. Common measures are related to system throughput or system states, such as the waiting time or number of customers in queueing systems, although cost could also be used as a performance measure. In general, the pattern of the performance characteristic's fluctuation around the target value will differ across these configurations. The cost of this fluctuation must be measured in order to optimize or improve the system. Since end-users will incur costs if system performance deviates from the target, the evaluation criterion is often philosophically referred to as the *loss to society*, or the *long term business loss*.

An ideal configuration would result in the performance characteristic's mean equal to τ and its variance equal to zero. Thus, a numerical method for trading off performance mean and variability is needed.

We utilize a quadratic loss function, which (in many cases) is a reasonable surrogate for the 'true' underlying loss function which may be difficult or impossible to specify exactly. Let $\mathbf{x}$ and $Y(\mathbf{x})$ denote a vector of decision factor settings and the associated performance characteristic respectively, and let Ω denote the noise factor space. Then, assuming that no loss is incurred when $Y(\mathbf{x})$ achieves the ideal state (τ), the quadratic loss function can be written as:

$$\ell\,(Y(\mathbf{x})) = c\,[Y(\mathbf{x}) - \tau]^2 \tag{1}$$

where the scaling constant c can be used to convert losses into monetary units to facilitate comparisons of systems with different capital costs. It follows from equation (1) that the expected loss associated with configuration $\mathbf{x}$ is

$$E(loss) = c\left[\sigma^2 + (\mu - \tau)^2\right]. \tag{2}$$

This loss function has several nice properties. It penalizes small deviations from τ only slightly, yet assesses a large penalty for responses far from the target. The expected loss is similar to a mean-squared-error loss, thus it has many desirable mathematical properties as well. However, other loss functions can be used if the true loss function is not approximated well by a quadratic.

While conceptually straightforward, the use of a loss function to incorporate system variability into the performance evaluation represented a major shift in perspective within the manufacturing community. No longer was it acceptable to think about optimizing mean performance without regard to performance variability: a "good" product was also robust. The quantification of robustness, instead of the 0/1 loss function often implicitly used to represent products which were within/outside specification limits, also provided impetus to management and manufacturing to continually improve product quality. We believe that for many types of applications, a similar change in perspective should occur within the discrete-event simulation community. If simulation is being used to identify "good" systems (e.g., plant layouts, scheduling and control mechanisms), where variability is not constant across alternative system designs, then a loss function such as that in equation (2) is a better descriptor of the system's desirability than the average (steady-state) performance. For example, a single-server queue will have the same mean waiting time for customers/jobs under the FIFO and LIFO queue disciplines, but the variability is quite different. In general, the configuration with the best mean need not be associated with the lowest loss.

3 EXPERIMENTAL DESIGNS

Choosing an experimental design means specifying the levels of all decision factors, noise factors, and artificial factors for each run of the simulation. An appropriate total sample size must also be determined. In order to evaluate the expected response variability across the noise space, a crossed decision$\times$noise factor plan can be used. This means that the same experimental plan for the noise factors is used for each run of the decision factor plan.

3.1 Basic Plans

Complete and fractional factorials are often used. Among these, two-level designs are popular choices because of their simplicity and efficiency. They permit the evaluation of the linear decision factor effects, as well as interaction or synergistic effects.

For a two-level factorial or fractional factorial experiment involving k decision factors, the factor levels should be chosen to cover the range of interest. For noise factor plans, the levels should be chosen so that the mean and variance of the two-point sampling distribution are equal to the mean and variance of the underlying distribution. In the case of two-level sampling of continuous factors (or discrete factors whose distributions can be closely approximated by continuous distributions), this corresponds to one standard deviation below and one standard deviation above the mean. In the case of equally likely Bernoulli outcomes, this corresponds to the two factor levels. For discrete distributions where $\mu \pm \sigma$ does not yield valid factor levels, the outcomes can be sampled (approximately) proportional to their probability of occurrence. If the factor is a mean estimated from data, then the upper and lower bounds of a 95% confidence interval ean be used (Wild and Pignatiello 1991).

Other orthogonal designs have been advocated for response surface metamodeling. For example, one might want to minimize the bias or mean-squared error of the regression coefficients (Donahue, Houck and Myers 1992). Central composite designs are good for fitting second-order metamodels. These designs are discussed for simulation experiments by Tew (1992) or Hood and Welch (1993); experimental design texts such as Box and Draper (1987), Box, Hunter, and Hunter (1988) or Montgomery (1991) contain details and alternative designs. Two-level plans are not sufficient if quadratic effects are anticipated.

3.2 Artificial Factor Plans

A well established field in simulation is that of variance reallocation (or variance reduction), where researchers have established methods of reducing the variance of the *estimators* of mean responses in order to increase power for hypothesis testing purposes. Unequal response variance at different system configurations is recognized as pervasive. It often influences the experiment design and analysis (e.g., varying run lengths for different system alternatives), but has rarely been incorporated into the system evaluation. In the robust design context, variance reallocation schemes hold promise for further increasing the efficiency of experimentation. Rather than using all independent random number streams, one can use a common/antithetic sampling strategy (Schruben and Margolin 1978, Tew and Wilson 1991, 1994). This reallocates variance among the coefficient estimates. The implications for robust design are that the artificial factor plan should be chosen in order to induce correlations which reallocate variance from the interesting terms (decision factors) to the uninteresting terms (noise factors) (Schruben et al. 1992). The artificial factor plan is typically embedded in the noise factor plan, e.g., through the choice of random number streams used during a simulation run.

3.3 Frequency Domain Plans

If the number of noise factors is large, even a saturated factorial plan for the noise factors may result in an unwieldy experimental design after crossing it with the decision factor plan. One way to cut down the size of the experiment is to first screen the noise factors and then employ a highly fractionated factorial design. Another efficient way to collect the data is to oscillate each noise factor sinusoidally within a simulation run at unique, carefully selected frequency. This allows examination of the system across a range of noise factor combinations without a prohibitively large number of runs (Moeeni et al. 1997). Such oscillation forms the basis of frequency domain experimentation in the simulation field (Schruben and Cogliano, 1987; Sanchez and Buss, 1987), although the analysis differs. Indexing by time, rather than by entity, is recommended (Mitra and Park, 1991).

For variance attribution, the analyst is interested in determining what portions of the total system variability can be attributed to the noise factors, and a frequency domain approach could be used for factor screening purposes. For robust design, we are interested in what the performance variability is at a particular decision factor configuration: the fact that noise factors are varying across the noise space is important, while estimates of their specific effects are not. In both cases, care should be taken to select driving frequencies which will not result in confounding and to choose frequencies resulting in cycles sufficiently long to affect the system response (Jacobson, Buss and Schruben 1991). Discrete factors can be handled either by oscillating their probabilities of realizing particular levels, or by discretizing the sinusoidal function (Sanchez and Sanchez, 1991).

3.4 Correlated Factor Plans

If the noise factors are correlated in the real world system, it might be that a factorial or fractional factorial design could not be conducted over the entire range of interest. For example, a queueing system might be unstable if all noise factors were held at their high levels. If this situation was unlikely to occur in practice because of correlation among the variables, then a sampling scheme which made use of the underlying dependence structure would seem more appropriate. If the noise factors are normally distributed, the analyst can sample at axial points on the elliptical contours of the joint distributions (Sanchez 1994b, Sanchez, Smith and Lawrence 1996, 2000).

3.5 Combined Array Plans

In some circumstances, a crossed decision×noise factor plan may not be the most efficient in terms of the total number of observations (runs) required. An alternative is a combined plan, where a single design matrix (such as a factorial) is used with columns divided among decision and noise factors. As Myers, Khuri and Vining (1992) suggest, this can be used if one can specify *a priori* which of the many possible interaction terms are potentially important. It may mean that the experiment can be conducted using a smaller total number of simulation runs than a crossed plan would require.

4 RESPONSE SURFACE METAMODELS

The response Y is a random function of the decision factors $\{X_i\}$, the noise factors $\{W_j\}$, the artificial factors $\{A_k\}$, and the inherent variability of the system. The form of the metamodels fit to the simulation outputs, and the metamodel uses, differ between the robust design and variance attribution stages.

4.1 Metamodels for Robust Design

For this analysis, we seek to characterize the system behavior as a function of the decision factors alone. For every combination of decision factor configuration i and noise factor configuration j, we first compute (after suitable truncation to remove initialization bias) the sample average $\overline{Y}_{ij}$ and sample variance s^2_{ij} for the run. Then we compute summary measures across the noise space for each decision factor configuration i:

$$\begin{aligned}
\overline{Y}_{i\cdot} &= \frac{1}{n_w}\sum_{j=1}^{n}\overline{Y}_{ij}, \\
\overline{V}_{i\cdot} &= \frac{1}{n_w-1}\sum_{j=1}^{n_w}\left(\overline{Y}_{ij}-\overline{Y}_{i\cdot}\right)^2+\frac{1}{n_w}\sum_{j=1}^{n}s^2_{ij}
\end{aligned}$$

where n_w is the number of points in the noise factor plan. Regression is used to fit two initial metamodels: one for the performance mean, and one for the performance variability (Sanchez et al. 1993, 1998; see also Vining and Myers, 1990). The terms in the initial metamodels depend on the experimental design used. For discrete-event simulation experiments we recommend a design which allows for fitting at least a quadratic effect. We obtain models such as

$$\begin{aligned} \mu &\approx \hat{\beta}_0 + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_k X_k + \hat{\beta}_{1,2} X_1 X_2 \\ &\quad + \ldots + \hat{\beta}_{k-1,k} X_{k-1} X_k + \ quadratic \end{aligned} \tag{3}$$

$$\begin{aligned} \log(\sigma^2) &\approx \hat{\gamma}_0 + \hat{\gamma}_1 X_1 + \ldots + \hat{\gamma}_k X_k \\ &\quad + \hat{\gamma}_{1,2} X_1 X_2 + \ldots + \hat{\gamma}_{k-1,k} X_{k-1} X_k \\ &\quad + \ quadratic \end{aligned} \tag{4}$$

The logarithmic transformation is used for stability purposes. If the quadratic is an important term in either metamodel, further experimentation is needed to determine the decision factor(s) from which it arises.

4.2 Metamodels for Variance Attribution

Metamodel construction is slightly Different for variance attribution purposes since all factors are treated as noise factors and we assume that the factor ranges are sufficiently small that a linear metamodels suffice. If we fit models of the response mean and standard deviation, then we obtain

$$\begin{aligned} \mu &\approx \hat{\beta}_0 + \hat{\beta}_1 W_1 + \ldots \hat{\beta}_k W_k, \\ s &\approx \hat{\gamma}_0 + \sum_{j=1}^{w} \hat{\gamma}_j W_j \end{aligned}$$

where the $\{\beta_j\}$ and $\{\gamma_j\}$ are the least-squares regression coefficients. By treating these coefficients as fixed, the overall variance can be approximated by

$$\sigma^2 \approx \hat{\gamma}_0^2 + \sum_{j=1}^{w} (\hat{\beta}_j^2 + \hat{\gamma}_j^2) \mathrm{Var}(W_j). \tag{5}$$

5 ANALYZING THE RESULTS

The initial metamodels constructed for either robust design or variance attribution should be assessed and may need to be refined. The experimental plans are typically unreplicated because of the cost of experimentation. This means that the analyst may have only a single degree of freedom for error in the initial regression metamodels, so heavy reliance should not be placed on the raw p-values or t-values. An option offered by many statistical packages (or which can be done manually) is a normal probability plot, which can be used to graphically assess whether or not any effects larger than the noise thresh-hold of the the experiment. Normal probability plots work well when 15 or more factor or interaction effects are estimated. If the regression metamodels can be simplified by eliminating unimportant terms, then pooling increases the degrees of freedom for the error estimate and allows formal tests of the statistical significance of the remaining metamodel coefficients.

5.1 Robust Design Analysis

The information resulting from the robust design metamodels of equations (3) and (4) can easily be combined using the quadratic loss function (equation (2)) to identify robust configurations. The metamodels themselves provide detailed information regarding the system performance: they indicate which decision factors affect the mean, which affect the variance, and which influence both aspects of performance.

For many simulation models, the presence of interaction terms and the relationship between the mean and the variability of the performance characteristic make it difficult to achieve the target value with the most robust product design. In these cases, contour plots may be useful for selecting candidate product designs. For example, one can first use the mean metamodel to identify several configurations for which the average performance characteristic is on target, and then use the the metamodel of $\log(\sigma^2)$ to select a configuration which is fairly robust.

Often the results suggest configurations that were not among those initially tested. In such cases, further experimentation is beneficial in order to confirm the performance characteristic's behavior before committing to a particular configuration. However, the secondary experiment may be much smaller than the initial experiment if several of the decision factors do not appear in the revised metamodels: they can be set at their most economical levels and screened from further experimentation.

We emphasize that the decision in the robust design framework can be very different than that made on the basis of mean performance alone. Even for queueing systems, where performance mean and variability tend to have high positive correlation, complex interactions among decision factors may affect that relationship. One example (Sanchez, Ramberg and Sanchez 1998) that will be illustrated in detail during the presentation concerns a job-shop simulation. The job-shop has three products, five machine groups, and varying processing time distributions, product mix percentages, etc. Experimentation showed that two configurations could have the same mean response to two decimal places, yet variances which differed by over a factor of two. The configurations corresponding to the best means were dramatically inferior to the low loss designs: one job shop configuration which was among the best in terms of mean performance had a 36% higher loss then the low-loss configuration, yet it used more machines.

The robust design philosophy and joint metamodeling approach have a synergistic relationship: they typically provide the analyst with more information than would result from either a loss comparison of only the configurations tested, or from a single metamodel which directly measures system loss or cost. In the latter case, if a metamodel shows that expected loss decreases as factor X increases, the root cause remains unknown. Perhaps the response mean is closer to the target. Perhaps the response variance is smaller. It could be that both the mean and variance improve, or that an improvement in one aspect is partially offset by a degradation in the other. However, separate construction of metamodels for the system mean and variability facilitate the identification of new designs which may be even better than those considered in the experimental framework.

5.2 Variance Attribution Analysis

Several types of questions can be addressed. First, one can assess the overall mean and variability for a particular configuration, e.g., that chosen at the end of the robust design stage. If the decision factors can be perfectly controlled at their chosen settings, then the overall mean and variance can just be estimated by the robust design metamodels. However, if variation in the decision factors' settings around their nominal values is anticipated, an additional experiment will provide a better picture of the system's capabilities.

Other questions concern the relative effects of the noise factors. The term $(\hat{\beta}_j^2 + \hat{\gamma}_j^2)\mathrm{Var}(W_j)$ in equation (5) is called the *transmitted variance* for noise factor W_j. This indicates the amount of variability in the noise factor which is passed along to variance in the response. Depending on the magnitudes of $\hat{\beta}_j$ and $\hat{\gamma}_j$, variation in W_j can be amplified or dampened by the system. The term $\hat{\gamma}_0^2$ in equation (5) is called the *inherent variance*: it is the smallest variance achievable if all noise factors investigated in the experiment have variances driven to zero. (For Monte Carlo simulations, all randomness has been removed during experimentation and every $\hat{\gamma}_i$ can be replaced by zero.)

Once transmitted variances have been computed for all noise factors, the relative importance of these sources of variation on the output is apparent. This information can be used to evaluate proposed changes to the system. For example, is it cost-effective to pay more for raw materials, machine maintenance, or training in order to improve the consistency of these factors? Alternatively, is it possible to relax controls slightly and allow more variation in the inputs without adversely affecting the system behavior? If the standard deviation of a noise factor W_j can be reduced by a factor of a_j without affecting the mean performance, then its transmitted variance is reduced by a factor of a_j^2 and the overall system variability is reduced by the amount $(1 - a_j^2)\mathrm{Var}(W_j)$. The conversion constant c from equation (2) can be used to express the overall performance change in dollars. A comparison with the cost of implementing the proposed change then shows whether or not such implementation would further improve the system performance. If changes in the noise factor variance also affect the system mean, then both the mean and variance components should be included in the cost assessment via equation (2).

Variance attribution can also aid in simulation modeling and validation. If the distributional characteristics used to generate random components of the simulation model are themselves estimates, and if the system is highly sensitive to those characteristics, then the simulation may not mimic the true system behavior adequately. Once again, the analyst can use variance attribution to obtain feedback regarding the modeling process. This allows model refinement efforts to be expended in accordance to factor sensitivity.

6 BENEFITS

The benefits of using a robust design approach can be substantial. Many of these apply regardless of the type of model (simulation, analytic, prototype) on which experiments are performed. First, because the chosen system configuration is robust, it is likely to work well across a variety of realizations of noise factor values. This means that there are likely to be fewer unpleasant surprises when the decision is implemented. A second benefit is improved communication between the analyst and client via expected loss. This also makes it possible to evaluate trade-offs between the costs of reducing noise and the benefits of improving performance quality. There are times when the insights gained from the robust design process will allow decision-makers to simultaneously improve performance and decrease costs. Finally, the fact that expected loss is calculated facilitates continuous improvement: even if target value is achieved on average, one can always seek to identify factors that can further reduce the variability of the response. In this manner, factors that were initially classified as noise factors may become decision factors at a later date.

There are some benefits of particular interest to simulators as well. First, the view that variability is a critical component of performance—not solely a nuisance to be overcome by taking larger samples—may improve the modeling and analysis process, and lend more credibility to the simulation results. Second, this structured approach lends itself to rapid model evaluation and scenario analysis. It is much more efficient than trial-and-error at identifying "good" decisions. Third, the variance attribution process gives the simulation analyst the ability to test whether or not the model performance is highly sensitive to input distributions and their parameters. This can reduce the time required to arrive at a functioning simulation model that captures the essential elements of the real-world process. If certain input parameters have little impact on the performance, then it is not necessary to spend a great deal of time or resources collecting data to build accurate empirical distributions. If, however, performance is found to vary greatly for small changes in distribution parameters, this is critical to identify in order to insure that good decisions are made. Finally, one may be able to assess *a priori* whether or not it would be valuable to add more complexity to component models. This will focus the modelers' efforts on refining the simulation in ways that add value, rather than simply adding to the run-time requirements.

7 CONCLUDING REMARKS

The approach outlined in this paper integrates the concepts of robust design with response surface metamodeling and system optimization efforts for discrete-event simulation. The simulation arena is amenable to analysis using robust design strategies since all factors are controllable by the analyst. The efficiency gained by designed experimentation is particularly beneficial for complex simulation models, since each realization of system performance corresponds to the results of a (potentially lengthy) run.

In summary, robust design can be a highly useful approach for analyzing models of complex systems for several reasons.

- *It is flexible*. Robust design can be applied to terminating or non-terminating simulation models analytic models, statistical models, or physical prototypes.
- *It is efficient*. Robust design can indicate when model components have sufficient detail. Sampling plans can incorporate simulation-specific artificial factors, and be chosen to keep either the data requirements or the analysts' time and effort low.
- *Solutions are realistic by design*. The suggested system configurations have already shown that they will behave well over a broad range of adverse conditions.
- *It facilitates continuous improvement*. Robust design clearly indicates important determinants of performance variation, and guides efforts for system 'optimization' and improvement by conveying hidden costs to the decision-makers.

ACKNOWLEDGMENTS

This work was supported in part by the National Research Council. An earlier version appeared in Sanchez (1994a).

REFERENCES

Box, G. E. P. 1988. Signal-to-noise ratios, performance criteria, and transformations (with discussions). *Technometrics* 30 (1): 1–40.

Box, G. E. P. and Draper, N. R. 1987. *Empirical model-building and response surfaces*. New York: John Wiley and Sons, Inc.

Box, G. E. P., W. G. Hunter and J. S. Hunter. 1978. *Statistics for experimenters*. New York: John Wiley and Sons, Inc.

Donohue, J. M., E. R. Houck and R. H. Meyers. 1992. Simulation designs for quadratic response surface models in the presence of model misspecification. *Management Science* 38 (12): 1765–1791.

Hood, S. J. and P. D. Welch. 1993. Response surface methodology and its application in simulation. In *Proc. of the 1993 Winter Simulation Conference*, ed. G. W. Evans, M. Mollaghasemi, E. C. Russell and W. E. Biles, 115–122. Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Jacobson, S., A. H. Buss and L. W. Schruben. 1991. Driving frequency selection for frequency domain simulation experiments. *Operations Research* 39 (6): 917–924.

Law, A. M. and W. D. Kelton. 2000. *Simulation modeling and analysis*. 3d ed. New York: McGraw-Hill.

Mitra, M. and S. K. Park. 1991. Solution to the indexing problem of frequency domain experiments. In *Proceedings of the 1991 Winter Simulation Conference*, ed. B. L. Nelson, W. D. Kelton and G. M. Clark, 907–915. Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Moeeni, F., S. M. Sanchez and A. J. Vakharia. 1997. A robust design methodology for kanban system design. *International Journal of Production Research* 35 (10): 2821-2838.

Montgomery, D.C. 1990. *Design and analysis of experiments*. New York: John Wiley and Sons, Inc.

Myers, R. H., A. I. Khuri and G. Vining. 1992. Response surface alternatives to the Taguchi robust parameter design approach. *The American Statistician* 46 (2): 131–139.

Nair, V. N. et al. 1992. Taguchi's parameter design: A panel discussion. *Technometrics*, 34 (2):127–161.

Pignatiello, J. J. Jr. and J.S. Ramberg. 1991. Top ten triumphs and tragedies of Genichi Taguchi. *Quality Engineering* 4 (2): 211–235.

Ramberg, J.S., J. J. Pignatiello, Jr. and S. M. Sanchez. 1992. A critique and enhancement of the Taguchi method. *ASQC Quality Congress Transactions* 491–498.

Ramberg, J. S., S. M. Sanchez, P. J. Sanchez and L. W. Hollick. 1991. Designing simulation experiments: Taguchi methods and response surface metamodels. In *Proceedings of the 1994 Winter Simulation Conference*, ed. B. L. Nelson, W. D. Kelton and G. M. Clark, 167–176. Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Sacks, J., W. J. Welch, T. J. Mitchell and H. P. Wynn. 1989. Design and analysis of computer experiments (with discussion). *Statistical Science* 4:409–435.

Sanchez, P. J. and A. H. Buss. 1987. A model for frequency domain experiments. In *Proceedings of the 1987 Winter Simulation Conference,* ed. A. Thesen, H. Grant, W. D. Kelton, 424–427. Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Sanchez, P. J. and S. M. Sanchez. 1991. Design of frequency domain experiments for discrete-valued factors. *Applied Mathematics and Computation* 42 (1): 1–21.

Sanchez, S. M. 1994a. A robust design tutorial. In *Proceedings of the 1994 Winter Simulation Conference,* ed. J. D. Tew, M. S. Manivannan, D. A. Sadowski and A. F. Seila, 106–113. Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Sanchez, S. M. 1994b. Experimental designs for system assessment when noise factors are correlated. In *Proceedings of the 1994 Winter Simulation Conference,* ed. J. D. Tew, M. S. Manivannan, D. A. Sadowski and A. F. Seila, 290–296. Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Sanchez, S. M., J. S. Ramberg, J. Fiero and J. J. Pignatiello, Jr. 1993. Quality by design. Ch. 10 in *Concurrent engineering: Automation, tools, and techniques*, ed. A. Kusiak, 235–286. New York: John Wiley and Sons, Inc.

Sanchez, S. M., P. J. Sanchez and J. S. Ramberg. 1998. A simulation framework for robust system design. Ch. 12 in *Concurrent Design of Products, Manufacturing Processes and Systems*, ed. B. Wang, 279-314. New York: Gordon and Breach.

Sanchez, S. M., L. D. Smith and E. C. Lawrence. 1996. Sensitivity and scenario analysis for simulation metamodels. *Proceedings of the 1996 Winter Simulation Conference,* eds. J. M. Charnes, D. J. Morrice, D. T. Brunner and J. J. Swain, 1440-1447. Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Sanchez, S. M., L. D. Smith and E. C. Lawrence. 2000. Tolerance design revisited: Assessing the impact of correlated noise factors. Working paper, Operation Research Department, Naval Postgraduate School, Monterey, CA.

Schruben, L. W. and V. J. Cogliano. 1987. An experimental procedure for simulation response surface model identification. *Communications of ACM*, 30 (8), 716–730.

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

Schruben, L. W., S. M. Sanchez, P. J. Sanchez and V. A. Czitrom. 1992. Variance reallocation in Taguchi's robust design framework. In *Proceedings of the 1992 Winter Simulation Conference*, ed. J. J. Swain, D. Goldsman, R. C. Crain and J. R. Wilson, 548–556. Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Taguchi, G. 1986. *Introduction to quality engineering.* White Plains, New York: UNIPUB/Krauss International.

Taguchi, G. 1987. *System of experimental design, vols. 1 and 2.* White Plains, New York: UNIPUB/Krauss International.

Taguchi, G. and Y. Wu. 1980. *Introduction to off-line quality control.* Nagoya, Japan: Central Japan Quality Association.

Tew, J. D. 1992. Using central composite designs in simulation experiments. In *Proceedings of the 1992 Winter Simulation Conference*, ed. J. J. Swain, D. Goldsman, R. C. Crain and J. R. Wilson, 529–538. Institute of Electrical and Electronic Engineers: Piscataway, NJ.

Tew, J. D. and J. R. Wilson. 1991. Validation of simulation analysis methods for the Schruben–Margolin correlation–induction strategy. *Operations Research* 40:87–103.

Tew, J. D. and Wilson, J. R. 1994. Estimating simulation metamodels using combined correlation-based reduction techniques. *IIE Transactions*, 26 (3): 2–16.

Vining, G. G. and R. H. Myers. 1990. Combining Taguchi and response surface philosophies: A dual response approach. *Journal of Quality Technology* 22:38–45.

Welch, W. J., T. K. Yu, S. M. Kang, and J. Sacks. 1990. Computer experiments for quality control by robust design. *Journal of Quality Technology* 22:15–22.

Wild, R. H. and J. J. Pignatiello, Jr. 1991. An experimental design strategy for designing robust systems using discrete-event simulation. *Simulation* 57 (6): 358–368.

AUTHOR BIOGRAPHY

SUSAN M. SANCHEZ is a Professor in the Operations Research Department at the Naval Postgraduate School, where she also holds a joint appointment in the Systems Management Department. She received a B.S. degree in Industrial & Operations Engineering from the University of Michigan, and M.S. and Ph.D. degrees in Operations Research from Cornell University. She is a member of INFORMS, DSI, ASA and ASQ. She serves as associate editor for *Operations Research* and *Naval Research Logistics*, and is currently the Vice President/President Elect of the INFORMS College on Simulation. Her email and web addresses are `<ssanchez@nps.navy.edu>` and `<http://diana.or.nps.navy.mil/ssanchez>`.