DESIGNING SIMULATION EXPERIMENTS: TAGUCHI METHODS AND RESPONSE SURFACE METAMODELS

John S. Ramberg
Susan M. Sanchez
Paul J. Sanchez
Ludwig J. Hollick

The University of Arizona
Tucson, Arizona 85721

ABSTRACT

Genichi Taguchi has made an innovative contribution to quality planning activities through the integrated use of loss functions and orthogonal arrays. In this tutorial, we focus on the improvement and implementation of certain of these techniques in the simulation arena. The orthogonal arrays advocated by Taguchi are related to classical experimental designs, which have played important tactical roles in the exploration of simulation model output and the construction of mathematical metamodels for the simulation response surface. However, the loss function and the associated robust design philosophy provide fresh insights into the process of optimizing or improving the simulation's performance. We use examples to illustrate concepts such as the simultaneous treatment of variability and mean of performance measures, strategies for achieving system robustness, and implementation of noise (uncontrollable variation) through factorial designs. We also discuss relationships to other issues in designing and analyzing simulation experiments, such as response surface metamodels and variance reduction. The tutorial is meant for both practitioners and researchers. We assume a knowledge base at the level of Chapters 11 and 12 of *Simulation Modeling and Analysis* (Law and Kelton, 1991), but will review essential elements in the presentation.

1 INTRODUCTION AND SCOPE

The exploration of simulation models and metamodels of their response surfaces has been assisted by experimental design methods and techniques developed specifically for simulation such as variance reduction and perturbation methods. We illustrate how techniques drawn from Taguchi's strategy for process improvement, developed for the purposes of design for quality in manufacturing, are applicable in simulation settings as well.

Taguchi's philosophies and teachings have sparked a great deal of both criticism and support in quality-engineering and related circles. Kackar (1985), Pignatiello (1988) and Ramberg and Pignatiello (1985, 1991), are just a few of the many articles in the literature that reflect this mixed reaction. One of Taguchi's notable contributions has been his strategy for incorporating variability measures into the evaluation of alternative systems. He pioneered this approach because he found that it was often more costly to control causes of manufacturing variation than to make a process insensitive to these variations. We refer the reader to Ramberg (1989) for manufacturing examples.

Taguchi's three-stage approach for quality improvement activities consists of *system design, parameter design*, and *tolerance design*. System design is the application of scientific and engineering knowledge to produce a functional prototype model. This prototype model defines the product/process design characteristics (parameters) and their initial settings. The goal of parameter design is the identification of settings that minimize variation in the performance characteristic and adjust its mean to an ideal value. Tolerance design is a method for scientifically assigning tolerances in order to minimize total product manufacturing and lifetime costs. In the simulation context, system design might correspond to building and validating a functional model of an existing real-world system or a prospective new facility, process, or product. Parameter design would be appropriate for attempting to "optimize" or "improve" performance of the simulation model by judiciously selecting settings for some of the decision factors in the model. Tolerance design can be viewed as a special case of parameter design for a single system configuration, with the goal of characterizing the simulation response and its sensitivity to randomly distributed inputs.

In Section 2, we discuss the concepts common to all three stages of the improvement process: perfor-

mance measures and their evaluation, the classification of controllable and uncontrollable variation, and experimental designs and analysis. Analytic methods, Monte Carlo distribution sampling, and a designed experiment approach are then used to illustrate tolerance design for an electrical circuit in Section 3. Static systems such as these have been the focus of tolerance design activities, but this simple example offers fresh insight into the process of meta-model building and system optimization. In Section 4, we present an example of parameter design for a dynamic stochastic simulation model. Our conclusions are provided in Section 5.

2   THE TAGUCHI FRAMEWORK

2.1   Performance Characteristics

The first step in designing an experiment is to identify the *performance characteristic(s)* of interest. In queueing system simulations, for example, the steady-state mean waiting time of customers in the queue is one widely-used measure. Percentiles of the steady-state distribution of waiting time are other performance characteristics which have received attention.

In order to determine the degree of satisfaction with the performance characteristic, an ideal state must be specified for comparison purposes. Such an ideal state is called the *target value* for the performance characteristic. Three types of targets, and associated goals, are considered. First, one could strive to make the performance characteristic as small as possible. For example, in a manufacturing simulation, one might seek to identify the system which minimizes the (average) time between order arrival and delivery of the product. Second, one could strive to make the performance characteristic as large as possible. For example, the goal of exploring a model of a nuclear power plant which runs until failure (a terminating simulation) might be to identify systems where the expected time until failure is quite large. Finally, one could strive to make the performance characteristic as close to some ideal state as possible. An economic simulation model, designed to assist the Federal Reserve Board in manipulating money supplies and interest rates to exercise control over the level of inflation, would fall into this category.

In systems where some stochastic variation is present, the performance characteristic will exhibit random fluctuation or variation around its target value. The cost of variation in the performance characteristic is thus something which must be measured. Taguchi advocates using *loss to society*, and has been commended for looking beyond manufacturing costs

and considering the costs incurred by end-users of the product. The *expected loss* is the expected value of the monetary losses that an arbitrary user of the product is likely to suffer during the product's life span due to performance variation. Unfortunately, it is usually difficult or impossible to specify an exact form for the underlying loss function. Instead, a quadratic loss is often utilized. Letting Y denote the performance characteristic (a random variable), τ denote the target value, and $\ell(\cdot)$ denote the loss function, we have

$$\ell(Y) = K(Y - \tau)^2 \qquad (1)$$

for some constant K. (Loss is assumed to be zero when Y achieves the ideal state.) It follows that the expected loss

$$
\begin{aligned}
L &= E[\ell(Y)] = KE[(Y - \tau)^2] & (2)\\
&= K\left[Var[Y] + (E[Y] - \tau)^2\right]\\
&\equiv K[MSE]
\end{aligned}
$$

Quadratic loss functions have been used in a variety of other settings, e.g., minimizing mean squared error loss is the basis for linear regression. It is also intuitively appealing: small deviations from the target value τ have little impact on the loss, while large deviations from τ will result in extremely large losses. Thus, in the context of system evaluation, the expected loss will tend to be low if most noise in the system inputs is not transmitted to response, and tend to be high for systems where the response variability is large.

Despite its apparent simplicity, the concept of expected loss redefines optimality in the context of simulation optimization. Attaining the ideal state in expectation is neither sufficient nor necessary for optimality: joint consideration of the mean *and variability* of the response is necessary. In contrast, the traditional emphasis of output analysis for simulation optimization efforts has been the valid estimation of expected performance, with variability viewed as a nuisance to be minimized as much as possible.

2.2   Parameter Design versus Noise Factors

In order to achieve systems/products with little variation in performance characteristics, several steps are necessary. First, one must identify factors in the system which are anticipated to affect the system response. Factors are classified as *parameter design factors* (hereafter called *parameters*) or *noise factors*, where the noise can result from sources either internal or external to the system. In a real-world setting, the

parameters are all those over which it is possible to exercise control, while the noise factors are not easily controllable or controllable only at great expense. Although all factors are truly controllable when the system under investigation is a computer simulation, the classification into parameter and noise factors can still be made on the basis of their controllability in the real-world system for which the simulation is a model.

2.3 Experimental Design and Analysis

Potential model configurations result from changing the settings of some or all parameters in the system (simulation model). Sometimes there may be managerial reasons for limiting investigation to a few alternatives. For example, the simulation might have been written to assess which station within a multi-server queueing network would yield the greatest improvement in processing time via the addition of another server. If, due to budgetary constraints, these are the only alternatives which can be considered (at least in the short term), then simulations of all alternatives could be conducted and evaluated to determine which is preferable.

Alternatively, experimental design techniques can also be utilized to observe how the expected value of the performance measure varies across several different system configurations. If the parameters are quantitative, then experimental design and analysis can lead to mathematical models of the simulation model (metamodels) of the system response. 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 parameter 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.

Traditionally, the parameter space has been the focus of experiment design, and analysts have assumed that all randomness between replications at a single design configuration is due to the influence of random, uncontrollable noise factors. Taguchi, on the other hand, incorporates the identifiable noise factors into the experimental design by crossing a parameter matrix with a noise factor matrix. Output from each configuration in the parameter space (commonly a simulation run) is then calculated across the noise factor space, in order to determine the parameter levels which result in the smallest sensitivity of

the response to the noise factors. For example, configurations C_1 and C_2 in Figure 1 have the same mean response resulting from variation of the noise factor around its mean μ_N. However, much less of the noise variability is transmitted to the response for C_2 than for C_1.
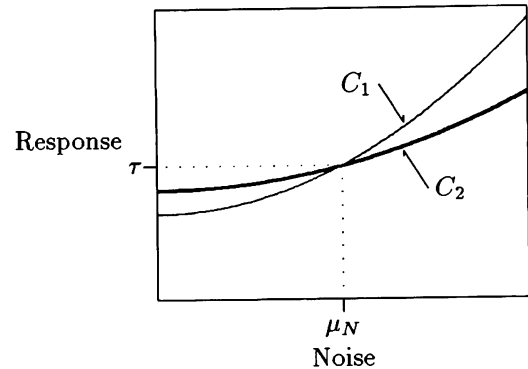

Figure 1: Response Sensitivity to Noise

Taguchi advocates the use of orthogonal arrays for the parameter and noise designs, and the use of an appropriately selected signal-to-noise ratio to systematically evaluate the tradeoffs between a performance characteristic's variability and its deviation from the target value. These tactical issues are the basis of most of the controversy concerning Taguchi's methods (Pignatiello and Ramberg, 1991). We implement the Taguchi strategy using the more conventional factorial designs. Typically, either two or three levels are chosen for each parameter: designs with two levels are simpler and require fewer experimental runs, but designs with three levels allow one to model quadratic as well as linear effects. Complete factorial designs (CFDs) allow estimation of all interaction effects (or all but one if a single replicate is used). Fractional factorial designs (FFDs) allow one to reduce the number of experimental runs if one is willing to assume that high order interactions are less important. While Taguchi combines the mean and variance of the performance characteristic into a signal-to-noise ratio, the drawbacks to that approach are well-documented in the literature (Box, 1988). We thus advocate joint analysis of the two measures.

3 TOLERANCE DESIGN

Ideally, tolerances for manufactured products are set to reflect the amount of variation in the performance characteristic anticipated to result from un-

controllable variation in the manufacturing process and component parts. In the simulation context, we can view tolerance design as a means of understanding system performance for a single configuration of parameters. Random fluctuation in the response is then attributable solely to noise factors.

Taguchi breaks down tolerance design into three phases (D'Errico and Zaino, 1988). In the first phase, *system evaluation*, a metamodel of the system response is constructed and the overall mean and variance of the response are determined. In the second phase, *noise factor assessment*, the sensitivity of the response to variation in each of the noise factors is determined. As we will show, these phases can be implemented using Monte Carlo distribution sampling or, as Taguchi advocates, designed experiments. The Taylor series approximation method is another alternative when the system model is analytical.

The final phase of tolerance design is *system optimization*. If some parameters are no longer considered fixed (i.e., alternative configurations are possible), then noise factor assessments can be combined with cost information to determine whether or not adjustments to the system will decrease expected loss.

3.1 Initial Circuit Configuration

Consider an electrical circuit with a performance measure of current I (in amps) and a target $\tau = 10$ amps. The analytical expression for the system is:

$$I = \frac{V}{\sqrt{R^2 + (2\pi f L)^2}} \quad (3)$$

where V is the voltage (in volts), R is the resistance (in ohms), f is the frequency (in Hertz), and L is the inductance (in Henries). In the first phase, systems design, L and R were specified as parameters of the circuit configuration. In the second phase, their nominal values were determined as part of the parameter design process in order to satisfy the application. Finally, in tolerance design, the variability of these components around their nominal values is evaluated. Thus L and R are treated as noise factors.

Table 1: Component Distribution Models

Component	Mean	Std. Dev.	Percent of Mean
L	0.004	0.0008	20%
R	10	1.0	10%
V	100	5.0	5%
f	50	5.0	10%

V and f are also sources of noise: their distributions should be established from information on the environment in which the circuit will operate.

We assume that all four factors are mutually independent, although the components might well be correlated due to the manufacturing process. Means and standard deviations for the factors are shown in Table 1. For the Monte Carlo approach, we also assume that the distributions are Normal.

3.2 System Evaluation

We first illustrate system evaluation using analytical methods. Letting $H(V, R, L, f)$ denote the right-hand side of Equation (3) and μ denote the vector $(\mu_L, \mu_R, \mu_V, \mu_f)$, one can approximate the performance characteristic I with the following Taylor series expansion:

$$
I \equiv H(\mu_L, \mu_R, \mu_V, \mu_f) + (L - \mu_L)\left.\frac{\partial H}{\partial L}\right|_\mu + \dots
$$
$$
+ (f - \mu_f)\left.\frac{\partial H}{\partial f}\right|_\mu + (L - \mu_L)^2 \left.\frac{\partial^2 H}{\partial L^2}\right|_\mu + \dots
$$
$$
+ (f - \mu_f)\left.\frac{\partial^2 H}{\partial f^2}\right|_\mu + Error
$$

Rearranging terms, we have a linear metamodel for the mean response as shown in Equation (4). (Generalization to higher-order models would be possible if higher-order moments were calculated.)

$$
I \approx \mu_I + \left.\frac{\partial I}{\partial L}\right|_\mu L + \left.\frac{\partial I}{\partial R}\right|_\mu R + \left.\frac{\partial I}{\partial V}\right|_\mu V + \left.\frac{\partial I}{\partial f}\right|_\mu f \quad (4)
$$

Solving for the derivatives and then substituting the appropriate component characteristics of Table 1 yields the metamodel in the first column of Table 2. Table 2 also provides the estimates of the system's overall mean and variance obtained by evaluating the Taylor series approximations below.

$$
\mu_I \approx H(\mu_L, \mu_R, \mu_V, \mu_f) \quad (5)
$$

$$
\sigma_I^2 \approx \left.\left(\frac{\partial I}{\partial L}\right)^2\right|_\mu \sigma_L^2 + \left.\left(\frac{\partial I}{\partial R}\right)^2\right|_\mu \sigma_R^2 \quad (6)
$$
$$
+ \left.\left(\frac{\partial I}{\partial V}\right)^2\right|_\mu \sigma_V^2 + \left.\left(\frac{\partial I}{\partial f}\right)^2\right|_\mu \sigma_f^2
$$

Table 2: System Evaluation for Circuit Model

	Metamodel	Mean	Variance
Taylor Series	$I \approx 9.922 - 38.56L - 0.9762R + 0.0992V - 0.00308f$	9.922	1.200
Monte Carlo (Metamodel)	$I \approx 10.325 - 35.15L - 1.002R + 0.1004V - 0.0039f$	10.018	1.257
Monte Carlo (Direct)		10.032	1.313
Designed Experiment	$I \approx 10.145 - 40.86L - 0.9853R + 0.1005V - 0.00329f$	10.014	1.225

Alternatively, system evaluation could be carried out using observational data (from the real system) or Monte Carlo distribution sampling in a simulation setting. These are commonly used when analytical models are not amenable to analysis, and may be particularly beneficial if higher-order models are desirable. Separate random number streams are used for each component (which will assist in noise factor assessment) and substituted into Equation (3) in order to generate n values of the response, I_i ($i = 1, \ldots, n$). A (linear) metamodel of the mean response is

$$I \approx \hat{\beta}_0 + \hat{\beta}_1 L + \hat{\beta}_2 R + \hat{\beta}_3 V + \hat{\beta}_4 F \qquad (7)$$

where the $\hat{\beta}$s are the regression coefficients. By treating the regression coefficients as constants, the following estimates of the overall response mean and variance result:

$$\mu_I \approx \hat{\beta}_0 + \hat{\beta}_1 \mu_V + \hat{\beta}_2 \mu_R + \hat{\beta}_3 \mu_L + \hat{\beta}_4 \mu_F \qquad (8)$$

$$\sigma_I^2 \approx \hat{\beta}_0^2 + \hat{\beta}_1^2 \sigma_V^2 + \hat{\beta}_2^2 \sigma_R^2 + \hat{\beta}_3^2 \sigma_L^2 + \hat{\beta}_4^2 \sigma_F^2 \qquad (9)$$

(Although the regression coefficients are actually random variables, their variances are small relative to the factor variances if n is large.) Alternatively, the response mean and variance could be estimated directly from the output:

$$\mu_I \approx \bar{I} \equiv \frac{1}{n} \sum_{i=1}^{n} I_i, \quad \sigma_I^2 \approx \frac{1}{n-1} \sum_{i=1}^{n} \left(I_i - \bar{I} \right)^2 \quad (10)$$

The metamodel constructed from a regression of 1000 simulated values of current on the associated values for the four noise factors is shown in the second row of Table 2. (The adjusted $R^2 = 0.982$, indicating that the linear metamodel is a good fit over the given ranges of component values.) Table 2 also provides the estimates of overall system performance computed both from the metamodel and directly.

We now demonstrate how system evaluation can be carried out using a designed experiment. Two-level designs are sufficient for fitting a linear metamodel comparable to those from the Taylor series and Monte Carlo distribution sampling analyses. Our design matrix (a 2^{4-1} FFD, or 'two-level half-fraction') is shown in Table 3; a coded level of '-1' in a column indicates that the corresponding factor is set at its low level, while a '+1' indicates the factor is set to its high level. The system configuration is summarized in the treatment combination column. Here, letters indicate which factors are set at their high levels; all four factors are at their low levels in the '(1)' row. (We remark that for this design matrix, main effects and three-factor interactions are confounded, and two-factor interactions are confounded with each other. A CFD could be used in order to avoid this confounding and include the above interaction terms in the metamodel. We refer the reader to Box, Hunter, and Hunter (1978) or Montgomery (1990) for detailed presentations of factorial designs.)

Table 3: Design Matrix in Coded Levels

Trtmt Comb.	Coded Levels			
	L	R	V	F
(1)	-1	-1	-1	-1
lf	+1	-1	-1	+1
rf	-1	+1	-1	+1
lr	+1	+1	-1	-1
vf	-1	-1	+1	+1
lv	+1	-1	+1	-1
rv	-1	+1	+1	-1
lrvf	+1	+1	+1	+1

The low and high levels must then be determined. For two-level factorial designs, these levels are set to one standard deviation below and one standard deviation above the mean so that each of these discrete distributions has a mean and standard deviation equal to the specifications. Table 4 gives the design matrix in terms of these natural values, along with the responses computed from evaluating Equation (3) at the corresponding configurations. We emphasize that all randomness has been removed from the experiment: we control the levels of the noise factors rather

Table 4: Results of Designed Experiment

Trtmt Comb.	Response: I	Natural Levels			
		L	R	V	F
(1)	10.5026	0.0032	9	95	45
lf	10.3807	0.0048	9	95	55
rf	8.5931	0.0032	11	95	55
lr	8.5714	0.0048	11	95	45
vf	11.5796	0.0032	9	105	55
lv	11.5362	0.0048	9	105	45
rv	9.5133	0.0032	11	105	45
lvrf	9.4387	0.0048	11	105	55

than allowing them to vary.

As for the Monte Carlo sampling approach, we can construct a metamodel for current from Equation (7) and estimates of overall system performance from Equations (8) and (9). These are provided in the third row of Table 2. (The adjusted R^2 for the fitted regression line was 0.994.) The results obtained from the designed experiment are comparable to those from the Taylor series expansion and Monte Carlo distribution sampling methods.

3.3 Noise Factor Assessment

Assuming independence, the system variance is the sum of the variances transmitted by each noise factor. From Equations (6) and (9), we see that these transmitted variances are the product of the component variance and the squared coefficient in the metamodel.

The transmitted variance can be assessed directly if Monte Carlo distribution sampling is used. All noise factors except the one of interest are set to their respective mean values, and a new set of response values are generated. The random number stream for the component of interest is seeded as it was during the system evaluation stage. The variance of this new response data is an estimate of the variance transmitted by the noise factor of interest.

The transmitted variances, in amps^2 and contributions as percents of MSE, are presented in Table 5. These results indicate that the inductor and frequency noise factors are not important: their variability is not transmitted to variability in the response. On this basis, one could simplify the system model (over the given ranges of factor levels) and approximate Equation (3) by a functional relation of the form $I = V/R$.

The importance/unimportance of the factors is not the same as statistical significance/insignificance for the factors in the metamodels determined using regression. The latter is largely dependent on sample size: for sufficiently large n one will find all factors to be significant (since the underlying system is nonlinear) while if n is small then it may be the case that no individual factors are statistically significant. Regardless of sample size, results from noise factor assessment provide valuable information regarding the attribution of variability in the response to the different components.

3.4 System Optimization

The quadratic loss function means that losses will be minimized when the response is equal to the target value: in our example, when a circuit produces a current $I = 10$ amps. The basis for system optimization is the quantification of expected loss in dollars, so the benefits of altering the initial system configuration can be evaluated. For example, suppose that the cost of a 4 amp deviation from the target is assessed at \$150.00 per unit. Substituting this into the right-hand side of Equation (1) and solving for the constant K yields $K = \$9.375$ per amp^2 per unit. The expected loss can then be evaluated from Equation (2) using estimates of μ_I and σ_I^2 obtained in the system evaluation phase.

Now suppose that enhancements to the system can be made at some cost. For our example, two potential upgrades are possible: the 20% inductor can be replaced by a 5% inductor at an additional cost of \$2.00/unit, and the 10% resistor can be replaced by a 5% resistor at an additional cost of \$1.00/unit. (Since we have shown that a linear metamodel is reasonable, we can assess the benefits of the upgrades separately.) Each would reduce the variability of one component: replacing the inductor would decrease σ_L^2 by a factor of 16, while replacing the resistor would decrease σ_R^2 by a factor of 4. The transmitted variances for these components would also decrease by factors of 16 and 4, respectively. Thus, we can recompute the overall system variance for the circuit with an upgraded inductor as follows:

$$\sigma_I^2 = \frac{1}{4}\sigma_{t,L}^2 + \sigma_{t,R}^2 + \sigma_{t,V}^2 + \sigma_{t,f}^2$$

where the transmitted variances are those determined in the noise factor assessment phase. Similarly, an estimate of the overall system variance can be computed for the circuit with an upgraded resistor. For the transmission variances estimated using the designed experiment, we find that the expected loss remains at \$9.10 if the upgraded inductor is used, resulting in a net loss of \$2.00 per unit. However, upgrading the resistor decreases the expected loss to

Table 5: Noise Factor Assessment for Circuit Model

	Transmitted Variances (% of MSE)						
	Inductor (L)		Resistor (R)		Voltage (V)		Frequency (f)
Taylor Series	0.001	(00.0)	0.953	(79.4)	0.246	(20.5)	0.000	(00.0)
Monte Carlo (Metamodel)	0.001	(00.0)	1.003	(79.8)	0.252	(20.1)	0.000	(00.0)
Monte Carlo (Direct)	0.001	(00.0)	1.058	(80.5)	0.241	(18.3)	0.000	(00.0)
Designed Experiment	0.001	(00.1)	0.971	(79.3)	0.253	(20.6)	0.000	(00.0)

$6.82, for a net savings of $5.82 per unit. Thus, although the relative improvement in the inductor is greater than that of the resistor, only the resistor upgrade is cost-effective.

3.5 Implications

First, we have shown the type of information that it is possible to obtain using tolerance analysis. This provides new insights into evaluation of simulation systems, namely the assignment of variability in the response to variability in the noise factors and the use of this information for system improvement/optimization.

Second, we have shown that the designed experiment gave comparable results to the other methods. However, it has benefits over both. It can be used in situations for which no analytical models are known, or Taylor series expansions of analytical models are difficult to compute. This is particularly true of complex systems: indeed, simulation models are often built because of the difficulty in obtaining analytical results. The designed experiment also requires many fewer runs than does the Monte Carlo distribution sampling approach. The benefits of designed experimentation are even more apparent when the "system model" under investigation is itself a simulation model, since each realization of system performance correspond to results of a (potentially lengthy) run.

4 PARAMETER DESIGN

The system used for this application is a variation of a dynamic, stochastic simulation model taken from Law and Kelton (1991). We describe the problem in Section 4.1, and discuss alternative model configurations and design considerations in Sections 4.2 and 4.3. In Section 4.4 we discuss issues related to the simulation metamodel and preparation of the simulation output for analysis. Results are summarized in Section 4.5: handouts containing more detailed results will be provided at the tutorial.

4.1 Initial Job Shop Simulation Model

A manufacturing shop consists of five groups of machines, and at present groups 1 through 5 consist of 3, 2, 4, 3, and 1 identical machines, respectively. Jobs are assumed to arrive at the shop according to i.i.d. exponential interarrival times with a mean of 0.40 hours. There are three types of jobs, and arriving jobs are of type 1, 2, or 3 with respective probabilities 0.3, 0.5, and 0.2.

Batch setup times, which we assume do not vary by job type, are fixed at 0.07, 0.10, 0.075, 0.10, and 0.03 hours for machine groups 1 through 5, respectively. Processing times are assumed to be independently distributed 2-Erlang random variables. The routings and mean processing times at a station differ by job type, and are provided in Table 6. For example, jobs of type 1 are first routed to machine group 3, where they require an average of 0.50 hours each in processing. They then proceed to machine group 1, and so forth. If a job arrives at a particular machine group and finds all machines in that group already busy, then the job joins a single FIFO queue at that machine group.

Table 6: Job Processing Requirements

Job Type	Machine Groups in Routing	Mean Times (hrs), Successive Tasks
1	3,1,2,5	0.50, 0.60, 0.85, 0.50
2	4,1,3	1.10, 0.80, 0.75
3	2,5,1,4,3	1.20, 0.25, 0.70, 0.90, 1.00

4.2 Alternative Model Configurations

The above description represents only one possible configuration of the stochastic job shop model, such as might be constructed to determine whether an existing job shop configuration would be capable of

responding to the pattern of demand described. A metamodel of this job shop simulation would provide additional insights into the system performance, by allowing the analyst to compare several different potential configurations (e.g., by changing the number of machines in one or more machine groups). Alternatively, experimental design techniques could be used to construct metamodels of the expected value of the system (simulation model) response (Kelton 1988; Law and Kelton, 1991).

As a first step, the performance characteristic of interest must be specified. We choose to examine the time elapsed from a job arrival to the completion of that job (hereafter referred to as the *time in system*) for job type 3. The average, steady-state value of this random variable will be our performance measure. This single measure was choosen over multiple response variables to simplify the analysis for exposition purposes. Job type 3 was selected because it requires processing at all five machine groups.

The job shop model has several additional factors which could be investigated in more detail. We list several of these in Table 7, and classify each as either a parameter (e.g., controllable variable in the real system) or a source of noise (internal/external). Although it is not the focus of this paper, it is important to recognize that when improvement efforts are an ongoing process, factors which are uncontrollable in the short term may be considered parameters in the long term. For our example, purchasing new equipment may affect the mean and/or variability of the required processing time, while maintenance scheduled off shift may reduce or eliminate the problem of machine breakdowns. Even external noise, such as the product mix, may be influenced in the long term by pricing and advertising strategies.

4.3   Experimental Design

For exposition purposes, we limit ourselves to the study of five parameter and two noise factors. The parameters are the number of machines in the first four groups (M_i, $i = 1, \ldots, 4$) and a common batch size B for all job type/machine group combinations, while the noise factors (PJ_1 and PJ_2) describe the job mix. (PJ_3 is fixed by the values of the other two.) The remaining factors were fixed at the original levels or distributions according to the problem statement.

As previously discussed, we will use conventional factorial designs. In our example with five parameter and two noise factors, a single replicate of a 2^k complete factorial design (CFD) would require $25 \times 22 = 128$ experimental runs, while a 3^k CFD

Table 7: Potential Factors for Job Shop Model

Factor	Description	Classification
M_i	Number of machines in group i	Parameter
B_{ij}	Batch size for jobs of type j at group i	Parameter
T_{ij}	Time to process job of type j at group i	Noise (Internal)
S_{ij}	Batch setup time for job type j at group i	Noise (Internal)
ID_i	Time between breakdowns of machines in group i	Noise (Internal)
D_i	Downtime for a machine in group i	Noise (Internal)
PJ_j	Proportion of jobs of type j arriving at the shop	Noise (External)
IA_j	Interarrival time for jobs of type j	Noise (External)

would require $35 \times 32 = 2187$ runs. Clearly, a 3^k design imposes a much greater data gathering burden than does a 2^k design. We select instead a 2^{5-1} resolution V design for the parameter space, which insures that main effects and two-factor interactions are not aliased with one another, and augment it with a centerpoint. (This allows us to assess lack-of-fit of the linear model with many fewer runs than would be required by a 3^k CFD.) The augmented FFD for the parameter space is crossed with a 2^2 CFD for the noise space. A total of 68 runs are then required for a single replication of the experiment.

The levels must then be set for each of the factors. Normally, the choice of noise factor level settings is accomplished by first establishing their respective mean and variability. For two-level factorial designs, the levels are set to one standard deviation above and one standard deviation below the mean (as we did in the tolerance design example). However, Kelton (1988) observes that the choice of parameter level settings is not nearly so clear, and requires some knowledge of how the system behaves. In general, settings ought to be selected such that the response surface contained within the parameter boundaries can be described by a linear or quadratic equation. The parameter/factor settings established for our example experiment are shown in Table 8.

Designing Simulation Experiments

Table 8: Factor Level Settings

		Low	Center	High
Parameters	M_1	3	4	5
	M_2	2	3	4
	M_3	4	5	6
	M_4	3	4	5
	B	1	3	5
Noise	PJ_1	0.25	0.30	0.35
	PJ_2	0.45	0.50	0.55

4.4 Simulation Coding and Data Collection

A program for the system described above was implemented in C++ and validated by a structured walkthrough. The simulation was designed so that either independent or common random number streams could be utilized. Decoupling the system means that 15 random number streams must be utilized.

The performance characteristic is the average steady-state time in the system. Initially, the simulation was run at a saturated design point for 25,000 simulated hours. Correlogram analysis indicated that observations 900 apart were virtually independent. This determined the truncation point for removing the initial transient in the production runs.

The mean response for parameter configuration i and noise configuration j, denoted by Y_{ij}, is estimated by the sample average of the output from the corresponding run (after truncation). A measure of the variability of the performance characteristic for each design point is needed. Since the data are correlated, standard statistical measures are not appropriate. Instead, we compute an unbiased estimate of σ_Y^2 by noting that observations 900 apart can be considered virtually independent. Let N_b denote the number of batches of size 900 (after deletion of the initial transient), and let Y_{ijkb} denote the kth observation in batch b for parameter configuration i and noise configuration j. Then for fixed k ($1 \le k \le 900$), the set of observations $\{Y_{ijkb}\ (b = 1, \ldots, N_b)\}$ can be used to construct a standard unbiased estimator of σ_Y^2 as follows:

$$s_{ijk}^2 = \frac{1}{N_b - 1} \sum_{b=1}^{N_b} \left(Y_{ijkb} - \overline{Y}_{ijk.} \right)^2$$

The s_{ijk}^2, although correlated, can then be averaged across k to obtain the following unbiased estimate of

the variability of the system response:

$$\overline{s}_{ij}^2 = \frac{1}{900} \sum_{k=1}^{900} s_{ijk}^2$$

We remark that this differs from standard simulation analysis, where estimates of the variability of the *mean response*, rather than that of the *system response*, are of interest in order to construct confidence intervals for the mean response. Mathematically, for a covariance-stationary stream of data of length n,

$$\sigma_{\overline{Y}}^2 = \frac{\sigma_Y^2}{n} \left[1 + \sum_{j=1}^{n-1} \left(1 - \frac{j}{n} \right) \rho_j \right] \quad (11)$$

where σ_Y^2 is the variance of a single observation and ρ_k is the lag k autocorrelation of the output. Methods for obtaining confidence intervals for the mean response, such as those based on batch means or replication, estimate $\sigma_{\overline{Y}}^2$ directly. However, we are unaware of a stable method for inverting Equation (11) to obtain an estimate of the system variance, σ_Y^2.

4.5 Results

Crossing of the parameter matrix and the noise matrix means that for each configuration of the parameters, runs are made for four configurations of the noise factors. For each parameter configuration i, we compute the following overall measures of performance mean and variation:

$$\overline{Y}_{i\cdot} = \frac{1}{4} \sum_{j=1}^{n} \overline{Y}_{ij}, \qquad \overline{s}_{i\cdot}^2 = \frac{1}{4} \sum_{j=1}^{n} \overline{s}_{ij}^2,$$

although we remark that other constructs of overall variability could have been used.

A regression of $\overline{s}_{i\cdot}$ on $\overline{Y}_{i\cdot}$ indicated that a square-root transformation of the $\overline{s}_{i\cdot}$ was appropriate. (The square-root transformation is often utilized for Poisson data to stabilize the variance.) Metamodels with linear terms and all two-way interactions were constructed for $\overline{y}_{i\cdot}$ and $\sqrt{\overline{s}_{i\cdot}}$ using linear regression. Investigation revealed that M_2, M_4, B and the M_2B interaction were the most important determinants of both the mean and the variability of the response. Residual analysis also showed that M_2 affected the mean performance in a quadratic manner. This term could be included in the metamodel for the mean response because of the additional degree of freedom afforded by the center point.

These results were obtained from runs where all random number streams were independently seeded.

Common random numbers can also be employed: using the same sets of random number seeds for each run within a replication of the design matrix is a way of "blocking" on the replication. Results obtained using common random numbers will be discussed at the presentation.

5   CONCLUSIONS

We have shown how joint investigation of the mean and variation in the performance measure may provide additional insights into system performance. We have related these strategies to response surface meta-models and illustrated how the results may be used to improve system performance. The simulation arena is particularly amenable to analysis using Taguchi's strategies, since all factors are controllable by the analyst. Our results indicate that models of system performance may be more easily constructed if some of the random variation, or noise, in the simulation inputs is controlled in an appropriate manner.

Consideration of variance as other than a nuisance represents a significant departure from traditional simulation output analysis, and restricts but does not eliminate the use of variance reduction techniques. Reducing the variability *between* runs (as we did by using common random number streams in our examples) may still be beneficial. However, any technique which attempts to reduce the variance *within* a simulation run eliminates the ability to estimate the system variance. This precludes noise factor assessment, as well as system optimization efforts based on expected loss.

ACKNOWLEDGMENTS

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., W.G. Hunter and J.S. Hunter. 1978. *Statistics for Experimenters*, John Wiley and Sons, Inc., New York, NY.

D'Errico, J.R. and N.A. Zaino. 1988. "Statistical Tolerancing Using a Modification of Taguchi's Method," *Technometrics*, 30(1), 397–405.

Kackar, R.N. 1985. "Off-Line Quality Control, Parameter Design, and the Taguchi Method," *Journal of Quality Technology* 17(4), 176–188.

Kelton, W.D. 1988. "Designing computer simulation experiments," *Proceedings of the 1988 Winter Simulation Conference*, eds. M.A. Abrams, P.L. Haigh, and J.C. Comfort, 15–18. Institute of Electrical and Electronics Engineers, San Francisco, California.

Law, A.M. and W.D. Kelton. 1991. *Simulation Modeling and Analysis*, McGraw-Hill, New York, NY.

Montgomery, D.C. 1990. *Design and Analysis of Experiments*, John Wiley and Sons, Inc., New York.

Pignatiello, J.J. 1988. "An Overview of the Strategy and Tactics of Taguchi," *IIE Transactions*, 20(3), 247–254.

Pignatiello, J.J. and J.S Ramberg. 1985. "Discussion of Off-Line Quality Control, Parameter Design, and the Taguchi Method," *Journal of Quality Technology*, 17(4), 198–206.

Pignatiello, J.J. and J.S Ramberg. 1991. "Top ten triumphs and tragedies of Genichi Taguchi," To appear in *Quality Engineering*.

Ramberg, J.S. 1989. "Improvements on Taguchi Methods for Semiconductor Design/Manufacture," *SME Semiconductor Manufacturing Conference Papers*, MS89-798, 1–16.

AUTHOR BIOGRAPHIES

JOHN S. RAMBERG is a professor in the Department of Systems and Industrial Engineering at The University of Arizona. His research interests include quality engineering, engineering statistics, and simulation. He is a fellow of ASA, ASQC, and IIE.

SUSAN M. SANCHEZ is an assistant professor in the MIS/Decision Sciences at The University of Arizona. Her research interests include applied statistics, selection and experimental design. She is a member of ORSA, TIMS, ASA, and ASQC.

PAUL J. SANCHEZ is an assistant professor in the Department of Systems and Industrial Engineering at The University of Arizona. His research interests include various aspects of simulation. He is a member of TIMS, and was the local arrangements chair for the 1991 Winter Simulation Conference.

LUDWIG J. HOLLICK is an Aerospace Engineering Officer in the Canadian Air Force pursuing a master's degree in Reliability Engineering at The University of Arizona.