Proceedings of the 2008 Winter Simulation Conference
S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds.

STOCHASTIC KRIGING FOR SIMULATION METAMODELING

Bruce Ankenman
Barry L. Nelson
Jeremy Staum

Department of Industrial Engineering & Management Sciences
Northwestern University
Evanston, IL, U.S.A.

ABSTRACT

We extend the basic theory of kriging, as applied to the design and analysis of deterministic computer experiments, to the stochastic simulation setting. Our goal is to provide flexible, interpolation-based metamodels of simulation output performance measures as functions of the controllable design or decision variables. To accomplish this we characterize both the intrinsic uncertainty inherent in a stochastic simulation and the extrinsic uncertainty about the unknown response surface. We use tractable examples to demonstrate why it is critical to characterize both types of uncertainty, derive general results for experiment design and analysis, and present a numerical example that illustrates the stochastic kriging method.

1 INTRODUCTION

Discrete-event simulation is a general-purpose tool for analyzing dynamic, stochastic systems. Virtually any level of detail can be modeled and any performance measure estimated, which explains simulation's popularity. However, simulation models are often tedious to build, need substantial data to parameterize, and require significant time to run, particularly when there are many alternatives to evaluate.

The objective of the methodology described in this paper is to get more benefit from a simulation investment. The specific context we have in mind is when time to exercise the simulation model in advance of the decision making it will support is relatively plentiful, but decision-making or decision-maker time is relatively scarce or expensive. Therefore, rather than executing a simulation run whenever a "what if" question is posed, or trying to anticipate every scenario of interest in advance, we use the simulation to "map" the performance response surfaces of interest as functions of the controllable design or decision variables. Ideally, these response surface maps provide the fidelity of the full simulation model with the ease of use of, say, a spreadsheet model.

Using simulation to construct metamodels (models of the simulation model) is not new (see Barton and Meckesheimer 2006 for a review). Starting with classical response-surface modeling in statistics (e.g., Myers and Montgomery 2002), simulation researchers have adapted experiment designs for linear regression models to account for dependence within a replication for steady-state simulations (e.g., Law and Kelton 2000); to permit the use of common random numbers (CRN) and antithetic variates across design points (e.g., Schruben and Margolin 1978, Nozari et al. 1987, Tew and Wilson 1992, 1994); and to compensate for the strong relationship between response variance and customer load in queueing simulations (e.g., Cheng and Kleijnen 1998, Yang, Ankenman and Nelson 2007). However, linear regression models (that are usually polynomials in the design variables and linear in their unknown coefficients) tend to fit well locally but do not provide the sort of robust global maps we desire. Nonlinear models based on queueing theory work very well for queueing simulations, but require domain knowledge of the problem context and specialized fitting algorithms.

We are interested in more general-purpose approaches that assume less structure than linear or queueing-specific nonlinear models; that tend to be more resistant to overfitting than general interpolators (e.g., neural networks, see for instance Sabuncuoglu and Touhami 2002); that facilitate sequential, adaptive experimental design rather than fixed, a priori designs; and that can provide statistical inference about when a good fit is obtained. We also want to account for the reality that the simulation output is stochastic, with variance that usually changes significantly across the design space.

To satisfy these requirements we extend the kriging methodology that is popular, and has been highly successful, in the design and analysis of (deterministic) computer experiments (DACE). DACE methodology is particularly well suited for systematically reducing uncertainty about the unknown response surface as experiments (computer runs at different design settings) are performed and leads to

978-1-4244-2708-6/08/$25.00

interpolation-based models. Our central contribution is to fully account for the sampling variability that is inherent to a stochastic simulation. We show that correctly accounting for both sampling and response-surface uncertainty has an impact on experiment design, response-surface estimation and inference.

In the next section we describe our extended metamodel under the special case that all model parameters are known; this setting allows us to demonstrate why the extension is critical without cluttering the discussion with estimation issues, which are resolved in Section 3. A numerical illustration and conclusions close the paper in Sections 4 and 5, respectively.

2 THE METAMODEL

We describe our approach by refining a sequence of models. We are interested in modeling an unknown performance-measure surface (or surfaces) $y(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2, \ldots, x_d)^\top$ is a vector of design variables and $y(\mathbf{x})$ is a deterministic function of $\mathbf{x}$. For instance, in a semiconductor fabrication simulation $\mathbf{x}$ might represent the release rates of d products and y could be the steady-state mean cycle time of product 1 (however, y need not be a mean).

The classical approach is to assume that the *observed* response obtained from the jth simulation replication at $\mathbf{x}$ is described by the model

$$Y_j(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \beta + \varepsilon_j(\mathbf{x}) \tag{1}$$

where $\mathbf{f}(\mathbf{x})$ is a vector of known functions of $\mathbf{x}$, β is a vector of unknown parameters of compatible dimension, and $\varepsilon_j(\mathbf{x})$ has mean 0 and represents the sampling variability inherent in a stochastic simulation. The distribution of $\varepsilon_j(\mathbf{x})$, and in particular its variance, may depend on $\mathbf{x}$, although this dependence is often ignored. We refer to ε as *intrinsic* uncertainty, because it comes from the nature of the stochastic simulation itself. An experiment design specifies settings of $\mathbf{x}$ at which to observe $Y(\mathbf{x})$, and the number of replications to obtain at each $\mathbf{x}$. In this paper we primarily address the replication setting (as opposed to the single-run experiment design sometimes used in steady-state simulation).

Now consider the following thought experiment: Suppose that the response $y(\mathbf{x})$ could be observed *without noise*, but we are still interested in developing a metamodel after observing $y(\mathbf{x})$ at a few design points $\mathbf{x}$. This problem is treated in the DACE literature (Kennedy and O'Hagan 2000, Sacks et al. 1989, Stein 1999, Santner et al. 2003). A remarkably successful approach is to cast this deterministic problem into a statistical framework by representing the unknown response surface as

$$\mathsf{Y}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \beta + \mathsf{M}(\mathbf{x}) \tag{2}$$

where M is a realization of a mean 0 *random field*; that is, we think of M as being randomly sampled from a space of functions mapping $\Re^d \to \Re$. The functions in this space are assumed to exhibit *spatial correlation*, which means that values $\mathsf{M}(\mathbf{x})$ and $\mathsf{M}(\mathbf{x}')$ will tend to be similar if $\mathbf{x}$ and $\mathbf{x}'$ are close to each other in space. We refer to the stochastic nature of M as *extrinsic* uncertainty, since it is imposed on the problem (not intrinsic to it) to aid in developing a metamodel. This paradigm embeds a deterministic problem into a probabilistic framework so that statistical concepts such as mean squared error (MSE) of estimation can be brought to bear. Statistical inference about $\mathsf{Y}(\mathbf{x})$ at values of $\mathbf{x}$ not simulated can aid experiment design and provide estimates of the metamodel's precision, a feature we want to exploit.

We argue that the following model is more useful than (1) or (2) for representing a stochastic simulation's output on replication j at design point $\mathbf{x}$:

$$\mathscr{Y}_j(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \beta + \mathsf{M}(\mathbf{x}) + \varepsilon_j(\mathbf{x}). \tag{3}$$

The intrinsic noise $\varepsilon_1(\mathbf{x}), \varepsilon_2(\mathbf{x}), \ldots$ at a design point $\mathbf{x}$ is naturally independent and identically distributed across replications, but we allow the possibility that $\mathsf{V}(\mathbf{x}) \equiv \mathrm{Var}[\varepsilon(\mathbf{x})]$ is not constant and that $\mathrm{Corr}[\varepsilon_j(\mathbf{x}), \varepsilon_j(\mathbf{x}')] > 0$ to model the effect of CRN. The intent of CRN is to reduce the variance of estimated differences through inducing positive correlation across design points by driving their simulations with the same sequence of pseudorandom numbers (see, for instance, Law and Kelton 2000). Later we propose simultaneously modeling M and V, which is a central contribution of this paper.

In our setting an experiment design consists of pairs $(\mathbf{x}_i, n_i), i = 1, 2, \ldots, k$, where n_i is the number of simulation replications taken at design setting $\mathbf{x}_i$. Let the sample mean at $\mathbf{x}_i$ be

$$\bar{\mathscr{Y}}(\mathbf{x}_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathscr{Y}_j(\mathbf{x}_i) \tag{4}$$

and let $\bar{\mathscr{Y}} = \left(\bar{\mathscr{Y}}(\mathbf{x}_1), \bar{\mathscr{Y}}(\mathbf{x}_2), \ldots, \bar{\mathscr{Y}}(\mathbf{x}_k)\right)^\top$.

We want a metamodel that predicts the response $\mathsf{Y}(\mathbf{x}_0) \equiv \mathbf{f}(\mathbf{x}_0)^\top \beta + \mathsf{M}(\mathbf{x}_0)$ at *any* $\mathbf{x}_0$, simulated or not. Until further notice we only consider the case $\mathbf{f}(\mathbf{x}_0)^\top \beta = \beta_0$ (that is, just a constant term representing the overall surface mean), because this model has tended to be the most useful in practice for DACE.

As is typical in spatial correlation models, we consider linear predictors of the form

$$\lambda_0(\mathbf{x}_0) + \boldsymbol{\lambda}(\mathbf{x}_0)^\top \bar{\mathscr{Y}} \tag{5}$$

where $\lambda_0(\mathbf{x}_0)$ and $\boldsymbol{\lambda}(\mathbf{x}_0)$ are weights that depend on $\mathbf{x}_0$ and are chosen to give the predictor good properties, such as minimum MSE for predicting $\mathsf{Y}(\mathbf{x}_0) = \beta_0 + \mathsf{M}(\mathbf{x}_0)$. Later, when we make Gaussian assumptions on the intrinsic and extrinsic uncertainty, this form drops out as the best predictor, linear or otherwise.

Let $\Sigma_\mathsf{M}(\mathbf{x},\mathbf{x}') = \mathrm{Cov}[\mathsf{M}(\mathbf{x}),\mathsf{M}(\mathbf{x}')]$ be the covariance implied by the extrinsic spatial correlation model, let Σ_M be the $k \times k$ covariance matrix across all design points $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_k$, and let $\Sigma_\mathsf{M}(\mathbf{x}_0,\cdot)$ be the $k \times 1$ vector $(\mathrm{Cov}[\mathsf{M}(\mathbf{x}_0),\mathsf{M}(\mathbf{x}_1)],\ldots,\mathrm{Cov}[\mathsf{M}(\mathbf{x}_0),\mathsf{M}(\mathbf{x}_k)])^\top$. Also let Σ_ε be the $k \times k$ covariance matrix with (h,i) element $\mathrm{Cov}\left[\sum_{j=1}^{n_h} \varepsilon_j(\mathbf{x}_h)/n_h, \sum_{j=1}^{n_i} \varepsilon_j(\mathbf{x}_i)/n_i\right]$ across all design points $\mathbf{x}_h$ and $\mathbf{x}_i$.

To illustrate the key issues, suppose that $\Sigma_\mathsf{M}, \Sigma_\varepsilon$ and β_0 are known (clearly, in a real application they need to be estimated, which is a contribution of our research). We can show that the MSE-optimal predictor of the form (5) is

$$\widehat{\mathsf{Y}}(\mathbf{x}_0) = \beta_0 + \Sigma_\mathsf{M}(\mathbf{x}_0,\cdot)^\top [\Sigma_\mathsf{M} + \Sigma_\varepsilon]^{-1} \left(\bar{\mathscr{Y}} - \beta_0 \mathbf{1}_k\right) \tag{6}$$

where $\mathbf{1}_k$ is the $k \times 1$ vector of ones. We refer to this predictor as *stochastic kriging*. Notice that the only computationally intensive operation in evaluating (6) is the matrix inversion, which is done once since it is independent of $\mathbf{x}_0$. If there were no intrinsic uncertainty due to simulation, Σ_ε would vanish and (6) would reduce to the standard kriging estimator that matches the data $\bar{\mathscr{Y}}$ at design points, and predicts $\mathsf{Y}(\mathbf{x}_0)$ by a weighted average of $\bar{\mathscr{Y}}$ elsewhere (e.g., Cressie 1993). Equation (6) clearly shows that the presence of intrinsic uncertainty impacts the prediction everywhere on the surface. We can also show that the optimal MSE is

$$\begin{aligned} \mathrm{MSE}^\star &= \Sigma_\mathsf{M}(\mathbf{x}_0,\mathbf{x}_0) - \Sigma_\mathsf{M}(\mathbf{x}_0,\cdot)^\top [\Sigma_\mathsf{M} + \Sigma_\varepsilon]^{-1} \Sigma_\mathsf{M}(\mathbf{x}_0,\cdot) \\ &= \left[\Sigma_\mathsf{M}(\mathbf{x}_0,\mathbf{x}_0) - \Sigma_\mathsf{M}(\mathbf{x}_0,\cdot)^\top \Sigma_\mathsf{M}^{-1} \Sigma_\mathsf{M}(\mathbf{x}_0,\cdot)\right] \\ &+ \Sigma_\mathsf{M}(\mathbf{x}_0,\cdot)^\top \Xi \Sigma_\mathsf{M}(\mathbf{x}_0,\cdot) \end{aligned} \tag{7}$$

where Ξ is a positive definite matrix that depends on Σ_ε and Σ_M. The term in brackets in (7) is the usual kriging MSE; the additional term is positive, showing that intrinsic uncertainty inflates MSE.

To actually estimate a stochastic kriging metamodel from data we need $\Sigma_\mathsf{M}(\cdot,\cdot)$ to have more structure. In particular, we will assume that M is second-order stationary, meaning that

$$\Sigma_\mathsf{M}(\mathbf{x},\mathbf{x}') = \tau^2 R_\mathsf{M}(\mathbf{x}-\mathbf{x}';\theta) \tag{8}$$

where τ^2 can be interpreted as the variance of $\mathsf{M}(\mathbf{x})$ for all $\mathbf{x}$, and R_M is the correlation which depends only on $\mathbf{x}-\mathbf{x}'$ and may be a function of some unknown parameters θ. Further, we will require that $R_\mathsf{M}(\mathbf{x}-\mathbf{x}';\theta) \to 0$ as the distance between $\mathbf{x}$ and $\mathbf{x}'$ goes to infinity, and $\mathbf{R}_\mathsf{M}(\mathbf{0};\theta) = 1$.

Use of kriging for metamodeling in stochastic simulation was first mentioned by Mitchell and Morris (1992), but has only been explored in depth by Kleijnen and his collaborators; the papers most closely related to our work are van Beers and Kleijnen (2003) and Kleijnen and van Beers (2005) (see also Biles et al. 2007 and van Beers and Kleijnen 2007). The central idea in these papers is to first model out any trend using least squares or generalized least-squares techniques, and then to apply kriging to some form of standardized residuals. They do not incorporate a model of the intrinsic uncertainty, which means that they cannot be used for the sort of adaptive design we desire that jointly considers the placement of design points and simulation effort. To illustrate the insights gained from our approach, we examine a tractable example in detail.

Consider the case of $k = 2$ design points $\mathbf{x}_1$ and $\mathbf{x}_2$ with equal numbers of replications $n_1 = n_2 = n$. Suppose that

$$\Sigma_\mathsf{M} = \tau^2 \begin{pmatrix} 1 & r_{12} \\ r_{12} & 1 \end{pmatrix} \quad \text{and} \quad \Sigma_\mathsf{M}(\mathbf{x}_0,\cdot) = \tau^2 \begin{pmatrix} r_0 \\ r_0 \end{pmatrix}.$$

The term $\tau^2 > 0$ represents the extrinsic variance of M, r_{12} is the extrinsic correlation between $\mathsf{M}(\mathbf{x}_1)$ and $\mathsf{M}(\mathbf{x}_2)$, and r_0 is the extrinsic correlation between the point to be predicted $\mathsf{Y}(\mathbf{x}_0)$ and each of the design points (these usually would not be equal). Typically we expect r_{12} and r_0 to be positive.

For the intrinsic uncertainty due to sampling at a design point, suppose

$$\Sigma_\varepsilon = \frac{\mathsf{V}}{n} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

where in this example the variance at the design points is a common $\mathsf{V} > 0$, and $-1 \le \rho \le 1$ represents intrinsic dependence between the design points; for instance, we would expect $\rho > 0$ if we used CRN. Substituting these into (6)–(7), the MSE-optimal predictor of $\mathsf{Y}(\mathbf{x}_0)$ is $\widehat{\mathsf{Y}}(\mathbf{x}_0) =$

$$\beta_0 + \frac{2\tau^2 r_0}{(1+r_{12})\tau^2 + (1+\rho)\mathsf{V}/n} \left(\frac{\bar{\mathscr{Y}}(\mathbf{x}_1) + \bar{\mathscr{Y}}(\mathbf{x}_2)}{2} - \beta_0\right) \tag{9}$$

with MSE

$$\mathrm{MSE}^\star = \tau^2 \left(1 - \frac{2\tau^2 r_0^2}{(1+r_{12})\tau^2 + (1+\rho)\mathsf{V}/n}\right). \tag{10}$$

Equation (9) shows that stochastic kriging is a bit like a control-variate estimator (e.g., Nelson 1990), where a correction term is applied to the mean based on the deviation

of the observed responses from their expectations and the strength of the correlation (r_0) between the design points and the response to be predicted.

The MSE (10) is even more revealing: MSE is decreasing in r_0^2, meaning the stronger the correlation between the design points and the response at $\mathbf{x}_0$, the smaller the MSE because the design points provide more information. However, MSE is increasing in r_{12}, since the more correlated the design points themselves are, the less additional information they provide. Intrinsic uncertainty, V, also increases MSE, but can be reduced by increasing the sample size n. Most interesting is that the assumed impact of CRN, which is to make $\rho > 0$, *increases* MSE relative to independent sampling. This may seem surprising because in standard linear regression models such as (1) the impact of CRN is to reduce the variance of the slope coefficients. However, the stochastic kriging predictor is a weighted average of the outcomes from the design points, and CRN inflates the variance of averages. In fact, (10) shows that antithetic variates (e.g., Law and Kelton 2000), which tries to induce $\rho < 0$, would reduce MSE.

There are two messages in this example: (i) In stochastic kriging there is an important interplay between the placement of design points (through their extrinsic correlation with each other) and the simulation effort at the design points (through their intrinsic variance); and (ii) CRN will not be helpful for predicting $\mathsf{Y}(\mathbf{x})$ *in general.*

3 PARAMETER ESTIMATION

To actually apply stochastic kriging for simulation metamodeling, a method for estimating the unknown parameters is required. The DACE literature contains several methods and refinements when there is only extrinsic uncertainty; see for instance Santner et al. (2003) and Fang et al. (2006). Here we focus on extending the most well-known method—maximum likelihood—to allow for intrinsic uncertainty.

Recall that our model for the simulation output is

$$\mathscr{Y}_j(\mathbf{x}) = \beta_0 + \mathsf{M}(\mathbf{x}) + \varepsilon_j(\mathbf{x}).$$

We now adopt the following

Assumption 1 *The random field* M *is a stationary Gaussian random field, and* $\varepsilon_1(\mathbf{x}_i), \varepsilon_2(\mathbf{x}_i), \ldots$ *are i.i.d.* $\mathsf{N}(0, \mathsf{V}(\mathbf{x}_i))$*, independent of* $\varepsilon_j(\mathbf{x}_h)$ *for all* j *and* $h \neq i$ *(i.e., no CRN), and independent of* M*.*

That M is a stationary Gaussian random field is a standard assumption in DACE. We refer the reader to, for instance, Santner et al. (2003, §2.3.2) for technical details, but in brief this assumption implies that for any finite collection of design points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ the random vector $(\mathsf{M}(\mathbf{x}_1), \mathsf{M}(\mathbf{x}_2), \ldots, \mathsf{M}(\mathbf{x}_k))$ has a multivariate normal distribution with constant marginal mean 0, variance $\tau^2 > 0$, and positive definite correlation matrix $\mathbf{R}_\mathsf{M}$ such that $\mathrm{Corr}(\mathsf{M}(\mathbf{x}_i), \mathsf{M}(\mathbf{x}_h))$ depends only on $\mathbf{x}_i - \mathbf{x}_h$. The normality of $\varepsilon_j(\mathbf{x})$ could be anticipated if, for instance, the output of each replication was itself the average of a large number of more basic random variables (e.g., the average of hundreds of individual product cycle times in the semiconductor fabrication example).

Under Assumption 1, $(\mathsf{Y}(\mathbf{x}_0), \bar{\mathscr{Y}}(\mathbf{x}_1), \ldots, \bar{\mathscr{Y}}(\mathbf{x}_k))$ is multivariate normal and the stochastic kriging predictor (6) is the conditional expectation of $\mathsf{Y}(\mathbf{x}_0)$ given $\bar{\mathscr{Y}}$, making it the minimum MSE predictor (Santner et al. 2003, Theorem 3.2.1).

We begin by assessing the impact of estimating the intrinsic variance Σ_ε, then derive the maximum likelihood estimators given Σ_ε and conclude by addressing experiment design.

3.1 Estimating the Intrinsic Variance

In this section we confront the fact that V is typically unknown. In summary, our approach is as follows:

Because we are interested in sequential experiment design, we need a model for V. To obtain it, we will assume V is also represented by a spatial correlation model

$$\mathsf{V}(\mathbf{x}) = \sigma^2 + \mathsf{Z}(\mathbf{x}) \tag{11}$$

where Z is a mean zero stationary random field that is independent of M. Denote the estimated model by $\widehat{\mathsf{V}}(\mathbf{x})$.

Since $\mathsf{V}(\mathbf{x}_i)$ is not observable, even at the design points, we let

$$\mathscr{S}^2(\mathbf{x}_i) = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} \left(\mathscr{Y}_j(\mathbf{x}_i) - \bar{\mathscr{Y}}(\mathbf{x}_i)\right)^2 \tag{12}$$

stand in for it. Under Assumption 1, $\mathscr{S}^2(\mathbf{x}_i)$ is strongly consistent for $\mathsf{V}(\mathbf{x}_i)$ and has a scaled chi-squared distribution.

Because we observe $\mathscr{S}^2$, not V, there is extrinsic and intrinsic uncertainty, just as in estimating $\beta_0 + \mathsf{M}$ from $\bar{\mathscr{Y}}$. However, since we are not interested in V except as it impacts our design and analysis, we will ignore the intrinsic uncertainty and fit model (11) using standard kriging as if $\mathscr{S}^2$ had no noise. Therefore, $\widehat{\mathsf{V}}(\mathbf{x}_i) = \mathscr{S}^2(\mathbf{x}_i)$ at design points $\mathbf{x}_i$ since standard kriging interpolates the response at the design points exactly. We will show that the consequences of estimating V in this way are slight as long as the n_i are not too small.

We do not describe estimation of model (11) from $\mathscr{S}^2(\mathbf{x}_1), \mathscr{S}^2(\mathbf{x}_2), \ldots, \mathscr{S}^2(\mathbf{x}_k)$ here, since no new ideas are introduced. In the numerical illustration in Section 4 we cite a specific approach.

Our first key result is that estimating Σ_ε in this way introduces no prediction bias.

Theorem 1 *Let* $\widehat{\Sigma}_\varepsilon = \mathrm{Diag}\left\{\widehat{\mathsf{V}}(\mathbf{x}_1)/n_1, \widehat{\mathsf{V}}(\mathbf{x}_2)/n_2, \ldots, \widehat{\mathsf{V}}(\mathbf{x}_k)/n_k\right\}$ *and define*

$$\widehat{\widehat{\mathsf{Y}}}(\mathbf{x}_0) = \beta_0 + \Sigma_{\mathsf{M}}(\mathbf{x}_0,\cdot)^\top \left[\Sigma_{\mathsf{M}} + \widehat{\Sigma}_\varepsilon\right]^{-1} \left(\bar{\mathscr{Y}} - \beta_0 \mathbf{1}_k\right). \quad (13)$$

If Assumption 1 holds, then $\mathrm{E}\left[\widehat{\widehat{\mathsf{Y}}}(\mathbf{x}_0) - \mathsf{Y}(\mathbf{x}_0)\right] = 0$.

As a consequence of Theorem 1, our key concern is how much variance inflation occurs when V is estimated. Clearly if the n_i are large enough there is little inflation. But how large do they have to be? To answer this question we consider another tractable example:

Suppose that

$$\Sigma_{\mathsf{M}} = \tau^2 \begin{pmatrix} 1 & r & \cdots & r \\ r & 1 & \cdots & r \\ \vdots & \vdots & \ddots & \vdots \\ r & r & \cdots & 1 \end{pmatrix},$$

$\Sigma_{\mathsf{M}}(\mathbf{x}_0,\cdot) = \tau^2(r_0, r_0, \ldots, r_0)^\top$ with $r_0, r \geq 0$, and $\Sigma_\varepsilon = (\mathsf{V}/n)\mathbf{I}$. This represents a situation in which the extrinsic correlations among the design points are all equal and the design points are equally correlated with the point we wish to predict, which might be (approximately) plausible if the design points are widely separated, say at the extremes of the region of interest, while $\mathbf{x}_0$ is central. Note that for the covariance matrix of $(\mathsf{Y}(\mathbf{x}_0), \bar{\mathscr{Y}}(\mathbf{x}_1), \ldots, \bar{\mathscr{Y}}(\mathbf{x}_k))^\top$ to be positive definite we must have $r_0^2 < 1/k + r(k-1)/k$. The structure of Σ_ε arises because we assume the intrinsic variance is the same across all design points and n replications have been allocated to each of them. Suppose also that we have an estimator $\widehat{\mathsf{V}} \sim \mathsf{V}\chi^2_{n-1}/(n-1)$, meaning that $(n-1)\widehat{\mathsf{V}}/\mathsf{V}$ has a chi-squared distribution. We use a common estimator of the intrinsic variance rather than estimating it at each design point individually to make the example tractable. Finally, let $\gamma = \mathsf{V}/\tau^2$ be the ratio of the intrinsic variance to the extrinsic variance, which is (roughly speaking) a measure of the sampling noise relative to the response surface variation.

For this example we can show that the MSE of $\widehat{\mathsf{Y}}(\mathbf{x}_0)$, the stochastic kriging predictor with V known, is

$$\mathrm{MSE}^\star = \tau^2\left(1 - \frac{kr_0^2}{1 + (k-1)r + \frac{\gamma}{n}}\right). \quad (14)$$

On the other hand, the MSE of $\widehat{\widehat{\mathsf{Y}}}(\mathbf{x}_0)$ obtained by substituting $\widehat{\mathsf{V}}$ for V is MSE $=$

$$\tau^2 \mathrm{E}\left[\left(1 + \frac{\left(1 + (k-1)r + \frac{\gamma}{n}\right)kr_0^2}{\left(1 + (k-1)r + \frac{\gamma}{n}\frac{\widehat{\mathsf{V}}}{\mathsf{V}}\right)^2} - \frac{2kr_0^2}{\left(1 + (k-1)r + \frac{\gamma}{n}\frac{\widehat{\mathsf{V}}}{\mathsf{V}}\right)}\right)\right]. \quad (15)$$

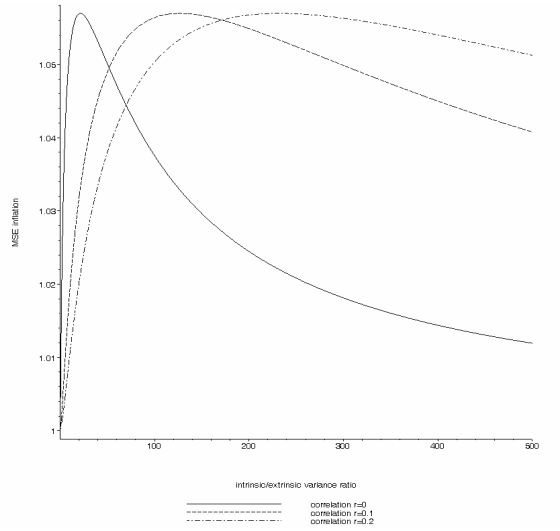

Figure 1: MSE inflation as a function of $\gamma = \mathsf{V}/\tau^2$ when $n = 10$ and correlation r_0 is 95% of its maximum possible value.

We assess the inflation by evaluating the ratio of (15) to (14) numerically. The ratio is largest when n is small and r_0 and r are large, so Figure 1 shows the inflation as a function of $\gamma = \mathsf{V}/\tau^2$ for $n = 10$, $r = 0, 0.1, 0.2$ and r_0 at 95% of the maximum value it can take. Even with this small value of n the inflation is slight over an extreme range of γ values. As n increases the inflation vanishes. This suggests that the penalty for estimating V will typically be small.

3.2 Maximum Likelihood Estimation

In this section we derive the maximum likelihood estimators of $(\beta_0, \tau^2, \theta)$ assuming Σ_ε is known. To reduce notation, let $\mathsf{V}_i \equiv \mathsf{V}(\mathbf{x}_i)/n_i$; thus, $\Sigma_\varepsilon = \mathrm{Diag}\{\mathsf{V}_1, \mathsf{V}_2, \ldots, \mathsf{V}_k\}$. Also define $\mathbf{R}_{\mathsf{M}}(\theta)$ to be correlation matrix of M across the design points.

For a fixed experiment design $\{(\mathbf{x}_i, n_i), i = 1, 2, \ldots, k\}$, and under Assumption 1, the log likelihood function of $(\beta_0, \tau^2, \theta)$ is

$$\begin{aligned} \ell(\beta_0, \tau^2, \theta) = & \quad (16) \\ & -\ln\left[(2\pi)^{k/2}\right] - \frac{1}{2}\ln\left[\left|\tau^2\mathbf{R}_{\mathsf{M}}(\theta) + \Sigma_\varepsilon\right|\right] \\ & -\frac{1}{2}\left(\bar{\mathscr{Y}} - \beta_0\mathbf{1}_k\right)^\top \left[\tau^2\mathbf{R}_{\mathsf{M}}(\theta) + \Sigma_\varepsilon\right]^{-1}\left(\bar{\mathscr{Y}} - \beta_0\mathbf{1}_k\right). \end{aligned}$$

If the Σ_ε terms are removed then this is the log likelihood function for kriging when M is a Gaussian random field. We have been intentionally vague about the covariance

function $\mathbf{R}_{\mathsf{M}}(\theta)$, because we want the results to be general, but when we apply stochastic kriging later we will use a standard model from the DACE literature.

Finding the maximum likelihood estimators requires simultaneously solving

$$\frac{\partial \ell(\beta_0,\tau^2,\theta)}{\partial \beta_0}=0 \quad \frac{\partial \ell(\beta_0,\tau^2,\theta)}{\partial \tau^2}=0 \quad \frac{\partial \ell(\beta_0,\tau^2,\theta)}{\partial \theta}=\mathbf{0} \tag{17}$$

for $(\widehat{\beta}_0,\widehat{\tau}^2,\widehat{\theta})$ which is no more computationally difficult than when Σ_ε is not present, and in fact is more likely to be numerically stable.

To summarize, given the data $\mathscr{Y}_j(\mathbf{x}_i), j=1,2,\ldots,n_i, i=1,2,\ldots,k$, a stochastic kriging metamodel is obtained as follows:

1. Estimate $\widehat{\mathsf{V}}$ as in Section 3.1 and let $\widehat{\Sigma}_\varepsilon = \mathrm{Diag}\left\{\widehat{\mathsf{V}}(\mathbf{x}_1)/n_1, \widehat{\mathsf{V}}(\mathbf{x}_2)/n_2, \ldots, \widehat{\mathsf{V}}(\mathbf{x}_k)/n_k\right\}$ where $\widehat{\mathsf{V}}(\mathbf{x}_i)=\mathscr{S}^2(\mathbf{x}_i)$.
2. Using $\widehat{\Sigma}_\varepsilon$ instead of Σ_ε, solve the likelihood equations (17) for $(\widehat{\beta}_0,\widehat{\tau}^2,\widehat{\theta})$.
3. Predict $\mathsf{Y}(\mathbf{x}_0)$ by the metamodel

$$\begin{aligned}\widehat{\mathsf{Y}}(\mathbf{x}_0) &= \widehat{\beta}_0 + \widehat{\tau}^2\mathbf{R}_{\mathsf{M}}(\mathbf{x}_0,\cdot;\widehat{\theta})^\top\left[\widehat{\tau}^2\mathbf{R}_{\mathsf{M}}(\widehat{\theta})+\widehat{\Sigma}_\varepsilon\right]^{-1} \\ &\quad\times\left(\bar{\mathscr{Y}}-\widehat{\beta}_0\mathbf{1}_k\right)\end{aligned} \tag{18}$$

with plug-in MSE estimate

$$\begin{aligned}\widehat{\mathrm{MSE}}(\mathbf{x}_0) &= \widehat{\tau}^2 \\ &\quad-\widehat{\tau}^4\mathbf{R}_{\mathsf{M}}(\mathbf{x}_0,\cdot;\widehat{\theta})^\top\left[\widehat{\tau}^2\mathbf{R}_{\mathsf{M}}(\widehat{\theta})+\widehat{\Sigma}_\varepsilon\right]^{-1} \\ &\quad\times\mathbf{R}_{\mathsf{M}}(\mathbf{x}_0,\cdot;\widehat{\theta}) \\ &\quad+\delta^\top\delta\left(\mathbf{1}_k^\top\left[\widehat{\tau}^2\mathbf{R}_{\mathsf{M}}(\widehat{\theta})+\widehat{\Sigma}_\varepsilon\right]^{-1}\mathbf{1}_k\right)^{-1}\end{aligned} \tag{19}$$

where $\delta = 1-\mathbf{1}_k^\top[\widehat{\tau}^2\mathbf{R}_{\mathsf{M}}(\widehat{\theta})+\widehat{\Sigma}_\varepsilon]^{-1}\mathbf{R}_{\mathsf{M}}(\mathbf{x}_0,\cdot;\widehat{\theta})\,\widehat{\tau}^2$. The last term on the right-hand side of (19) accounts for the variability due to estimating β_0.

3.3 Experiment Design

In this section we describe an approach to obtain experiment designs with low integrated MSE (IMSE). Our results assume that the extrinsic covariance function $\Sigma_{\mathsf{M}}(\cdot,\cdot)$ and the extrinsic variance function $\mathsf{V}(\cdot)$ are known; later in the section we describe how we might use the results when these functions are estimated.

Let $\mathscr{X}$ be the d-dimensional experiment design space of interest, and suppose that we have k fixed design points $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_k$ to which we want to allocate N replications. Let $\mathbf{n}^\top=(n_1,n_2,\ldots,n_k)$. Then our goal is to

$$\begin{aligned}\text{minimize IMSE}(\mathbf{n}) &= \int_{\mathbf{x}_0\in\mathscr{X}}\mathrm{MSE}(\mathbf{x}_0;\mathbf{n})\,d\mathbf{x}_0 \quad (20)\\ \text{subject to:}& \\ &\mathbf{n}^\top\mathbf{1}_k\le N \quad (21)\\ &n_i\in\mathscr{Z}^+ \quad (22)\end{aligned}$$

where the integrand $\mathrm{MSE}(\mathbf{x}_0;\mathbf{n}) = \Sigma_{\mathsf{M}}(\mathbf{x}_0,\mathbf{x}_0) - \Sigma_{\mathsf{M}}(\mathbf{x}_0,\cdot)^\top[\Sigma_{\mathsf{M}}+\Sigma_\varepsilon(\mathbf{n})]^{-1}\Sigma_{\mathsf{M}}(\mathbf{x}_0,\cdot)$ and $\Sigma_\varepsilon(\mathbf{n}) = \mathrm{Diag}\{\mathsf{V}(\mathbf{x}_1)/n_1,\mathsf{V}(\mathbf{x}_2)/n_2,\ldots,\mathsf{V}(\mathbf{x}_k)/n_k\}$. In words, we minimize the IMSE for the MSE-optimal stochastic kriging estimator as a function of the number of replications allocated to each design point. To obtain an approximate solution to this problem, we relax the integrality constraint (22) and assume only that $n_i\ge 0$. Since we will have repeated need of it, let $\Sigma(\mathbf{n})=\Sigma_{\mathsf{M}}+\Sigma_\varepsilon(\mathbf{n})$.

Assuming M is second-order stationary, as in (8), we can let $\Sigma_{\mathsf{M}}(\mathbf{x}_i,\mathbf{x}_0)=\tau^2 r_i(\mathbf{x}_0)$. We can then show that the optimal solution $\mathbf{n}^\star$ to (20), with integrality relaxed, satisfies $n_i^\star\propto\sqrt{\mathsf{V}(\mathbf{x}_i)C_i(\mathbf{n}^\star)}$ where

$$C_i(\mathbf{n}) = \left[\Sigma(\mathbf{n})^{-1}\mathbf{W}\Sigma(\mathbf{n})^{-1}\right]_{ii}$$

and $\mathbf{W}$ is the $k\times k$ matrix with elements

$$W_{ij}=\int_{\mathbf{x}_0\in\mathscr{X}} r_i(\mathbf{x}_0)r_j(\mathbf{x}_0)\,d\mathbf{x}_0.$$

To gain some insight into this result, suppose that N is large enough that $\Sigma(\mathbf{n})\approx\Sigma_{\mathsf{M}}$ so that

$$C_i(\mathbf{n})\approx C_i=\left[\Sigma_{\mathsf{M}}^{-1}\mathbf{W}\Sigma_{\mathsf{M}}^{-1}\right]_{ii}.$$

Then

$$n_i^\star\approx N\frac{\sqrt{\mathsf{V}(\mathbf{x}_i)C_i}}{\sum_{j=1}^k\sqrt{\mathsf{V}(\mathbf{x}_j)C_j}}. \tag{23}$$

Notice that C_i is only a function of the extrinsic correlation structure, and V is the intrinsic variance. Expression (23) shows how the response surface, as represented by its correlation structure, distorts the allocation of replications from one that is proportional to only the extrinsic standard deviation at the design point; it tends to favor design points that are centrally located because they do more to reduce MSE throughout the design space. This further emphasizes that both intrinsic and extrinsic uncertainty matter in the experiment design.

In practice neither $\Sigma_{\mathsf{M}}(\cdot,\cdot)$ nor $\mathsf{V}(\cdot)$ are known in advance, and the design points are not given. One way to use these results is via a two-stage design strategy:

1. In Stage 1, select a space-filling design of m predetermined design points $\mathbf{x}_1,\ldots,\mathbf{x}_m$ and allocate n_0 replications to each.
2. Fit $\widehat{\mathsf{V}}$ and $\widehat{\tau}^2\mathbf{R}_{\mathsf{M}}(\cdot,\cdot;\widehat{\theta})$ as described above.
3. In Stage 2, jointly select $k-m$ additional design points $\mathbf{x}_{m+1},\ldots,\mathbf{x}_k$ from a larger set and optimally allocate the $N-mn_0$ additional replications among $\mathbf{x}_1,\ldots,\mathbf{x}_k$ to minimize IMSE using $\widehat{\mathsf{V}}$ and $\mathbf{R}_{\mathsf{M}}(\cdot,\cdot;\widehat{\theta})$ in place of the true functions.

4 ILLUSTRATION

To illustrate the methodology developed in this paper, we consider the steady-state mean number in an M/M/1 queue. The statistic we record from each replication is the average number of customers in the system from time 0 to T. For the M/M/1 queue we can initialize each replication in steady state by independently sampling the number in the system at time 0 from the steady-state distribution. We keep the run length per replication T the same for all arrival rates x so that we entirely control intrinsic variance through the number of replications. We do not employ CRN. For fitting the mean and variance models we assume a Gaussian correlation structure of the form $\mathbf{R}_{\mathsf{M}}(x_i,x_j;\theta_{\mathsf{M}})=\exp(-\theta_{\mathsf{M}}(x_i-x_j)^2)$ and $\mathbf{R}_{\mathsf{V}}(x_i,x_j;\theta_{\mathsf{V}})=\exp(-\theta_{\mathsf{V}}(x_i-x_j)^2)$, respectively, with the θ's unknown. All of the simulation and fitting of the metamodels was done using our own code written in S-PLUS; fitting was via maximum likelihood.

To illustrate stochastic kriging, we consider an experiment that starts with four design points, $x=0.3,0.5,0.7,0.9$, making 20 replications of length $T=1000$ time units at each of them (80 replications total). Based on the results we allocate a total of $N=500$ replications among these four design points, plus 3 additional points $x=0.4,0.6,0.8$, using the approximately optimal allocation formula (23), and view the final fit.

Figure 2 shows the results for the mean number in queue metamodel $\widehat{\widehat{\mathsf{Y}}}(x_0)$ from the first-stage experiment. In the plot a circle represents an estimated response from the simulation (the data points); the solid-line curve is the stochastic kriging metamodel, which is surrounded by $\pm\sqrt{\widehat{\mathrm{MSE}}}$ intervals at a fine grid of points; and the dashed-line curve is the true surface. Since this is stochastic kriging, as opposed to ordinary kriging, the fitted surface need not pass through the data points (see especially at $x=0.9$), and the $\pm\sqrt{\widehat{\mathrm{MSE}}}$ intervals account both for intrinsic and extrinsic uncertainty about the surface. Notice that the true surface is within the $\pm\sqrt{\widehat{\mathrm{MSE}}}$ bounds on the fitted surface.

The fitted variance curve $\widehat{\mathsf{V}}(x_0)$ is shown in Figure 3. Since we use ordinary kriging for this model the fitted curve passes through the data points, and it is clear that the simulation provided a particularly poor estimate of $\mathsf{V}(0.9)$.

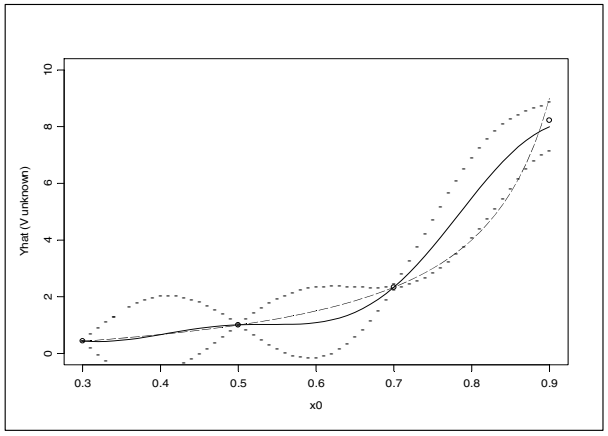

Figure 2: Fitted via stochastic kriging (solid line) and true (dashed line) expected number in an M/M/1 queue from the first-stage experiment.

For reference we also plot the known variance function $\mathsf{V}(x)/T=2x(1+x)/(T(1-x)^4)$ (Whitt 1989).

Using the results from the first-stage experiment (in particular $\widehat{\theta}_{\mathsf{M}}$ and $\widehat{\mathsf{V}}(x)$) we apply (23) to obtain the optimal allocation of $N=500$ replications to the full set of design points $x=0.3,0.4,0.5,0.6,0.7,0.8,0.9$. The variance model is required as the full design includes design points that were not simulated in the first-stage experiment. The estimated optimal allocation is $n=2,80,11,81,33,165,128$, respectively. That design points 2 and 4 (0.4 and 0.6) receive relatively large allocations relative to design points 1, 3 and 5 (0.3, 0.5 and 0.7) results mostly from their variance being overestimated by $\widehat{\mathsf{V}}$. More interesting is that $x=0.8$ receives a larger allocation than $x=0.9$, even though the standard deviation at 0.9 is predicted to be substantially greater than at 0.8 by $\widehat{\mathsf{V}}$. This occurs because our optimal allocation considers not only the relative standard deviations at the design points, but also their range of influence in the metamodel; $x=0.8$ is closer to more points in the design than 0.9 and therefore is more valuable.

Since several of the design points have already received more replications than optimal—always a danger when the initial sample size has to be selected arbitrarily—we reran the experiment allocating the 500 replications optimally (in practice we would not discard the data we already have and would instead allocate as close to the optimal design as possible). Figure 4 shows the result. The most important thing to notice is not the close fit to the true curve as much as the nearly constant $\pm\sqrt{\widehat{\mathrm{MSE}}}$ intervals surrounding the fitted curve.

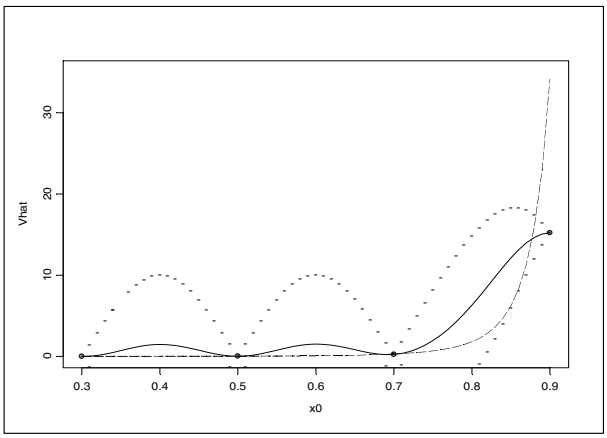

Figure 3: Fitted via ordinary kriging (solid line) and true (dashed line) variance of average number in an M/M/1 queue from the first-stage experiment.

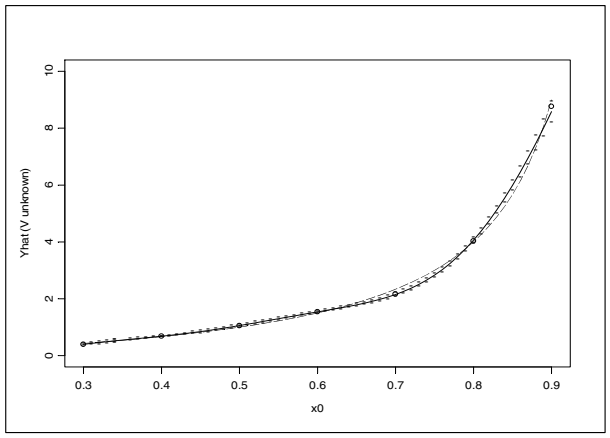

Figure 4: Fitted via stochastic kriging (solid line) and true (dashed line) expected number in an M/M/1 queue from the second-stage experiment.

5 CONCLUSIONS

This paper provides a mathematical foundation for stochastic kriging, a method that extends the power of kriging metamodeling for deterministic computer experiments to modeling responses from stochastic simulations. To realize the full potential of this technique we need to, and are, addressing these follow-up issues:

Our initial results on experimental design should lead to methods for sequential, adaptive design that places design points and allocates simulation effort as we learn more about the response surface being modeled. The ability to capture intrinsic and extrinsic uncertainty in the design is a strength of stochastic kriging.

In our limited experiments it appeared that the Gaussian random field model with Gaussian correlation structure did not work as well for representing estimator variance as it did for the response mean. Other alternative models should be explored, as well as whether there is any benefit from fitting a joint model for (M,V).

We largely ignored the possibility of including a trend term, $\mathbf{f}(\mathbf{x})^\top\boldsymbol{\beta}$, in our metamodel. Clearly there are applications for which the form of such a term is known or suspected and including it may leads to better fits. The presence of a trend term may make the use of CRN worthwhile.

The examples in this paper employed only a one-dimensional design variable x, but the theory is for general d-dimensional $\mathbf{x}$. In addition to the numerical issues that can arise in fitting high-dimensional kriging models, there is also a practical matter of visualizing and exploring the fitted surface. Tools such as ATSV (Stump et al. 2007) may be particularly helpful in this regard.

ACKNOWLEDGMENTS

This paper is based upon work supported by the National Science Foundation under Grant No. DMI-0555485, by the Semiconductor Research Corporation under Grant No. 2004-OJ-1225, and by General Motors R&D. The authors also acknowledge helpful advice from Dan Apley, Russell Barton, Thomas Santner and Tim Simpson.

REFERENCES

Barton, R. R., and M. Meckesheimer. 2006. Metamodel-based simulation optimization. In *Elsevier Handbooks in Operations Research and Management Science: Simulation*, ed. S. G. Henderson and B. L. Nelson, 535–574. New York: Elsevier.

Biles, W. E., J. P. C. Kleijnen, W. C. M. van Beers and I. Nieuwenhuyse. 2007. Kriging metamodeling in constrained simulation optimization: An exploratory study. *Proceedings of the 2007 Winter Simulation Conference*, ed. S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew and R. R. Barton, 355–362. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Cheng, R. C. H., and J. P. C. Kleijnen. 1998. Improved design of queueing simulation experiments with highly heteroscedastic responses. *Operations Research* 47:762–777.

Cressie, N. A. C. 1993. *Statistics for spatial data*. New York: John Wiley.

Fang, K. T., R. Li and A. Sudjianto. 2006. *Design and modeling for computer experiments*. Boca Raton, FL: Chapman & Hall/CRC.

Kennedy, M. C. and A. O'Hagan. 2000. Predicting the output from a complex computer code when fast approximations are available. *Biometrika* 87:1–13.

Kleijnen, J. P. C. and W. C. M. van Beers. 2005. Robustness of Kriging when interpolating in random simulation with heterogeneous variances: Some experiments. *European Journal of Operational Research* 165:826–834.

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

Mitchell, T. J. and M. D. Morris. 1992. The spatial correlation function approach to response surface estimation. In *Proceedings of the 1992 Winter Simulation Conference*, ed. J. J. Swain, D. Goldsman, R. C. Crain and J. R. Wilson, 565–571. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Myers, R. H. and D. C. Montgomery. 2002. *Response surface methodology*, 2nd ed. New York: John Wiley.

Nelson, B. L. 1990. Control-variate remedies. *Operations Research* 38:974–992.

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

Sabuncuoglu, I. and S. Touhami. 2002. Simulation metamodeling with neural networks: An experimental investigation. *International Journal of Production Research* 40:2483–2505.

Sacks, J., W. J. Welch, T. J. Mitchell and H. P. Wynn. 1989. Design and analysis of computer experiments. *Statistical Science* 4:409–423.

Santner, T. J., B. J. Williams and W. I. Notz. 2003. *The design and analysis of computer experiments*. New York: Springer.

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

Stein, M. L. 1999. *Interpolation of spatial data: Some theory for Kriging*. New York: Springer.

Stump, G., S. Lego, M. Yukish, T. W. Simpson and J. A. Donndelinger. 2007. Visual steering commands for trade space exploration: User-guided sampling with example. In *ASME Design Engineering Technical Conferences–Design Automation Conference*, ed. F. Liou. ASME DETC2007/DAC-34684.

Tew, J. D. and J. R. Wilson. 1992. Validation of simulation analysis methods for the Schruben-Margolin correlation-induction strategy. *Operations Research* 40:87–103.

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

van Beers, W. C. M. and J. P. C. Kleijnen. 2003. Kriging for interpolation in random simulation. *Journal of the Operational Research Society* 54:255–262.

van Beers, W. C. M. and J. P. C. Kleijnen. 2007. Customized sequential designs for random simulation experiments: Kriging metamodeling and bootstrapping. *European Journal of Operational Research*, forthcoming.

Whitt, W. 1989. Planning queueing simulations. *Management Science* 35:1341–1366.

Yang, F., B. E. Ankenman and B. L. Nelson. 2007. Efficient generation of cycle time-throughput curves through simulation and metamodeling. *Naval Research Logistics* 54:78–93.

AUTHOR BIOGRAPHIES

BRUCE ANKENMAN is an Associate Professor in the Department of Industrial Engineering & Management Sciences at Northwestern University. His research interests include the statistical design and analysis of experiments. Although much of his work has been concerned with physical experiments, recent research has focused on computer simulation experiments. Professor Ankenman is currently the director of the Masters of Engineering Management Program and the director of the Manufacturing and Design Engineering Program. He co-directs the freshman engineering and design course (EDC), and is the director of undergraduate programs for the Segal Design Institute. His e-mail and web addresses are <ankenman@northwestern.edu> and <users.iems.northwestern.edu/~bea/>.

BARRY L. NELSON is the Charles Deering McCormick Professor of Industrial Engineering and Management Sciences at Northwestern University and is Editor in Chief of *Naval Research Logistics*. His research centers on the design and analysis of computer simulation experiments on models of stochastic systems. His e-mail and web addresses are <nelsonb@northwestern.edu> and <www.iems.northwestern.edu/~nelsonb/>.

JEREMY STAUM is Associate Professor of Industrial Engineering and Management Sciences at Northwestern University. His research interests include risk management and simulation in financial engineering. Staum is Associate Editor of *ACM Transactions on Modeling and Computer Simulation*, *Naval Research Logistics*, and *Operations Research*, and was Risk Analysis track coordinator at the 2007 Winter Simulation Conference. His e-mail and web addresses are <j-staum@northwestern.edu> and <users.iems.northwestern.edu/~staum/>.