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

METAMODELS FOR SIMULATION INPUT-OUTPUT RELATIONS

Russell R. Barton
Department of Industrial and Management Systems Engineering
The Pennsylvania State University
University Park, Pennsylvania 16802, U.S.A.

ABSTRACT

The simulation community has used *metamodels* to study the behavior of computer simulations for over twenty-five years. The most popular techniques have been based on parametric polynomial response surface approximations. In this state of the art review, we present recent developments in this area. We also discuss seven alternative modeling strategies that are active topics in the current literature.

1 INTRODUCTION

Complex computer simulation models of proposed or existing real systems are often used to make decisions on changes to the system design. Analysts use the simulation model as a surrogate because it is impractical to construct multiple prototype versions of the real system, or because cost or other constraints prohibit experimentation with the real system. These models themselves may be quite complex, and so simpler approximations are often constructed; models of the model, or *metamodels*. Although this term has recent origins (Kleijnen, 1987), the simulation community has used *metamodels* to study the behavior of computer simulations for over twenty years (e.g., Mihram (1970), Racite, and Lawlor (1972), Biles (1974)). Interest in metamodeling issues continues today (see Sargent (1991)).

Metamodels have several uses in simulation. The simple form of a metamodel can reveal the general characteristics of behavior of the more complex simulation model. The insight provided by the simpler metamodel may be used for verification and validation of the complex parent model. It may also be used to identify the system parameters that most affect system performance (i.e. factor screening). Since it uses fewer computer resources, the metamodel can be run iteratively many times for repeated 'what if' evaluation for multi-objective systems or for design optimization. This is particularly important when the output of the simulation is a random quantity. Substitution of metamodel code is also an important strategy when the original model is just one component of a complex system model (Racite and Lawlor (1972)). In this case, the system model may be impractically slow and/or large without using metamodels for some or all components.

If we accept the definition of a metamodel as any 'model of a model', then the field of metamodeling encompasses much of the research in simulation methodology. For example, Little's Law, $L = \lambda W$, is a model of a queue *or* a metamodel of a simulation of a queue. Similarly, a complex simulation model may itself be modeled as a Jackson network of queues to gain insight on parent model behavior, for example for validation. Rapid modeling techniques (Suri and Diehl (1985, 1987) and Anderson (1987)) are generally employed *before* the complex parent simulation is built, but also offer the same uses as other metamodels. Perturbation analysis (Ho and Li (1988)) and likelihood ratio methods (Rubinstein (1986), Glynn (1987)) can be used to build Taylor approximation metamodels of the simulation input-output function based on a single simulation run.

This state of the art review will focus more narrowly on *general purpose mathematical approximations to input - output functions*. The 'general purpose' excludes metamodels such as Little's law and approximations based on perturbation analysis or likelihood ratios. The mathematical representation of a simulation model input - output function will be represented as

$$\mathbf{y} = \mathbf{g}(\mathbf{v}) . \tag{1}$$

Here, $\mathbf{y}$ and $\mathbf{v}$ are vector valued, and will usually include random components. The $\mathbf{v}$ vector for a manufacturing simulation might include the following components: the number of machines, machine processing times, machine breakdown time probability distribution parameters, and perhaps all the pseudorandom quantities used in the simulation run. The vector $\mathbf{y}$ might include the average work in process, the average daily throughput, and the average daily operating expenses.

Metamodels are typically developed separately for each component of $\mathbf{y}$, that is, for each coordinate function of $\mathbf{g}$. For most of the discussion, we will restrict out attention to input - output models where: i) y has one component, ii) the random component, if present, is additive, and iii) the list of parameters is restricted to those that will be in the argument list of the metamodel:

$$y = g(\mathbf{x}) + \varepsilon . \tag{2}$$

The metamodeling task involves finding ways to model g and ways to model ε. We will generally denote the metamodel as f and the predicted output responses as $f(\mathbf{x})$ or $\hat{y}$.

$$g(\mathbf{x}) \approx f(\mathbf{x}) = \hat{y} \tag{3}$$

The major issues in metamodeling include: i) the choice of a functional form for f, ii) the design of experiments, i.e., the selection of a set of **x** points at which to observe y (run the full model) to adjust the fit of f to g, the assignment of random number streams, the length of runs, etc., and iii) the assessment of the adequacy of the fitted metamodel (confidence intervals, hypothesis tests, lack of fit and other model diagnostics). The functional form will generally be described as a linear combination of basis functions from a parametric family. So there are choices for families (e.g., polynomials, sine functions, piecewise polynomials, wavelets, etc.) and choices for the way to pick the 'best' representation from within a family (e.g. least squares, maximum likelihood, cross validation, etc.). The issues of experiment design and metamodel assessment are related since the selection of an experiment design will be determined in part by its effect on assessment issues.

This review will focus primarily on the issue of choosing the functional form for f, with only occasional remarks in the other two areas. The most popular techniques for constructing f have been based on parametric polynomial response surface approximations. While we review recent developments in this area, we also discuss seven alternative modeling approaches (eleven distinct models) from the current literature:

- Taguchi models
- generalized linear models
- four methods based on splines (tensor product, interaction splines, MARS, and Π)
- radial basis functions
- kernel smoothing
- spatial correlation models, and
- frequency-domain (Fourier and wavelet) approximations.

(Space limitations do not allow the discussion of robust regression methods - outliers are less likely to cause significant problems for simulation model outputs.)

With so many techniques available, a word is in order on how one might choose from among them. There are many criteria that can be considered. Some of these are listed below.

1 The ability to gain insight from the form of the metamodel.
2 The ability to capture the shape of arbitrary smooth functions based on observed values which may be perturbed by stochastic components with general distribution.
3 The ability to characterize the accuracy of the fit through confidence intervals, etc.
4 The robustness of the prediction away from observed (**x**, y) pairs.
5 The ease of computation of the approximant function f.
6 The numerical stability of the computations, and consequent robustness of predictions to small changes in the parameters defining f.

And, finally, a practical concern for simulation modelers,

7 Does software exist for computing the metamodel, characterizing its fit, and using it for prediction?

The review is organized to cover each of the modeling techniques in sequence, beginning with traditional response surface methodology. The discussion will focus on the nature of the approximating functions and their characteristics via criteria 1-7.

The design of simulation experiments has received extensive attention for the polynomial form of the general linear model, and therefore it is discussed in that section Much development remains to be done for the other modeling techniques in the areas of design of experiments and model assessment. This is an exciting opportunity, because the evidence indicates that some of these techniques offer significant improvements over polynomial metamodels, particularly for providing a better global fit (Schumaker (1981), Härdle (1990)).

2 POLYNOMIAL RESPONSE SURFACE METAMODELING

Based on the work of Box (1954) and Hunter (1958-1959), response surface methods have been used effectively for over thirty years as metamodels. These methods are the topic of entire texts (see Box and Draper (1987), Khuri and Cornell (1987), Myers (1976)), but our review must be brief. In Myers, Khuri, and Carter (1989), the authors define response surface models as

"... a collection of tools that enhance the exploration of a region of design variables in one or more responses ... the function can be approximated in some region of the x's by a polynomial model."

Polynomial regression models were developed for the 'exploitation' of response surfaces (1), that is, for optimization. This approach fits first or second order polynomial models to y, the system response. The model is of the form (3) with y a scalar and ε a scalar, although these quantities are often viewed as vectors by considering multiple observations simultaneously. For the remainder of the section we will use $\mathbf{y} = (y_1, ..., y_n)'$ to represent a set of (univariate) outputs of the simulation model run under input conditions $\mathbf{x}_1, ..., \mathbf{x}_n$, respectively. The ε_i for the multiple observations are assumed to be independent, identically distributed Gaussian quantities with variance σ^2. The basis functions are usually taken as the products of the power functions, $1, x_j, x_j^2, ...$, giving

$$f(\mathbf{x}) = \Sigma \beta_k p_k(\mathbf{x}) \tag{4}$$

For example, $p_k(\mathbf{x})$, a product of power functions, might be (x_1), $(x_1)^2$, $(x_3)^2(x_4)$, etc. Alternatively, the basis may be orthogonal polynomials, $\varphi_k(\mathbf{x})$, providing the same polynomial for f but a different representation:

$$f(\mathbf{x}) = \Sigma \alpha_k \varphi_k(\mathbf{x}) . \tag{5}$$

The coefficients β_k or α_k are estimated from observed $(\mathbf{x}_i, y_i)$ data points, i=1,..., n via least squares or maximum likelihood estimation, which are identical procedures for Gaussian errors. The resulting estimates can be thought of as random quantities that depend on the random observations.

The advantage of (5) over (4) is that the coefficient estimates for the α_k's will be uncorrelated and will be robust to small changes in the observed data.

2.1 Design of Experiments

The recent developments for polynomial response surface models have been in the area of experimental design. To introduce these advances, we first describe the design problem. The coefficient vector β in (4) is determined by

$$\beta = (X'X)^{-1}X'y, \tag{6}$$

where $X = (\mathbf{1}, \mathbf{x}_1, ..., \mathbf{x}_n)'$ for a first degree (linear) polynomial, and includes products of these columns for higher order polynomials. From (2), we see that, since $\mathbf{y}$ is a random vector, β will be random.

Some recent research relates to two properties of β. First, one would like to minimize the variance of β. This will make the approximating function f less sensitive to the random perturbations introduced by ε. Second, one may want to estimate some of the coefficients in the β vector without making the number of simulation runs needed to estimate all of the coefficients in β. By leaving terms out of the metamodel (4), the fitting process may produce biased estimates for the remaining coefficients. Both of these properties are affected by the choice of the experimental design strategy. Each is discussed briefly below.

With independent ε_i values the variance-covariance matrix for the coefficient vector β is

$$\Sigma_\beta = \sigma^2(X'X)^{-1}. \tag{7}$$

When the ε_i values are dependent, with covariance matrix Σ_ε, the variance-covariance matrix for β is

$$\Sigma_\beta = (X'X)^{-1}X'\Sigma_\varepsilon X(X'X)^{-1}. \tag{8}$$

Schruben and Margolin (1978) exploited (8) to produce a reduced variance-covariance matrix for β by inducing correlation in the ε_i values. This is not generally possible in regression modeling, but in simulation metamodeling, the random number streams used for the simulation runs can be controlled to induce both positive and negative correlation between runs. The Schruben-Margolin strategy induces positive correlation between runs within a block, and negative correlations between blocks. The usual statistical analysis must be modified for this strategy, as described by Nozari, Arnold, and Pegden (1987) and Tew and Wilson (1992). Tew and Crenshaw (1990) discuss the implications when *all* of the random number streams are used as common or antithetic streams across the experiment (no pure error term remains).

The second experiment design issue receiving attention in simulation designs is bias. If there is concern that higher order terms may be present in (2) that are not modeled in (4), then simulation runs over the design space must be chosen differently. Draper and Guttman (1986) discuss how to choose optimal run conditions over a variety of design region shapes. Donohue, Houck and Myers (1990) combine these ideas with common streams and the Schruben-Margolin strategy to produce response surface designs with low bias (scaled central composite and Box-Behnken designs were found to be generally superior to factorial and small composite designs).

A third issue receiving attention is the reduction in the variance of β by removing (statistically) some of the variation in y before estimating β. This is possible when another response, y', has a known relation to $\mathbf{x}$ and a significant correlation with y. This y' is called a control variate. Recent research in control variate analysis and applications include Porta Nora and Wilson (1989), Avramidis and Wilson (1990), Cheng (1990), and Tew (1992). Tew and Wilson (1988) and Tew (1992) discuss the results of combining the Schruben-Margolin strategy with the control variates strategy for reducing the variance of β.

2.2 Properties of Polynomial Metamodels

First and second order polynomial models have standard interpretations based on the coefficients of the terms in (4) (Box and Youle (1955), Mason, Gunst and Hess (1989)). A large coefficient for a linear term indicates that the corresponding component of $\mathbf{x}$ has a significant effect on the outcome being modeled, y. A large coefficient for a quadratic term indicates a nonlinear response. A large coefficient for a cross product ($x_p x_q$) term is interpreted as a change in the (linear) effect of x_p as a function of the value of x_q, and similarly for q vs. p.

Polynomial models do well for criteria 3-7. The calculation of the estimates for β from (6) are simple and numerically stable for any reasonable design matrix X. Using the normal theory, confidence intervals and goodness of fit tests can be constructed. See Box, Hunter, and Hunter (1978) for example. Software abounds.

For low-order polynomial fits, the accuracy of the predicted value does not degrade rapidly as one moves away from any experimental observation. This is not true for high order polynomial fits, however. While the interpretation of these models is straightforward, the range of shapes that they can approximate is limited. For response surface optimization this is usually not a problem, because models are repeatedly built for approximation in a small $\mathbf{x}$ region, and then discarded as the search for the optimum moves on. When there are multiple objectives, however, it is unlikely that a small $\mathbf{x}$ region will contain all of the non-dominated solutions, and so (global) polynomial approximations may be a poor choice in this case.

The classic polynomial regression model requires restrictive assumptions on the ε term in (2) for the observed simulation outputs (i.i.d. normal). We next review two extensions of the linear model, Taguchi's model and the generalized linear model. Both relax the assumption that the ε_i are i.i.d. Gaussian quantities. The generalized linear model also deals with the inflexibility of (4).

3 ALTERNATIVE METAMODELING STRATEGIES

3.1 Taguchi Models

The Taguchi parameter design approach to product improvement actually involves several concepts, including

signal to noise ratios, linear graphs for selecting experiment designs, and noise factors (see Nair, et. al. 1992 for a comprehensive review). This section concentrates on the generalization of metamodels that occurs by allowing the variance of ε to depend explicitly on the parameter vector $\mathbf{x}$. This is applied in the simulation context by Ramberg et al. (1991).

The Taguchi approach (Phadke 1989) provides two modifications to the model described by (4). First, the errors, ε_i, are independent, but the variance is modeled as explicitly depending on $\mathbf{x}$, i.e. $\mathrm{Var}(\varepsilon_i) = \sigma^2(\mathbf{x})$. This situation is often the case practically, and can be addressed (though not in an entirely satisfactory way) for polynomial response surface models by i) transforming y, ii) using weighted least squares, or iii) varying the length of the simulation run at $\mathbf{x}_i$ in inverse proportion to the variance of ε_i. These alternatives are presented nicely in the simulation context by Welch (1990). Second, Taguchi's parameter design methodology incorporates this relation in the model not by providing a second model for $\sigma^2(\mathbf{x})$, but by combining the two (models of $y(\mathbf{x})$ and $\sigma^2(\mathbf{x})$) through the *signal to noise ratio*.

$$10*\log[E(y^2)/\sigma^2) = \Sigma\gamma_k p_k(\mathbf{x}) \qquad (9)$$

Typically the model involves only terms linear in $\mathbf{x}$. Taguchi expects to find some components of $\mathbf{x}$ that have small γ_k values in (9) but large β_k values in (4). After adjusting other parameters to maximize (9), these insensitive components of $\mathbf{x}$ are adjusted to move the expected response, $f(\mathbf{x})$ to the desired value.

3.1.1 Experiment Design for Taguchi Metamodels

Taguchi proposed designs for fitting (4) and (9) that were Kronecker products of: i) small-perturbation fractional factorial designs on the parameters affecting variance (the 'outer array') and, ii) fractional factorial designs on the full set of all parameters (the 'inner array'). Other authors have proposed more compact fractional designs in which the outer and inner arrays are not crossed, but combined in a more economical fashion. See Montgomery (1991) and Nair et al. (1992) for a full discussion.

3.1.2 Properties of Taguchi Metamodels

Taguchi metamodels share many properties with the polynomial response models of the previous section. The estimates for the β and γ vectors are computed in the same way and share the same (good) numerical properties. The low order polynomial form for the models (4) and (9) make interpretations straight forward. The signal/noise structure of (9) destroys the simple procedures for assessment offered by (4), unfortunately. Furthermore, the inherent limitations of polynomials still apply here, now in (9) as well as (4). The Taguchi metamodel can be fitted and studied effectively using the linear modeling capabilities of standard statistical packages. While software for choosing appropriate fractional factorial experiment designs is not generally available (but see Shoemaker and Tsui's concluding remarks in Nair, et al. 1992), graphical methods ease this task (see Taguchi 1959 (960), Phadke 1989, Kacker and Tsui 1990, and Wu and Chen 1992 for details).

3.2 The Generalized Linear Model

The model in (4) is often referred to as the *general linear model* (Scheffé 1954). We have refrained from using that name to reduce confusion with the yet more general *generalized linear model*. Both are referred to by the acronym GLM. We will use this term in this review only for the generalized linear model of McCullagh and Nelder (1989). GLM's are a generalization of general linear models in two ways. First, the distribution of the ε_i values may come from an exponential family other than the Normal/Gaussian. This includes any probability density of the form $f_{\theta,\phi}(\varepsilon) = \exp\{[(\varepsilon\theta - b(\theta)) / a(\phi)) + c(\varepsilon,\phi)\}$. For the Gaussian distribution, $\theta = \mu$, $\phi = \sigma^2$, $a(\phi) = \phi$, $b(\theta) = \theta^2/2$, and $c(\varepsilon,\phi) = -(1/2)[\varepsilon^2/\sigma^2 + \log(2\pi\sigma^2)]$. Other distributions in this class include the Poisson, binomial, gamma, and inverse Gaussian distributions.

The second generalization has to do with the form of the model that replaces (4). For the classical model in (4), $E(y) = f(\mathbf{x})$. For the GLM, this connection is made more flexible through the 'link function' or its inverse, λ, which allows $E(y) = \lambda(f(\mathbf{x}))$. Any monotonic differentiable function is allowed for λ. (The link function is denoted g by Nelder.) The link function adds another degree of flexibility to GLM metamodels. Probit and logit analyses are now included as special cases, via special link functions. For probit analysis, λ is the standard normal cumulative distribution function.

3.2.1 Properties of Generalized Linear Metamodels

The generalized linear model provides a formal framework for simulation metamodels when the Central Limit Theorem cannot provide an adequate approximation for simulation model outputs within run lengths that are practical. This model has been suggested an alternative for analyzing Taguchi experiments (see Nelder and Lee 1991). It provides a stronger statistical foundation and a more flexible form. Its strengths are in its flexibility. This flexibility is not without cost, however. The assessment tools for criterion 3 that are available for generalized linear models are based on the asymptotic theory of maximum likelihood estimators. These are weaker than the corresponding techniques for the classic linear models of (4). The topics of design of experiments, hypothesis testing, and analysis of variance are not well developed for this class as a whole, although residuals and an 'analysis of deviance' are discussed by McCullagh and Nelder (1989). Fitting and analysis software is available (GLIM: see Payne 1986), including macros for analysis of Taguchi-type experiments, (available from Nelder).

3.3 Modeling Methods Based on Splines

Any polynomial approximation represented by (4) can be constructed from linear combinations of the functions $\prod x_{jk}$, where the product is over k, and the index jkmay take the same value more than once. This choice for a basis has drawbacks, as mentioned earlier. Runge (1901) gives a clear example of the failure of polynomial basis functions to approximate a simple response function. This example is presented in many

texts, for example in Schumaker (1981). Figure 1 reproduces the fit of a 15th order (degree 14) polynomial f(x) to the function $g(x) = 1/(1+x^2)$ over the interval [-5,5). The problem in the tails does not go away for increasing samples, increasing polynomial degree, and increasingly fine sampling interval. Runge shows that as the fineness and order of the approximation increases, so does the maximum error, without bound. Design of the experiment plays a critical role here. If, rather than evenly, the points are selected according to the zeros of the Chebychev polynomials, the polynomial approximation has decreasing error as the sample size increases. The rate of decrease in the error can still be slow, however. See deBoor (1978) for a discussion.

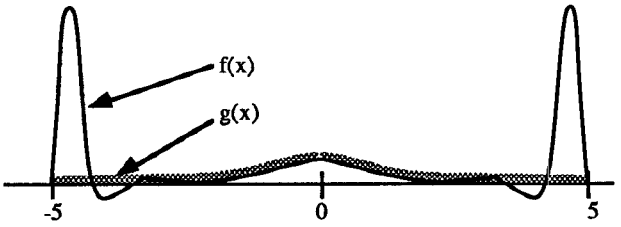

Figure 1. Fit of 14th degree polynomial to the function $g(x) = 1/(1+x^2)$. The fit deteriorates for $|x| > 3.64$, for a fit based on uniformly spaced experimental points as the number and degree increase.

The high order polynomial achieves a good fit by adjusting coefficients to achieve cancellation of large oscillations over most of the range. This reliance on cancellation makes high order polynomial fits non-robust. If a quadratic approximation to the function in figure 1 is adequate, then global polynomial basis functions can be used to build the approximating metamodel. If a more accurate representation is needed, the simulation modeler should consider other basis functions from which to build the metamodel. The remainder of this section discusses the 'state of the art' in the choice of basis functions.

3.3.1 Univariate Spline Metamodels

The difficulties with polynomial basis functions are avoided if: i) they are applied to a small region and, ii) only low order polynomials are used. This is the motivation for metamodels based on piecewise polynomial basis functions. When continuity restrictions are applied to adjacent pieces, the piecewise polynomials are called splines. The metamodel can be written as

$$f(x) = \sum c_j B_j(x) \qquad (10)$$

where the B_j are the quadratic or cubic piecewise polynomial basis functions. The basis functions can be described most simply for the univariate case. The domain is divided into intervals $[t_1, t_2), [t_2, t_3) \ldots [t_{n-1}, t_n)$ whose endpoints are called knots. Two sets of spline basis functions are commonly used, the truncated power function basis and the B-spline basis (de Boor 1978). The truncated power function basis for cubic splines consists of 1, x, and $\{(x-t_j)_+^3\}$, where $(x-t_j)_+^3$ denotes the one-sided function that is $(x-t_j)^3$ for $x > t_j$ and 0 for $x \le t_j$.

This set of basis functions are easy to describe and understand, but computationally not robust. Numerical difficulties arise when there are many intervals and consequently many basis function coefficients to estimate. Also, the basis functions have infinite support. A change in one observed value will affect the coefficients of all the basis functions. Further, some of the shape control is still effected by cancellation, resulting in the lack of robustness problems described for polynomial models.

The second set of basis functions, called B-splines, are more difficult to describe. They can be derived from divided differences of the power function spline basis elements. Their great advantage is that they have nonzero support over only k intervals: $B_j(x) = 0$ unless $x \in [t_j, t_{j+k})$. The parameter k is determined by the order of polynomials used in the spline model; k = 4 for cubic splines. Using a B-spline basis, estimation of the coefficients c_j in (10) is trivial, because $\sum B_j = 1$. A natural choice for c_j is $g(t_j)$, giving $f(t_j) = g(t_j)$, i.e. an interpolating function. An alternative experiment design is to choose the design points $t_j^* = (t_j + t_{j+1} + \ldots + t_{j+k})/k$ and set $c_j = g(t_j^*)$. This produces variation diminishing splines: "..the spline approximation to g crosses any particular straight line at most as many times as does g itself" (de Boor (1978)). This implies that if g is nonnegative so is f, and f is convex if g is. This choice makes sense for metamodels of deterministic functions, but it must be modified if the functions are expected to include a random component.

Since most simulation model output functions will not be deterministic, another approach is necessary to estimate the spline coefficients. The motivation for *smoothing splines* is based on an explicit tradeoff between the fit or accuracy of the approximation at known points and smoothness of the resulting metamodel. The fit term is represented as a sum of squared differences of the metamodel and simulation model responses at each of the experimental runs. The smoothness is represented by an integral of the square of some derivative over the region of validity of the metamodel. This is shown in figure 2.

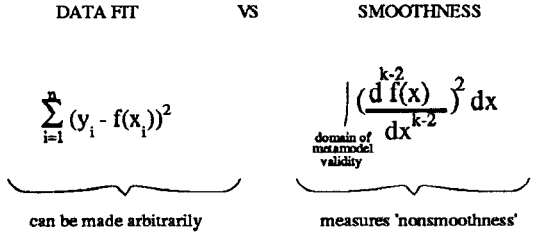

Figure 2. The tradeoff between accuracy and smoothness.

The spline functions arise as solutions to the following optimization problem, where the relative importance of fit vs. smoothness is controlled by the smoothing parameter λ:

$$\min_{f_\lambda \in C^{k-2}} \quad \Sigma(y_i - f_\lambda(x_i))^2 + \lambda \int (f_\lambda^{(k-2)})^2 \, dx \,. \qquad (11)$$

The function that minimizes (11) will be a spline of order k, which is in C^{k-2} (continuous derivatives up to the $(k-2)^{th}$ derivative) and is a piecewise polynomial with terms up to x^{k-1}. The knots will occur at points in x corresponding to the observed data, x_j. An example of smoothing splines with three different values for λ is shown in figure 3.

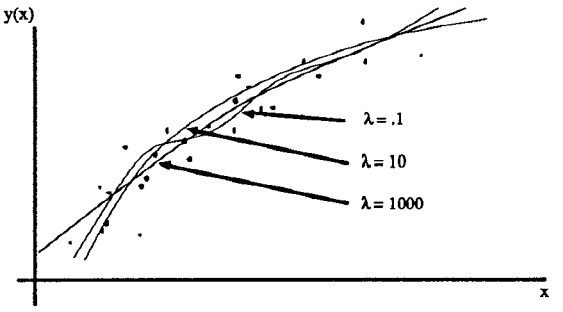

Figure 3. Three smoothing splines fitted to a set of data $\{(x_j, y(x_j))\}$. Knots are not shown, but they will occur at some of the x_j values (for smoothing splines), and nowhere else.

An important issue is the selection of the value for the smoothing parameter λ. The value may be chosen by visual examination of the fit (e.g. figure 3), or by minimizing cross validation (like residual sum of squares), or generalized cross validation (GCV) (an adjusted residual sum of squares). These approaches are discussed by Eubank (1988) and Craven and Wahba (1979). Li (1985) gives consistency results for smoothing splines (and other linear estimates) based on GCV estimates.

Three classes of spline metamodels can be described as solutions to special cases of the objective (11): spline smoothing, spline interpolation (described earlier), and least squares splines. The key differences are summarized below.

Smoothing Splines: k is chosen by the user, knots are not pre-specified, but they will occur at the x_j values in the optimal solution (i.e., $t_j = x_j$), λ can be chosen based on the user's preference (e.g., visual check of several trial values for λ) or by cross-validation, or generalized cross validation.

Spline Interpolation: k is chosen by the user, knots are not pre-specified, but they will occur at the x_j values in the optimal solution, $\lambda = 0$.

Least Squares Splines: k is chosen by the user, preferably near local maxima/minima and inflection points, knots are pre-specified, $\lambda = 0$.

3.3.2 Properties of Spline Metamodels

As the number of experiments increase, spline metamodels provide increasingly accurate approximation to any metamodel function g. Using the B-spline basis, deBoor (1978) shows that, for x in the interval $[t_j, t_{j+1})$

$$|g(x)-f(x)| \le \max |g(x)-g(t_i)|, \qquad j-k < i \le j \qquad (12)$$

The result is stronger for continuous functions, where the error decreases as the length of the maximum knot interval, say d to the k^{th} power:

$$\max|g(x)-f(x)| \le \xi_k d^k \|g^{(k)}\|, \qquad (13)$$

where the maximum is over the entire approximation interval [a,b), ξ_k represents a constant depending on k, and $\|g^{(k)}\|$ is the maximum modulus of the k^{th} derivative of g over [a,b).

3.3.3 Multivariate Spline Metamodels: Tensor Product, Interaction Splines, MARS, and Π

The extension of the univariate spline metamodels to multivariate situations has been an active area of recent research. Tensor products of univariate splines can be used for multivariate metamodels (deBoor 1978). Tensor product approximation requires a full factorial experiment design to estimate the parameters of the metamodel. Univariate splines are fit for each factor, for each level of every other factor. There is no requirement for equal numbers of levels across all design factors, nor equal spacing within one factor. Grosse (1980) provides efficient calculation of the tensor product spline coefficients based on the univariate coefficients. Finding the tensor product spline coefficients involves solving a linear system with a condition number that is on the order of $25^{(\#\text{ of factors})}$.

Because tensor product splines require many experimental runs on a complete rectangular grid, and because there are numerical difficulties in calculating the spline coefficients for metamodels with many input parameters, several alternative multivariate spline models have been proposed. The first, interaction splines, were presented by Wahba (1986). These models are linear combinations of products of at most two univariate splines. Gu (1990) gives an application to further generalize the generalized linear model by replacing (9) with an interaction spline metamodel.

Multivariate Adaptive Regression Spline models (Friedman (1990)) use a stepwise procedure to recursively partition the simulation input parameter space. The univariate product degree and the knot sequences are determined in a stepwise fashion based on the GCV score. The MARS model uses truncated power basis functions, which are not as numerically robust as B-splines. Friedman provides an ANOVA-like decomposition of the MARS fit that helps one gain insight from the metamodel. While our review has focused primarily on static metamodels, these techniques can be applied to study dynamic phenomena as well. Lewis and Stevens (1991) develop a nonlinear time series model of sunspot numbers using MARS. The data have a periodic structure which is successfully identified by the MARS regression.

The Π model (Breiman 1991) also uses a stepwise procedure for selecting a linear combination of products of univariate spline functions to be included in the metamodel. This method begins with a large number of knots for each variable, and uses a forward stepwise procedure based on the GCV score to select terms for the product, and to select the number of products. The backwards elimination step is also based on the GCV, and is used to delete knots (or, equivalently, univariate basis elements). Breiman provides plots of the univariate spline basis functions to provide some insight on the structural properties of the model.

These methods all consider products of univariate splines, where the order of the products and knot locations are

determined by different strategies. All three methods use the GCV score to identify appropriate components of the spline model in a stepwise fashion. All three have software implementations available from the authors (Gu, Friedman, and Breiman, respectively). All three can be expected to have the attractive properties of spline metamodels: the ability to provide accurate, computationally robust metamodels over the entire domain of the simulation model inputs. All three have means for gaining insight based on the metamodel: the (two-factor) interaction splines can be understood using techniques similar to those described for the Π metamodel. And all can be used for static or time series modeling. Finally, we note that the authors assume that the set of data values $\{(\mathbf{x}_i, y_i)\}$ to be fit are given. There is no discussion about the design of the simulation experiment to provide the best fit of f to g over some region of interest.

3.4 Radial Basis Functions

Radial basis functions provide an alternative approach to multivariate metamodeling. The original development by Hardy (1971) introduced, among others, simple 'multiquadric' basis functions

$$f(\mathbf{x}) = \sum a_i \, \|\mathbf{x}-\mathbf{x}_i\| \, , \qquad (14)$$

where the sum is over the observed set of system responses, $\{(\mathbf{x}_i, y_i)\}$ and $\|\bullet\|$ represents the Euclidean norm. The coefficients a_i are found simply by replacing the left hand side of (14) with $g(\mathbf{x}_i)$, i=1, ..., n, and solving the resulting linear system. Hardy showed that these functions provided a good fit to irregular topographic contours.

Unfortunately, the condition number of the linear system deteriorates rapidly with increasing dimension and increasing numbers of data values to be fitted. Also, since this is an interpolation method, its direct application to simulation metamodeling is limited. Dyn, Levin, and Rippa (1986) and Dyn (1987) solve both of these problems by finding effective preconditioners for the linear system, and by using smoothed global basis functions to fit scattered noisy data.

Radial basis functions also arise for a class of spline functions. The so called thin plate splines have radial basis functions of $\|\mathbf{x}-\mathbf{x}_i\|^2 \log\|\mathbf{x}-\mathbf{x}_i\|$. Like smoothing splines, the radial basis functions, as well as their coefficients in the metamodel, depend on the location of the observed values $\mathbf{x}_i$.

3.5 Kernel Smoothing Regression Models

All of the estimation methods described above produce predicted values, $f(\mathbf{x})$, that are linear functions of the observed y_i values, with coefficients determined by the basis functions and their coefficients. The kernel smoothing metamodel uses this representation explicitly, without developing an explicit representation for f in terms of basis functions. A value, $f(\mathbf{x})$, is computed directly as a weighted sum of the observed y_i values, where the weights are determined by a *kernel function.*

For smooth functions g, simulation responses near $\mathbf{x}$ contain information about $E(y(\mathbf{x}))$. Consider a simple example where the response, y, is a function of a single design parameter, x. As shown in figure 4, for an x near x_1, we expect a response near y_1. On the other hand, for an x near x_2 the response should be more like y_2. Kernel smoothing captures this idea through a weight function that depends on the distance separating x from points x_i where the (random) response y_i has been observed via one or more simulation runs. One might imagine a triangular weight function superimposed at the point x as shown in figure 5. The point x is nearer x_1 and so the weight given to y_1 is greater than the weight given to y_2.

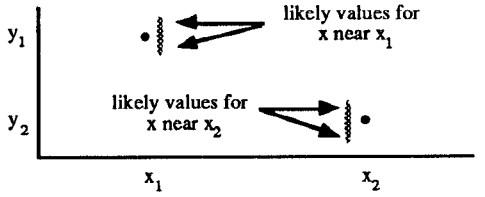

Figure 4. Motivation for local (kernel) smoothing.

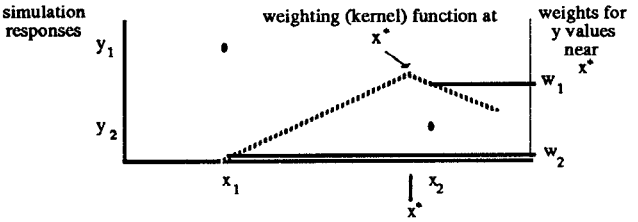

Figure 5. Weighting the nearby observed y values to estimate y(x).

There are many forms that this weighting or kernel function may take, and there are several ways to use the weighting function to calculate $f(\mathbf{x})$. To simplify the discussion, we will first discuss kernel smoothing in the setting of a single design parameter, i.e., $\mathbf{x} = x$. We present only one way to use the weighting function to compute f(x), the Nadaraya-Watson formula, because it is popular and easy to understand, and it reduces the bias of the kernel metamodel near the borders of the region over which model outputs have been computed (see the discussion below). Details on other kernels and kernel smoothing formulas are in Eubank (1988) and Härdle(1990).

3.5.1 The Nadaraya-Watson Formula

Given a set of completed simulation runs with data (x_i, y_i) the N-W formula for the metamodel is

$$f_\lambda(x) = \frac{\sum_{i=1}^{n} w((x-x_i)/\lambda) y_i}{\sum_{i=1}^{n} w((x-x_i)/\lambda)} \qquad (15)$$

where $w(\bullet)$ is the kernel function. Common choices for the kernel include

uniform	$w(u) = 1/2$	$-1 \le u \le 1$
triangular	$w(u) = 1 - \|u\|$	$-1 \le u \le 1$
quadratic	$w(u) = (3/4)(1-u^2)$	$-1 \le u \le 1$
quartic	$w(u) = (15/16)(1-u^2)^2$	$-1 \le u \le 1$

The approximation formula also depends on a smoothing parameter λ, which controls the size of the neighborhood over which y values are averaged. When λ is small, few points will be included in the range of u, producing a nonsmooth metamodel f(x). When λ is large, many points are included in the weighted average, and f(x) will be a slowly varying function, with greater bias.

The natural extension of (15) to the multivariate case would replace $(x-x_i)/\lambda$ with $\|\mathbf{x}-\mathbf{x}_i\|/\lambda$ ($\|\bullet\|$ is the Euclidean norm). This form is symmetric about $\mathbf{x}_i$. As a consequence, asymmetric boundary modifications of the kernel are not possible. Furthermore, individual λ_j are not possible. Instead, $(x-x_i)/\lambda$ is replaced by $\prod w((x_j-x_{ij})/\lambda)$ or, more generally, by $(\prod_{j=1}^{d} w((x_j-x_{ij})/\lambda_j))$:

$$f_\lambda(x) = \frac{\sum_{i=1}^{n} (\prod_{j=1}^{d} w((x_j-x_{ij})/\lambda_j))y_i}{\sum_{i=1}^{n} \prod_{j=1}^{d} w((x_j-x_{ij})/\lambda_j)} \qquad (16)$$

3.5.2 Bias in Kernel Smoothing Metamodels

Figure 6 shows the bias for two fitted metamodels to a hypothetical simulation response function. Near local minima, the metamodel will consistently overestimate the response because it is averaging values that are all greater than the local minimum. A similar problem of underestimation exists near local maxima. Reducing λ reduces the size of the bias but does not eliminate it. This problem is handled by modifying the kernel function for evaluations that are near the boundary of the data (Rice 1984).

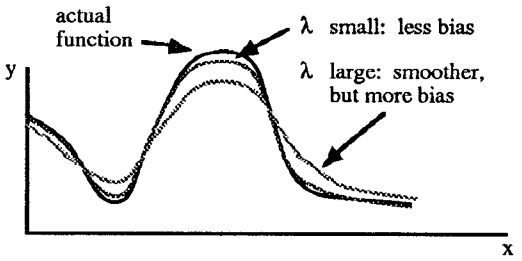

Figure 6. Bias as a function of λ for kernel smoothing.

The value of the smoothing parameter λ affects both smoothness and bias, and so must be chosen to balance these properties of the fitted metamodel. The method of least squares might be applied to choose the value of λ. However, $\lim_{\lambda \to 0} f_\lambda(x_i) = y_i$ so that least squares will drive the choice of λ to zero. An alternative to eliminate this behavior is to leave (x_i, y_i) out of the metamodel when calculating the difference between the metamodel and y_i. If we denote metamodel that is fit to all data *except* (x_i, y_i) as $f_{\lambda i}(x)$, then the sum of squared deviations between y_i and $f_{\lambda i}(x)$ is called cross validation mean sum of squares. The cross validation estimate for λ minimizes this quantity. It is just a least squares estimate, with f_λ replaced by $f_{\lambda i}$. This quantity can be minimized using an iterative optimization technique. The optimality properties of this approach are discussed by Härdle and Marron (1985).

Wahba proposed another technique for choosing λ called *generalized cross validation* (GCV). This includes an adjustment to the sum of squared residuals, and is discussed in Wahba (1990) and Eubank (1988).

3.6 A Spatial Correlation Model

Sacks, Welch, Mitchell, and Wynn (1989) and numerous references therein develop a parametric regression modeling approach that shares some common features with spline smoothing and kernel metamodeling. The expected smoothness of the function is captured in a spatial correlation function. The model assumption is

$$y(x) = g(x) + Z(x). \qquad (17)$$

Z is assumed to be a Gaussian stochastic process with spatial correlation function

$$\mathrm{Cov}(Z(\mathbf{u}), Z(\mathbf{v})) = R(\mathbf{u},\mathbf{v}) = \exp(-\Sigma\theta_j(u_j-v_j)^p). \qquad (18)$$

The value of p is sometimes fixed at 2, and $g(\mathbf{x})$ is usually approximated by a constant, or a linear function of $\mathbf{x}$. The values θ_j are estimated by maximum likelihood, and are used to calculate approximate expected values of (17) to provide the metamodel $f(\mathbf{x})$. This metamodel family has been used to model deterministic simulation models, but Sacks, et al. suggest the addition of a stochastic term for nondeterministic simulation metamodeling.

Currin et al. (1991) discuss the design of simulation experiments for estimating the p and θ_j parameters in (18), and discuss linear model and cubic spline forms for ln(R). Their examples show a much better fit than linear regression or cubic splines. Sacks, et al. consider initial Latin hypercube experiment designs followed by the sequential addition of points to minimize mean squared error integrated over the region of interest.

3.7 Frequency Domain Basis Functions

Viewing variations of g over its domain in terms of spatial correlation leads naturally to the idea of Fourier basis functions for representing an approximation to g in (1). While such an approach is possible, it is prone to difficulties (as is the global polynomial model) because the Fourier decomposition is based on basis functions with global support. Close approximations of g by a metamodel using a Fourier basis depends heavily on cancellation to achieve the desired form. This results in a method that fails to satisfy criteria 4 and 6.

This is less of an issue when modeling dynamic phenomena. Schruben and Cogliano (1987) use Fourier

decomposition to determine steady state input output structure by deliberately varying input parameters sinusoidally. There have been a series of papers since then discussing the design of experiments for this class of metamodels (see for example Morrice 1991, Buss 1990, and Jacobson et al. 1992). Yet even for dynamic models, Fourier basis representation still has the weakness associated with global basis functions. See Lewis and Stevens (1991) for a discussion and example.

Wavelet basis functions provide a decomposition in both location and frequency, providing local rather than global basis functions. The wavelet basis elements have finite support, and are adjusted by dilation factors to achieve a good fit (Daubechies (1988)). This methodology is still in the early stages of development. At present, applications of wavelet models have been limited to functions of one or two variables; in particular, to the construction of a smoothed visual image from noisy image intensity data. Strang (1989) gives an overview of wavelet models. While the software and applications are limited at present, this class of metamodels show great promise, since they offer a frequency-based interpretation of simulation model response, but do not suffer the limitations of global basis functions.

4 CONCLUSION

This is an exciting period in metamodeling. Recent and continuing increases in computing power allow new more flexible models to be computed and used for prediction and insight. Graphical representation of metamodel response contours for up to three variables can be computed and displayed automatically for the user to examine and adjust.

The weakness of polynomial models is their inability to provide a global fit to smooth response functions of arbitrary shape. This is particularly important for multicriteria optimization, where the response region of interest will never be reduced to a 'small neighborhood' by optimization. Alternative models such as kernel smoothing, radial basis functions, and multivariate splines are robust and provide good fits to arbitrary smooth response functions. On the other hand, they are computationally intensive, and in some cases the estimation problems are numerically ill-conditioned. For some of these models, the sign and size of the model coefficients provide less insight than a polynomial model. An additional drawback of the alternative models is that it is not clear how to select appropriate experiment designs for fitting such models.

The advantage of global polynomial metamodels of the form (4) is that techniques for experiment design, calculation, interpretation, and assessment are all well developed. We can expect to see similar developments for alternative metamodels in the near future.

ACKNOWLEDGMENTS

Joan Donohue read a preliminary draft and made several suggestions to improve the accuracy and clarity of this paper. Part of the research has been supported through NSF grant DDM-9118846.

REFERENCES

Anderson, K.R. 1987. A method for planning analysis and design simulation of CIM systems. *Proceedings of the 1987 Winter Simulation Conference*, A. Thesen, H. Grant, W. D. Kelton eds. 715-720.

Avramidis, T.N. and Wilson, J.R. 1990. Control variates for stochastic network simulation. *Proceedings of the 1990 Winter Simulation Conference*, O. Balci, R.P. Sadowski, and R.E. Nance eds. 323-332.

Biles, W.E. 1974. A gradient-regression search procedure for simulation experimentation. *Proceedings of the 1974 Winter Simulation Conference*, H.J. Highland ed. 490-497.

Box, G.E.P. 1954. The exploration and exploitation of response surfaces: some general considerations and examples. *Biometrics*, 10, 16-60.

Box, G.E.P. and Draper, N.R. 1987. *Empirical Model-Building and Response Surfaces*. New York: John Wiley and Sons.

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

Box, G.E.P. and Youle, P.V. 1955. The exploration and exploitation of response surfaces: an example of the link between the fitted surface and the basic mechanism of the system. *Biometrics*, 11, 287-323.

Breiman, L. 1991. The Π method for estimating multivariate functions from noisy data. *Technometrics*, 33, 125-160 (with discussion).

Buss, A.H. 1990. System Identification Using Frequency Domain Methodology. *Proceedings of the 1990 Winter Simulation Conference*, O. Balci, R.P. Sadowski, and R.E. Nance eds. 360-363.

Cheng, R.H. 1990. Fitting parametric models by conditional simulation. *Proceedings of the 1990 Winter Simulation Conference*, O. Balci, R.P. Sadowski, and R.E. Nance eds. 333-336.

Craven, P. and Wahba, G. 1979. Smoothing noisy data with spline functions. *Numerische Mathematik*, 31, 377-403.

Currin, C., Mitchell, T.J., Morris, M., and Ylvisaker, D. 1991. Bayesian prediction of deterministic functions, with application to the design and analysis of computer experiments. *Journal of the American Statistical Association*, 86, 953-963.

Daubechies, I. 1988. Orthonormal bases of compactly supported wavelets. *Communications in Pure and Applied Mathematics*, 41, 909-996.

deBoor, C. 1978. *A Practical Guide to Splines*. New York: Springer-Verlag.

Donohue, J.M., Houck, E.M., and Myers, R.H. 1990. Some optimal simulation designs for estimating quadratic response surface functions. *Proceedings of the 1990 Winter Simulation Conference*, O. Balci, R.P. Sadowski, and R.E. Nance eds. 337-343.

Draper, N.R. and Guttman, I. 1986. Response surface designs in flexible regions. *Journal of the American Statistical Association*, 81, 1089-1094.

Dyn, N. 1987. Interpolation of scattered data by radial functions. *Topics in Multivariate Approximation*, C.K. Chui and L.L. Schumaker eds. Boston: Academic Press.

Dyn, N., Levin, D., and Rippa, S. 1986. Numerical procedures

for surface fitting of scattered data by radial functions. *SIAM Journal of Scientific and Statistical Computing*, 7, 639-659.

Eubank, R.L. 1988. *Spline Smoothing and Nonparametric Regression.* New York: Marcel Dekker.

Friedman, J.H. 1991. Multivariate adaptive regression splines. *The Annals of Statistics*, 19, 1-141.

Glynn, P.W. 1987. Likelihood ratio gradient estimation: an overview. *Proceedings of the 1987 Winter Simulation Conference* A. Thesen, H. Grant, W.D. Kelton eds. 366-375.

Grosse, E. H. 1980. Tensor spline approximation. *Linear Algebra and Its Applications*, 34, 29-41.

Gu, C. 1990. Adaptive spline smoothing in non-gaussian regression models. *Journal of the American Statistical Association*, 85, 801-807.

Gu, C., Bates, D.M., Chen, Z., and Wahba, G. 1989. The computation of GCV functions through householder tridiagonalization with application to the fitting of interaction spline models. *SIAM Journal on Matrix Analysis and Applications*, 10, 457-480.

Härdle, W. 1990. *Applied Nonparametric Regression.* Cambridge: Cambridge University Press.

Härdle, W. and Marron, J.S. 1985. Optimal bandwidth selection in nonparametric regression function estimation. *The Annals of Statistics*, 13, 1465-1481.

Ho, Y.C. and Li, S. 1988. Extensions of infinitesimal perturbation analysis. *IEEE Transactions on Automatic Control*, AC-33, 427-438.

Hunter, J.S. 1958-59. Determination of optimal operating conditions by experimental methods. *Industrial Quality Control*, 15, Part II-1 (Dec. 1958), Part II-2 (Jan. 1959), and 16, Part II-3 (Feb. 1959).

Jacobson, S., Buss, A.H., and Schruben, L.W. 1992. Driving frequency selection for frequency domain experiments. *Operations Research*(to appear).

Kacker, R.N. and Tsui, K. 1990. Interaction graphs: graphical aids for planning experiments. *Journal of Quality Technology*, 22, 1-14.

Khuri, A.I. and Cornell, J.A. 1987. *Response Surfaces, Designs, and Analyses.* New York: M. Dekker.

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

Lewis, P.A.W. and Stevens, J.G. 1991. Nonlinear modeling of time series using multivariate adaptive regression splines (MARS). *Journal of the American Statistical Association*, 86, 864-877.

Mason, R.L., Gunst, R.F., and Hess, J.L. 1989. *Statistical Design and Analysis of Experiments with Applications to Engineering and Science.* New York: John Wiley and Sons.

McCullagh, P. and Nelder, J.A. 1989. *Generalized Linear Models.* Second Edition, London: Chapman and Hall.

Mihram, G.A. 1970. An efficient procedure for locating the optimal simular response. *Proceedings of the Fourth Conference on the Applications of Simulation*, New York: Association for Computing Machinery.

Montgomery, D.C. 1991. Using Fractional factorial designs for robust design process development. *Quality Engineering*, 3, 193-205.

Morrice, D.J. 1991. The effects of batching on the power of the test for frequency domain methodology. *Proceedings of the 1991 Winter Simulation Conference*, B.L. Nelson, W. D. Kelton, G.M. Clark eds. 1018-1022.

Myers, R.H. 1976. *Response Surface Methodology.* Boston: Allyn and Bacon.

Myers, R.H., Khuri, A.I., and Carter, W.H. Jr. 1989. Response surface methodology: 1966-1988. *Technometrics*, 31, 137-157.

Nair, V.N. *et. al.* 1992. Taguchi's parameter design: a panel discussion. *Technometrics*, 34, 127-161.

Nelder, J.A. and Lee, Y. 1991. Generalized linear models for the analysis of Taguchi-type experiments. *Applied Stochastic Models and Data Analysis*, 7, 107-120.

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

Payne, C.D. ed. 1986. *The GLIM Manual, release 3.77.* Oxford: Numerical Algorithms Group (NAG).

Phadke, M. 1989. *Quality Engineering Using Robust Design.* Englewood Cliffs, NJ: Prentice-Hall.

Porta Nora, A. M. and Wilson, J.R. 1989. Estimating multiresponse simulation metamodels using control variates. *Management Science*, 35, 1316-1333.

Racite, M.P. and Lawlor, F.D. 1972. Techniques for the development of empirical subsystem models. *Proceedings of the 1972 Summer Simulation Conference*, M.K. Horn ed. 155-162.

Ramberg, J.S., Sanchez, S.M., Sanchez, P.J., and Hollick, L.J. 1991. Designing simulation experiments: Taguchi methods and response surface metamodels. *Proceedings of the 1991 Winter Simulation Conference*, B.L. Nelson, W. D. Kelton, G.M. Clark eds. 167-176.

Rice, J. 1984. Boundary modification for kernel regression. *Communications in Statistics - Theory and Methodology* (Part A), 13, 893-900.

Rubinstein, R.Y. 1986. *Monte Carlo Optimization, Simulation and Sensitivity of Queueing Networks.* New York: John Wiley and Sons.

Runge, C. 1901. Uber die darstellung willkürlicher functionen und die interpolation zwischen aquidistanten ordinaten. *Z. Angew. Math. Phys.* 46, 224-243.

Sacks, J., Welch, W.J., Mitchell, T.J., and Wynn, H.P. 1989. Design and analysis of computer experiments. *Statistical Science*, 4, 409-435.

Sargent, R.G. 1991. Research issues in metamodeling. *Proceedings of the 1991 Winter Simulation Conference*, B.L. Nelson, W. D. Kelton, G.M. Clark eds.), 888-893.

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

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

Schumaker, L.L. 1981. *Spline Functions: Basic Theory.* New York: John Wiley and Sons.

Sheffé, H. 1954. *The Analysis of Variance.* New York: John Wiley and Sons.

Strang, G. 1989. Wavelets and dilation equations: a brief

introduction. *SIAM Review*, 31, 614-627.

Suri, R. and Diehl, G.W. 1985. MANUPLAN: a precursor to simulation for complex manufacturing systems. *Proceedings of the 1985 Winter Simulation Conference.*

Suri, R. and Diehl, D.W. 1987. Rough cut modelling: an alternative to simulation. *CIM Review*, 3, 25-32.

Taguchi, G. 1959. Linear graphs for orthogonal arrays and their application to experimental design, with the aid of various techniques. *Reports of Statistical Application Research, Union of Japanese Scientists and Engineers*, 6, 4, 1-43.

Taguchi, G. 1960. Tables of orthogonal arrays and linear graphs. *Reports of Statistical Application Research, Union of Japanese Scientists and Engineers*, 6, 5, 1-52.

Tew, J.D. 1992. Correlated replicates for first-order metamodel estimation in simulation experiments. *Transactions of The Society for Computer Simulation*, 8, 218-244.

Tew, J.D., and Crenshaw, M.D. 1990. Heuristic diagnostics for the presence of pure error in computer simulation models. *Proceedings of the 1990 Winter Simulation Conference*, O. Balci, R.P. Sadowski, and R.E. Nance eds. 347-356.

Tew, J.D. and Wilson, J.R. 1988. Estimating simulation metamodels using integrated variance reduction techniques. *1988 Proceedings of the Statistical Computing Section*, Alexandria, Va: American Statistical Association, 28-32.

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

Thompson, J. R. and Tapia, R.A. 1990. *Nonparametric Function Estimation, Modeling, and Simulation.* Philadelphia: Society for Industrial and Applied Mathematics.

Welch, P.D. 1990. Simulation and regression: dealing with the assumption of a common error variance. *Proceedings of the 1990 Winter Simulation Conference*, O. Balci, R.P. Sadowski, and R.E. Nance eds. 392-394.

Wu, C.F.J. and Chen, Y.Y. 1992. A graph-aided method for planning two-level experiments when certain interactions are important. *Technometrics*, 34, 162-175.

AUTHOR BIOGRAPHY

RUSSELL R. BARTON is an Associate Professor in the Department of Industrial and Management Systems Engineering at The Pennsylvania State University. He received a B.S. in electrical engineering from Princeton University in 1973 and a Ph.D. in operations research from Cornell University in 1978. His current research interests include graphical methods for experiment design, optimization and metamodeling methods for simulation models, and statistical issues in modeling manufacturing yield.