Proceedings of the 1994 Winter Simulation Conference
ed. J. D. Tew, S. Manivannan, D. A. Sadowski, and A. F. Seila

METAMODELING: A STATE OF THE ART REVIEW

Russell R. Barton

Department of Industrial and
Manufacturing Engineering
The Pennsylvania State University
University Park, PA 16802

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, with a particular focus on new developments in the experimental designs employed.

1 INTRODUCTION

Complex computer simulation models of proposed or existing real systems are often used to make decisions on changes to the system design. These models themselves may be quite complex, and so simpler approximations are often constructed; models of the model, or *metamodels* (Kleijnen, 1987).

This state of the art review will follow the focus of Barton (1992) 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, **y** and **v** are vector valued, and will usually include random components. The **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 **y** might include the average work in process, the average daily throughput, and the average daily operating expenses.

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(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 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 augment the results reported in Barton (1992), with a focus primarily on recent research on experimental designs and the choice of smoothing parameters. The most popular techniques for constructing f have been based on parametric polynomial response surface approximations. While we review recent developments for polynomial metamodels, we also review research related to alternative modeling approaches from the current literature:

- splines,
- radial basis functions,
- kernel smoothing,
- spatial correlation models, and
- frequency-domain approximations.

The review is organized to cover each of the modeling techniques in sequence, beginning with traditional response surface methodology. The discussion will focus on recent advances: for more background and additional metamodeling techniques, see Barton (1993).

2 RESPONSE SURFACE METAMODELING

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.

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.

2.1 Mathematical Form for RS Models

Let $\mathbf{y} = (y_1, ..., y_n)'$ 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

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

Here $p_k(\mathbf{x})$ is a product of univariate power functions, such as (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. 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.

2.2 Design of Experiments for RS Models

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 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. 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) and Tew (1994) 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. Donohue et al. (1993b) develop two-level factorial designs that protect against bias in the presence of polynomials of order two. Optimal designs are found for three pseudorandom number generation strategies: independent streams, common random numbers, and Schruben/Margolin CRN/ARN in orthogonal blocks.

Response surface metamodels for discrete event simulation models must often contend with nonhomogeneous variance. Cheng and Kleijnen (1994) develop optimal design of experiments for fitting metamodels when the response is some output function of a nearly saturated queue. Kleijnen and Van Groenendaal (1994) develop sequential experimental designs for weighted least squares regression metamodels.

Donohue et al. (1993a) propose a sequential strategy for fitting polynomial metamodels to discrete event

simulation models. Stage zero is a pilot study to check the magnitude and homogeneity of variance. The stage one design permits fitting a first order model and estimating bias from unfitted quadratic terms. Stage two augments the first order design to minimize the mean squared error of response or of slope. If bias error is still significant, stage three further augments the design to permit the estimation of a full second order model. These designs are considered for the three pseudorandom number assignment strategies mentioned above. Ordinary least squares and generalized least squares models are considered, and generalized least squares with the antithetic pseudorandom assignment strategy performs best on a Monte-Carlo study of an open Jackson queueing network.

3 SPLINE METAMODELS

Any polynomial approximation represented by (4) can be constructed from linear combinations of the functions $\prod x_{j_k}$, where the product is over k, and the index j_k may take the same value more than once. This choice for a basis has drawbacks, as mentioned earlier.

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 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.

3.1 Mathematical Form for Spline Models

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

$$f(x) = \sum c_j B_j(x) \qquad (9)$$

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 (deBoor 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 $x \leq t_j$. 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. Since the basis functions have infinite support, a change in one observed value will affect the coefficients of all the basis functions. 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. A natural choice for c_j is $g(t_j)$, giving $f(t_j) = g(t_j)$, i.e. an interpolating function.

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 smoothing 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:

$$\min_{f_\lambda \in C^{k-2}} \quad \Sigma(y_i - f_\lambda(x_i))^2 + \lambda \int (f_\lambda^{(k-2)})^2 \, dx \, . \qquad (10)$$

The function that minimizes (10) 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 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).

Three classes of spline metamodels can be described as solutions to special cases of the objective (10): spline smoothing, spline interpolation (described earlier), and least squares or regression splines. The key differences are summarized below.

Smoothing Splines: k is chosen by the user, knots 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 or by generalized cross validation.

Spline Interpolation: k is chosen by the user, knots will occur at the x_j values in the optimal solution, $\lambda = 0$.

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

3.2 Multivariate Splines

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. 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 (MARS) 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.

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).

For all of these regression spline methods, 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. There has been recent work, however, on the placement of the knots for optimal fit.

3.3 Optimal Knot Distribution for Regression Splines

Dyn and Yad-Shalom (1991) consider the optimal distribution of knots for tensor product spline approximations. They find distributions for each axis that are asymptotically optimal as the number of knots approaches infinity. For thin plate regression splines, the knots need not occur at data points. McMahon and Franke (1992) select knot locations to minimize the sum of squared distances from each data point to the nearest knot point. This knot selection method can be applied to the location of multiquadrics for radial basis function approximation discussed in the next section.

4 RADIAL BASIS FUNCTION METAMODELS

Radial basis functions (RBF) provide an alternative approach to multivariate metamodeling. In an empirical comparison, Franke (1982) found radial basis functions to be superior to thin plate splines, cubic splines and B-splines, and several others.

4.1 Mathematical Form for RBF Models

The original development by Hardy (1971) introduced, among others, simple 'multiquadric' basis functions

$$f(\mathbf{x}) = \sum a_i \, \|\mathbf{x}-\mathbf{x}_i\| \, , \qquad (11)$$

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 (11) with $g(\mathbf{x}_i)$, i=1, ..., n, and solving the resulting linear system.

4.2 Design of Experiments for RBF Models

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 executing only the first few iterative steps in solving the system of equations to provide a smooth fit to noisy data.

The issue of solvability has been addressed recently by Ball et al. (1992) and Sun (1993). Ball et al. provide upper and lower bounds on the l_2 norm of the matrix of equation coefficients (Hardy matrix), and Sun gives necessary and sufficient conditions on the location of the design points for the Hardy matrix to be nonsingular.

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 "best fit" metamodel, depend on the location of the observed values $\mathbf{x}_i$.

5 KERNEL SMOOTHING METAMODELS

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.*

5.1 Mathematical Form for Kernel Models

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).

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

$$f_\lambda(x) = \frac{\sum_{i=1}^{n} w((x-x_i)/\lambda)y_i}{\sum_{i=1}^{n} w((x-x_i)/\lambda)} \tag{12}$$

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

uniform	$w(u) = 1/2$	$-1 \le u \le 1$,
triangular	$w(u) = 1 - \lvert u \rvert$	$-1 \le u \le 1$,
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 (12) 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 $(\prod_{j=1}^{d} w((x_j-x_{ij})/\lambda_j))$:

$$f_\lambda(x) = \frac{\sum_{i=1}^{n} (\prod_{j=1}^{d} w((x_j-x_{ij})/\lambda_j))y_i}{\sum_{i=1}^{n} \prod_{j=1}^{d} w((x_j-x_{ij})/\lambda_j)} \tag{13}$$

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, $\lim_{\lambda \to 0} f_\lambda(\mathbf{x}_i) = y_i$ so that least squares will drive the choice of λ to zero. An alternative to eliminate this behavior is to leave $(\mathbf{x}_i, y_i)$ out of the metamodel when calculating the difference between the metamodel and y_i. The cross validation estimate for λ minimizes this quantity. 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).

5.2 Recent Developments for Kernel Models

Since kernel models are straightforward functions of the observed data, it is possible to construct confidence intervals for the true underlying function. Eubank and Wang (1994) develop asymptotically valid confidence intervals when the underlying function satisfies a Lipschitz condition, and the the choice of the smoothing parameter satisfies certain requirements. The GCV estimator satisfies these requirements.

An alternative to the GCV method for estimating λ was proposed by Taylor (1994). A bootstrapping method was used to choose the smoothing parameter value, and the results were applied to image reconstruction with a Gaussian kernel.

6 SPATIAL CORRELATION METAMODELS

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.

6.1 Mathematical Form for Spatial Models

The model assumption is

$$y(\mathbf{x}) = g(\mathbf{x}) + Z(\mathbf{x}) . \tag{14}$$

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

$$\mathrm{Cov}(Z(\mathbf{u}), Z(\mathbf{v})) = R(\mathbf{u},\mathbf{v}) = \exp(-\Sigma\theta_j(u_j-v_j)^p). \quad (15)$$

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 (15) 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. Mitchell and Morris (1992) discuss this extension, as well as other correlation functions.

6.2 Design of Experiments for Spatial Models

Currin et al. (1991) discuss the design of simulation experiments for estimating the p and θ_j parameters in (15). Factorial designs are not appropriate for fitting these parameters. In the case of a factorial design on r factors, if there are fewer than r factors active in the model, the design will be projected effectively on the remaining factors, giving duplicate points. For the spatial correlation model, this leads to difficulties: the covariance matrix R will not be full rank, and the likelihood function will be impossible to maximize. Latin hypercube designs avoid this problem, but often provide a poor coverage of the space. Sacks, et al. (1989) consider initial Latin hypercube experiment designs followed by the sequential addition of points to minimize mean squared error integrated over the region of interest.

The spatial correlation model provides a very good fit with relatively small designs. Orthogonal arrays of strength r (see Owen 1992) are an attractive class of sparse designs because they provide balanced (full factorial) designs for any projection onto r factors. Unfortunately, this kind of balance can lead to problems for spatial correlation models. Tang (1993) replaces the h entries which have value = k in a particular column of the orthogonal array with a permutation of the set of h values equally spaced between $(k-1)+1/h$ and k. Balance is maintained, but the design projection is no longer a factorial; in fact it is a Latin hypercube.

Alternative approaches to improve on the coverage of Latin hypercube designs are proposed by Handcock (1992), Salagame and Barton (1993) and Corsten and Stein (1994). All are hierarchical designs, in which the design space is first subdivided into regions to maintain balance, and sub-designs are constructed for a subset of the regions.

Morris et al. (1993) expand the spatial correlation model to consider the case where function and derivative information is available. Considered example with latin hypercube, D-optimal, and two hybrid design procedures designed to have the properties of both latin hypercube and D-optimality. One of the hybrid methods provided the smallest prediction error.

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, the Fourier decomposition is based on basis functions with global support. Close approximations of g by using a Fourier basis depends heavily on cancellation to achieve the desired form, so the model form may not be robust.

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, Jacobson et al. 1992, and Morrice and Schruben 1993).

For static metamodels, 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.

8 CONCLUSION

Developments for both response surface models and nontraditional models provide increased efficiency and applicability for these methods. In particular, recent work in the areas of spatial correlation and radial basis functions has clarified the importance of experimental design for non-traditional models. While the fitting capability of these alternative methods is exciting, at the present time it is based on a small set of examples (see Ilgenstein and Sargent 1994, and Laslett 1994). A more extensive computational comparison of the methods is needed, but this will have to wait for more generally available computer codes for the newer metamodeling methods.

ACKNOWLEDGMENTS

Mr. Ching-Hsin Tu found a number of references that were included in this paper. Part of this research was supported through NSF grant DDM-9308492.

REFERENCES

Ball, K., Sivakumar, N., and Ward, J.D. 1992. On the sensitivity of radial basis interpolation to minimal data separation distance. *Constructive Approximation* 8, 401-426.

Barton, R.R. 1992. Metamodels for simulation input-output relations. *Proceedings of the 1992 Winter Simulation Conference* J.J. Swain, D. Goldsman, R.C. Crain, and J.R. Wilson eds. 289-299.

Barton, R.R. 1993. New tools for simulation metamodels. Working Paper 93-110, Industrial and Management Systems Engineering, The Pennsylvania State University.

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.

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.C.H. and Kleijnen, J.P.C. 1994. Optimal design of simulation experiments with nearly saturated queues. Department of Information Systems and Auditing Technical Report, School of Management and Economics, Tilburg University, May 1994.

Corsten, L.C.A. and Stein, A. 1994. Nested sampling for estimating spatial semivariograms compared to other designs. *Applied Stochastic Models and Data Analysis*, 10, 103-122.

Cox, M.G., Harris, P.M., and Jones, H.M. 1990. A knot placement strategy for least squares spline fitting based on the use of local polynomial approximations. In *Algorithms for Approximation, II* (based on the proceedings of the Second International Conference on Algorithms for Approximation, held at Royal Military College of Science, Shrivenham, July 1988), 37-45.

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.

Donohue, J.M., Houck, E.C., and Myers, R.H. 1993a. A sequential experimental design procedure for the estimation of first- and second-order simulation metamodels. *ACM Transactions on Modeling and Computer Simulation*, 3, 190-224.

Donohue, J.M., Houck, E.C., and Myers, R.M. 1993b. Simulation designs and correlation induction for reducing second-order bias in first-order response surfaces. *Operations Research*, 41, 880-902.

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

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.

Dyn, N. and Yad-Shalom, I. 1991. Optimal distribution of knots for tensor product spline approximations. *Quarterly of Applied Mathematics*, 49, 19-27.

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

Eubank, R.L. and Wang, S. 1994. Confidence regions in nonparametric regression. *Scandinavian Journal of Statistics*, 21, 147-158.

Franke, Richard 1982. Scattered data interpolation: tests of some methods. *Mathematics of Computation*, 38, 181-200.

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

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.

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

Handcock, M.S. 1992. On cascading Latin hypercube designs and additive models for experiments. Paper presented at the Joint Statistical Meetings, Boston, August 1992.

Ilgenstein, U. and Sargent, R.G. 1994. An investigation into metamodeling. CASE Center Technical Report No. 9405, Syracuse University, March 1994.

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

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.

Kleijnen, J.P.C. and Van Groendendaal, W. 1994. Two stage versus sequential sample size determination in

regression analysis of simulation experiments. Department of Information Systems and Auditing Technical Report, School of Management and Economics, Tilburg University, May 1994.

Laslett, G.M. 1994. Kriging and splines: an empirical comparison of their predictive performance in some applications (with discussion). *Journal of the American Statistical Association*, 89, 391-409.

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.

McMahon, J. R. and Franke, R. 1992. Knot selection for least squares thin plate splines. *SIAM Journal on Scientific and Statistical Computing* 13, 484-498.

Mitchell, T.J. and Morris, M.D. 1992. The spatial correlation function approach to response surface estimation. *Proceedings of the 1992 Winter Simulation Conference*, J.J. Swain, D. Goldsman, R.C. Crain, and J.R. Wilson, eds. 565-571.

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.

Morrice, D.J. and Schruben, L.W. 1993. Simulation factor screening using harmonic analysis. *Management Science*, 39, 1459-1476.

Morris, M.D., Mitchell, T.J., and Ylvisaker, D. 1993. Bayesian design and analysis of computer experiments: use of derivatives in surface prediction. *Technometrics*, 35, 243-255.

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

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.

Owen, A. 1992. Orthogonal arrays for computer experiments, integration, and visualization. *Statistica Sinica*, 2, 439-452.

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

Salagame, R.R. and Barton, R.R. 1993. Factorial hypercube designs for spatial correlation regression. Working Paper 93-138, Industrial and Management Systems Engineering, The Pennsylvania State University.

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.

Sun, Xingping 1993. Solvability of multivariate interpolation by radial or related functions. *Journal of Approximation Theory*, 72, 252-267.

Tang, Boxin 1993. Orthogonal array-based Latin hypercubes. *Journal of the American Statistical Association*, 88, 1392-1397.

Taylor, C.C. 1994. Simulation methods to estimate smoothing parameters in image reconstruction. *Journal of Statistical Computation and Simulation*, 49, 161-177.

Tew, J.D. 1994. Where is the error in simulation metamodels? *TIMS College on Simulation Newsletter*, 18, Spring 1994.

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. 1992. Validation of simulation analysis methods for the Schruben-Margolin correlation induction strategy. *Operations Research*, 40, 87-103.

Wahba, G. 1986. Partial and interaction splines for the semi-parametric estimation of functions of several variables. In *Computer Science and Statistics: Proceedings of the 18th Symposium on the Interface*, T.J. Boardman, ed., Washington, D.C.: American Statistical Association, 75-80.

Wahba, G. 1990. *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics, 59, Philadelphia: SIAM.

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.