COMPARISON OF RESPONSE SURFACE AND KRIGING MODELS FOR MULTIDISCIPLINARY DESIGN OPTIMIZATION

Timothy W. Simpson[*]
G. W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0405, U.S.A.

Timothy M. Mauery[†]
Department of Civil and Environmental Engineering
Brigham Young University
Provo, Utah 84602, U.S.A.

John J. Korte[‡]
Multidisciplinary Optimization Branch
NASA Langley Research Center
Mail Stop 159, Hampton, Virginia 23681, U.S.A.

Farrokh Mistree[§]
G. W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0405, U.S.A.

ABSTRACT

In this paper, we compare and contrast the use of second-order response surface models and kriging models for approximating non-random, deterministic computer analyses. After reviewing the response surface method for constructing polynomial approximations, kriging is presented as an alternative approximation method for the design and analysis of computer experiments. Both methods are applied to the multidisciplinary design of an aerospike nozzle which consists of a computational fluid dynamics model and a finite-element model. Error analysis of the response surface and kriging models is performed along with a graphical comparison of the approximations, and four optimization problems are formulated and solved using both sets of approximation models. The second-order response surface models and kriging models—using a constant underlying global model and a Gaussian correlation function—yield comparable results.

NOMENCLATURE

β - constant underlying global portion of kriging model
β_i, β_{ij}, β_{ii}, - linear, interaction, and quadratic coefficients of polynomial equation in response surface
DOE - design of experiments
GLOW - gross lift off weight
MSE - mean square error
n_s - number of sample points
$\mathbf{R}$ - correlation matrix in kriging model
$R(\mathbf{x}^i,\mathbf{x}^j)$ - correlation function between points $\mathbf{x}^i$ and $\mathbf{x}^j$
RS - response surface
$\hat{\sigma}^2$ - variance estimate
θ_k - correlation parameters in kriging model
$\hat{y}$ - predicted response value at untried $\mathbf{x}$

---

[*] NSF Graduate Research Fellow and student member AIAA.

[†] Undergraduate Researcher and student member AIAA.

[‡] Senior Research Engineer and senior member AIAA.

[§] Professor and senior member AIAA. **Corresponding author.** Email: farrokh.mistree@me.gatech.edu.

1 FRAME OF REFERENCE

Current engineering analyses rely heavily on running expensive and complex computer codes. Despite the steady and continuing growth of computing power and speed, the complexity of these codes maintains pace with computing advances. Statistical techniques are widely used in engineering design to construct *approximations* of these analysis codes; these approximations are then used in place of the actual analysis codes, offering the following benefits:

- They yield insight into the relationship between output responses, $\mathbf{y}$, and input design variables, $\mathbf{x}$.
- They provide fast analysis tools for optimization and design space exploration since the inexpensive-to-run approximations are used in lieu of the expensive-to-run computer analyses.
- They facilitate the integration of discipline dependent analysis codes.

A common method for building approximations of computer analyses is to apply design of experiments (DOE), response surface (RS) models, and regression analysis to build polynomial approximations of the computationally expensive analyses. For example, the Robust Concept Exploration Method[1] has been developed to facilitate quick evaluation of different design alternatives, identify important design drivers, and generate robust top-level design specifications using DOE, RS models, and the compromise Decision Support Problem; it has been successfully applied to the multiobjective design of a high speed civil transport,[2] a family of General Aviation aircraft,[3] a turbine lift engine,[4] and a flywheel.[5] In other work, the Variable Complexity Response Surface Modeling (VCRSM) method[6] uses analyses of varying fidelity to reduce the design space to the region of interest and build response surface models of increasing accuracy. The VCRSM method employs DOE and RS modeling techniques and has been successfully applied to the multidisciplinary wing design of a high speed civil transport,[7] to the analysis and design of composite

American Institute of Aeronautics and Astronautics

curved channel frames,[8] and to reduce numerical noise in structural analyses.[9,10] A recent review of several applications of response surface models in engineering design is given in Ref. 11; for applications of approximations in structural design, see Ref. 12.

Since computer experiments typically lack random error, a more appropriate and perhaps more "statistically sound" method for approximating deterministic computer experiments is through the use of kriging which is also referred to as the Design and Analysis of Computer Experiments (DACE) models.[13,14] The validity of the kriging model is not dependent on the existence of random error and may be better suited for applications involving computer experiments because it can either "honor the data," providing an exact interpolation, or "smooth the data."[15]

Booker[16] contrasts traditional DOE and RS modeling with DACE models. In the "classical" design and analysis of physical experiments, random variation is accounted for by spreading the sample points out in the design space and by taking multiple data points (replicates), see Figure 1. Sacks, et al.[14] state that the "classical" notions of blocking, replication, and randomization are irrelevant when it comes to deterministic computer experiments; thus, sample points should be chosen to fill the design space.
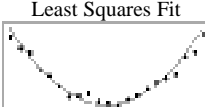
	DOE/RS Modeling for Physical Experiments	DACE/Kriging for Computer Experiments
Experimental Design Input settings at which to obtain output	Account for Variability	Space Filling
Models Estimate output at untried input	Least Squares Fit	Maximum Likelihood Estimate
Validation Determine fit accuracy	t-tests, F-statistics R^2, R^2_{adj}; Residual plots (see, e.g., Ref. 17)	Cross-validation Integrated MSE (see, e.g., Ref. 14)

Figure 1. Comparison of DOE/RS Modeling and DACE/Kriging[16]

As noted in Figure 1, response surface modeling typically employs least squares regression to fit a polynomial model to the sampled data while kriging models are chosen to interpolate the data and are fit using maximum likelihood estimation.[13] Validation of RS models is based on: (a) testing statistical hypothesis (t-tests and F-statistics) derived from error estimates of the variability in the data, (b) plotting and checking the residuals, and (c) computing R^2, the ratio of the model sum of squares to the total sum of squares, and R^2_{adj}, which is R^2 adjusted for the number of parameters in the

model.[17] Meanwhile, Sacks, et al.[14] and Welch, et al.[18] both state that statistical testing is inappropriate when it comes to deterministic computer experiments which lack random error; therefore, cross-validation and integrated mean square error (MSE) are often employed to assess the accuracy of a kriging model.

Kriging and DACE models have found limited application in engineering design perhaps because of the lack of readily available software to fit kriging models, the added complexity of fitting a kriging model, or the additional effort required to use a kriging model. To clarify this last point, prediction with a kriging model requires the inversion and multiplication of several matrices, and the kriging model does not exist as a "closed-form" polynomial equation; this is further clarified in Section 2.2. Meanwhile, RS model prediction requires computation of a simple polynomial equation once the model has been fit. The goal in this paper is to examine the added computational expense required to perform kriging and compare the predictive capability of kriging and RS models.

In Section 2 an overview of the statistical and mathematical foundations of RS modeling and kriging is given. In Section 3 the multidisciplinary aerospike nozzle design example is introduced; it serves as a test problem to compare RS and kriging models for approximation. In Section 4 the RS and kriging models are constructed and validated. In Section 5 four optimization problems are formulated and solved using the RS and kriging models, and Section 6 contains a discussion of ongoing work.

2 STATISTICAL APPROXIMATIONS FOR COMPUTER EXPERIMENTS

Building approximations of computer analyses typically involves: (a) choosing an experimental design to sample the computer analysis code, (b) choosing a model to represent the data, and (c) fitting the model to the observed data. There are a variety of options for each of these choices, and several of the advantages and disadvantages of each—with emphasis on response surface methodology, neural networks, inductive learning and kriging—are discussed in Ref. 11.

In this work, we are primarily concerned with the model choice and model fitting portion of building approximations. In particular, we focus on response surface models (Section 2.1) and kriging (Section 2.2).

2.1 Overview of Response Surface Modeling

2.1.1 Mathematics of Response Surface Modeling
Response surface modeling techniques were originally developed to analyze the results of physical experiments and create empirically-based models of the observed

response values. Response surface modeling postulates a model of the form:

$$y(\mathbf{x}) = f(\mathbf{x}) + \varepsilon \qquad (1)$$

where $y(\mathbf{x})$ is the unknown function of interest, $f(\mathbf{x})$ is a known polynomial function of $\mathbf{x}$, and ε is random error which is assumed to be normally distributed with mean zero and variance σ^2. The individual errors, ε_i, at each observation are also assumed to be independent and identically distributed. The polynomial function, $f(\mathbf{x})$, used to approximate $y(\mathbf{x})$ is typically a low order polynomial which is assumed to be either linear, Eqn. (2), or quadratic, Eqn. (3).

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i \qquad (2)$$

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_i \sum_{j>i} \beta_{ij} x_i x_j \qquad (3)$$

The parameters, β_0, β_i, β_{ii}, and β_{ij}, of the polynomials in Eqns. (2) and (3) are determined through least squares regression which minimizes the sum of the squares of the deviations of predicted values, $\hat{y}(\mathbf{x})$, from the actual values, $y(\mathbf{x})$. The coefficients of Eqns. (2) and (3) used to fit the model can be found using least squares regression given by Eqn. (4):

$$\beta = [\mathbf{X'X}]^{-1}\mathbf{X'y} \qquad (4)$$

where $\mathbf{X}$ is the design matrix of sample data points, $\mathbf{X'}$ is its transpose, and $\mathbf{y}$ is a column vector containing the values of the response at each sample point. For more details on least squares regression or polynomial RS modeling see, e.g., Ref. 17.

2.1.2 General RS Modeling Approach The general approach for building polynomial response surface models is shown in Figure 2. A three step process involving screening, model building, and model exercising is typically employed.

As shown in Figure 2, the first step involves screening which may be employed if there are a large number of factors to reduce the design space to an appropriate region of interest. In the second step, the approximation models are built from sample data which is obtained from an appropriately chosen experimental design; if there are noise factors in the design, robustness models of the mean and variance of each response would also be created. If the models are sufficiently accurate, the model is exercised in the last stage of the process to search the design space and find improved or robust solutions; the reader is referred to, e.g., Ref. 17 for more information on response surface modeling.
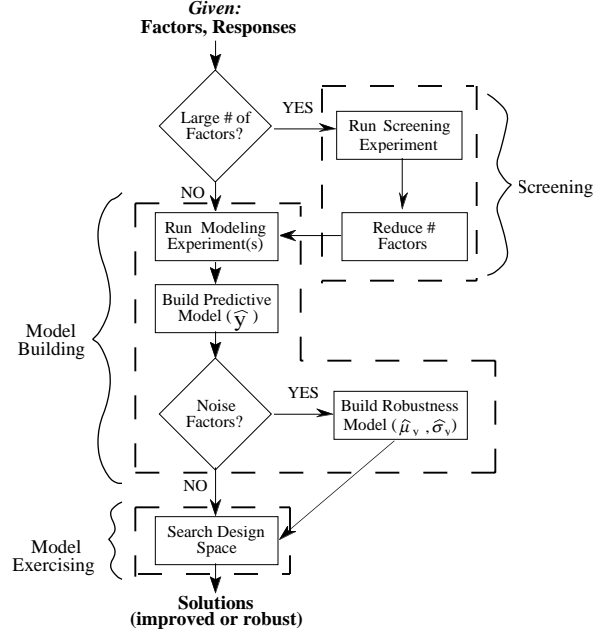

Figure 2. General RS Modeling Approach[11]

2.2 Overview of Kriging

2.2.1 Mathematics of Kriging Kriging postulates a combination of a global model plus departures:

$$y(\mathbf{x}) = f(\mathbf{x}) + Z(\mathbf{x}) \qquad (5)$$

where $y(\mathbf{x})$ is the unknown function of interest, $f(\mathbf{x})$ is a known (usually polynomial) function of $\mathbf{x}$, and $Z(\mathbf{x})$ is the realization of a stochastic process with mean zero, variance σ^2, and non-zero covariance. The $f(\mathbf{x})$ term in Eqn. (5) is similar to the polynomial model in a response surface and provides a "global" model of the design space. In many cases $f(\mathbf{x})$ is simply taken to be a constant term β (see, e.g., Refs. 13, 14, 18); we use a constant term for $f(\mathbf{x})$ in the example in Section 4.

While $f(\mathbf{x})$ "globally" approximates the design space, $Z(\mathbf{x})$ creates "localized" deviations so that the kriging model interpolates the n_s sampled data points. The covariance matrix of $Z(\mathbf{x})$ is given by Eqn. (6).

$$\text{Cov}[Z(\mathbf{x}^i), Z(\mathbf{x}^j)] = \sigma^2 \, \mathbf{R}([R(\mathbf{x}^i, \mathbf{x}^j)]) \qquad (6)$$

In Eqn. (6), $\mathbf{R}$ is the correlation matrix, and $R(\mathbf{x}^i, \mathbf{x}^j)$ is the correlation function between any two of the n_s sampled data points $\mathbf{x}^i$ and $\mathbf{x}^j$. $\mathbf{R}$ is a $(n_s \times n_s)$ symmetric matrix with ones along the diagonal. The correlation function $R(\mathbf{x}^i, \mathbf{x}^j)$ is specified by the user; Sacks, et al.[14] and Koehler and Owen[13] discuss several correlation functions which may be used. In this work, we employ a Gaussian correlation function of the form:

$$R(x^i, x^j) = \exp[-\sum_{k=1}^{n_{dv}} \theta_k |x_k^i - x_k^j|^2] \qquad (7)$$

where n_{dv} is the number of design variables, θ_k are the unknown correlation parameters used to fit the model, and the x_k^i and x_k^j are the kth components of sample points $\mathbf{x}^i$ and $\mathbf{x}^j$. In some cases using a single correlation parameter gives sufficiently good results (see, e.g., Refs. 19, 20, and 14).

Predicted estimates, $\hat{y}(\mathbf{x})$, of the response $y(\mathbf{x})$ at untried values of $\mathbf{x}$ are given by:

$$\hat{y} = \hat{\beta} + \mathbf{r}^T(\mathbf{x})\mathbf{R}^{-1}(\mathbf{y} - \mathbf{f}\hat{\beta}) \qquad (8)$$

where $\mathbf{y}$ is the column vector of length n_s which contains the sample values of the response, and $\mathbf{f}$ is a column vector of length n_s which is filled with ones when $f(\mathbf{x})$ is taken as a constant. In Eqn. (8), $\mathbf{r}^T(\mathbf{x})$ is the correlation vector of length n_s between an untried $\mathbf{x}$ and the sampled data points $\{\mathbf{x}^1, ..., \mathbf{x}^{n_s}\}$:

$$\mathbf{r}^T(\mathbf{x}) = [R(\mathbf{x},\mathbf{x}^1),\ R(\mathbf{x},\mathbf{x}^2),\ ...,\ R(\mathbf{x},\mathbf{x}^{n_s})]^T. \qquad (9)$$

In Eqn. (8), $\hat{\beta}$ is estimated using Eqn. (10).

$$\hat{\beta} = (\mathbf{f}^T\mathbf{R}^{-1}\mathbf{f})^{-1}\mathbf{f}^T\mathbf{R}^{-1}\mathbf{y} \qquad (10)$$

The estimate of the variance, $\hat{\sigma}^2$, between the underlying global model $\hat{\beta}$ and $\mathbf{y}$, is estimated as:

$$\hat{\sigma}^2 = \frac{(\mathbf{y} - \mathbf{f}\hat{\beta})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{f}\hat{\beta})}{n_s} \qquad (11)$$

where $f(\mathbf{x})$ is assumed to be the constant $\hat{\beta}$. The maximum likelihood estimates (i.e., "best guesses") for the θ_k in Eqn. (7) used to fit the model are found by maximizing Eqn. (12) over $\theta_k > 0$ (see, e.g., Ref. 19).

$$-\frac{[n_s \ln(\hat{\sigma}^2) + \ln | \mathbf{R} |]}{2} \qquad (12)$$

Both $\hat{\sigma}^2$ and $|\mathbf{R}|$ are functions of θ_k. While *any* values for the θ_k create an interpolative model, the "best" kriging model is found by solving the k-dimensional unconstrained non-linear optimization problem given by Eqn. (12).

2.2.2 Engineering Applications of Kriging DACE and kriging models have found limited use in engineering design applications since its introduction into the literature by Sacks, et al.[14] Giunta[21] performs a preliminary investigation into the use of DACE modeling for the multidisciplinary design optimization of a High Speed Civil Transport aircraft. He explores a 5 and a 10 variable design problem, observing that the DACE and response surface modeling approaches yield similar results due to the quadratic trend of the responses. Booker, et al.[22] solve a 31 variable helicopter rotor structural design problem; Booker[16] expands the problem to include 56 variables to examine the aeroelastic and dynamic response of the rotor. Osio

and Amon[20] use a multistage DACE modeling strategy to design an embedded electronic package which has 5 design variables. Finally, some researchers have employed DACE modeling strategies specifically for numerical optimization (see, e.g., Refs. 23 and 24).

3 AEROSPIKE NOZZLE EXAMPLE

The multidisciplinary design of an aerospike nozzle has been selected as the test problem for comparing the predictive capability of RS and kriging models. The linear aerospike rocket engine is the propulsion system proposed for the VentureStar[25] Reusable Launch Vehicle (RLV) which is illustrated in Figure 3.
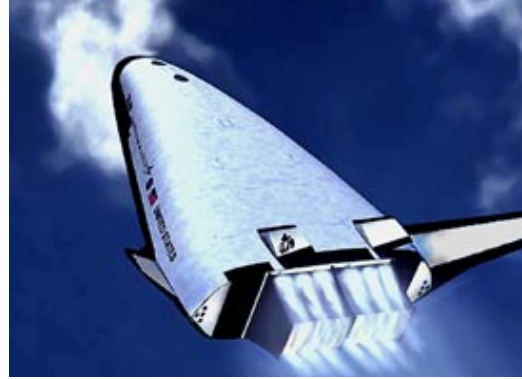

Figure 3. VentureStar RLV with Linear Aerospike Propulsion System[26]

The aerospike rocket engine consists of a rocket thruster, cowl, aerospike nozzle, and plug base regions as shown in Figure 4. The aerospike nozzle is a truncated spike or plug nozzle that adjusts to the ambient pressure and integrates well with launch vehicles.[27] The flow field structure changes dramatically from low altitude to high altitude on the spike surface and in the base flow region.[28-30] Additional flow is injected in the base region to create an aerodynamic spike[31] which gives the aerospike nozzle its name and increases the base pressure and contribution of the base region to the aerospike thrust.
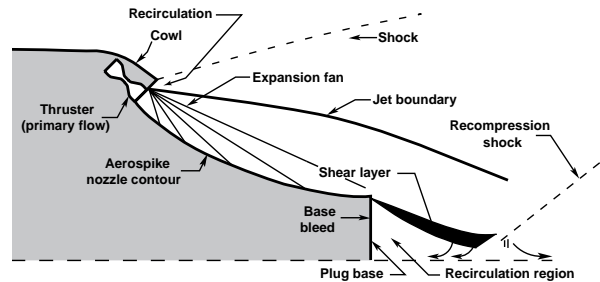

Figure 4. Aerospike Components and Flow Field Characteristics[26]

The analysis of the aerospike nozzle involves two disciplines: aerodynamics and structures; there is an interaction between the structural displacements of the nozzle surface and the pressures caused by the varying aerodynamic effects. Thrust and nozzle wall pressure calculations are made using computational fluid dynamics analysis and are linked to a structural finite-element analysis model for determining nozzle weight and structural integrity. A mission average engine specific impulse and engine thrust/weight ratio are calculated and used to estimate vehicle gross-lift-off-weight (GLOW). The corresponding multidisciplinary domain decomposition is illustrated in Figure 5. Korte, et al.[26] provide additional details on the aerodynamic and structural analyses for the aerospike nozzle.
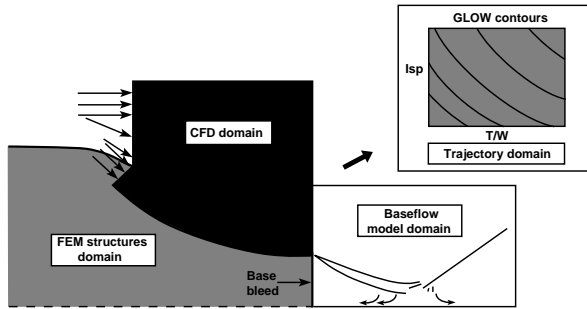

Figure 5. Multidisciplinary Domain Decomposition[26]

For this study, we consider three design variables for the multidisciplinary design of the aerospike nozzle: (starting) thruster angle, (exit) height, and length as shown in Figure 6.
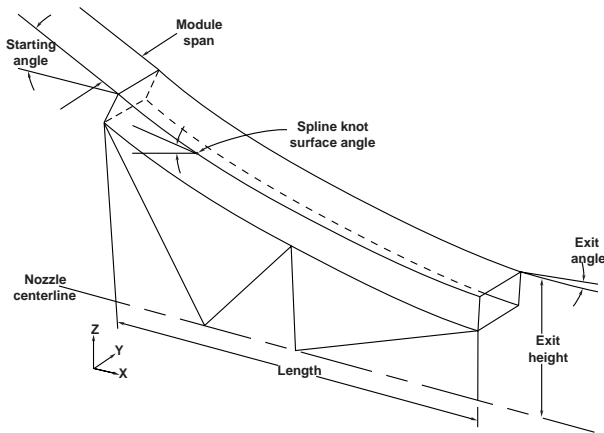

Figure 6. Nozzle Geometry[26]

The thruster angle is the entrance angle of the gas from the combustion chamber onto the nozzle surface; the height and length refer to the solid portion of the nozzle itself. A quadratic curve defines the aerospike nozzle surface profile based on the values of thruster angle, height, and length.

Bounds for the design variables are set to produce viable nozzle profiles from the quadratic model based on all combinations of thruster angle, height, and length within the design space. Second-order response surface models and kriging models are developed for each response (thrust, weight, and GLOW) in the next section; optimization of the aerospike nozzle using the RS and kriging models for different objective functions is performed in Section 5.

4 APPROXIMATIONS FOR THE AEROSPIKE NOZZLE PROBLEM

The data used to fit the RS and kriging models is obtained from a 25 point random orthogonal array;[32] The use of these orthogonal arrays is based, in part, on the work in Ref. 19 and the discussion in Ref. 33. The sample points are illustrated in Figure 7 and are scaled to fit the design space defined by the bounds on the thruster angle (a) , height (h), and length (l).
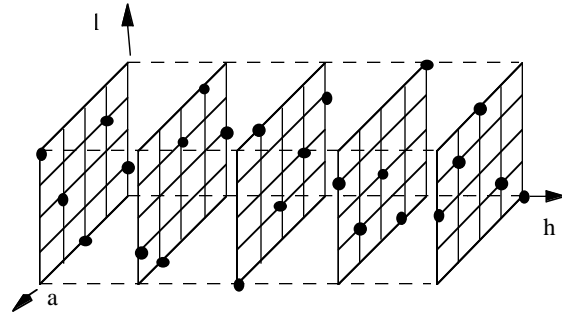

Figure 7. Orthogonal Array Sample Points

In Section 4.1 we discuss the RS models which are fit to the data and in Section 4.2, the kriging models. Error analysis of the RS and kriging models is discussed in Section 4.3, and a graphical comparison of the approximations is given in Section 4.4.

4.1 Response Surface Models

Second-order RS models for weight, thrust, and GLOW are obtained using ordinary least squares regression techniques.[34] The corresponding RS models are given in the Eqns. (13)-(15). The equations have been scaled against the baseline design to protect the proprietary nature of some of the data.

$$\text{Weight} = 0.810 - 0.116a + 0.121h \qquad (13)$$
$$+ 0.152l + 0.065a^2 - 0.025ah + 0.0013h^2$$
$$- 0.0539al - 0.0131hl + 0.0301l^2$$

$$\text{Thrust} = 0.9968 + 0.00031a + 0.0019h \qquad (14)$$
$$+ 0.0060l - 0.00175a^2 + 0.00125ah - 0.0011h^2$$
$$+ 0.00125al - 0.00198hl - 0.00165l^2$$

$$GLOW = 0.9930 - 0.0270a + 0.0065h \quad (15)$$
$$- 0.0265l + 0.0307a^2 - 0.0163ah + 0.0100h^2$$
$$- 0.0226al + 0.0151hl + 0.0195l^2$$

The resulting R^2, R^2_{adj}, and root MSE values for each RS model are given in Table 2. The root MSE is:

$$\text{root MSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (16)$$

where n is the number of sample points, y_i is the actual value of the response, and $\hat{y}_i$ is the predicted value. As evidenced by the high R^2 and R^2_{adj} values and low root MSE values, the second-order RS models appear to capture a large portion of the observed variance.

Table 2. Model Diagnostics of RS Models

| | Response | | |
Measure	Weight	Thrust	GLOW
R^2	0.986	0.998	0.971
R^2_{adj}	0.977	0.996	0.953
root MSE	1.12%	0.01%	0.25%

4.2 Kriging Models for the Aerospike Nozzle Problem

Kriging models are built from the same 25 sample points used to fit the response surface models in Section 4.1. We chose to model the data using a constant term for the global model and a Gaussian correlation function, Eqn. (7), for the local departures determined by the correlation matrix, **R**.

Initial investigations revealed that a single θ parameter was insufficient to accurately model the data due to scaling of the design variables (a similar problem is discussed in Ref. 21). Therefore, an exhaustive grid search with a refinable step size was used to find the maximum likelihood estimates for the three θ parameters needed to obtain the "best" kriging model. The resulting maximum likelihood estimates for three θ parameters for the weight, thrust, and GLOW models are summarized in Table 3; these values are for the scaled sample points.

Table 3. θ Parameters for Kriging Models

| | Response | | |
	Weight	Thrust	GLOW
$\theta_{angle} =$	0.548	0.30	3.362
$\theta_{height} =$	1.323	0.50	2.437
$\theta_{length} =$	2.718	0.65	0.537

With these parameters for the Gaussian correlation function and the 25 sample data points, the kriging models are fully specified. A new point is predicted using these values in combination with Eqns. (8)-(10).

4.3 Error Analysis of Response Surface and Kriging Models

An additional 25 randomly selected validation points are used to verify the accuracy of the RS and kriging models. Error is defined as the difference between the actual response from the computer analysis and the predicted value from the RS or kriging model. The maximum absolute error, the average absolute error, and the root MSE—from Eqn. (16) where n (= 25) is the number of validation points—for the 25 randomly selected validation points are summarized in Table 4.

Table 4. Error Analysis of Approximations

Second-Order Response Surface Models

	Weight	Thrust	GLOW
Max ABS(error)	19.57%	0.032%	3.68%
Avg ABS(error)	2.44%	0.012%	0.53%
root MSE	4.54%	0.015%	0.90%

Kriging Models (w/Constant and Gaussian Cor. Fcn.)

	Weight	Thrust	GLOW
Max ABS(error)	17.23%	0.048%	3.43%
Avg ABS(error)	2.51%	0.012%	0.59%
root MSE	4.37%	0.018%	0.89%

For the weight and GLOW responses, the kriging models have lower maximum absolute errors and lower root MSEs than the RS models; however, the average absolute error is slightly larger for the kriging models. As for thrust, the RS models are slightly better than the kriging models according to the values in the table; the maximum absolute error and root MSE are slightly less while the average absolute errors are essentially the same. It is not surprising that the RS model predicts thrust better; it has a very high R^2 value (0.998) and low root MSE (0.01%). It is reassuring to note, however, that the kriging model, despite using a constant term and a Gaussian correlation function, is only slightly less accurate than the corresponding RS model. In summary, it appears that both models predict well with the kriging models having a slight advantage in accuracy because of the lower root MSE values.

4.4 Graphical Comparison of RS and Kriging Models

In addition to the error analysis of Section 4.3, a graphical comparison of the RS and kriging models is performed to visualize differences in the two approximations. In Figures 8-11, three-dimensional contour plots of thrust, weight, and GLOW as a function of angle, length, and base height are given. In each figure, the same contour levels are used for the RS and kriging models so that the shapes of the contours can be compared directly.
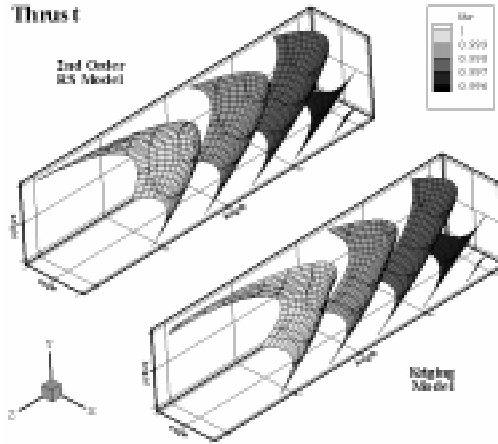
Figure 8.   Thrust Approximation Contours

In Figure 8, the contours of thrust for the RS and kriging models are very similar. As evidenced by the high R^2 and low root MSE values, we expect the RS models to fit the data quite well. It is again reassuring to note that the kriging models resemble the RS models even through the underlying global model for the kriging models is just a constant term.
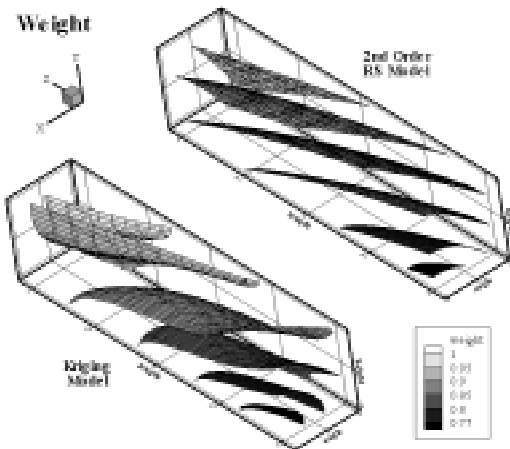

Figure 9.   Weight Approximation Contours

The contours of the RS and kriging models in Figure 9 are also very similar, but we begin to see localized perturbations caused by the Gaussian correlation function in the kriging model for weight. The error analysis from the previous section indicated that the kriging model for weight is slightly more accurate than the RS model which might be attributed to these small non-linear localized variations.
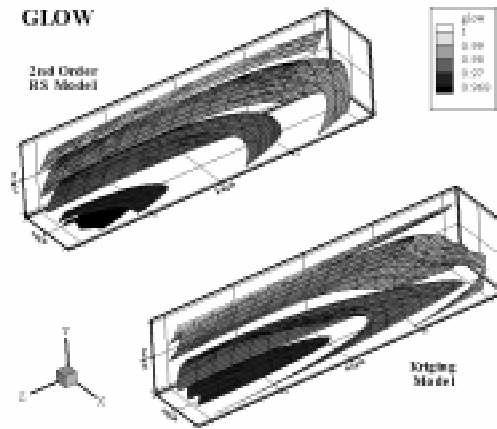

Figure 10.   GLOW Approximation Contours

The general shape of the GLOW contours is the same in Figure 10; however, the size and shape of the different contours, particularly along the length axis, are quite different. The end view along the length axis in Figure 11 further highlights the differences between the two models. Notice also in Figure 11 that the kriging model predicts a minimum GLOW located within the design space centered around Height = -0.8, Angle = 0, along the axis defined by $0.2 \leq$ Length ≤ 0.8; this minimum was verified through additional experiments.
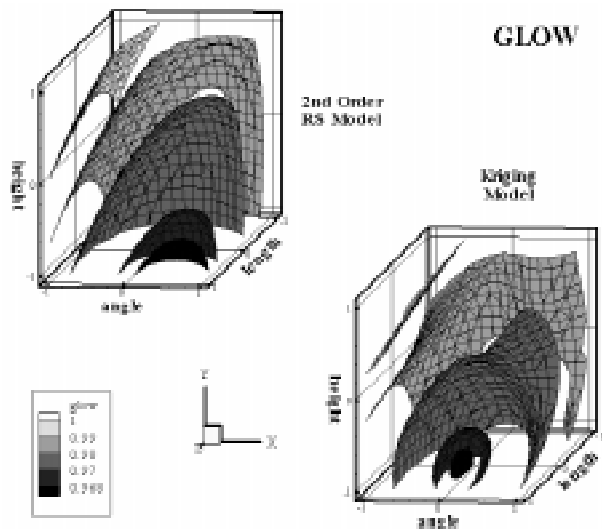

Figure 11.   GLOW Approximation Contours—End View

The true test of the accuracy of the RS and kriging models comes when the approximations are used during optimization. This is performed in the next section.

American Institute of Aeronautics and Astronautics

5 OPTIMIZATION USING RESPONSE SURFACE AND KRIGING MODELS

It is paramount that any approximations used in optimization are reasonably accurate, lest they lead the optimization algorithm into regions of bad designs. Trust Region approaches (see, e.g., Ref. 35) and the Model Management framework (see e.g., Refs. 36 and 37) are being developed to ensure that optimization algorithms are not led astray by inaccurate approximations. In this work the focus has been on developing the approximation models, particularly the kriging models, and not on the optimization itself.

We formulate and solve four different optimization problems to compare the accuracy of the RS and kriging models: (1) maximize thrust, (2) minimize weight, (3) minimize GLOW, and (4) maximize thrust/weight ratio. The first two objective functions represent traditional single objective, single discipline optimization problems. The second two objective functions are more characteristic of multidisciplinary optimization; minimizing GLOW or maximizing the thrust/weight ratio requires trade-offs between the aerodynamics and structures disciplines. For each objective function, constraint limits are placed on the remaining responses; for instance, constraints are placed on the maximum allowable weight and GLOW and the minimum allowable thrust/weight ratio when maximizing thrust. However, none of the constraints are active in any of the final results. The optimization results are summarized in Table 5.

As shown in Table 5, each optimization problem is solved using: (a) the RS model approximations and (b) the kriging model approximations for thrust, weight, and GLOW. The optimization is performed using the Generalized Reduced Gradient (GRG) algorithm in OptdesX.[38] Three different starting points are used for each objective function (the lower, middle, and upper bounds of the design variables) to assess the average number of analysis and gradient calls to the approximations that is necessary to obtain the optimum design within the given design space. The same parameters (i.e., step size, constraint violation, etc.) are used within the GRG algorithm for each optimization. Design variable and response values have been scaled as a percentage of the baseline design to protect the proprietary nature of some of the data.

Table 5. Optimization Results using Response Surface and Kriging Models

	Avg. # of Analysis Calls	Avg. # of Gradient Calls	Optimum Design		Predicted Optimum		Verified Optimum[¶]	% Error[#]
Maximize Thrust								
RS Models	27	4	Angle	0.096	Thrust	1.0016	1.0013	0.02%
			Height	-0.433	Weight	0.9450	0.9476	-0.27%
			Length	1.000	Thr/Wt	1.0141	1.0134	0.07%
					GLOW	0.9724	0.9759	-0.36%
Kriging Models	62	5	Angle	0.656	Thrust	1.0016	1.0014	0.02%
			Height	-0.627	Weight	0.9385	0.9155	2.51%
			Length	1.000	Thr/Wt	1.0157	1.0210	-0.51%
					GLOW	0.9690	0.9683	0.08%
Minimize Weight								
RS Models	29	3	Angle	0.800	Thrust	0.9957	0.9957	-0.01%
			Height	-1.000	Weight	0.7584	0.7496	1.18%
			Length	-1.000	Thr/Wt	1.0533	1.0555	-0.21%
					GLOW	0.9936	0.9906	0.30%
Kriging Models	43	4.67	Angle	1.000	Thrust	0.9965	0.9956	0.08%
			Height	-0.873	Weight	0.7725	0.7443	3.79%
			Length	-1.000	Thr/Wt	1.0506	1.0568	-0.59%
					GLOW	0.9824	0.9914	-0.90%
Minimize GLOW								
RS Models	30.67	3.33	Angle	0.616	Thrust	1.0013	0.9957	0.56%
			Height	-1.000	Weight	0.8969	0.8617	4.09%
			Length	1.000	Thr/Wt	1.0251	1.0286	-0.34%
					GLOW	0.9660	1.0146	-4.79%
Kriging Models	57.67	6.33	Angle	0.764	Thrust	1.0009	1.0006	0.04%
			Height	-0.833	Weight	0.9060	0.8732	3.75%
			Length	0.676	Thr/Wt	1.0228	1.0302	-0.72%
					GLOW	0.9675	0.9680	-0.05%
Maximize Thrust/Weight Ratio								
RS Models	27	4	Angle	0.096	Thrust	1.0016	0.9959	0.57%
			Height	-0.433	Weight	0.9450	0.9073	4.16%
			Length	1.000	Thr/Wt	1.0141	1.0173	-0.31%
					GLOW	0.9724	1.0228	-4.93%
Kriging Models	62	5	Angle	0.656	Thrust	1.0016	1.0014	0.02%
			Height	-0.627	Weight	0.9385	0.9063	3.56%
			Length	1.000	Thr/Wt	1.0157	1.0231	-0.73%
					GLOW	0.9690	0.9666	0.25%

[¶]*The predicted optimum value is obtained by using the values of angle, height, and length (from the optimum design) in the actual analysis code.*
[#]*A (+) error term indicates that the model is over-predicting; a (-) indicates under-predicting.*

The following observations are made based on the data in Table 5.

- *Average number of analysis and gradient calls:* In general, the optimization requires fewer analysis and gradient calls to the RS models than the kriging models. This can be attributed to the fact that the RS models are simple second-order polynomials whereas the kriging models are more complex, non-linear functions.

- *Convergence rates:* Although not shown in the table, optimization using the RS models tends to converge more quickly than when using kriging models. This can be inferred from the number of gradient calls which is one to three calls fewer for the RS models than the kriging models.

- *Optimum designs:* The optimum designs obtained from the RS and kriging models are essentially the same for each objective function. The largest discrepancy is the length when minimizing GLOW; RS models predict the optimum GLOW occurs at the upper bound on length (+1) while the kriging models yield 0.676. This difference is evident in Figures 10 and 11.

- *Predicted optima and prediction errors:* To check the accuracy of the predicted optima, the optimum design values for angle, height, and length are used as inputs into the original analysis codes and the percentage difference between the actual and predicted values is computed. The prediction error is less than 5% for all cases and is 0.5% or less in three quarters of the results.

In summary, the RS and kriging approximations yield comparable results with minimal difference in predictive capability. It is worth noting that *the kriging models perform as well as the second-order RS models even though the global portion of the kriging model is only a constant.* Ongoing work to further improve the accuracy of the kriging models is discussed next.

6 CLOSING REMARKS

This simple, yet realistic, engineering example of the design of an aerospike nozzle has been utilized to demonstrate the use of kriging models as an alternative approximation technique to second-order response surface models for multidisciplinary design optimization. There are several research issues to address for the application of kriging to other (and larger) engineering design problems.

- *Selecting a kriging model:* In this example, we use a constant for the global portion of the kriging model. However, using a global, low-order polynomial model for f($\mathbf{x}$) in Eqn. (5) may further improve the accuracy of the kriging model. Giunta[21] performs a

preliminary investigation of such an approach; his results indicate that minimal improvement in model accuracy is obtained.

- *Fitting a kriging model:* For small problems with relatively few sample points, fitting a kriging model by optimizing Eqn. (12) is rather trivial. However, as the size of the problem increases and the number of sample points increases, the added effort needed to obtain the "best" kriging model may begin to outweigh the benefit of building the approximation.

- *Predicting with a kriging model:* Unlike RS model prediction, prediction with a kriging model requires the inversion and multiplication of several matrices which grow with the number of sample points. Hence, for large problems prediction with a kriging model may become computationally expensive as well. The kriging software we are developing will facilitate fitting, building, and validating kriging models, increasing their attractiveness for engineering applications.

- *Validating a kriging model:* Since kriging models interpolate the data, R^2 values and residual plots cannot be used to assess model accuracy. In this example we use an additional 25 random validation points to check model accuracy; however, other approaches exist. Yesilyurt and Patera[39] and Otto, et al.[40] have developed a Bayesian-validated surrogate approach which systematically uses additional validation points to make quantitative assessments of the quality of the approximation model and provide theoretical bounds on the largest discrepancy between the model and the actual computer analysis. An alternative method which does not require additional points is leave-one-out cross validation,[41] but it is uncertain how well this measure correlates with model accuracy.

- *Design of experiments for building kriging models:* As discussed in Section 1, "space filling" experimental designs may be better suited for computer experiments. In this example, we use orthogonal arrays, but several other experimental designs exist. Koehler and Owen[13] discuss a wide variety of designs including Latin hypercubes, minimax/maximin designs and orthogonal arrays.

Future work on the aerospike nozzle design problem includes adding more design variables and responses and investigating the impact of decomposing the problem into its disciplines by building separate approximations of each discipline and examining the effects of different multidisciplinary design formulations on the solution.

REFERENCES

[1] Chen, W., Allen, J. K., Mavris, D. and Mistree, F., "A Concept Exploration Method for Determining Robust Top-Level Specifications," *Engineering Optimization,* Vol. 26, No. 2, 1996, pp. 137-158.

[2] Koch, P. N., Allen, J. K., Mistree, F. and Mavris, D., "The Problem of Size in Robust Design," *Advances in Design Automation,* Sacramento, CA, ASME, Sept. 14-17, 1997, Paper No. DETC97/DAC-3983.

[3] Simpson, T. W., Chen, W., Allen, J. K. and Mistree, F., "Conceptual Design of a Family of Products Through the Use of the Robust Concept Exploration Method," *6th AIAA/USAF/NASA/ ISSMO Symposium on Multidisciplinary Analysis and Optimization,* Bellevue, WA, AIAA, Vol. 2, Sept. 4-6, 1996, pp. 1535-1545. AIAA-96-4161.

[4] Koch, P. N., Allen, J. K. and Mistree, F., "Configuring Turbine Propulsion Systems Using Robust Concept Exploration," *Advances in Design Automation (Dutta, D., ed.),* Irvine, CA, ASME, Aug. 18-22, 1996, Paper No.96-DETC/DAC-1285.

[5] Lautenschlager, U., Eschenauer, H. A. and Mistree, F., "Multiobjective Flywheel Design: A DOE-Based Concept Exploration Task," *Advances in Design Automation (Dutta, D., ed.),* Sacramento, CA, ASME, Sept. 14-17, 1997, Paper No. DETC97/DAC-3961.

[6] Giunta, A. A., Balabanov, V., Haim, D., Grossman, B., Mason, W. H. and Watson, L. T., "Wing Design for a High-Speed Civil Transport Using a Design of Experiments Methodology," *6th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization,* Bellevue, WA, AIAA, Vol. 1, Sept. 4-6, 1996, pp. 168-183.

[7] Kaufman, M., Balabanov, V., Burgee, S. L., Giunta, A. A., Grossman, B., Mason, W. H. and Watson, L. T., "Variable-Complexity Response Surface Approximations for Wing Structural Weight in HSCT Design," *34th Aerospace Sciences Meeting and Exhibit,* Reno, NV, Jan. 15-18, 1996, AIAA-96-0089.

[8] Mason, B. H., Haftka, R. T. and Johnson, E. R., "Analysis and Design of Composite Channel Frames," *5th AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization,* Panama City, FL, AIAA, Vol. 2, Sept. 7-9, 1994, pp. 1023-1040. AIAA-94-4363.

[9] Giunta, A. A., Dudley, J. M., Narducci, R., Grossman, B., Haftka, R. T., Mason, W. H. and Watson, L. T., "Noisy Aerodynamic Response and Smooth Approximations in HSCT Design," *5th AIAA/USAF/NASA/ISSMO Symposium on Multi-disciplinary Analysis and Optimization,* Panama City, FL, AIAA, Vol. 2, Sept. 7-9, 1994, pp. 1117-1128. AIAA-94-4376-CP.

[10] Venter, G., Haftka, R. T. and Starnes, J. H., Jr., "Construction of Response Surfaces for Design Optimization Applications," *6th AIAA/USAF/ NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization,* Bellevue, WA, AIAA, Vol. 1, Sept. 4-6, 1996, pp. 548-564.

[11] Simpson, T. W., Peplinski, J., Koch, P. N. and Allen, J. K., "On the Use of Statistics in Design and the Implications for Deterministic Computer Experiments," *Design Theory and Methodology - DTM'97,* Sacramento, CA, ASME, Sept. 14-17, 1997, Paper No. DETC97/DTM-3881.

[12] Barthelemy, J.-F. M. and Haftka, R. T., "Approximation Concepts for Optimum Structural Design - A Review," *Structural Optimization,* Vol. 5, 1993, pp. 129-144.

[13] Koehler, J. R. and Owen, A. B., "Computer Experiments," *Handbook of Statistics (Ghosh, S. and Rao, C. R., eds.),* Elsevier Science, New York, 1996, pp. 261-308.

[14] Sacks, J., Welch, W. J., Mitchell, T. J. and Wynn, H. P., "Design and Analysis of Computer Experiments," *Statistical Science,* Vol. 4, No. 4, 1989, pp. 409-435.

[15] Cressie, N. A. C., *Statistics for Spatial Data,* John Wiley & Sons, New York, 1993.

[16] Booker, A. J., "Case Studies in Design and Analysis of Computer Experiments," *Proceedings of the Section on Physical and Engineering Sciences,* American Statistical Association, 1996.

[17] Myers, R. H. and Montgomery, D. C., *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, John Wiley & Sons, New York, 1995.

[18] Welch, W. J., Yu, T.-K., Kang, S. M. and Sacks, J., "Computer Experiments for Quality Control by Parameter Design," *Journal of Quality Technology,* Vol. 22, No. 1, 1990, pp. 15-22.

[19] Booker, A. J., Conn, A. R., Dennis, J. E., Frank, P. D., Trosset, M. and Torczon, V., "Global Modeling for Optimization: Boeing/IBM/Rice Collaborative Project," *ISSTECH-95-032,* The Boeing Company, Seattle, WA, 1995.

[20] Osio, I. G. and Amon, C. H., "An Engineering Design Methodology with Multistage Bayesian Surrogates and Optimal Sampling," *Research in Engineering Design,* Vol. 8, 1996, pp. 189-206.

[21] Giunta, A., Watson, L. T. and Koehler, J., "A Comparison of Approximation Modeling Techniques: Polynomial Versus Interpolating Models," *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis & Optimization,* St. Louis, MI, AIAA, Sept. 2-4, 1998, AIAA-98-4758.

[22] Booker, A. J., Conn, A. R., Dennis, J. E., Frank, P. D., Serafini, D., Torczon, V. and Trosset, M., "Multi-Level Design Optimization: A Boeing/IBM/ Rice Collaborative Project," *ISSTECH-96-031,* The Boeing Company, Seattle, WA, 1996.

[23] Trosset, M. W. and Torczon, V., "Numerical Optimization Using Computer Experiments," *Report No. TR97-02,* Dept. of Comp. and App. Math., Rice University, Houston, TX, 1997.

[24] Cox, D. D. and John, S., "SDO: A Statistical Method for Global Optimization," *Proceedings of the ICASE/NASA Langley Workshop on Multidisciplinary Optimization (Alexandrov, N. M. and Hussaini, M. Y., eds.),* Hampton, VA, SIAM, March 13-16, 1995, pp. 315-329.

[25] Sweetman, B., "VentureStar: 21st Century Space Shuttle", *Popular Science*, Oct. 1996, pp. 42-47

[26] Korte, J. J., Salas, A. O., Dunn, H. J., Alexandrov, N. M., Follett, W. W., Orient, G. E. and Hadid, A. H., "Multidisciplinary Approach to Aerospike Nozzle Design," *NASA-TM-110326,* NASA, Hampton, VA, 1997.

[27] Rao, G. V. R., "Recent Developments in Rocket Nozzle Configurations," *ARS Journal,* Sept., 1961, pp. 1488-1494.

[28] Hagemann, G., Schley, C.-A., Odintsov, E. and Sobatchkine, A., "Nozzle Flowfield Analysis with Particular Regard to 3D-Plug-Cluster Configurations," July, 1996, AIAA-96-2954.

[29] Mueller, T. J. and Sule, W. P., "Basic Flow Characteristics of a Linear Aerospike Nozzle Segment," ASME, Nov., 1972, Paper No. 72-WA/Aero-2.

[30] Rommel, R., Hagemann, G., Shley, C., Krülle, G. and Manski, D., "Plug Nozzle Flowfield Calculations for SSTO Applications," July, 1995, AIAA-95-2784.

[31] Iacobellis, S. F., Larson, V. R. and Burry, R. V., "Liquid-Propellant Rocket Engines: Their Status and Future," *Journal of Spacecraft and Rockets,* Vol. 4, 1967, pp. 1569-1580.

[32] Owen, A. B., "Orthogonal Arrays for Computer Experiments, Integration and Visualization," *Statistica Sinica,* Vol. 2, 1992, pp. 439-452.

[33] Barton, R. R., "Metamodeling: A State of the Art Review," *Proceedings of the 1994 Winter Simulation Conference (Tew, J. D., et al., eds.),* IEEE, Dec. 11-14, 1994, pp. 237-244.

[34] SAS, *JMP® User's Guide, Version 3.1*, SAS Institute, Inc., Cary, NC, 1995.

[35] Rodriguez, J. F., Renaud, J. E. and Watson, L. T., "Trust Region Augmented Lagrangian Methods for Sequential Response Surface Approximation and Optimization," *Advances in Design Automation (Dutta, D., ed.),* Sacramento, CA, ASME, Sept. 14-17, 1997, Paper No. DETC97/DAC-3773.

[36] Alexandrov, N., Dennis, J.E., Jr., Lewis, R. M., and Torczon, V., "A Trust Region Framework for Managing the Use of Approximation Models in Optimization," *NASA CR-201745, ICASE Report No. 97-50,* NASA Langley Research Center, Hampton, VA, October 1997.

[37] Dennis, J. E. and Torczon, V., "Managing Approximation Models in Optimization," *Proceedings of the ICASE/NASA Langley Workshop on Multidisciplinary Design Optimization (Alexandrov, N. M. and Hussaini, M. Y., eds.),* Hampton, VA, SIAM, March 13-16, 1995, pp. 330-347.

[38] Design Synthesis, Inc., "OptdesX™: A Software System for Optimal Engineering Design," *User's Manual, Release 2.0.4,* Provo, UT, 1998.

[39] Yesilyurt, S. and Patera, A. T., "Surrogates for Numerical Simulations; Optimization of Eddy-Promoter Heat Exchangers," *Computer Methods in Applied Mechanics and Engineering,* Vol. 121, No. 1-4, 1995, pp. 231-257.

[40] Otto, J. C., Landman, D. and Patera, A. T., "A Surrogate Approach to the Experimental Optimization of Multielement Airfoils," *6th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization,* Bellevue, WA, AIAA, Vol. 2, Sept. 4-6, 1996, pp. 1323-1333. AIAA-96-4138.

[41] Mitchell, T. J. and Morris, M. D., "Bayesian Design and Analysis of Computer Experiments: Two Examples," *Statistica Sinica,* Vol. 2, 1992, pp. 359-379.