AIAA JOURNAL
Vol. 51, No. 4, April 2013

Comparison of Adaptive Sampling Methods for Generation of Surrogate Aerodynamic Models

T. J. Mackman* and C. B. Allen†
University of Bristol, Bristol, England BS8 1TR, United Kingdom
M. Ghoreyshi‡
U.S. Air Force Academy, Colorado 80840
and
K. J. Badcock§
University of Liverpool, Liverpool, England L69 3GH, United Kingdom

DOI: 10.2514/1.J051607

A surrogate modeling strategy, using effective interpolation and sampling methods, facilitates a reduction in the number of computational fluid dynamics simulations required to construct an aerodynamic model to a specified accuracy. In this paper, two adaptive sampling strategies are compared for generating surrogate models, based on Kriging and radial basis function interpolation, respectively. The relationships between the two model formulations are discussed, and three test cases are considered, including analytic functions and recovery of aerodynamic coefficients for two example applications: longitudinal flight mechanics analysis for the DLR-F12 aircraft and structural loads analysis of an RAE2822 airfoil. For the airfoil example, models of C_L, C_D, and C_M were constructed with the two sampling strategies using Euler/boundary-layer-coupled computational fluid dynamics and a three-dimensional flight envelope of incidence, Mach, and Reynolds number. The two sampling approaches direct some samples toward exploration of the domain by minimizing model uncertainty and some toward refinement of local nonlinearities, by adapting to model curvature or extrema. The results provide some evidence that, for certain functions, in certain scenarios, each update scheme could be useful. Both methods were at least better than traditional space-filling sampling for all the test cases presented.

I. Introduction

IN RECENT years, there has been increasing adoption of surrogate models in the aerospace community for generating databases of computational fluid dynamics (CFD) data [1,2]. A surrogate modeling strategy, using effective interpolation and sampling methods, facilitates a reduction in the number of simulations required to construct an aerodynamic database to a specified accuracy. Interpolation allows a continuous model to be reconstructed from a finite set of simulations, and an efficient method for sampling the relevant parameters allows the interesting nonlinear aerodynamics to be represented as well as possible, using a small budget of CFD runs.

For aircraft structural loads analysis, a model of the global forces and moments on the aircraft, and their corresponding distributions on the components, is required across the flight envelope, which may relate to a range of Reynolds numbers, Mach numbers, incidence and sideslip angles, and control surface deflections [3]. Without surrogate modeling, this parameter space is traditionally populated using a factorial design with regular intervals in each dimension, which leads to a large number of required simulations. A surrogate modeling approach that reduces the amount of effort spent on sufficiently populating the space allows critical loads cases to be identified earlier, and allows a greater amount of effort to be spent on further simulations of these points.

Furthermore, as computational resources and simulation capability increase, there is a growing desire to use physics-based simulations for additional aspects of aircraft design. For example, the design of unstable configurations could be helped by the use of physics-based simulations to reduce the lead time on production of high-quality aerodynamic models toward development of control laws [4]. A model of global forces and moments with respect to aircraft state and control variables is required for flight mechanics analysis, which is typically fulfilled by a full factorial lookup table of wind tunnel measurements taken relatively late in the design cycle. This can be replaced by a surrogate model of CFD simulations taken earlier in the design cycle, using an appropriate sampling strategy to reduce the potentially significant computational cost of running a new simulation for every entry in the table.

A comparison of sampling strategies for generating surrogate models is presented in this paper, considering three test cases, including analytic functions, recovery of aerodynamic coefficients toward longitudinal stability and control analysis, and recovery of aerodynamic coefficients toward structural loads analysis. Two adaptive sampling methods [4,5], one based on Kriging [6] and one based on radial basis function (RBF) interpolation [7], are compared against a standard space-filling Latin hypercube approach, and full factorial designs and their relative merits are discussed for the different test problems. Adaptive sampling methods, which incrementally add points based on interim models of the data, are a natural accompaniment to surrogate modeling, and especially so for problems involving expensive simulations in which it is not possible to generate all the required data simultaneously. The two approaches presented both aim to direct a portion of the update samples to refinement of local nonlinearities and the remainder to exploration of the parameter space.

First, the relevant background is discussed, then the model formulations are presented, followed by the sampling strategies and a discussion of results for the three test cases. In the results section, the two-dimensional analytic test functions are presented first, followed

Presented as Paper 2011-1171 at the 49th AIAA Aerospace Sciences Meeting, Orlando, Florida, 4–7 January 2011; received 9 September 2011; revision received 29 August 2012; accepted for publication 1 September 2012; published online 8 February 2013.

*Ph.D. Student, Department of Aerospace Engineering. Student Member AIAA.
†Professor of Computational Aerodynamics, Department of Aerospace Engineering. Senior Member AIAA.
‡Research Fellow. Senior Member AIAA.
§Professor of Computational Aerodynamics, School of Engineering. Senior Member AIAA.

by the simple stability and control analysis application, which uses data from the U.S. Air Force Stability and Control Digital DATCOM [8], and finally the structural loads analysis application, which uses a set of simulation data for the RAE2822 airfoil generated using the MSES flow solver [9], and which incorporates some highly nonlinear features.

II. Background

Much of the current thinking on the design of experiments for surrogate modeling of computer simulations stems from the paper by Sacks et al. [10], who presented a statistical framework for building surrogate models and determining sample locations by minimizing an associated uncertainty measure. They noted that, for deterministic data, modeling error can be minimized without any knowledge of the data by placing samples in a space-filling manner. Thus, it has become customary to use space-filling designs for constructing global surrogate models where good accuracy is required throughout the whole parameter space. They also made the distinction between one-stage sampling methods, sequential sampling methods, and sequential sampling with adaption to the data, depending on how the chosen sampling criterion is evaluated. Adaptive sampling offers a potential improvement to customary one-stage or sequential space-filling methods by using information from successive metamodels as the number of points is increased.

The most notable space-filling methods include factorial designs, Latin hypercube sampling (LHS), low-discrepancy sequences, and designs based on statistical optimality criteria (A-, D-, and G-optimal designs), which are discussed in the detailed review papers by Giunta et al. [11], Koehler and Owen [12], Simpson et al. [13], and Chaloner and Verdinelli [14]. Factorial designs are extremely simple to construct and convey results intuitively, hence they have traditionally been used for computer simulations when no surrogate is constructed, despite having relatively poor space-filling properties. Latin hypercube designs are probably the most popular method for constructing surrogate models, but in their original form, as presented by McKay et al. [15], space-filling behavior is not guaranteed. The Latin hypercube requirement is simply that all one-dimensional projections of the design have the maximal number of intervals, and it is necessary to then select from the possible Latin hypercube designs for a given number of points using a separate criterion that quantifies how space-filling different arrangements are. This can be done by maximizing the minimum distance between samples (a *maximin* distance design [16]) or by minimizing the "potential energy" of the sample set (an Audze–Eglais design [17]), either of which is quite a challenging optimization problem in practice ([18] pp. 17–27), because the number of possible permutations of sample coordinates increases dramatically with dimensionality and the number of samples. An alternative might be to use a deterministic (hence, cheap to compute) low-discrepancy sequence, where *discrepancy* is a measure of the uniformity of the samples produced. One of the best known sequences, in terms of the rate of convergence of the discrepancy with respect to the number of samples, is the Sobol sequence (also the LP_τ sequence, or (t, s) sequence of base two) [19,20].

The Sobol sequence employs a somewhat similar concept of stratification to LHS, whereby the interval $[0, 1]$ is split into successive uniform intervals using special binary fractions, called direction numbers, providing coordinates for each dimension [21,22]. Unlike LHS, however, careful formulation of the direction numbers for each dimension and imposed uniformity properties produce a deterministic sequence that can be used to generate space-filling points very efficiently, with a discrepancy convergence rate of $\mathcal{O}(\log^n N)$ for N points in n dimensions, which is believed to be as good as possible for an infinite sequence [20].

In general, adaptive methods are either multiple-criteria methods or statistical methods, with criteria that reflect both the spread of the data and the function values themselves. Jin et al. [23] provide a fairly comprehensive description of many existing methods and some proposals of their own. They propose a two-criteria method based on the leave-one-out cross-validation (LOOCV) error,[¶] evaluated throughout the domain, and scaled by the distance to the nearest existing sample point. This is a promising form of approach, also used by Aute et al. [24] and others, but the LOOCV error can be a particularly poor indicator for adaptive sampling if the initial model contains too few data points. Also, the computational cost of calculating the LOOCV error explicitly at arbitrary locations is of $\mathcal{O}(N^4)$ for N samples, hence an interpolated model of the LOOCV error at the data sites is often used [24] (for which $\mathcal{O}(N^3)$ can be achieved [25]), but this introduces an additional level of approximation.

Another two-criteria approach worth mentioning is by Tang et al. [26], who developed a strategy for sampling a three-dimensional flight envelope for a crew transfer vehicle, which achieved good results. They divided the domain into bins with only one point allowed per bin, placed randomly within it, and then chose only to add points in bins with large gradients or Laplacian. Other authors have also used a function gradient criterion and a separate space-filling criterion, notably Turner et al. [27] and, more recently, Jakobsson et al. [28], who used a rational RBF model and based the space-filling criterion on the power function, which is an uncertainty measure for RBF models.

Statistical criteria, using the framework of Sacks et al. [10], are particularly popular owing to the widespread use of Kriging models. Maximizing the entropy, or amount of information provided by a sample, is the simplest and most intuitive of these methods, involving successively adding sample points at the locations with the largest value of error predicted by the Kriging mean squared error (MSE) function [29]. This will be referred to as the entropy or MaxMSE criterion. MaxMSE samples are primarily space-filling but adapt to the relative variability of each coordinate direction as well, because the MSE function is dependent on the both the sample positions and the Kriging model parameters, which are optimized to fit the data. A number of statistical criteria have also been developed for the purpose of global optimization, which attempt to both explore the domain and refine minima or maxima [30]. This is a slightly different objective to refining nonlinearities everywhere in the domain, but optimization criteria can be useful for building complete surrogate models as part of a larger sampling strategy, ensuring that the absolute values of peaks and troughs are well predicted. A popular statistical optimization criterion is the expected improvement function (EIF), described in detail by Jones et al. [31].

Adaptive sampling strategies developed by the authors have used the MaxMSE and EIF criteria to develop Kriging models for tables of aerodynamic data for stability and control analysis [4], and used the model Laplacian and a sample separation function to build RBF models of analytic functions and aerodynamic data [5,32]. In this work, the RBF sampling approach is given a space-filling criterion based on the power function, and the two methods are compared for the chosen test problems, with factorial designs and Audze-Eglais-optimized Latin hypercube sampling used as baseline data-independent methods.

To further reduce the cost of generating models of aerodynamic data, it is possible to use gradient information for constructing the surrogate [33], or to merge data from both low- and high-fidelity simulations [3,4,34,35], or from simulation and experiment [36]. Variable fidelity modeling such as this also introduces the possibility of sampling based on the difference between the high- and low-fidelity data. However, for simplicity, the scope of this comparison study is limited to regular surrogate models.

III. Methods

The formulations of the Kriging and RBF models are presented here. Let $y(\mathbf{x})$ be the original function to be modeled, and let y_i be the

[¶]The LOOCV error at $\mathbf{x}$ for a given sample i is the difference between the full model and the model with point i removed, i.e., $\hat{y}_{-i}(\mathbf{x}) - \hat{y}(\mathbf{x})$. For adaptive sampling, this must be evaluated at arbitrary $\mathbf{x}$ throughout the domain. LOOCV can also provide a single global error metric for the whole model by taking the norm (usually L^1 or L^2) of these errors at all sample locations.

values at N sample points where $\mathbf{x}_i$ is the vector of coordinates $\mathbf{x}_i = \{x_{i,1}, \ldots, x_{i,n}\}$ for the ith sample point. The samples can be described as a set $\mathcal{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ confined to the domain Ω in n-dimensional space, which is normalized to the unit hypercube $[0, 1]^n$ for interpolation. The label y denotes a scalar function, and multivariate responses are considered by constructing separate models for each response.

The Kriging model was implemented using the DACE toolbox developed by Lophaven et al. [37,38], and the RBF model was implemented by the authors, using the theory presented by Wendland [39] and Fasshauer [40] as reference.

A. Kriging Model

A Kriging model approximates the target function at an untried site $\mathbf{x}_*$ as

$$\hat{y}(\mathbf{x}_*) = \mu + \epsilon \tag{1}$$

where μ is the mean value and ϵ is a normally distributed error term with zero mean and variance σ^2. Different types of Kriging models based on different assumptions about the mean are discussed by Goovaerts [41]. *Universal* Kriging assumes only that the mean value μ is a linear combination of known functions $f_1(\mathbf{x}_i), \ldots, f_M(\mathbf{x}_i)$. It is common to use linear functions $f_k(\mathbf{x}_i) = 1, x_{i,1}, \ldots, x_{i,n}$, such that the mean is a linear regression model, but higher-order terms can be used. A universal Kriging model is written as

$$\hat{y}(\mathbf{x}_*) = \sum_{k=1}^{M} \gamma_k f_k(\mathbf{x}_*) + \epsilon \tag{2}$$

where $\gamma = \{\gamma_1, \ldots, \gamma_M\}$ is the M-dimensional vector of regression coefficients, and $M = n + 1$ for a linear mean.

The error term ϵ is assumed to be a random process, because it is in a pure regression model, but is assumed to be a continuous function $\epsilon(\mathbf{x})$ with correlated values $\epsilon_i, i = 1, \ldots, N$ depending on their distance apart. To estimate the correlation for the error term, a spatially weighted distance formula between samples $\mathbf{x}_i$ and $\mathbf{x}_{ii}$ is generally defined as

$$d(\mathbf{x}_i, \mathbf{x}_{ii}) = \sum_{j=1}^{n} \theta_j |x_{i,j} - x_{ii,j}|^{p_j}, \qquad (\theta_j \geq 0, p_j \in [1, 2]) \tag{3}$$

where the parameter θ_j expresses the importance of the jth coordinate and the exponent p_j is related to the smoothness of the function in coordinate direction j. A correlation matrix $\mathbf{R}$ of all pairwise correlations can then be written as

$$\mathbf{R} = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,N} \\ \phi_{2,1} & \phi_{2,2} & \cdots & \phi_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{N,1} & \phi_{N,2} & \cdots & \phi_{N,N} \end{bmatrix} \tag{4}$$

where $\phi_{i,ii}(\mathbf{x}_i, \mathbf{x}_{ii})$ is the correlation function. The form of ϕ can be chosen, and for this work a standard exponential of the form $\phi_{i,ii} = \mathrm{e}^{-d(\mathbf{x}_i, \mathbf{x}_{ii})}$ was used, with all p_j fixed to two. This is an intuitive choice because $\epsilon(\mathbf{x})$ is then a Gaussian process, and the correlation function ϕ has a radial basis.

To compute the Kriging model, values must be estimated for the regression coefficients $\gamma_1, \ldots, \gamma_M$, in conjunction with the model parameters $\theta_1, \ldots, \theta_n$ (and $p_1, \ldots, p_n$ if used). The parameters can be quantified by an optimization process called maximum likelihood estimation, as discussed in detail by Toal et al. [42] and Martin and Simpson [43], and a generalized least-squares (GLS) fit for the mean gives

$$\gamma = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{y} \tag{5}$$

The predictions at an unsampled location $\mathbf{x}_*$ can then be obtained by expanding Eq. (1). By writing the correlation and regression terms as follows,

$$\mathbf{r} = \begin{bmatrix} \phi_{*,1} \\ \phi_{*,2} \\ \vdots \\ \phi_{*,N} \end{bmatrix}, \qquad \mathbf{f} = \begin{bmatrix} f_1(\mathbf{x}_*) \\ f_2(\mathbf{x}_*) \\ \vdots \\ f_M(\mathbf{x}_*) \end{bmatrix},$$
$$\mathbf{F} = \begin{bmatrix} f_1(\mathbf{x}_1) & f_2(\mathbf{x}_1) & \ldots & f_M(\mathbf{x}_1) \\ f_1(\mathbf{x}_2) & f_2(\mathbf{x}_2) & \ldots & f_M(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(\mathbf{x}_N) & f_2(\mathbf{x}_N) & \ldots & f_M(\mathbf{x}_N) \end{bmatrix} \tag{6}$$

the standard form of the Kriging predictor for point $\mathbf{x}_*$ is given by

$$\hat{y}(\mathbf{x}_*) = \mu(\mathbf{x}_*) + \epsilon(\mathbf{x}_*) \quad = \mathbf{f}^T \gamma + \mathbf{r}^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{F}\gamma) \tag{7}$$

An alternative way of writing Eq. (7) is as follows:

$$\hat{y}(\mathbf{x}_*) = \mathbf{f}^T \gamma + \mathbf{r}^T \beta, \qquad \beta = \mathbf{R}^{-1} (\mathbf{y} - \mathbf{F}\gamma) \tag{8}$$

$$= \mu(\mathbf{x}_*) + \sum_{i=1}^{N} \beta_i \phi(\mathbf{x}_*, \mathbf{x}_i) \tag{9}$$

which shows that a Kriging model can be described as a linear combination of basis functions, scaled by the coefficients β_i, plus a contribution from the mean [31]. Equation (8) is useful in practice, because the vectors β and γ are fixed and may be stored after their initial calculation.

An uncertainty measure can also be calculated for the predictor. The Kriging variance, or *mean squared error* as it is usually called [10], is given by

$$s^2(\mathbf{x}_*) = \sigma^2 (1 + \mathbf{u}^T (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{u} - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r}),$$
$$\mathbf{u} = \mathbf{F}\mathbf{R}^{-1}\mathbf{r} - \mathbf{f} \tag{10}$$

where σ^2 is the sample variance with respect to the mean μ, given by

$$\sigma^2 = \frac{1}{n} (\mathbf{y} - \mathbf{F}\gamma)^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{F}\gamma) \quad = \frac{1}{n} \beta^T \mathbf{R} \beta \tag{11}$$

If $\mathbf{x}_*$ is close to a sample point, there is a small amount of uncertainty in the model and, conversely, if $\mathbf{x}_*$ is far away from all sample points, there is a large amount of uncertainty. The MSE function reflects this and has a value of zero at the data sites and an increasing value away from them, dependent on the correlation between $\mathbf{x}_*$ and the other points, which is a function of distance and the model parameters, but has no direct dependence on the response values.

B. Radial Basis Function Interpolation

An RBF model is a linear combination of basis functions, whose argument is the Euclidean distance between the evaluation point $\mathbf{x}_*$ and all points in the training dataset, known as the basis function centers. If ϕ is the chosen basis function and $\| \cdot \|$ is used to denote the Euclidean norm, the interpolation model has the form

$$\hat{y}(\mathbf{x}_*) = \sum_{i=1}^{N} \beta_i \phi(\|\mathbf{x}_* - \mathbf{x}_i\|) + p(\mathbf{x}_*) \tag{12}$$

where $\beta_i, i = 1, \ldots, N$ are the model coefficients, and $p(\mathbf{x}_*)$ is an optional polynomial. The coefficients are found by requiring exact recovery of the original data $\hat{\mathbf{y}}_{\mathcal{X}} = \mathbf{y}$ for all training data points $\mathcal{X}$.

The model is then a sum of contributions from the polynomial $p(\mathbf{x}_*)$ and the basis functions, which form a set of departures. The general form for a polynomial is given by

$$p(\mathbf{x}) = \sum_{k=1}^{M} \gamma_k f_k(\mathbf{x}) \tag{13}$$

where $f(\mathbf{x})$ are the monomial components, and γ_k are the M additional coefficients that must be solved for. A side constraint is introduced to provide the additional equations required to solve the system. This ensures that the polynomial is orthogonal to the basis functions, which is the same as requiring a GLS approximation. The side constraint is written as

$$\sum_{i=1}^{N} \beta_i f_k(\mathbf{x}_i) = 0, \qquad k = 1, \ldots, M \tag{14}$$

The problem can be cast as an augmented linear system in matrix form, which must be solved before any new approximations can be made. Writing the training data vector $\mathbf{y}$, the coefficient vectors β and γ, and the matrices of basis function and monomial terms $\mathbf{R}$ and $\mathbf{F}$, respectively, the system may be written as follows:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{F} \\ \mathbf{F}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \beta \\ \gamma \end{bmatrix} \tag{15}$$

where

$$\mathbf{R} = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,N} \\ \phi_{2,1} & \phi_{2,2} & \cdots & \phi_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{N,1} & \phi_{N,2} & \cdots & \phi_{N,N} \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} f_{1,1} & \cdots & f_{1,M} \\ f_{2,1} & \cdots & f_{2,M} \\ \vdots & \ddots & \vdots \\ f_{N,1} & \cdots & f_{N,M} \end{bmatrix} \tag{16}$$

and where $\phi_{i,ii} = \phi(\|\mathbf{x}_i - \mathbf{x}_{ii}\|)$, $f_{i,k} = f_k(\mathbf{x}_i)$. The points $\mathbf{x}_i$ and $\mathbf{x}_{ii}$ are both data sites, and $\mathbf{R}$ contains basis function values for all pairwise combinations.

The solution of the system is given by

$$\gamma = (\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F})^{-1}\,\mathbf{F}^T\mathbf{R}^{-1}\mathbf{y} \quad \beta = \mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\gamma) \tag{17}$$

$$= (\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{F}(\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F})^{-1}\,\mathbf{F}^T\mathbf{R}^{-1})\mathbf{y} \tag{18}$$

This allows the response at an arbitrary evaluation point $\mathbf{x}_*$ to be calculated using

$$\hat{y}(\mathbf{x}_*) = \mathbf{r}^T\beta + \mathbf{f}^T\gamma \tag{19}$$

where $\mathbf{r}$ and $\mathbf{f}$ are vectors of basis function and monomial terms for the point $\mathbf{x}_*$, respectively.

The form of basis function chosen governs the stability of the system, the smoothness of the interpolation, and its predictive capability. It was decided to use Wendland's C^4 function $\phi_{3,2} = (1-r)_+^6(35r^2 + 18r + 3)$ for this work [39]. Wendland's functions $\phi_{n,k}$ are compact functions of minimal degree for a stated number of continuous derivatives C^{2k} in n dimensions, and they offer a good compromise between matrix conditioning and modeling behavior. They decay to zero at a given distance from the center, known as the support radius R. In addition to the support radius, it was decided to use weighting parameters w_j for each dimension to allow the norm to be biased in certain parameter directions. Thus, the basis function argument is a scaled and weighted Euclidean distance given by

$$\|\mathbf{x}_* - \mathbf{x}_i\|_{\text{scaled}} = \frac{1}{R}\sqrt{\sum_{j=1}^{n} w_j^2(x_{j,*} - x_{j,i})^2} \tag{20}$$

The weights (and, therefore, R) can be tuned to give the best possible model according to some accuracy measure. For this work, w_j values were determined using a particle swarm optimizer [44], with minimizing LOOCV error at the data sites as the objective, in a similar approach to that of Rippa, who developed an efficient algorithm for optimizing a single basis function shape parameter [25].

A standard uncertainty measure can also be constructed for RBF interpolants, based on the power function and the *native space norm* [28,40]. The power function $P_{\phi,\mathcal{X}}(\mathbf{x}_*)$ quantifies the component of interpolation error that is dependent on the basis function, model parameters, and sample locations, and is given by

$$P_{\phi,\mathcal{X}}(\mathbf{x}_*) = \sqrt{\phi(0) + \mathbf{u}^T(\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F})^{-1}\,\mathbf{u} - \mathbf{r}^T\mathbf{R}^{-1}\mathbf{r}}, \quad \mathbf{u} = \mathbf{F}\mathbf{R}^{-1}\mathbf{r} - \mathbf{f} \tag{21}$$

This can be scaled by the native space norm $\|\hat{y}_{\mathbf{y},\mathcal{X}}\|_{\mathcal{N}_\phi}$, which quantifies the variability of the training data, to give the more generic uncertainty measure $P_{\phi,\mathcal{X}}(\mathbf{x}_*)\|\hat{y}_{\mathbf{y},\mathcal{X}}\|_{\mathcal{N}_\phi}$, although this quantity still has no direct spatial dependence on the function values. The native space norm has a scalar value, given by

$$\|\hat{y}_{\mathbf{y},\mathcal{X}}\|_{\mathcal{N}_\phi} = \beta^T\mathbf{R}\beta \tag{22}$$

It should be apparent that there is a great deal of similarity between the Kriging and RBF model formulations and their error estimates. In fact, the only significant difference between the methods as implemented is the choice of basis function and model parameters, which reflect different logical choices (and conventions) arising from the different branches of mathematics used for derivation.

C. Maximum Mean Squared Error and Expected Improvement Function Sampling Criteria

As mentioned in Sec. I, the aims of the sampling strategy are to explore the domain and refine regions of local nonlinearity to produce an accurate model. The Kriging MSE can be exploited to achieve the first of these goals and, to a small extent, the second. As a measure of model uncertainty that is dependent on the distance to existing samples and the model parameters, $s(\mathbf{x})$ can be used directly as a sampling criterion by simply adding a point at the maximum value. This is the MaxMSE criterion, which gives predominantly space-filling points, with emphasis on certain parameter directions according to the optimized θ_j.

The EIF is a statistical criterion, developed for efficient global optimization [31]. The EIF leads to points that maximize the expectation of improvement upon the global minimum or maximum of the current predictor and can be used to refine significant peaks and troughs in the model. An estimate of the optimum value of the function is first found based on the available samples, for example, $y_{\min} = \min(y_1, y_2, \ldots, y_N)$. The Kriging predictor at any point can then be regarded as a random variable with mean $\hat{y}$ and variance s. Viewed in this way, a probability can be computed that the value at any point will fall below the current minimum (or above the maximum). The EIF is obtained by weighting the possible improvements by these probability densities and, for minimization, is written as

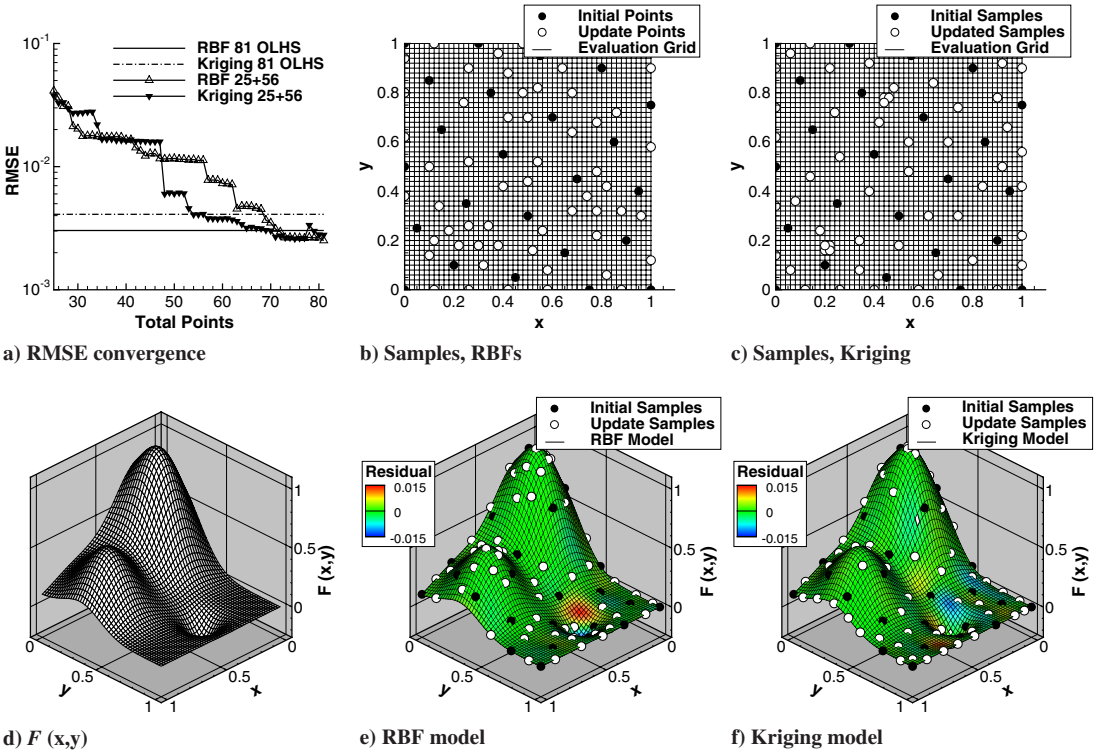
$$E[I(\mathbf{x})] = \begin{cases} (y_{\min} - \hat{y})\Phi\left(\frac{y_{\min} - \hat{y}}{s}\right) + s\varphi\left(\frac{y_{\min} - \hat{y}}{s}\right) & s > 0 \\ 0 & s = 0 \end{cases} \tag{23}$$

where φ and Φ are the probability density and cumulative distribution functions. New samples are located by finding max $E[I(\mathbf{x})]$.

For modeling aerodynamic loads data, refinement of the minimum and maximum coefficient values is desirable in addition to overall modeling accuracy. Moreover, accurate prediction of extrema values can also be expected to improve the model fit and therefore average error. A sampling strategy based on initial use of the MaxMSE criterion, followed by a number of EIF samples directed toward finding both the global minima and maxima in turn, was proposed by the authors [4], and the same strategy is used for this comparison study. The ratio of MaxMSE and EIF samples used is decided for each problem, and the method requires a small set of space-filling samples to build an initial interpolation.

D. Power Function and Curvature Sampling Criteria

A sampling criterion for RBF interpolation has been developed by the authors, using response surface curvature and a sample separation function to achieve the two aims of local refinement and exploration of the domain [5,45]. This places more emphasis on improving global accuracy than the Kriging MaxMSE/EIF strategy, but does not explicitly refine minima and maxima.



a) RMSE convergence b) Samples, RBFs c) Samples, Kriging

d) F (x,y) e) RBF model f) Kriging model

Fig. 1 Franke's function, interpolation error, sample placements, and models; 25 + 56 points.

The method requires an initial set of samples to first build an interpolation, and then update points are added based on a combined criterion C, defined as the product of the modeled Laplacian and the separation function. The separation function is a continuous function that grows with distance from the data sites and must have zero value when that distance is zero. For this work, the square of the power function and the native space norm were used for the separation function, which is then directly comparable to the MaxMSE criterion. This differs from previous published work by the authors [5], which primarily used an artificial function, created by interpolating values of one at the data sites using a small support radius. The criterion is given by

$$C = (|\nabla^2 \hat{y}| + \varepsilon) P^2_{\phi,\mathcal{X}}(\mathbf{x}) \|\hat{y}_{\mathbf{y},\mathcal{X}}\|_{\mathcal{N}_\phi} \tag{24}$$

where ε is an offset parameter to ensure a nonzero value when $|\nabla^2 \hat{y}| = 0$. The largest values of C indicate promising new sample locations, and the formulation given in Eq. (24) achieves a balance between adding points in locations where the data are nonlinear and adding points in unsampled regions of the domain. This is different to the Kriging MaxMSE/EIF strategy, in that the two effective criteria are switched on all the time.

The Laplacian of the interpolation model can be calculated by

$$\nabla^2 \hat{y} = \sum_{j=1}^{n} \frac{\partial^2 \hat{y}}{\partial x_j^2} \tag{25}$$

and model derivatives for a given direction x_j can be evaluated at arbitrary locations by applying the following differentiated form of Eq. (19):

$$\frac{\partial^2 \hat{y}}{\partial x_j^2} = \frac{\partial^2 \mathbf{r}^T}{\partial x_j^2}\beta + \frac{\partial^2 \mathbf{f}^T}{\partial x_j^2}\gamma \tag{26}$$

The interpolation coefficients β and γ for a dataset $\mathcal{X}$ remain the same for derivative evaluations, which are no more expensive than additional function evaluations.

IV. Results

A. Analytic Test Functions

Previous work demonstrated that sampling based on RBF model curvature worked well for two-dimensional analytic functions [5]. It was decided to use two of the more challenging of these test functions as an initial comparison for this work, herein referred to as Franke's function [46] $F(x, y)$ and the droplet function $D(x, y)$, given by

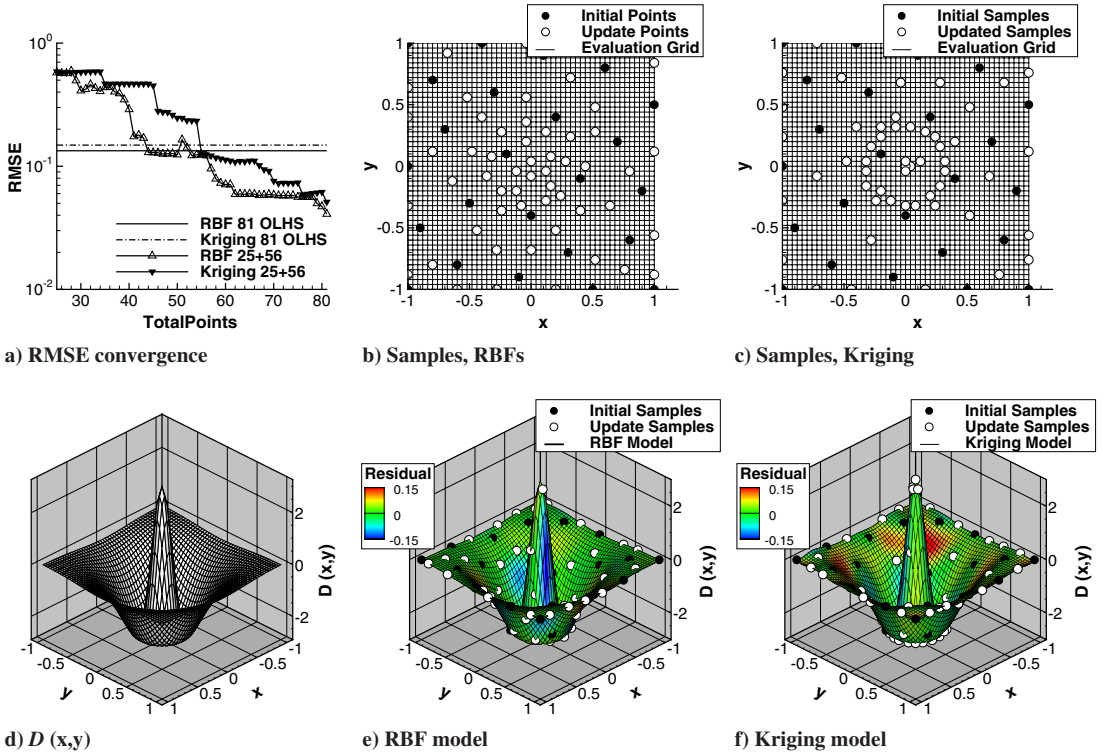
$$F(x,y) = \frac{3}{4}e^{-\frac{1}{4}((9x-2)^2+(9y-2)^2)} + \frac{3}{4}e^{-\frac{1}{49}(9x+1)^2-\frac{1}{10}(9y+1)^2} + \frac{1}{2}e^{-\frac{1}{4}((9x-7)^2+(9y-3)^2)} - \frac{1}{5}e^{-(9x-4)^2-(9y-7)^2}, \Omega = [0,1]^2 \tag{27}$$

$$D(x,y) = -4e^{-\frac{25}{8}(x^2+y^2)} + 7e^{-\frac{125}{4}(x^2+y^2)}, \qquad \Omega = [-1,1]^2 \tag{28}$$

For these examples, optimization of the sampling criteria was done using a search on a fine grid of 51×51 (51^2) evaluation points in accordance with the original results, and sampling was stopped when a predetermined budget of samples had been exhausted. The two methods could then be compared with each other and with equivalent nonadaptive full factorial (FF) and optimized Latin hypercube (OLHS) designs** for reference. Two different numbers of points were chosen for this purpose, 49 and 81, allowing comparison between the two adaptive methods using 4^2 initial points plus 33 updates, with a 7^2 full factorial design and 49 OLHS points as reference, and similarly for $5^2 + 56$ points with 9^2 FF and 81 OLHS points as reference. One-third of the total points, or as close as the FF designs will allow, have been allocated as initial points, which is the guideline strategy suggested by Forrester et al. ([18] pp. 102–103).

Figures 1 and 2 show the results for methods using $5^2 + 56$ points and OLHS nonadaptive samples. The original test function shape, root mean squared error (RMSE) convergences, and the sample positions and final models with residuals $[y(\mathbf{x}_*) - \hat{y}(\mathbf{x}_*)]$ for the adaptive methods are shown. Table 1 also summarizes the final results for configurations not plotted. A naming convention has been

**All Latin Hypercube designs for this work had the corner points added, inclusive of the total number of points.



a) RMSE convergence b) Samples, RBFs c) Samples, Kriging

d) D (x,y) e) RBF model f) Kriging model

Fig. 2 Droplet function, interpolation error, sample placements, and models; 25 + 56 total points.

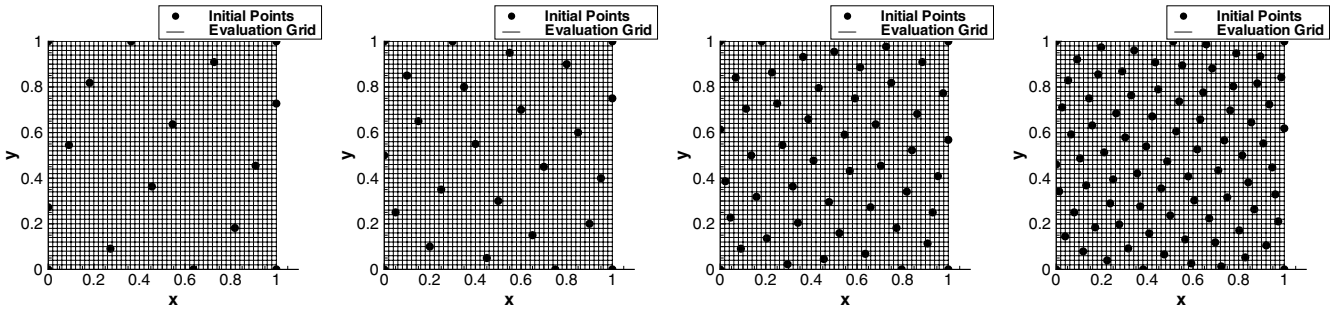
used here, first indicating whether an RBF (R-) or a Kriging (K-) model has been used, then indicating the type of nonadaptive samples (FF or OLHS), and, if applicable, followed by the adaptive criterion used, either C (C) or MaxMSE/EIF (M/E).

For fixed sets of points (e.g., 7^2 FF, 9^2 FF), the Kriging models should theoretically be slightly better suited to these test functions, because the smoothness of the Gaussian correlation function matches the response, but it can be seen that there is no clear advantage evident in the results. The existing differences are a necessary observation to put further results into context, however.

Both adaptive sampling strategies perform well for Franke's function, although the RMSE convergence graphs show that the error in the Kriging model reduces to a low level faster than the RBF model. This is perhaps surprising because, up until the final 10% of samples for this test case, the Kriging model uses MaxMSE samples, whereas the RBF model uses the directly equivalent power function combined with second derivative information. In fact, the space-filling MaxMSE criterion has first sampled the minimum for Franke's function (causing a sharp drop in RMSE) earlier than the curvature-based criterion, which emphasizes improvement of the known features more than exploration of the domain. Note also that the Kriging MaxMSE criterion produces space-filling samples slightly denser in y than x, due to differences in optimized θ_x and θ_y. A further difference in the sample locations is that the EIF has led to refinement of the global minimum and maximum, but curvature-based samples have caused refinement of all three prominent features, including the additional local maximum. Looking at the results table, the minimum and maximum values for Franke's function are determined very

Table 1 Final results for the analytic test functions, all configurations

Method	Total points	$F(x, y)$			$D(x, y)$		
		RMSE	Min	Max	RMSE	Min	Max
R-FF (7^2)	49	0.01156	−0.12502	1.03517	0.27284	−2.60185	3.00000
K-FF (7^2)	49	0.01200	−0.12506	1.03583	0.21952	−2.61051	3.00000
R-FF/C ($4^2 + 33$)	49	0.00779	−0.17727	1.03909	0.12090	−2.89483	3.00000
K-FF/M/E ($4^2 + 33$)	49	0.00846	−0.19664	1.03924	0.16851	−2.65103	3.00000
R-OLHS (49)	49	0.00896	−0.19187	1.03428	0.36075	−2.43084	0.00366
K-OLHS (49)	49	0.01130	−0.19163	1.03749	0.36514	−2.43535	0.00884
R-OLHS/C (16 + 33)	49	0.00714	−0.19747	1.04015	0.15565	−2.77027	2.90975
K-OLHS/M/E (16 + 33)	49	0.00853	−0.19747	1.03924	0.17548	−2.66233	3.00000
R-FF (9^2)	81	0.00425	−0.16522	1.03696	0.06149	−2.77838	3.00000
K-FF (9^2)	81	0.00420	−0.17167	1.04134	0.05455	−2.58278	3.00000
R-FF/C ($5^2 + 56$)	81	0.00277	−0.19327	1.03888	0.04230	−2.68516	3.00000
K-FF/M/E ($5^2 + 56$)	81	0.00343	−0.19747	1.03924	0.05197	−2.62098	3.00000
R-OLHS (81)	81	0.00303	−0.19408	1.03896	0.13359	−2.93167	2.30904
K-OLHS (81)	81	0.00410	−0.19427	1.03946	0.14848	−2.78564	2.29828
R-OLHS/C (25 + 56)	81	0.00251	−0.18566	1.03925	0.04091	−2.70317	2.91463
K-OLHS/M/E (25 + 56)	81	0.00279	−0.19747	1.03924	0.05126	−2.62555	3.00000
Target			−0.19747	1.03924		−2.61926	3.00000



a) 16 points b) 25 points c) 49 points d) 81 points

Fig. 3 Optimized Latin hypercube samples on the domain [0,1], plus the domain corners.

closely using only 10% EIF samples for the Kriging models, whereas curvature-based sampling does not capture these values quite as well, particularly for the smaller case with 49 points.

For the droplet function, the EIF was used to generate half of the total samples for the Kriging models, leading to very well-defined rings of samples around the maximum and minimum annulus in Fig. 2. However, this distribution of samples is relatively clustered, and the curvature-based sampling leads to faster error convergence. Slight increases in RMSE are due to the combined effect of adaptive sampling and parameter optimization, hence it might be suggested that the parameters should be fixed. However, preliminary investigations showed that overall an optimized interpolation leads to quicker error convergence; this has also been investigated by Toal et al. [42], who draw the same conclusion.

In general, OLHS points give improved results compared with FF designs for the initial samples. However, note that for the 5^2, 7^2, and 9^2 FF designs, the center of the domain is included, which captures the maximum peak of the droplet function and improves the result. The center point does not appear in any of the OLHS designs, as shown in Fig. 3. For the case with 81 total points, the adaptive methods initialized with 25 OLHS points produce better RMSE than those initialized by the 5^2 FF design, despite this. An important motivation for considering an FF design for the initial points is the inclusion of points on the boundary to mitigate edge oscillations or extrapolation; in general, interpolation errors are known to be greater near boundaries than in the interior of the domain for constant sample spacing [47]. It can be seen in Figs. 1 and 2 that the power function/MSE component of the adaptive criteria places a significant number of samples on the domain boundary when starting with OLHS samples. The addition of corner points to the OLHS designs was also intended to aid on this issue.

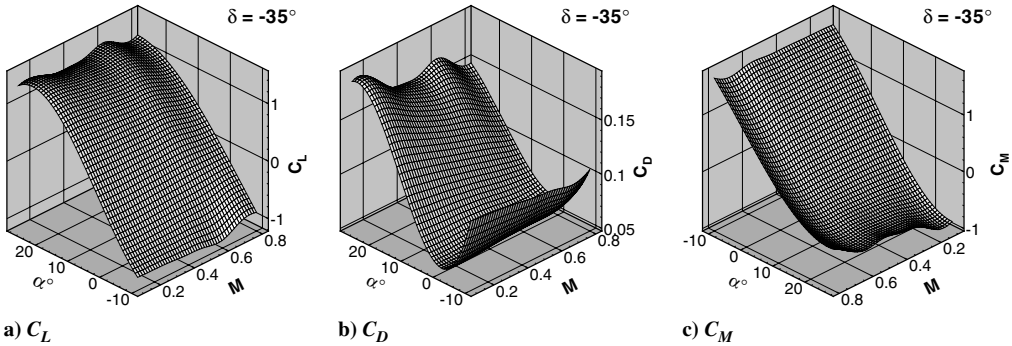
B. Aerodynamic Loads Model for DLR-F12

The second test case is an example representative of generating a lookup table of forces and moments for stability and control analysis, with a limited selection of input and output variables to simplify the problem. For analysis of slow rate maneuvers, a model of the quasi-steady aerodynamic behavior of an aircraft is required, which involves knowing the variation of six-degree-of-freedom forces and moments with respect to aircraft state variables (e.g., Mach, incidence and sideslip angles and their time rates, roll, pitch, and yaw rates) and control variables (e.g., aileron, elevator, rudder, and flap angles) [4,34]. To reduce the dimensionality of the problem, coupled effects can be ignored for all but Mach number and incidence, reducing the size from one model with 12 dimensions to 10 models with three dimensions. The example presented here includes incidence α, Mach number M, and elevator angle δ to illustrate the problem, and the longitudinal coefficients lift C_L, drag C_D, and pitching moment C_M are modeled.

The configuration considered is based on the DLR-F12 wind tunnel model, which is a 1:40 scale development model of a commercial passenger jet. A lookup table of aerodynamic data for this configuration was generated using DATCOM and interpolated to increase resolution. The data encompass α, M, and δ values in the ranges $[-10, +27$ deg], [0.1, 0.8], and $[-35, +35$ deg] respectively, and 49^3 points were used in the lookup table to give a fine evaluation grid for subsampling using the adaptive criteria, and for accurately calculating the RMSE. A two-stage evaluation grid of 24^3, followed by 3^3 points, was used to search for optimum sampling criterion values to decrease computation. In practice, an optimization algorithm may be used, such as that for optimization of the interpolation parameters [44], but a direct search was used for robustness in this comparison study.

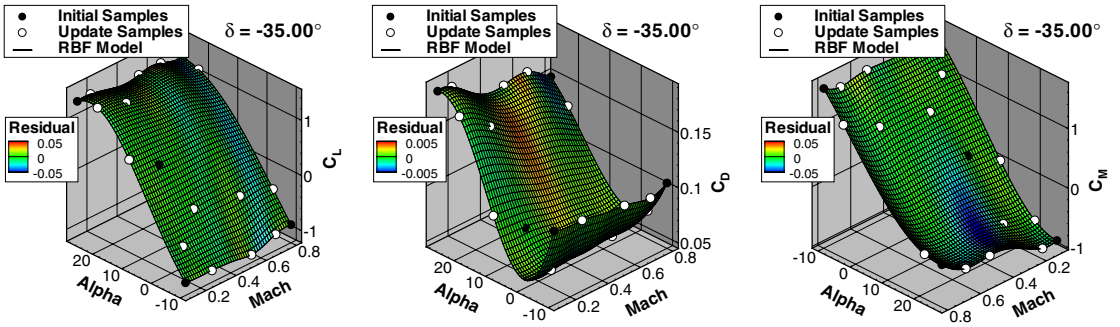
The reference data are shown in Fig. 4. The underlying physics, and therefore absolute accuracy of the test data, is limited, but for the purpose of this comparison, the function may be considered a black box, with only the relative accuracy of the surrogate compared with the original data being important. Within the parameter ranges chosen, the forces and moment exhibit almost monotone trends in incidence and elevator angle, with some nonlinearity in the Mach number direction, hence, it was decided to use a limited number of points to replicate them with the surrogate models. One hundred twenty-five total points were used, with 42 OLHS + 83 updates for the adaptive methods compared against 5^3 and 125 OLHS designs. For the MaxMSE/EIF method, 10% of the total points were dedicated toward expected improvement, split evenly between minimization and maximization of the C_L, C_D, and C_M responses.

To find optimal sample placements for all of the N_r responses simultaneously, the MaxMSE and RBF sampling criteria were modified as follows:



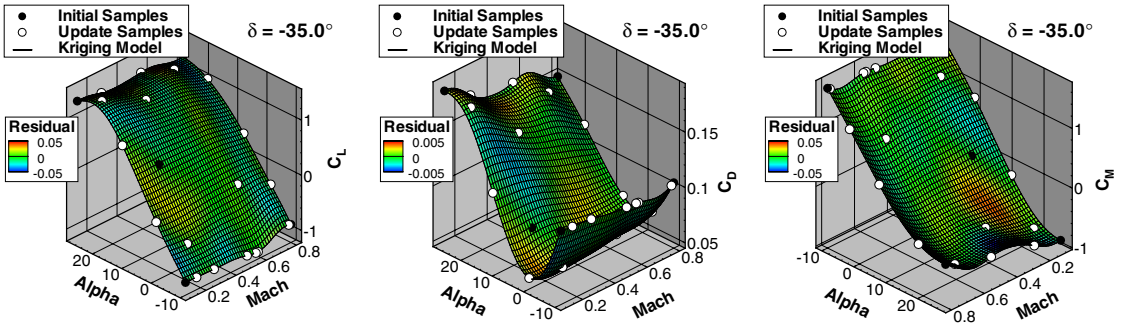
a) C_L b) C_D c) C_M

Fig. 4 Reference data for the DLR-F12 test case, $\delta = -35$ deg.



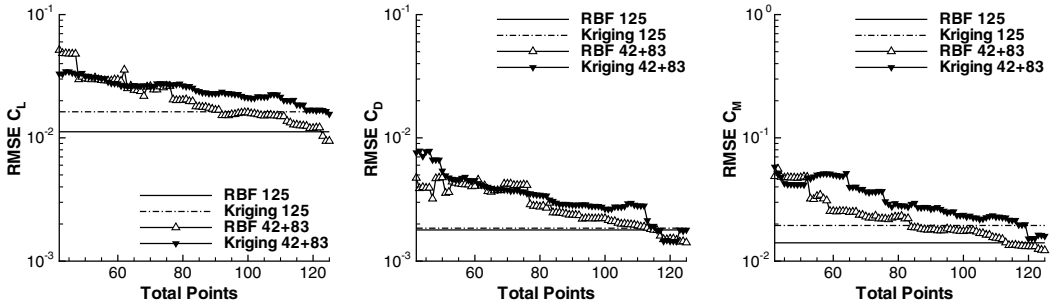
a) C_L b) C_D c) C_M

Fig. 5 RBF models and residuals for the DLR-F12 test case, $\delta = -35$ deg, 42 + 83 points.



a) C_L b) C_D c) C_M

Fig. 6 Kriging models and residuals for the DLR-F12 test case, $\delta = -35$ deg, 42 + 83 points.



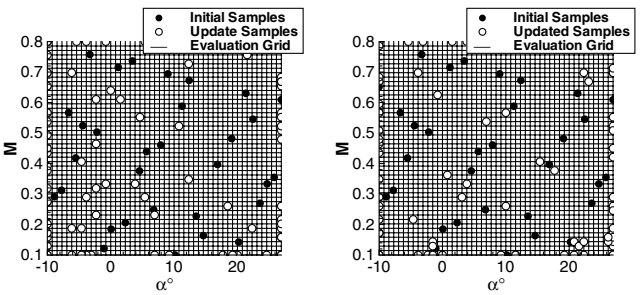
a) C_L b) C_D c) C_M

Fig. 7 RMSE convergence for the DLR-F12 test case, OLHS and 42 + 83 points.

$$\text{MaxMSE} = \max\left(\sum_{k=1}^{N_r} s_k(\mathbf{x})\right), \qquad C = \max\left(\sum_{k=1}^{N_r} C_k(\mathbf{x})\right) \quad (29)$$

Figures 5 and 6 show the resulting RBF and Kriging models and residuals for $\delta = 35$ deg, and Fig. 7 shows the RMSE convergence for the two adaptive sampling approaches, with respective OLHS interpolations as reference. In both cases, the final result is a slight improvement over the space-filling design, with slightly faster RMSE convergence achieved with RBF interpolation and curvature-based samples than Kriging and MaxMSE/EIF. Looking at the sample positions, shown in Fig. 8, the behavior of the two adaptive methods appears quite similar, with the effect of the curvature criterion being quite subtle, due to fairly even curvature in the responses at both high and low incidence, Mach number, and elevator angle. Both methods have again placed a number of points on the boundary, due to the use of the power function and MSE function, which have larger values on the boundary than the interior of the domain for a given sample separation.

Tables 2 and 3 list the remaining results. From this, it can be seen that the interpolations using OLHS points comprehensively outperform those with full factorial samples in terms of RMSE and maximum error, and it was decided not to run additional adaptive designs with full factorial initial points. Kriging MaxMSE/EIF makes a much closer approximation of the minimum and maximum values, but fails to capture all six target minimum and maximum values exactly, owing to the limited number of samples available overall.



a) RBF b) Kriging

Fig. 8 Flattened samples (showing all δ) for the DLR-F12 test case, 42 + 83 points.

Table 2 Final results for the DLR-F12 test case, all configurations

Method	RMSE			Max error		
	C_L	C_D	C_M	C_L	C_D	C_M
R-FF (5^3)	0.02896	0.00350	0.04655	0.16451	0.02337	0.23258
R-OLHS (125)	0.01120	0.00179	0.01407	0.09937	0.01704	0.15275
K-FF (5^3)	0.03107	0.00375	0.04090	0.18250	0.02399	0.20772
K-OLHS (125)	0.01627	0.00185	0.01946	0.12497	0.01787	0.19189
R-OLHS/C (42 + 83)	0.00943	0.00141	0.01221	0.05750	0.01013	0.07206
K-OLHS/M/E (42 + 83)	0.01560	0.00179	0.01596	0.07961	0.07206	0.08554

Table 3 Final results for the DLR-F12 test case, all configurations

Method	Min value			Max value		
	C_L	C_D	C_M	C_L	C_D	C_M
R-FF (5^3)	−1.05946	0.00785	−2.31276	1.79934	0.23572	1.64971
R-OLHS (125)	−1.05286	0.00942	−2.32175	1.80695	0.23594	1.64749
K-FF (5^3)	−1.05635	0.00892	−2.31236	1.80486	0.23594	1.64749
K-OLHS (125)	−1.07571	0.00833	−2.31236	1.83251	0.23594	1.66974
R-OLHS/C (42 + 83)	−1.07772	0.00885	−2.35607	1.83658	0.23798	1.64749
K-OLHS/M/E (42 + 83)	−1.08903	0.00844	−2.36126	1.83966	0.23779	1.64765
Target	−1.08903	0.00844	−2.36126	1.83447	0.23779	1.64819

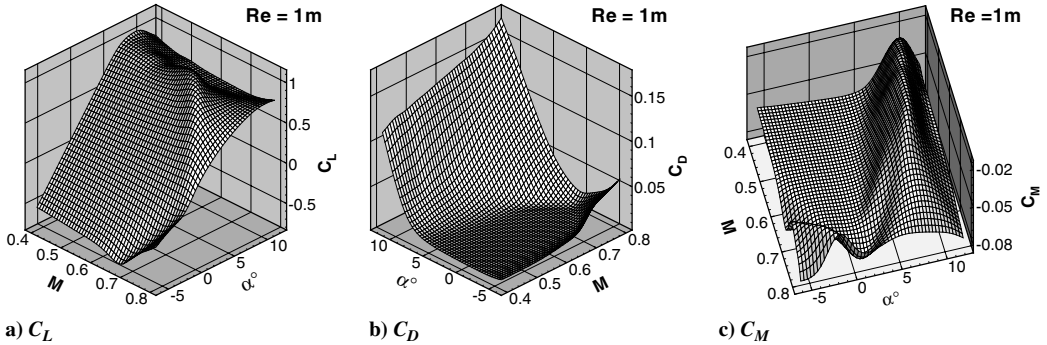
C. Aerodynamic Loads Model for RAE2822

The final test case is another example of aerodynamic loads modeling, this time representative of a structural loads analysis application. A rigid structure is considered here for simplicity, although unsteady loads might be considered by applying a secondary correction. Also, the problem of modeling the global loads (integrated over the whole geometry) on an airfoil is considered, which again is a notable simplification over the real problem, which requires modeling distributed loads for components such as the fuselage, wingbox, flaps, and slats, sufficient for a beam stick model for a full three-dimensional configuration. For three-dimensional configurations, either a separate approximation for the relationship between global and distributed loads would need to be defined or separate models constructed for different components (which would likely not lead to consistent results).

The reference data are derived from MSES, a two-dimensional Euler solver with coupled boundary-layer method that offers fast execution time. The geometry is an RAE2822 airfoil [48], and C_L, C_D, and C_M have been mapped for incidence, Mach number, and Reynolds number values in the ranges $[-6, +12$ deg], [0.4, 0.8], and [10^6, 10^7]. As mentioned, this is a somewhat reduced problem, but the data include transonic conditions and incidences beyond stall, leading to some particularly nonlinear features in the responses. A regression model was fitted to the CFD data (which included 11,000 simulations) to remove noise and obtain a complete dataset, filling in for many post-stall incidences where the solver would not converge. Again, the reference data were given a resolution of 49^3 points for accurate RMSE calculation.

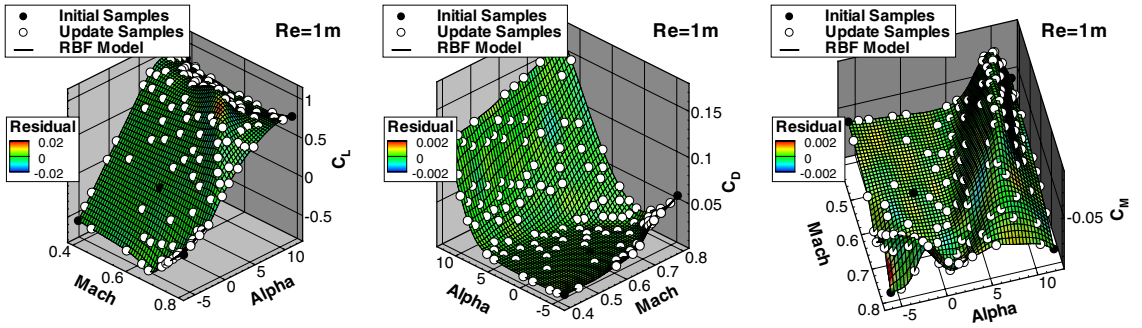
The lift, drag, and pitching moment coefficients are shown in Fig. 9 for a Reynolds number of 10^6. Variation in Re leads to a slight change in $C_{L\,\max}$ and $C_{M\,\min}$, but the responses are significantly less sensitive with respect to this parameter direction compared with either incidence or Mach. Because of the nonlinearity of the data, a budget of roughly 500 points was chosen, which was an amount sufficient to achieve a satisfactory model with a stratified factorial (SF) design using $13 \times 8 \times 5$ points (α, M, Re) biased in incidence and Mach using knowledge of the shape of the response, which was then used to compare against 500 OLHS points. The adaptive methods were started with 83 and 166 initial OLHS points, with the MaxMSE/EIF method using 50 (10%) update points for expected improvement, split evenly between minimization and maximization of C_L, C_D, and C_M, but excluding maximization of C_D because this maximum C_D occurs in the domain corner at maximum Mach and incidence. Also, to reduce the computational cost of generating the models, the RBF and Kriging model parameters were tuned only once every 10 update points.

Figures 10 and 11 show the final RBF and Kriging models and residuals, and Figure 12 shows the RMSE convergence of the different methods. It can be seen that the curvature-based sampling achieves notably faster error convergence than MaxMSE sampling for this test case. However, the final 50 EIF samples significantly reduce the RMSE of the Kriging models, such that the final result is very similar, or even slightly better. Both adaptive methods are significantly more accurate than the pure OLHS designs, which themselves are significantly more accurate than the manually stratified designs. Figure 13 shows the sample locations for the



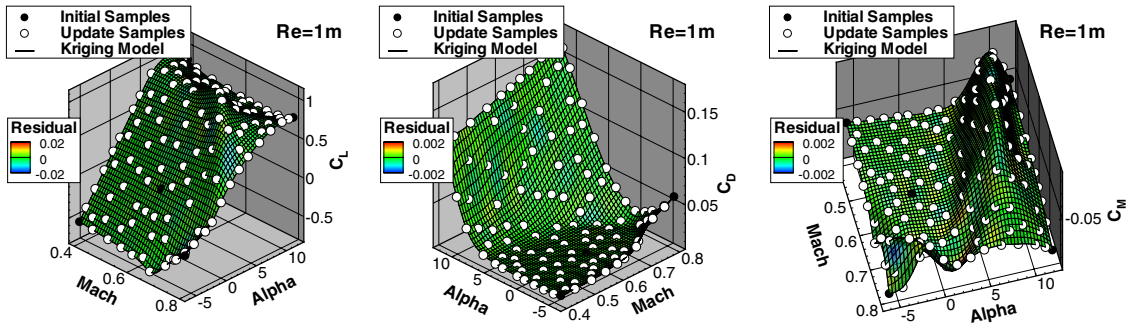
a) C_L b) C_D c) C_M

Fig. 9 Reference data for the RAE2822 test case, $Re = 10^6$.



a) C_L b) C_D c) C_M

Fig. 10 RBF models and residuals for the RAE2822 test case, $Re = 10^6$, 166 + 334 points.



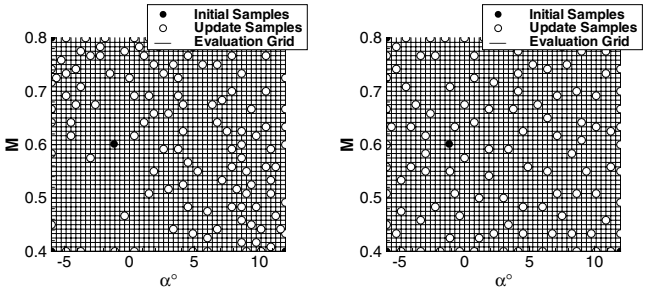
a) C_L b) C_D c) C_M

Fig. 11 Kriging models and residuals for the RAE2822 test case, $Re = 10^6$, 166 + 334 points.

minimum of the Reynolds number range. This gives a good indication of the overall sampling behavior with respect to α and M, and it can be easily visualized. The curvature-based sampling gives a particularly nice sample distribution, with many points concentrated at high Mach number, where there is a sharp feature in pitching moment due to shock movement, and at high incidence around stall.

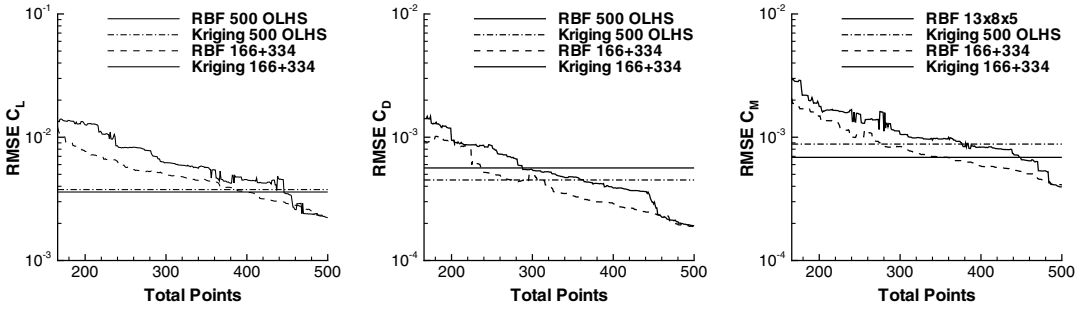
Tables 4 and 5 show results for the final interpolations for all sample configurations. As with the previous test cases, the prediction of the minimum and maximum values was significantly improved for the Kriging models by adding a small number of EIF samples, but the adaptive RBF models also perform very well in this regard for this test case. It can be seen that the results using 83 initial points, or one-sixth of the total number, are slightly worse than those using 166 points for both the RBF and Kriging models. This indicates that there is indeed an optimal number to choose, and that the recommendation of Forrester et al. ([18] pp. 102–103) to use a fraction of one-third seems valid here without conducting a more rigorous investigation.

Finally, another way of assessing the performance is to consider the percentage savings in number of samples required to match the accuracy of a traditional space-filling approach, using the adaptive methods. This relates to the amount by which the sample budget could be reduced, or the potential time saved in building a suitable model, for real world applications. According to Fig. 12, the RBF and Kriging methods afford a savings of 20 and 13%, respectively, over the OLHS design. This is a notable improvement over the traditional approach.



a) RBF b) Kriging

Fig. 13 Samples in the plane $Re = 10^6$ for the RAE2822 test case, 166 + 334 points.



a) C_L b) C_D c) C_M

Fig. 12 RMSE convergence for the RAE2822 test case, 500 OLHS and 166 + 334 points.

Table 4 Final results for the RAE2822 test case, all configurations

Method	Total points	RMSE			Max error		
		C_L	C_D	C_M	C_L	C_D	C_M
R-SF ($13 \times 8 \times 5$)	520	0.01004	0.00066	0.00117	0.06738	0.00821	0.00543
R-OLHS (500)	500	0.00359	0.00056	0.00069	0.05139	0.00930	0.00688
K-SF ($13 \times 8 \times 5$)	520	0.02081	0.00079	0.00223	0.12421	0.00889	0.01285
K-OLHS (500)	500	0.00376	0.00045	0.00088	0.05126	0.00662	0.01185
R-OLHS/C ($83 + 417$)	500	0.00262	0.00022	0.00054	0.01898	0.00118	0.00411
K-OLHS/M/E ($83 + 417$)	500	0.00245	0.00018	0.00039	0.02141	0.00143	0.00230
R-OLHS/C ($166 + 334$)	500	0.00227	0.00019	0.00041	0.02255	0.00121	0.00359
K-OLHS/M/E ($166 + 334$)	500	0.00222	0.00019	0.00040	0.01330	0.00117	0.00228

Table 5 Final results for the RAE2822 test case, all configurations

Method	Total points	Min value			Max value		
		C_L	C_D	C_M	C_L	C_D	C_M
R-SF ($13 \times 8 \times 5$)	520	−0.74355	0.00581	−0.10562	1.16445	0.17550	−0.00957
R-OLHS (500)	500	−0.74101	0.00609	−0.10374	1.15691	0.17550	−0.00895
K-SF ($13 \times 8 \times 5$)	520	−0.74343	0.00490	−0.10554	1.17030	0.17550	−0.01278
K-OLHS (500)	500	−0.75511	0.00524	−0.10398	1.16634	0.17550	−0.00969
R-OLHS/C ($83 + 417$)	500	−0.75282	0.00642	−0.10572	1.16574	0.17550	−0.00949
K-OLHS/M/E ($83 + 417$)	500	−0.75252	0.00677	−0.10566	1.16614	0.17550	−0.00953
R-OLHS/C ($166 + 334$)	500	−0.75274	0.00680	−0.10514	1.16376	0.17550	−0.00952
K-OLHS/M/E ($166 + 334$)	500	−0.75252	0.00678	−0.10566	1.16614	0.17550	−0.00953
Target		−0.75252	0.00678	−0.10566	1.16614	0.17550	−0.00953

V. Conclusions

For the purpose of aircraft loads analysis for structural design, or stability and control analysis, the ability of a surrogate model to predict aerodynamic coefficients across the flight envelope with consistent accuracy, and the ability to predict critical loads cases likely (but not exclusively) appearing at the minima and maxima, are objectives which should be addressed by the sampling approach. Two different adaptive sampling strategies for surrogate model generation have been discussed and compared for a range of test cases, including analytic functions and two aerodynamic loads modeling examples.

The results give some evidence that curvature-based sampling, coupled with a theoretical estimate of modeling error [the radial basis function (RBF) power function], which grows with distance from the data sites, leads to models with improved root mean squared error (RMSE) convergence compared with space-filling designs and maximum mean squared error/expected improvement function (MaxMSE/EIF) sampling. This was true for all of the example problems except in one case, for the analytic Franke's function, in which MaxMSE sampling converged faster due to faster exploration of the domain. There is also some evidence that sampling based on theoretical modeling error (the Kriging MSE function) with the addition of a few points allocated to the EIF leads to models with good RMSE compared with space-filling designs and gives an improvement in prediction of minimum and maximum values. This was the case for all of the test cases conducted. Both adaptive methods performed better than traditional full factorial and optimized Latin hypercube designs over all of the problems conducted, sometimes by a considerable margin, which is a necessary condition that these methods could be of use in certain circumstances.

The ratio of MaxMSE to EIF samples deserves further investigation and, potentially, an automatic method for determining the most appropriate ratio could be developed based on the data. Also, the number of initial and update samples for both methods could be investigated further to verify the stated recommendation of using one-third of the total points as initial points.

A further consideration would be the presence of numerical noise in the simulation data. To overcome this, the RBF and Kriging models could be adapted by including a regularization parameter (subtracting a small quantity from the diagonal of the interpolation or correlation matrix). The MSE/power function and EIF/curvature derivations could then be adjusted accordingly, and the criteria applied in the same way. However, the regularization parameter ought to be optimized to give good results, and it is not clear how the convergence of the update schemes would be affected coupled with reoptimization, or if one-off optimization at the beginning of the process would be sufficient. Hence, there are notable challenges to overcome.

Based on the studies conducted, a suggested strategy for similar applications might be to use a combination of curvature-based sampling, incorporating the RBF power function/Kriging MSE function to explore the domain, and expected improvement function points, added at the end to improve the responses values at extremum values and further reduce RMSE.

References

[1] Simpson, T. W., Toropov, V., Balabanov, V., and Viana, F. A. C., "Design and Analysis of Computer Experiments in Multidisciplinary Design Optimization: A Review of How Far We Have Come—or Not," *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, AIAA Paper 2008-5802, 2008.

[2] Quiepo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., and Tucker, P. K., "Surrogate-Based Analysis and Optimization," *Progress in Aerospace Sciences*, Vol. 41, No. 1, 2005, pp. 1–28. doi:10.1016/j.paerosci.2005.02.001

[3] Han, Z.-H., Zimmermann, R., and Görtz, S., "A New Cokriging Method for Variable-Fidelity Surrogate Modeling of Aerodynamic Data," *48th AIAA Aerospace Sciences Meeting*, AIAA Paper 2010-1225, Jan. 2010.

[4] Ghoreyshi, M., Badcock, K. J., and Woodgate, M. A., "Accelerating the Numerical Generation of Aerodynamic Models for Flight Simulation," *Journal of Aircraft*, Vol. 46, No. 3, 2009, pp. 972–980. doi:10.2514/1.39626

[5] Mackman, T. J., and Allen, C. B., "Investigation of an Adaptive Sampling Method for Data Interpolation Using Radial Basis Functions," *International Journal for Numerical Methods in Engineering*, Vol. 83, No. 7, 2010, pp. 915–938. doi:10.1002/nme.2885

[6] Krige, D. G., "A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand," *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, Vol. 52, No. 6, 1951, pp. 119–139.

[7] Hardy, R. L., "Multiquadric Equations of Topology and Other Irregular Surfaces," *Journal of Geophysical Research*, Vol. 76, No. 8, 1971,

pp. 1905–1915.
doi:10.1029/JB076i008p01905
[8] Williams, J. E., and Vukelich, S. R., "The USAF Stability and Control Digital DATCOM," Air Force Flight Dynamics Lab., Wright–Patterson AFB TR-AFFDL-TR-79-3032, 1979.
[9] Drela, M., "A User's Guide to MSES 3.04," Dept. of Aeronautics and Astronautics, Massachusetts Inst. of Technology TR, 2006.
[10] 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.
[11] Giunta, A. A., Wojtkiewicz, S. F. Jr., and Eldred, M. S., "Overview of Modern Design of Experiments Methods for Computational Simulations," *41st Aerospace Sciences Meeting and Exhibit*, AIAA Paper 2003-649, 2003.
[12] Koehler, J. R., and Owen, A. B., "Computer Experiments," Ghosh, S., and Rao, C. R. eds., *Handbook of Statistics*, Elsevier, New York, 1996, pp. 261–308.
[13] Simpson, T. W., Lin, D. K. J., and Chen, W., "Sampling Strategies for Computer Experiments: Design and Analysis," *International Journal of Reliability and Applications*, Vol. 2, No. 3, 2001, pp. 209–240.
[14] Chaloner, K., and Verdinelli, I., "Bayesian Experimental Design: A Review," *Statistical Science*, Vol. 10, No. 3, 1995, pp. 273–304.
doi:10.1214/ss/1177009939
[15] McKay, M. D., Beckman, R. J., and Conover, W. J., "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, Vol. 21, No. 2, 1979, pp. 239–245.
[16] Johnson, M. E., Moore, L. M., and Ylvisaker, D., "Minimax and Maximin Distance Designs," *Journal of Statistical Planning and Inference*, Vol. 26, No. 2, 1990, pp. 131–148.
doi:10.1016/0378-3758(90)90122-B
[17] Bates, S. J., Sienz, J., and Langley, D. S., "Formulation of the Audze–Eglais Uniform Latin Hypercube Design of Experiments," *Advances in Engineering Software*, Vol. 34, No. 8, 2003, pp. 493–506.
doi:10.1016/S0965-9978(03)00042-5
[18] Forrester, A. I. J., Sóbester, A., and Keane, A. J., *Engineering Design via Surrogate Modelling: A Practical Guide*, Wiley, Hoboken, NJ, 2008.
[19] Sobol, I. M., "On the Distribution of Points in a Cube and the Approximate Evaluation of Integrals," *U.S.S.R. Computational Mathematics and Mathematical Physics*, Vol. 7, No. 4, 1967, p. 86112.
doi:10.1016/0041-5553(67)90144-9
[20] Bratley, P., and Fox, B. L., "Algorithm 659: Implementing Sobol's Quasirandom Sequence Generator, *ACM Transactions on Mathematical Software*, Vol. 14, No. 1, 1988, pp. 88–100.
doi:10.1145/42288.214372
[21] Kollig, T., and Keller, A., "Efficient Multidimensional Sampling," *Computer Graphics Forum*, Vol. 21, No. 3, 2002, p. 557563.
doi:10.1111/cgf.2002.21.issue-3
[22] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., *Numerical Recipes: The Art of Scientific Computing*, 3rd ed., Cambridge Univ. Press, Cambridge, U.K., 2007, pp. 404–406.
[23] Jin, R., Chen, W., and Sudjianto, A., "On Sequential Sampling for Global Metamodelling in Engineering Design," *Proceedings of DETC'02. ASME 2002 Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 2: 28th Design Automation Conference*, Paper DETC2002/DAC-34092, American Society of Mechanical Engineers, New York, NY, 2002.
[24] Aute, V. C., Abdelaziz, O., and Radermacher, R., "Cross-Validation Based Single Response Adaptive Design of Experiments for Deterministic Computer Simulations," *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, AIAA Paper 2008-6067, 2008.
[25] Rippa, S., "An Algorithm for Selecting a Good Value for the Parameter *c* in Radial Basis Function Interpolation," *Advances in Computational Mathematics*, Vol. 11, Nos. 2–3, 1999, pp. 193–210.
doi:10.1023/A:1018975909870
[26] Tang, C. Y., Gee, K., and Lawrence, S. L., "Generation of Aerodynamic Data Using a Design of Experiment and Data Fusion Approach," *43rd AIAA Aerospace Sciences Meeting and Exhibit*, AIAA Paper 2005-1137, 2005.
[27] Turner, C. J., Campbell, M. I., and Crawford, R. H., "Generic Sequential Sampling for Metamodel Approximations," *Proceedings of DETC'03. ASME 2003 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 1: 23rd Computers and Information in Engineering Conference, Parts A and B*, Paper DETC2003/CIE-48230, American Society of Mechanical Engineers, New York, NY, 2003.
[28] Jakobsson, S., Anderson, B., and Edelvik, F., "Rational Radial Basis Function Interpolation with Applications to Antenna Design, *Journal of Computational and Applied Mathematics*, Vol. 233, No. 4, 2009, pp. 889–904.
doi:10.1016/j.cam.2009.08.058
[29] Shewry, M. C., and Wynn, H. P., "Maximum Entropy Sampling," *Journal of Applied Statistics*, Vol. 14, No. 2, 1987, pp. 165–170.
doi:10.1080/02664768700000020
[30] Keane, A. J., "Statistical Improvement Criteria for Use in Multiobjective Design Optimization," *AIAA Journal*, Vol. 44, No. 4, 2006, pp. 879–891.
doi:10.2514/1.16875
[31] Jones, D. R., Schonlau, M., and Welch, W. J., "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, Vol. 13, No. 4, 1998, pp. 455–492.
doi:10.1023/A:1008306431147
[32] Mackman, T. J., and Allen, C. B., "Aerodynamic Data Modelling Using Multi-Criteria Adaptive Sampling," *13th AIAA/ISSMO Multidisciplinary Analysis Optimization Conference*, AIAA Paper 2010-9194, 2010.
[33] Yamazaki, W., Rumpfkeil, M. P., and Mavriplis, D. J., "Design Optimization Utilizing Gradient/Hessian Enhanced Surrogate Model," *28th AIAA Applied Aerodynamics Conference*, AIAA Paper 2010-4363, 2010.
[34] Ghoreyshi, M., Badcock, K. J., Da Ronch, A., Marques, S., Swift, A., and Ames, N., "Framework for Establishing Limits of Tabular Aerodynamic Models for Flight Dynamics Analysis," *Journal of Aircraft*, Vol. 48, No. 1, 2011, pp. 42–55.
doi:10.2514/1.C001003
[35] Laurenceau, J., and Sagaut, P., "Building Efficient Response Surfaces of Aerodynamic Functions with Kriging and Cokriging," *AIAA Journal*, Vol. 46, No. 2, 2008, pp. 498–507.
doi:10.2514/1.32308
[36] Kuya, Y., Takeda, K., Zhang, X., and Forrester, A. I. J., "Multifidelity Surrogate Modeling of Experimental and Computational Aerodynamic Data Sets," *AIAA Journal*, Vol. 49, No. 2, 2011, pp. 289–298.
doi:10.2514/1.J050384
[37] Lophaven, S. N., Nielsen, H. B., and Søndergaard, J., "DACE, A Matlab Kriging Toolbox, Version 2.0," Technical Univ. of Denmark TR-IMM-TR-2002-12, Lyngby, Denmark, 2002.
[38] Lophaven, S. N., Nielsen, H. B., and Søndergaard, J., "Aspects of the Matlab Toolbox DACE," Technical Univ. of Denmark TR-IMM-TR-2002-13, Lyngby, Denmark, 2002.
[39] Wendland, H., *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge Univ. Press, Cambridge, England, U.K., 2005.
[40] Fasshauer, G. E., *Meshfree Approximation Methods with Matlab*, Interdisciplinary Mathematical Sciences, Vol. 6, World Scientific, Hackensack, NJ, 2007.
[41] Goovaerts, P., *Geostatistics for Natural Resources Evaluation*, Oxford Univ. Press, New York, 1997, pp. 147–152.
[42] Toal, D. J. J., Bressloff, N. W., and Keane, A. J., "Kriging Hyperparameter Tuning Strategies," *AIAA Journal*, Vol. 46, No. 5, 2008, pp. 1240–1252.
doi:10.2514/1.34822
[43] Martin, J. D., and Simpson, T. W., "Use of Kriging Models to Approximate Deterministic Computer Models," *AIAA Journal*, Vol. 43, No. 4, 2005, pp. 853–863.
doi:10.2514/1.8650
[44] Carlisle, A., and Dozier, G., "An off-the-Shelf PSO," *Proceedings of the Particle Swarm Optimization Workshop*, Purdue School of Engineering & Technology IUPUI, Indianapolis, 2001, pp. 1–6.
[45] Mackman, T. J., and Allen, C. B., "Multidimensional Adaptive Sampling for Global Metamodelling," *48th AIAA Aerospace Sciences Meeting*, AIAA Paper 2010-1418, 2010.
[46] Franke, R., "Scattered Data Interpolation: Tests of Some Methods," *Mathematics of Computation*, Vol. 38, No. 157, 1982, pp. 181–200.
[47] Fornberg, B., Driscoll, T. A., Wright, G., and Charles, R., "Observations on the Behavior of Radial Basis Function Approximations near Boundaries," *Computers and Mathematics with Applications*, Vol. 43, Nos. 3–5, 2002, pp. 473–490.
doi:10.1016/S0898-1221(01)00299-1
[48] Advisory Group for Aerospace Research and Development, "Experimental Data Base for Computer Program Assessment: Report of the Fluid Dynamics Panel Working Group 04," NATO TR-AGARD-AR-138, 1979.

R. Haftka
Associate Editor