AIAA AVIATION Forum
June 25-29, 2018, Atlanta, Georgia
2018 Applied Aerodynamics Conference

10.2514/6.2018-4199

Adaptive sampling techniques for surrogate modeling to create high-dimension aerodynamic loading response surfaces

Andrew L. Kaminsky[1,2]
The University of Tennessee, Knoxville, TN, 37996, USA
and CFD Research Corporation, Huntsville, AL, 35806, USA

Yi Wang[3]
University of South Carolina, Columbia, SC, 29208, USA

Kapil Pant[4]
CFD Research Corporation, Huntsville, AL, 35806, USA

And

Wendy N. Hashii[5] and Abraham Atachbarian[6]
Edwards Air Force Base, CA, 93524, USA

Surrogate models can be used to approximate complex systems at a reduced cost and are frequently used when data collection is expensive or time consuming. The accuracy of these models is typically dependent on the samples used to create them. Adaptive sampling procedures have been shown to significantly reduce the number of samples required to build accurate surrogate models by leveraging response information to identify promising sample locations. However, adaptive sampling techniques have a cost associated with determining the ideal sample locations, and this cost typically grows with the sample count.

This work seeks to reduce the cost associated with the adaptive sampling procedure and surrogate model creation. First, improvements to two prominent adaptive sampling techniques: the local linear approximation (LOLA)-Voronoi and the cross-validation (CV)-Voronoi adaptive sampling techniques are proposed to reduce the cost associated with determining optimum sample locations. A new K-fold cross-validation (KFCV)-Voronoi adaptive sampling technique is proposed, to reduce the sample selection costs by adding a global KFCV filter to the CV-Voronoi technique. Adaptive sampling costs are also reduced through an innovative Voronoi batch sampling technique. The cost reductions make it feasible to aggregate the Voronoi sampling techniques into a "mixture of experts." This aggregate technique is examined to determine whether utilizing multiple adaptive sampling techniques can improve the consistency and robustness of the adaptive sampling techniques. We evaluate these adaptive sampling acceleration techniques using benchmark functions with increasing parameter space dimension and aerodynamic loading data.

[1] Graduate student, Department of Mechanical, Aerospace, and Biomedical Engineering, 1512 Middle Dr. Knoxville TN, 37996 AIAA Student member
[2] Research Engineer. Biomedical and Energy Technologies, 701 McMillan Way NW, Huntsville, AL 35806
[3] Associate Professor, Department of Mechanical Engineering, 300 Main Street Columbia, SC 29208, AIAA Member, Corresponding Author: yiwang@cec.sc.edu
[4] Vice President Biomedical and Energy Technologies, 701 McMillan Way NW, Huntsville, AL 35806, non AIAA member, Corresponding Author: kapil.pant@cfdrc.com
[5] Technical Expert, Structures Flight, 307 East Popson Ave., EAFB, CA 93524, non-AIAA member
[6] Program Manager, Small Business Innovation Research, 307 East Popson Ave., EAFB, CA 93524, non-AIAA member

Nomenclature

A	Adhesion
B	Neighborhood score
β^*	Regression coefficients
C	Cohesion
E	Non-linearity metric
e_L	Cross-validation error
e_{KF}	K-fold cross-validation error
F	Design matrix
f	Regression function
G	Aggregate score
g	Gradient
H	Hybrid score
m	Neighborhood size
N	Neighborhood
p	Number of regression functions
Q	Set of test points
R	Stochastic-process correlation matrix
S	Set of samples
$S\backslash sr$	Set of samples without s_r
s	Sample
t	Number of test points
V	Relative volume
x	Input variable vector
Y	Set of responses
y	Response
$\hat{y}$	Surrogate response surface
Subscript	
i	Variable number
r	Reference

I. Introduction

Surrogate models are used to closely approximate complex functional relationships at a considerably reduced computational expense. Consequently, surrogate modeling is becoming increasingly prevalent in engineering analysis. Surrogate models are particularly valuable when data collection is time consuming, expensive, or difficult. Surrogate modeling has been applied to a wide-variety of applications including: heat exchanger design [1], aerodynamic loading [2], and modeling printed circuit-boards [3].

Surrogate models are built from discrete evaluations of responses to inputs within a parameter space of interest. The surrogate model is used to evaluate and predict the behavior of the remaining parameter space, where no sample is present. This can significantly reduce the number of required simulations or experiments. However, the accuracy of the surrogate model is dependent upon how well the response behavior is captured by the samples used to create the model. As a result, the sampling process plays a key role in the quality of the surrogate response surface.

In general, sampling techniques can be classified as either one-shot or adaptive. One-shot sampling approaches designate the sample set size and locations in a single stage, prior to the sample collection. One-shot approaches include Latin hypercube design (LHD) [4] [5], orthogonal arrays [6], or uniform sampling. In contrast, adaptive sampling approaches dynamically vary the sample set size and location based on prior results. The adaptive sampling techniques use information from the response surface built using the current samples to select the next sample in a region that will offer the maximum improvement in model accuracy. Compared to the one-shot techniques, adaptive sampling techniques obtain response surface models of the same accuracy with significantly fewer samples [7].

Effective adaptive sampling techniques balance two concepts: exploration and exploitation. Exploration is performed by sampling regions far from existing points, i.e., where the prediction uncertainty is high. The one-shot procedures are generally based on this premise. Conversely, exploitation is performed by using the information of the response surface constructed from the initial samples to identify regions of interest where additional sampling would provide the most useful information for resolving the complex functional behavior. Many different approaches have

been used to identify these ideal regions, including gradient-, committee-, variance-, and cross-validation-based methods [8].

Gradient-based adaptive sampling approaches use the gradient and Hessian of the response to identify regions that are difficult to model, i.e. regions with large gradients or Hessians. Samples are then concentrated in these regions to ensure that the underlying behavior is adequately resolved. The Hessian matrix has been employed by the local linear approximation (LOLA)-Voronoi approach [7] and Mackman [9].

The committee strategy uses some combination of the many types of surrogate models: polynomial, support vector regression [10], proper orthogonal decomposition [11] [12], Kriging [13], and radial basis functions [14] [15], among others. A committee of one or more surrogate modeling techniques is formed, and each is used individually to build separate response surfaces. The point at which the committee members disagree the most is selected as the new sample location. Douak et al. [16] used a query-by-committee (QBC) strategy consisting of 3 regression models, and Hendrickx and Dhaene [17] created a committee of three metamodels to select a point where two of the models have the largest disagreement.

Variance-based adaptive sampling can be used with Kriging or Gaussian process-based surrogate models. Using information from previously sampled points, these surrogate modeling techniques provide the prediction response and a measure of the variance, which is also known as the mean square error (MSE). The MSE has been used frequently within the literature to identify future sample locations. Jin et al. [18] proposed sampling a new point by maximizing the mean square error (MMSE); Sacks et al. [19] considered the integrated mean square error (IMSE); Shewry and Wynn [20] proposed the maximum entropy (ME) criterion; and Jones et al. [21] suggested sampling procedures based on the expected improvement (EI) function.

Finally, the cross-validation (CV) approach attempts to locate the next sample position by partitioning the existing sample set into a training and testing set, to respectively build and evaluate the model accuracy. The next sample is placed in the region with the largest error. In leave-one-out cross-validation (LOOCV), a single sample is used to test the surface built from the remaining points. Cross-validation is typically performed for each sample point and the next sample is selected in the region where the response surface is most sensitive to the sample. The cross-validation approach has been used by Aute et al. [1], Li et al. [22], and Xu et al. [23].

These adaptive techniques have been developed to minimize sampling costs by ensuring the next sample provides the most information. Determining ideal sample locations, means that each sample now has a cost associated with their evaluation and selection. The goal of the present work is to reduce the cost associated with evaluating future sample locations. We refashion prominent adaptive sampling techniques to reduce their computational cost, so that they can be applied economically to high-dimension parameter spaces. Changes to the LOLA-Voronoi and CV-Voronoi are proposed to significantly reduce the computational time associated with the sample selection procedure while retaining sampling accuracy. A new K-fold cross-validation (KFCV)-Voronoi scheme is developed. The KFCV-Voronoi significantly accelerates the adaptive sampling process, and provides samples that can be used to form surrogate models of accuracy equal to or better than existing Voronoi techniques. A novel Voronoi batch sampling method is also implemented for each adaptive sampling technique to reduce the number of calls to the adaptive sampling techniques, and to enable testing resources to be used in parallel. This paper is organized as follows: Section II outlines surrogate modeling; Section III introduces the adaptive sampling cost reduction techniques; Section IV presents a method to evaluate surrogate model accuracy; and Section V demonstrates the efficacy of the adaptive sampling cost reduction techniques.

II. Kriging-Based Surrogate Modeling

There are a number of prominent surrogate modeling methods, and an in-depth examination of surrogate modeling techniques can be found in the work of Wang and Shan [24]. In this work, the Kriging technique was implemented using the MATLAB Design and Analysis of Computer Experiments (DACE) toolbox [25]. Kriging is a method of interpolation that models the response via a Gaussian process governed by prior covariances. Compared to other interpolation schemes, Kriging provides high accuracy from a smaller database size. However, implementation is moderately difficult and requires significant pre-processing.

Given a set of n sample sites:

$$\boldsymbol{S} = [s_1, s_2, \ldots, s_n]^T, \tag{1}$$

and responses:

$$\boldsymbol{Y} = [y_1, y_2, \ldots, y_n]^T, \tag{2}$$

the Kriging predictor develops a model, $\hat{\boldsymbol{y}}(x)$, that predicts the response at an unsampled point by:

$$\hat{\boldsymbol{y}}(x) = \boldsymbol{f}(x)^T \boldsymbol{\beta}^* + \boldsymbol{r}(x)^T \boldsymbol{\gamma}^*. \tag{3}$$

The first term of Eq. (3) is a regression model which is a low-order polynomial formed by p inputs:

$$\boldsymbol{f}(x) = \left|f(x_1), f(x_2), \dots, f\left(x_p\right)\right|^T. \tag{4}$$

For the set of samples $\boldsymbol{S}$ the $n \times p$ design matrix, $\boldsymbol{F}$, can be built by substituting each sample into f so that $\boldsymbol{F_{ij}} = \boldsymbol{f_j}(\boldsymbol{s_i})$,

$$\boldsymbol{F}_{ij} = \begin{bmatrix} f_1(s_1) & f_2(s_1) & \dots & f_p(s_1) \\ f_1(s_2) & f_2(s_2) & \dots & f_p(s_2) \\ \vdots & \vdots & & \vdots \\ f_1(s_n) & f_2(s_n) & \dots & f_p(s_n) \end{bmatrix}. \tag{5}$$

The coefficients $\boldsymbol{\beta}^*$ are found by solving the generalized least squares solution (with respect to $\boldsymbol{R}$):

$$\boldsymbol{\beta}^* = (\boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{Y}. \tag{6}$$

Here $\boldsymbol{R}$ is the matrix of stochastic-process correlations between sample points,

$$\boldsymbol{R}_{ij} = \boldsymbol{\mathcal{R}}\left(\theta, s_i, s_j\right), \quad i, j = 1, \dots, n. \tag{7}$$

The second term of Eq. (3) is dependent on the vector of correlations between sampled points S and an unsampled point x:

$$\boldsymbol{r}(x) = [\boldsymbol{\mathcal{R}}(\theta, s_1, x), \dots, \boldsymbol{\mathcal{R}}(\theta, s_n, x)]^T. \tag{8}$$

Thus, the Kriging method builds a model that expresses the response through a regression model and random function (stochastic process) [25].

III. Proposed Acceleration Techniques for Adaptive Sampling

The sampling process can have a large impact on the number of samples needed to obtain a Kriging model of sufficient accuracy. High-dimension parameter spaces often require large sample quantities to adequately resolve the response using surrogate modeling techniques. The process of evaluating candidate sample locations can become intractable as the number of samples increases for several adaptive sampling techniques. For both the LOLA-Voronoi and CV-Voronoi techniques, the sampling cost grows superlinearly with the number of samples. In this work, alterations to the Voronoi adaptive sampling procedures are presented to reduce the adaptive sampling cost, and hence, render them more amenable to surrogate modeling in high-dimensional parameter space.

A. Local Linear Approximation Voronoi

The LOLA-Voronoi technique [7] is a gradient-based technique that builds a hybrid exploration-exploitation metric, called the hybrid score, to evaluate which regions should be sampled next. This technique first partitions the domain into Voronoi cells using the sample points. The Voronoi cells correspond to the region that is closer to a particular sample than any other sample. Fig. 1 presents an example of Voronoi tessellations.

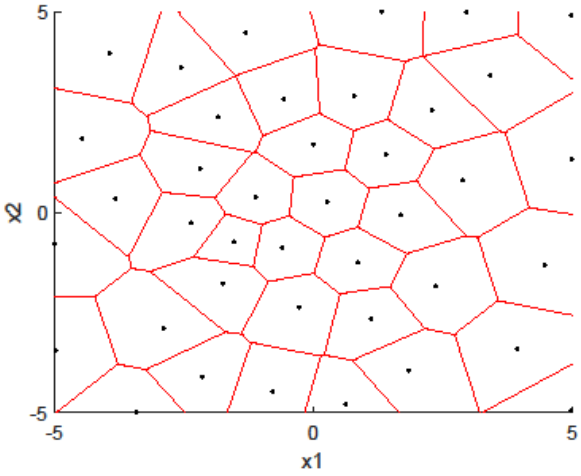


Fig. 1 Voronoi cells consist of the region that is closer to a given point than any other sample

The exploration portion of the hybrid score for a Voronoi cell depends on its relative volume, V. Cells with a large relative volume have a large exploration score and vice versa. Thus, to calculate the exploration score, the volume of each cell must be calculated. Crombecq et al. [7] suggest that the cost associated with calculating the volumes can be reduced by approximating the volumes through a Monte Carlo approach. This approach approximates the volume by generating a large number of random test points and calculating the distance from each test point to the samples. The test points are then assigned to the nearest sample, and the relative volume of each tessellation is calculated based on the number of test points assigned to it.

The exploitation portion is determined by building a LOLA to estimate the local gradient which is then used to calculate a non-linearity measure of each cell. To build the LOLA, a set of neighboring points, called a "neighborhood," must be found for each sample. For a given reference sample, s_r, the ideal neighborhood, $N(s_r)$, consists of the set of m remaining samples that give the best representation of the region around s_r, where $(m = 2 \times dimension)$. This is accomplished through optimization of two inversely related fundamental properties: cohesion and adhesion. An ideal neighborhood should consist of neighbors that are as close to the reference sample as possible, which can be quantified by a measure of Cohesion, C. It is defined mathematically as the average distance of all neighbors, s_{r_i}, from the reference point:

$$C\big(N(s_r)\big) = \frac{1}{m}\sum_{i=1}^{m} ||s_r - s_{r_i}||. \tag{9}$$

Conversely, Adhesion, A, is used to quantify how spread out the neighborhood points are. In an ideal neighborhood, neighbors should be situated as far as possible from each other to ensure information is present from all sides of the reference sample. Adhesion is measured by calculating the average minimum distance of neighbors from each other:

$$A\big(N(s_r)\big) = \frac{1}{m}\sum_{i=1}^{m} \min\left(||s_{r_i} - s_{r_j}|| \quad i \neq j\right). \tag{10}$$

Cohesion and adhesion can be combined into a single metric called the neighborhood score:

$$B\big(N(s_r)\big) = \frac{A\big(N(s_r)\big)}{\sqrt{2}\ C\big(N(s_r)\big)^2}. \tag{11}$$

The neighborhood is selected as the subset of samples:

$$N(s_r) = \{s_{r_1}, s_{r_2}, \dots, s_{r_m}\} \in S \backslash s_r \ , \tag{12}$$

that maximizes the neighborhood score. Fig. 2 shows an example of neighborhoods found using different techniques for a given sample set. The neighborhood score method produces a neighborhood that provides information from all directions, while remaining as close to the sample of interest as possible.

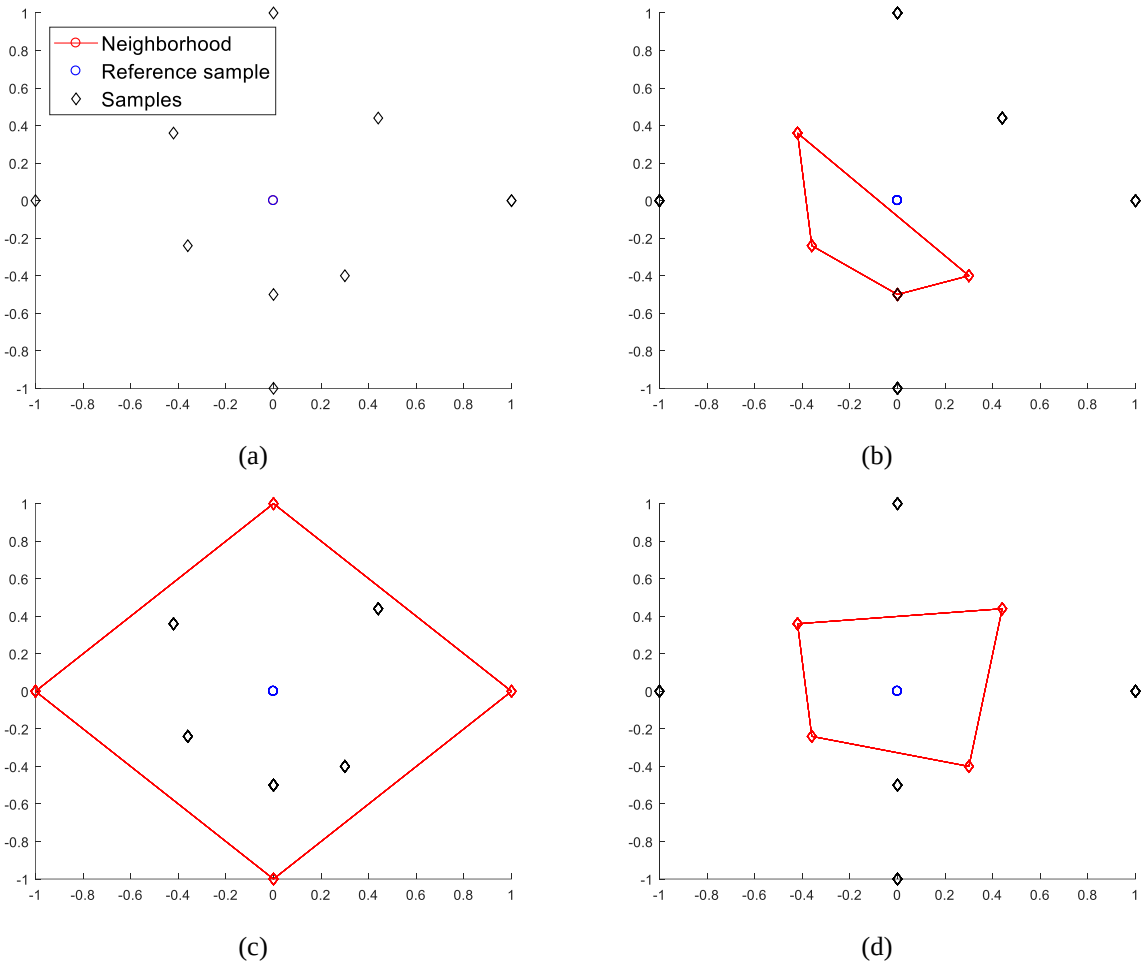


Fig. 2 (a) Given the reference sample (blue), a set of neighborhood samples (red) were identified using (b) a purely cohesive metric, (c) a purely adhesive metric, and (d) the neighborhood score.

The neighborhood N_{s_r} for each sample point s_r is then used to estimate the gradient by fitting a hyper plane through s_r and its neighbors using least squares

$$\begin{bmatrix} s_{r_1}^{(1)} - s_r^{(1)} & s_{r_1}^{(2)} - s_r^{(2)} & \cdots & s_{r_1}^{(d)} - s_r^{(d)} \\ s_2^{(1)} - s_r^{(1)} & s_{r_2}^{(2)} - s_r^{(2)} & \cdots & s_{r_2}^{(d)} - s_r^{(d)} \\ \vdots & \vdots & & \vdots \\ s_{r_m}^{(1)} - s_r^{(1)} & s_{r_m}^{(2)} - s_r^{(2)} & \cdots & s_{r_m}^{(d)} - s_r^{(d)} \end{bmatrix} \begin{bmatrix} g_r^{(1)} \\ g_r^{(2)} \\ \vdots \\ g_r^{(d)} \end{bmatrix} = \begin{bmatrix} y(s_{r_1}) \\ y(s_{r_2}) \\ \vdots \\ y(s_{r_d}) \end{bmatrix}, \tag{13}$$

where s_{r_i} is the ith neighbor of s_r with response $y(s_{r_i})$ and g is the gradient being calculated. This gradient approximation is then used to estimate the non-linearity of the system around the reference sample by calculating the difference between the response at the neighbors from the local linear approximation:

$$E(s_r) = \sum_{i=1}^{m} \left| f(s_{r_i}) - \left(f(s_r) + g \cdot (s_{r_i} - s_r) \right) \right|. \tag{14}$$

The exploration and exploitation metrics are then combined to form the hybrid score:

$$H(s_i) = V(s_i) + \frac{E(s_i)}{\sum_{j=1}^{n} E(s_j)} \quad \text{for } s_i \in S \ . \tag{15}$$

The sample with the maximum hybrid score is identified, and the next point is chosen as the point within this sample's Voronoi cell that is furthest from the sample.

1. Preparation of LOLA-Voronoi for high-dimensions

The original LOLA-Voronoi procedure, proposed by Crombecq et al. [7], works well when only a few variables are considered, but the computational cost increases dramatically as the number of samples and parameter space dimension grows. The number of potential neighborhoods is given by an n-choose-k binomial coefficient,

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}, \tag{16}$$

where n is the number of samples and k is equal to the number samples in a neighborhood. Fig. 3 shows how rapidly the number of potential neighborhoods increases when considering increasing dimensions and sample quantities. Evaluation of the neighborhoods is computationally inexpensive, but once the number of neighborhoods approaches 1×10^7 the wall clock time becomes appreciable.

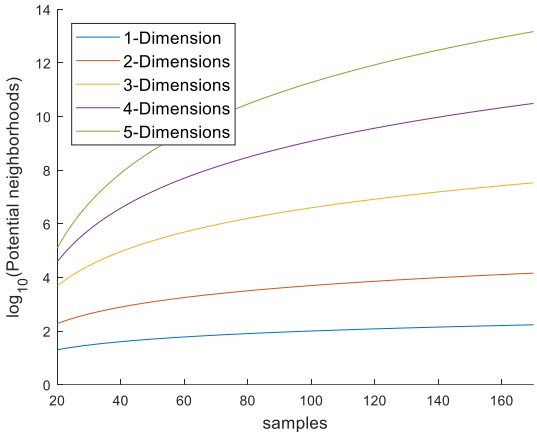


Fig. 3 Potential neighborhoods for different sample quantities and dimensions

It should also be noted that an appropriate neighborhood must be built for each sample point. Furthermore, there is also a large neighborhood evaluation cost associated with adding an additional sample to a sample set; because, a neighborhood must be found for the new sample, and the neighborhoods of the other samples must be updated.

To limit the neighborhood evaluation costs, we propose a multi-faceted approach. First, a typical neighborhood evaluation considers neighborhoods created from every combination of the remaining samples, but here we propose the quantity of potential neighborhoods can be limited by only considering neighborhoods made from the j nearest neighbors. This process limits the computational cost of neighborhood evaluation of the initial set. The cost of adding new samples can be limited by observing that the optimal neighborhood of the initial samples is only likely to change if the new sample is near that sample. As a result, we can limit the neighborhood re-evaluation to the j samples closest to the added sample. The neighborhoods of the j initial samples are re-evaluated by replacing each neighbor with the new sample and calculating the neighborhood score. The neighborhood with the maximum score is selected. The remaining samples' neighborhoods are not updated and the initial neighborhoods are retained. This approach drastically reduces the neighborhood evaluation cost when adding new samples.

B. Cross-Validation Voronoi

The CV-Voronoi technique, developed by Xu et al. [23], selects samples in regions with the largest prediction error. This is achieved by again breaking up the domain into Voronoi tessellations surrounding each sample. The error is then evaluated using a leave-one-out cross-validation (LOOCV) technique, which is performed by: (1) removing a single point from the sample set, (2) building a new response surface using the remaining points, (3) evaluating the new response surface at the removed point, and (4) calculating the CV-error, which is taken as the difference between the true value and the response surface with the missing point:

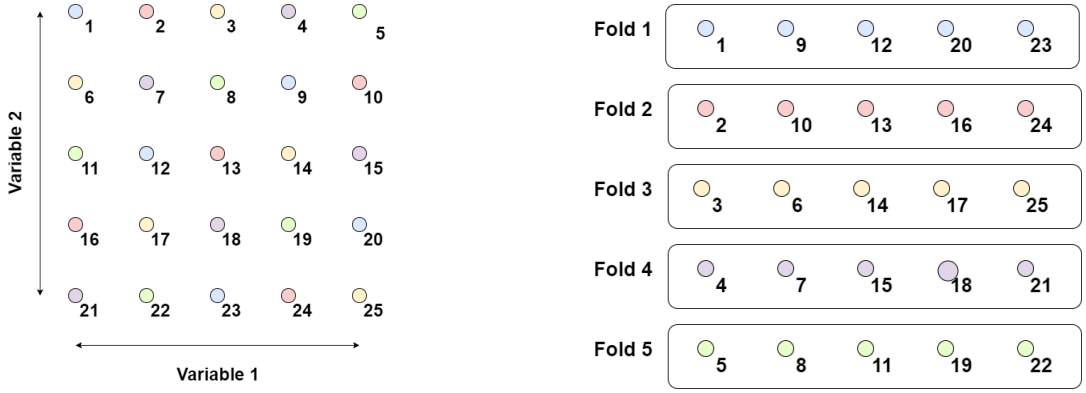
$$e_{LOO}^{i} = \left| y(s_i) - \hat{y}_{S\setminus s_i}(s_i) \right|. \tag{17}$$

Here, $y(s_i)$ is the measured response at s_i, and $\hat{y}_{S\setminus s_i}(s_i)$ is the response predicted at s_i by the surrogate model constructed without s_i . The new sample location is found by identifying the Voronoi cell with the largest prediction error and selecting the point within the Voronoi cell that is furthest from the existing sample point.

The CV-Voronoi technique has been shown to perform very well for several benchmark functions [23]. However, the computational load increases superlinearly with the number of samples; because according to Eq. (17), a response surface must be created for each sample. As higher dimension parameter spaces are considered, more samples are required, and the number of response surface models required can become computationally taxing.

C. K-fold Cross-Validation Voronoi

To reduce the computational cost, a novel K-fold cross-validation (KFCV)-Voronoi adaptive sampling technique is proposed here. The KFCV-Voronoi technique reduces the cost of sample selection by employing a K-fold cross-validation procedure to decrease the number of response surface models built during the sampling process. KFCV breaks the sample set into K randomly selected subsets. Fig. 4(a). shows this for a two-dimensional parameter space with 25 samples, which are assigned randomly to one of 5 folds (for this example $K = 5$). The fold assignment is portrayed by the sample color, and each fold is shown in Fig. 4(b).



(a) (b)

Fig. 4 (a) The original samples in a 2D parameter space and (b) their random assignment to K-folds.

The K-folds are then used for cross-validation by assigning a single fold for testing, and using the remaining folds to train the surrogate model. Fig. 5 shows the evaluation of the first fold. The training set consists of folds 2, 3, 4, and 5, and fold 1 makes up the testing set.

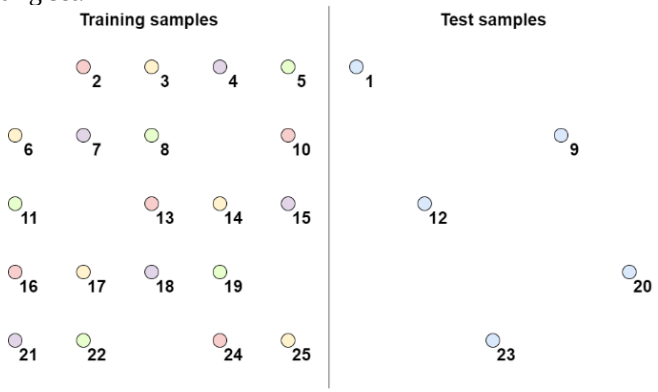


Fig. 5 KFCV utilizes a single fold as a test set for the remaining folds

The error associated with each fold is calculated in a manner similar to the LOOCV error metric, Eq. (17), and it is defined as

$$e_{KF}^{i} = \sum_{j \in kf_i}^{n_K} \left| y(s_j) - \hat{y}_{S\setminus kf_i}(s_j) \right| , \tag{18}$$

where n_K is the number of samples per fold, $y(s_j)$ is the true response at s_j, and $\hat{y}_{S\setminus kf_i}(s_j)$ is the response predicted at s_j using the surrogate model built without fold i. The KFCV process is repeated K times so that each fold is evaluated

as the test set. The folds are then sorted by error magnitude. The folds with the largest error contain the samples that have the biggest impact on the accuracy of the response surface, as a result the Voronoi cells of these samples are identified as regions of interest. To identify where additional points are needed within these regions of interest, LOOCV is performed for the samples that belong to the l folds with the largest errors. The next sample is then selected as the point within the Voronoi cell with the largest CV error that is furthest from the sample. The procedure is summarized in Algorithm 1. This procedure limits the number of surrogate response surfaces required to $K + l\,\frac{n_{samp}}{K}$.

Algorithm 1. KFCV-Voronoi Adaptive sample procedure

1. Randomly sort existing samples into K-folds
2. Perform K-fold cross-validation (KFCV)
3. Select the l folds with the largest KFCV error
4. Use leave-one-out cross-validation (LOOCV) to evaluate each sample within the l folds
5. Identify the sample with the highest LOOCV error
6. Select the point within this sample's Voronoi cell that is furthest from the existing sample

D. Aggregate Voronoi a Mixture of Experts

Within the literature, it has been observed that when multiple forecasting models are available, a combination of the forecasts can lead to improved performance [26]. Methods using multiple approximation sources have already been used in conjunction with surrogate modeling to identify optimum model parameters [27], leverage multiple fidelity solutions [28] [29], and improve performance [30]. By reducing the cost associated with our adaptive sampling techniques, it becomes more feasible to consider multiple adaptive sampling techniques. Here a mixture of experts is proposed to combine multiple adaptive sampling technique, to improve the robustness of the adaptive sampling technique. Specifically, the hybrid score from the modified LOLA-Voronoi and the newly proposed KFCV-Voronoi are combined to form the aggregate score:

$$G(s_i) = \frac{H(s_i)}{\sum_{j=1}^{n} H(s_j)} + \frac{e_{LOO}^{i}}{\sum_{j=1}^{n} e_{LOO}^{j}(s_j)}. \tag{19}$$

The scores within a particular method are normalized by the sum of the scores for that method, so that a sample that has a high relative weight within a given method will maintain that weight in the aggregate score. In general, different scaling coefficients can be applied to the scores from different methods, however this requires a trial-and-error evaluation. Therefore, in this paper, all the scores are treated equal, i.e., weighting coefficients of 1 are used for all scores as shown in Eq. 19.

E. Batch Sampling

Adaptive sampling can be further sped up by selecting samples in batches rather than one at a time. In the literature, the LOLA-Voronoi and CV-Voronoi techniques have only been presented sampling a single point at a time. In this work, a new Voronoi batch sampling approach is evaluated. A naïve batch sampling approach can lead to samples that are overly clustered, as shown in Fig. 6. These samples are selected in a region of high error, but the information provided by these samples is likely redundant.

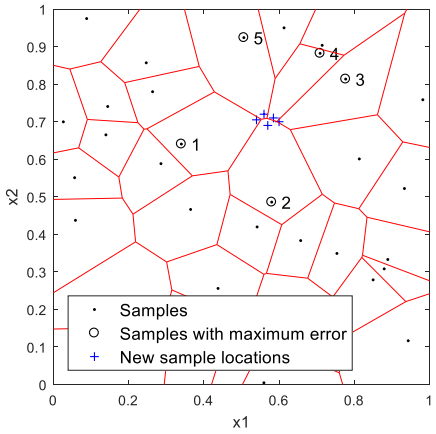


Fig. 6 Future samples from a naïve batch sampling approach.

To prevent this, the Voronoi batch sampling approach proposed here calculates the first sample of the batch by identifying the sample with the largest error metric and sampling the point within the Voronoi cell that is furthest from the existing sample, as shown in Fig. 7(a).

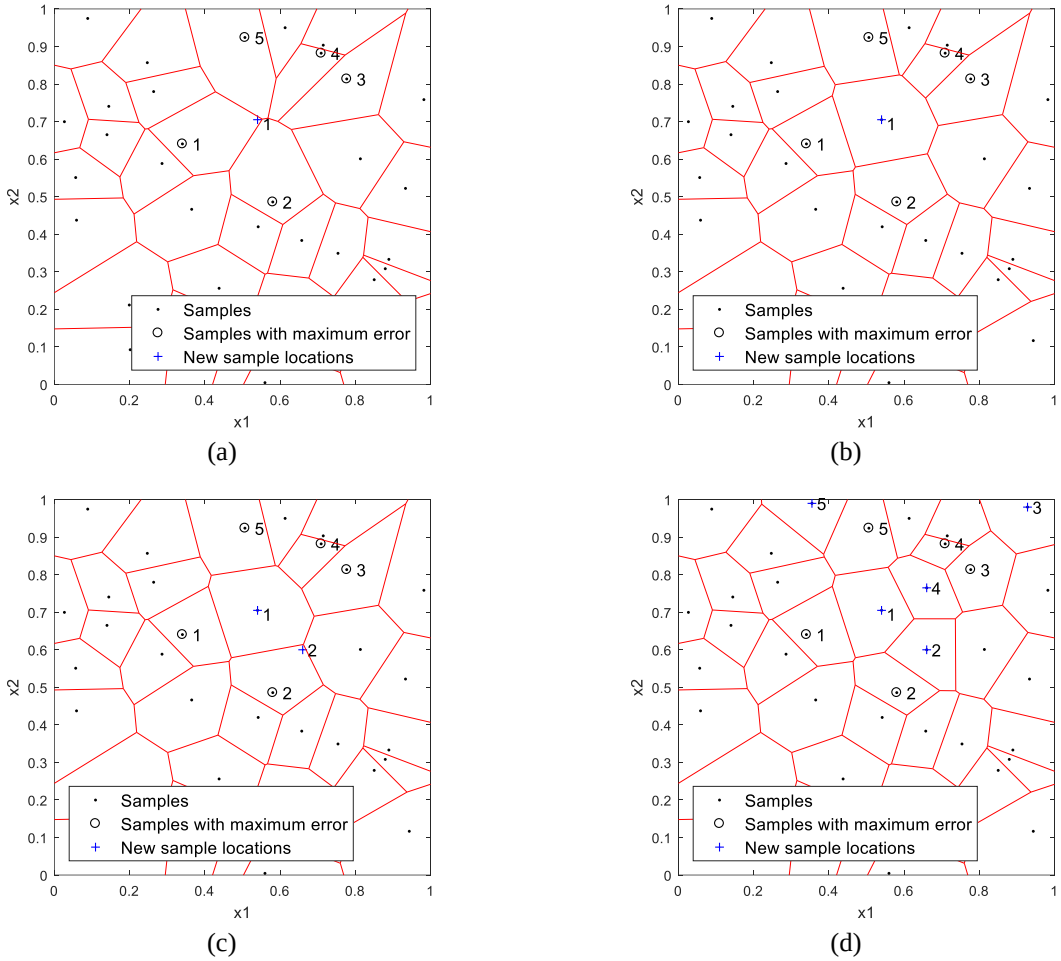


Fig. 7 The Voronoi batch sampling procedure: (a) the first sample is identified, (b) the Voronoi cells are reassigned, (c) the second sample is selected within the new Voronoi cells, (d) the procedure is repeated until the full batch is found.

The Voronoi regions are then reassigned with the new point, as illustrated in Fig. 7(b). The second sample is then selected within the updated Voronoi region associated with the second highest score, this is demonstrated in Fig. 7(c). This procedure is then repeated until the number of samples in the batch is reached. Fig. 7(d) presents the resulting 5-sample batch. This reassignment leads to samples that are better distributed than the naïve batch.

By selecting samples in batches, we reduce the number of calls to our adaptive sampling technique by a factor equivalent to the batch size. Furthermore, selecting samples in batches grants the ability to evaluate multiple responses in parallel which can greatly speed up the wall-clock time. Thus, batch sampling can greatly reduce the time associated with surrogate modeling

IV. Response surface accuracy evaluation

The quality of a given response surface must be assessed so that the performance of the different adaptive sampling techniques can be compared. Many techniques have been proposed to evaluate response surface accuracy. Within this work, we evaluate the response surfaces using the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{t}\sum_{i=1}^{t}\left(y(q_i) - \hat{y}(q_i)\right)^2}. \tag{20}$$

Here y is the true response, $\hat{y}$ is the response predicted by the surrogate model, and t represents a number of test points, $Q = [q_1,\ q_2, \ldots, q_t]$. As t increases, the RMSE gives an increasingly accurate evaluation of the surrogate response surface. The RMSE approach requires collecting an additional set of samples to be used solely for verifying the accuracy of a surrogate model.

V. Results

The adaptive sampling techniques were evaluated using several test cases. First, analytic benchmark functions were considered. Because the benchmark functions can be evaluated rapidly, the assessment of the adaptive sampling techniques is only limited by the sampling technique run time. Therefore, the benchmark functions enable easy analysis of computational cost savings. Benchmark functions with increasing variables of interest were used to examine performance in higher-dimension parameter spaces. After scrutinizing the performance using these benchmark functions, the adaptive sampling techniques were applied to aerodynamic loading data generated by CFD simulation of a NACA 0012 airfoil. All the adaptive sampling techniques are compared with uniform distributions and translational propagation Latin hyper cube design (TPLHD) [5], which is a technique for obtaining optimum Latin Hypercube Design distributions through translation of building blocks in the hyperspace.

A. 2D Benchmark functions

1. Peaks function

The 2D peaks function was the first benchmark function to be considered, and it is defined as

$$y = 3(1-x_1)^2 e^{(-x_1^2-(x_2+1)^2)} - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right)e^{(-x_1^2-x_2^2)} - \frac{1}{3}e^{(-(x+1)^2-x_2^2)} \text{ for } x_{1,2} \in [-5,5]\,. \quad (21)$$

Fig. 8 presents the peaks function which behaves non-linearly near the origin but remains constant away from the origin.

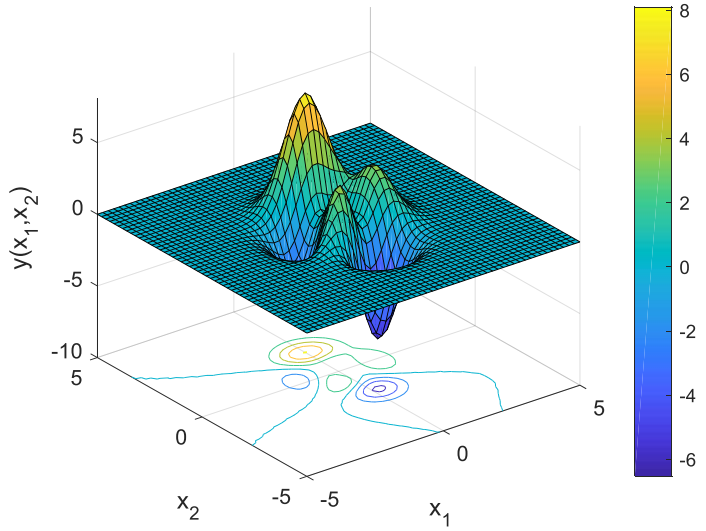


Fig. 8 The peaks function has complex behavior near the origin

The peaks function was chosen because it represents the ideal situation to demonstrate the value of an adaptive sampling approach. The constant region away from the origin can be correctly modeled using very few samples. The non-linear region near the origin is more difficult to model. As a result, more samples are required to approximate this region accurately. Exploitative adaptive sampling techniques can outperform one-shot sampling techniques by concentrating samples in regions that are harder to model. Fig. 9 presents the RMSE history of the proposed adaptive sampling techniques applied to the peaks function.

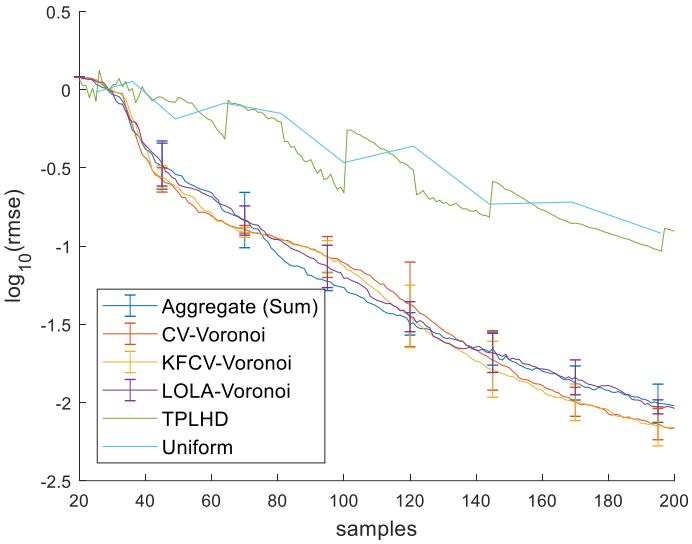
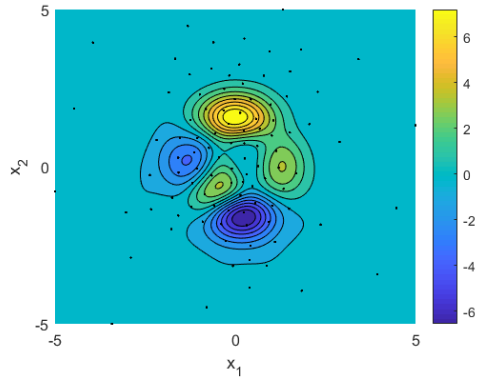
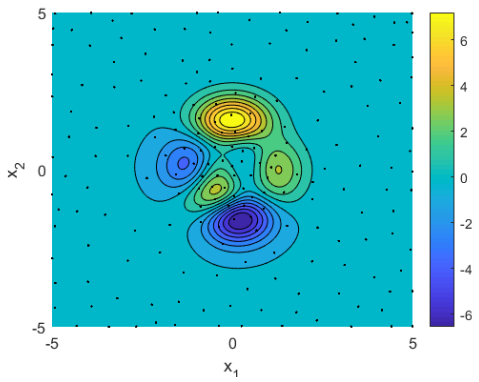
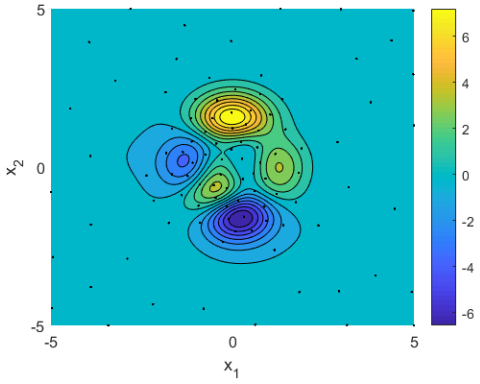
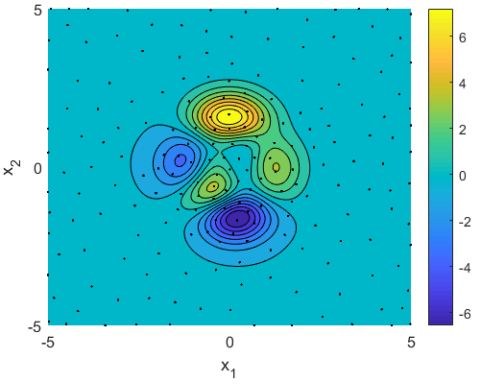
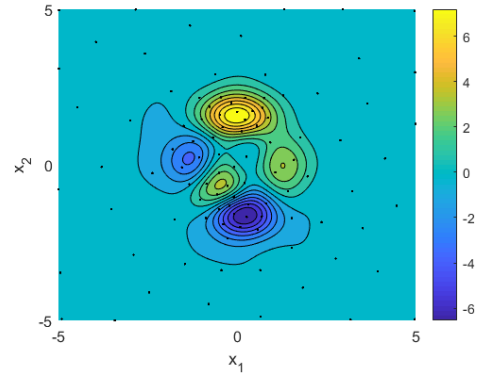
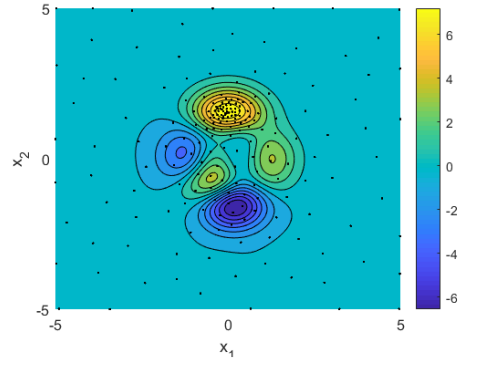
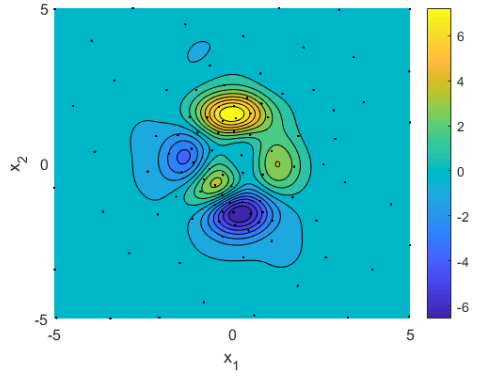
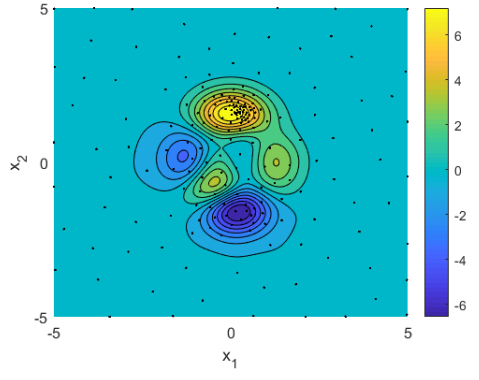


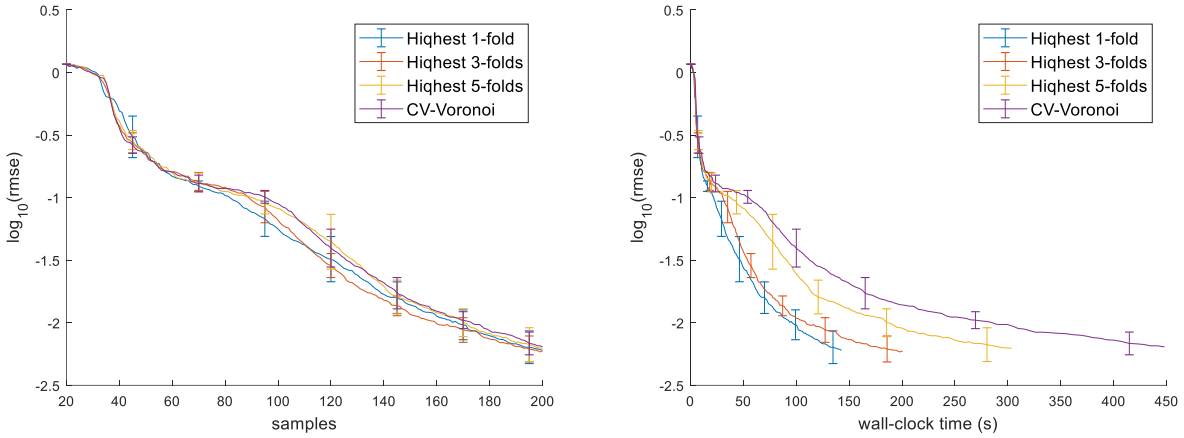
Fig. 9 Convergence history for surrogate response surfaces built using adaptive and one-shot sampling approaches applied to the peaks function.

As expected, all the adaptive sampling procedures dramatically outperform the one-shot methods. Initially, the adaptive sampling techniques perform analogously. This is followed by a region where the aggregate techniques perform best. Finally, the KFCV- and CV-Voronoi techniques overtake the aggregate methods and continue to outperform the other methods. Examination of the sample distribution in Table 1 for each sampling method at different points during the sampling procedure provides insight into why the RMSE profiles behave this way. The CV-Voronoi approach initially clusters points in the rapidly-varying regions and then explores the remaining domain after a reasonable resolution has been reached. The KFCV-Voronoi samples appear to be slightly more explorative than the CV-Voronoi samples, but the distributions are quite similar at 200 samples. Conversely, the LOLA-Voronoi approach appears to initially sample in a more exploratory manner and then clusters in the areas that are harder to resolve. Unfortunately, the LOLA-Voronoi technique appears to oversample these regions, which leads to stalled performance. Examining Eq. (15), it can be seen that as the number of samples increases, the second term is going to play an increasingly small role because individual Voronoi cell volumes will decrease with the number of samples. As a result, exploration-based sampling becomes devalued. The aggregate acts as a compromise between CV-Voronoi and LOLA-Voronoi, though it appears to be skewed towards the LOLA-Voronoi behavior. This could be overcome by assigning different weights to the contributions of each sampling technique during the adaptive sampling process.

Table 1 Sample distributions for adaptive sampling techniques at different sample quantities

	100 Samples	200 samples
CV-Voronoi		
KFCV-Voronoi		
Aggregate-Voronoi		
LOLA-Voronoi		

The KFCV-Voronoi RMSE profile in Fig. 9 was generated by using 10-fold KFCV-Voronoi to identify the 3 folds with the highest error followed by LOOCV for the samples within these folds. Within this work the K is taken to be 10. The peaks function is now used to evaluate the performance of the KFCV-Voronoi technique with LOOCV applied to different fold quantities $l = 1, 3,$ and 5. Fig. 10 shows that for a given number of sequentially generated samples, the KFCV-Voronoi approach slightly outperforms the CV-Voronoi technique. Furthermore, the KFCV-Voronoi approach reduces the number of surrogate response surfaces that must be formed to select the next point because the KFCV procedure limits the number of samples that must be evaluated with LOOCV. This decreases the computational time associated with the sampling procedure considerably, as shown in Fig. 10(b).

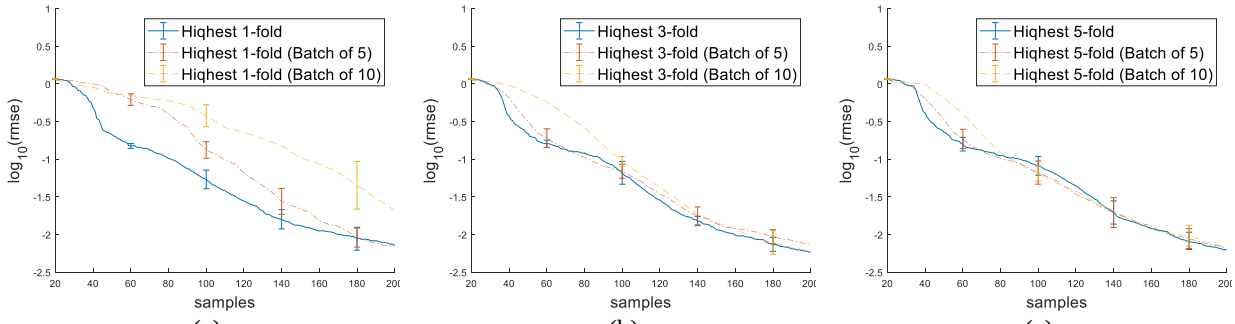


(a) (b)

Fig. 10 Convergence history for the surrogate response surfaces built using KFCV-Voronoi applied to the peaks function with different fold quantities (a) compared by sample count and (b) by time.

The KFCV-Voronoi technique was initially proposed purely as a method to reduce the computational cost associated with selecting the next sample location. KFCV-Voronoi could potentially misidentify the best Voronoi cell to sample if it was lumped with relatively unimportant points. Therefore, it was expected to be less accurate than the CV-Voronoi technique, and overcome the accuracy decrease through the reduced sample selection cost. It is interesting to note that KFCV-Voronoi adaptive sampling technique provides samples that yield surrogate models of accuracy equal to or better than those found using the CV-Voronoi technique. The improved sample selection is potentially due to a global filtering effect that occurs when performing KFCV, as the error in Eq. 18 represents the sum of the samples in a fold. In other words, the KFCV initially considers the global impact and identifies a region (rather than a single sample) with large prediction errors, whereas the LOOCV only considers the local impact of a sample. By taking into account both the global and local impact, the KFCV-Voronoi technique can outperform the CV-Voronoi technique.

The sampling process can be further accelerated by generating a batch of samples simultaneously, instead of a single sample at a time. By generating batches of samples, it is possible to evaluate multiple responses in parallel. Fig. 11 presents the RMSE history for the KFCV-Voronoi based adaptive sampling procedure performing LOOCV for different fold quantities and batch sizes.



(a) (b) (c)

Fig. 11 Convergence history for surrogate response surfaces built using KFCV-Voronoi applied to the peaks function with increasing batch sizes and LOOCV fold counts l : (a) 1-fold, (b) 3-folds, and (c) 5-folds

As the batch size increases, KFCV-Voronoi with LOOCV for only a single fold selects samples poorly. When LOOCV is performed for only the fold with the highest error, the number of Voronoi cells that have been identified as needing additional sampling is small, particularly early in the adaptive sampling process. Thus, the best sampling locations may not have been identified, and the batch procedure underperforms. Increasing the batch size is expected to diminish performance because samples are being selected with less information. However, this decrease in performance can be reduced by performing LOOCV for additional folds, as shown by the 3-fold and 5-fold cases in Fig. 11 (b) and Fig. 11 , respectively. The cost increase required to consider additional folds is less than the cost savings from the reduction in the number of calls to the adaptive sampling procedure. By increasing the batch size to five samples, the adaptive sampling procedure only needs to be called one fifth as frequently as a procedure adding a single point at a time.

2. *Shubert function*

The next benchmark is the Shubert function, which is a two-dimensional uniformly periodic function defined by

$$y = \left(\sum_{i=1}^{5} i\cos\big((i+1)x_1 + i\big)\right) \times \left(\sum_{i=1}^{5} i\cos\big((i+1)x_2 + i\big)\right) \text{ for } x_{1,2} \in [1,3]\,. \tag{22}$$

The Shubert function is a test case that emphasizes the importance of exploratory sampling over exploitative sampling. Due to the uniform oscillations, shown in Fig. 12, if a single peak or valley is missed, a relatively large RMSE will exist. Therefore, in addition to leveraging the exploitative sampling, the adaptive sampling procedure must also give adequate value to exploratory sampling.

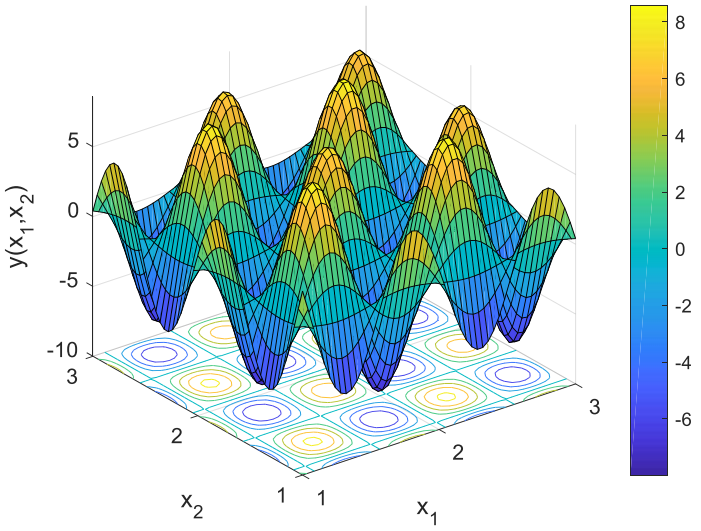


Fig. 12 The Shubert function has a periodic behavior repeated uniformly over the entire domain.

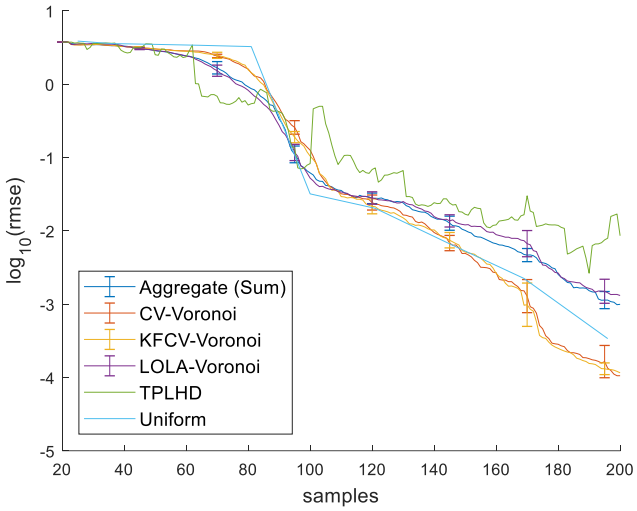


Fig. 13 Convergence history for surrogate response surfaces built using adaptive and one-shot sampling approaches applied to the Shubert function.

The RMSE histories of the adaptive sampling techniques and the one-shot approaches are presented in Fig. 13. The one-shot approaches perform significantly better for this benchmark function, and the uniform sampling approach outperforms the LOLA-Voronoi and aggregate approaches over a significant portion of the sample sizes considered. It is interesting to note that the LOLA- and aggregate-Voronoi techniques initially outperform the other adaptive sampling techniques. This is likely due to the initial exploration-based behavior as shown in Table 1. However, the KFCV- and CV- Voronoi overtake them around 120 samples and once again perform best for larger sample quantities. This shows that the CV-based Voronoi approaches value both exploitative and exploratory sampling.

B. 3D Benchmark functions

1. *Hartmann 3 function*

The Hartmann 3 function is a highly nonlinear function with three inputs and is given by:

$$y = -\sum_{i=1}^{4} c_i e^{\left|-\sum_{j=1}^{3} a_{ij}\left(x_j - p_{ij}\right)^2\right|}, x_j \in [0,1], j = 1,2,3, \tag{23}$$

where

$$\left[a_{ij}\right]_{j=1,\ldots,3} = \begin{bmatrix} 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3.0 & 10 & 30 \\ 0.1 & 10 & 36 \end{bmatrix}, \tag{24}$$

$$c_i = [1 \quad 1.2 \quad 3 \quad 3.2]^T, \tag{25}$$

and

$$\left[p_{ij}\right]_{j=1,\ldots,3} = \begin{bmatrix} 0.3689 & 0.1170 & 0.2673 \\ 0.4699 & 0.4387 & 0.7470 \\ 0.1091 & 0.8732 & 0.5547 \\ 0.0381 & 0.5743 & 0.8828 \end{bmatrix}. \tag{26}$$

The slices of the Hartmann 3 function over the parameter space are included in Fig. 14 .

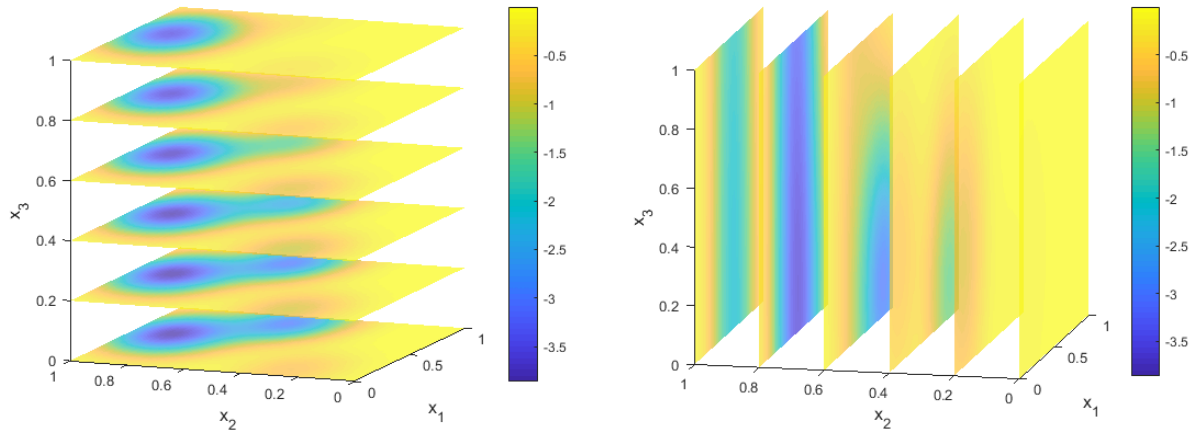


Fig. 14 The Hartmann 3 function is displayed using (a) x_3 slices and (b) x_2 slices.

The RMSE history of the sampling techniques is included in Fig. 15, and again the adaptive sampling techniques outperform the one-shot approaches. The performance of the adaptive sampling methods is similar, but the KFCV-Voronoi typically performs best for a given sample quantity.

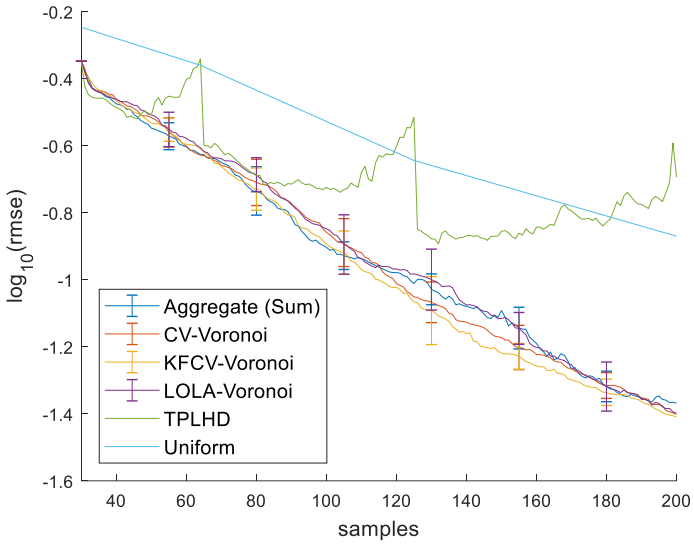


Fig. 15 Convergence history for surrogate response surfaces built using adaptive and one-shot sampling approaches applied to the Hartmann 3 functions.

The computational cost of evaluating the next sample is not constant between adaptive sampling approaches. A comparison of the wall clock time required to achieve a specified RMSE is included in Fig. 16. The alterations proposed here significantly reduce the computational time associated with each adaptive sampling technique. The KFCV-Voronoi requires less than half the computational time of the CV-Voronoi technique, and the altered LOLA-Voronoi technique features the lowest sampling cost.

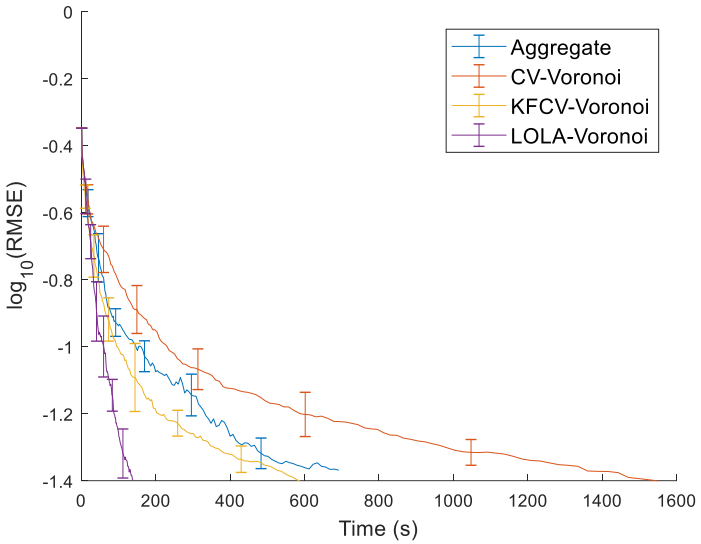


Fig. 16 Convergence history over time for surrogate response surfaces built using adaptive and one-shot sampling approaches applied to the Hartmann 3 function.

However, only the cost associated with finding the next sample point is considered here. The cost associated with calculating the response at each sample must also be considered when selecting the appropriate method. Depending on the ratio of the computational cost between the adaptive sampling selection and response evaluation, the best sampling procedure can vary. The KFCV-Voronoi has been shown to have less error when the number of samples is held constant. If evaluation of the response is considerably more expensive than the adaptive sample selection process, the KFCV-Voronoi adaptive sampling technique would be preferable.

C. 6D Benchmark functions

1. *Hartmann 6 function*

The need to reduce the cost associated with the adaptive sampling procedure becomes evident as the number of required samples increases. Larger sample sets become necessary to adequately resolve the response surfaces in a high-dimension parameter space. Therefore, the final benchmark function is the Hartmann 6 function, which is a highly nonlinear function with six inputs given by:

$$y = \sum_{i=1}^{4} c_i e^{\left[-\sum_{j=1}^{6} a_{ij}\left(x_j - p_{ij}\right)^2\right]}, \qquad x_j \in [0,1], j = 1, \ldots, 6\,, \tag{27}$$

where

$$\left[a_{ij}\right]_{j=1,\ldots,6} = \begin{bmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{bmatrix}, \tag{28}$$

$$c_i = [1 \quad 1.2 \quad 3 \quad 3.2]^T\,, \tag{29}$$

and

$$\left[p_{ij}\right]_{j=1,\ldots,6} = \begin{bmatrix} 0.1312 & 0.1696 & 0.5569 & 0.0124 & 0.8283 & 0.5886 \\ 0.2329 & 0.4139 & 0.8307 & 0.3736 & 0.1004 & 0.9991 \\ 0.2348 & 0.1451 & 0.3522 & 0.2883 & 0.3047 & 0.6550 \\ 0.4047 & 0.8828 & 0.8732 & 0.5743 & 0.1091 & 0.0381 \end{bmatrix}. \tag{30}$$

Only the CV-based Voronoi techniques have been examined in this example, and the RMSE history is included in Fig. 17. Again, the KFCV-Voronoi is competitive with the CV-Voronoi technique with respect to accuracy, at a dramatically reduced computational cost.

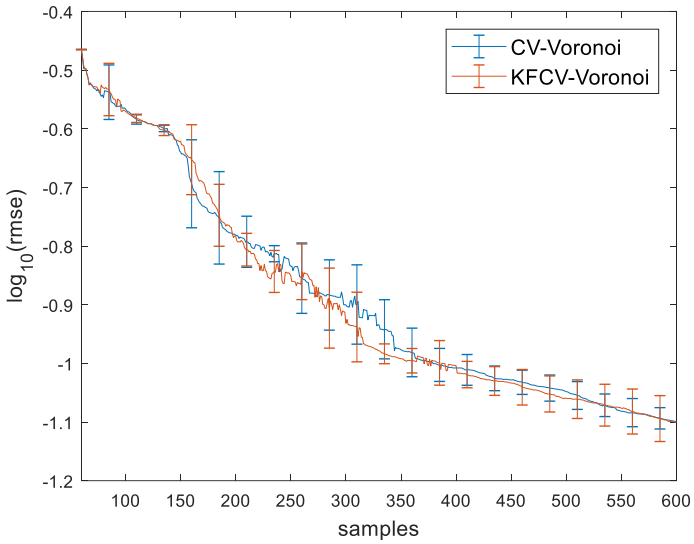


Fig. 17 Convergence history for surrogate response surfaces applied to the six-dimension Hartmann 6 function using the CV- and KFCV-based adaptive sampling approaches

D. Aerodynamic loading for NACA 00012

The CV-Voronoi and KFCV-Voronoi adaptive sampling techniques are now used to build surrogate response models for aerodynamic loading of a NACA 0012 airfoil. Here each sample corresponds to a CFD simulation. The two-dimensional parameter space considered freestream Mach numbers from 0.4 to 1.2 and angles of attack between 0° and 16°. The flow fields were solved using FUN3D [31] considering the RANS equations to solve for fully turbulent viscous flow fields. The lift coefficient is selected as the parameter of interest. Initially, the lift coefficient is sampled in a 17x17 uniform grid with variations of 0.05 in the freestream Mach and 1° in the angle of attack. These uniform samples are collected to evaluate the accuracy of the model built using the adaptive sampling techniques. A surrogate model was built for the 289 uniformly distributed samples and it is shown in Fig. 18. The lift coefficient clearly has regions with steep gradients and changes in gradients, particularly for a freestream Mach number around 0.85. Therefore, the expectation is that surrogate models developed using the adaptive sampling techniques will be more accurate than those developed using a one-shot technique.

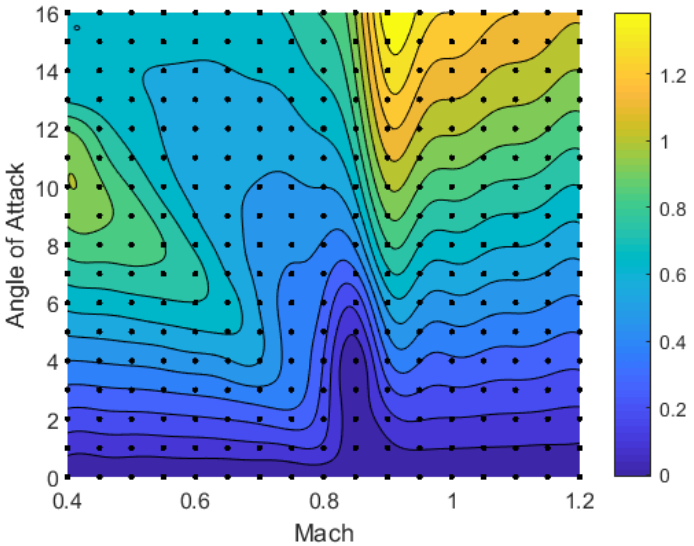


Fig. 18 A Uniform sampling process can be used to approximate the lift coefficient of a NACA 0012 within the parameter space.

The CV-Voronoi and KFCV-Voronoi adaptive sample techniques were evaluated for this case. Initially a surrogate model was built for 17 samples generated using TPLHD. The two adaptive sampling methods were applied to select a batch of 5 sample locations. This process was repeated to reduce the RMSE of the predicted values at the test points. A plot of the convergence history of the RMSE for increasing sample quantities selected by the two adaptive sampling techniques is included in Fig. 19 .

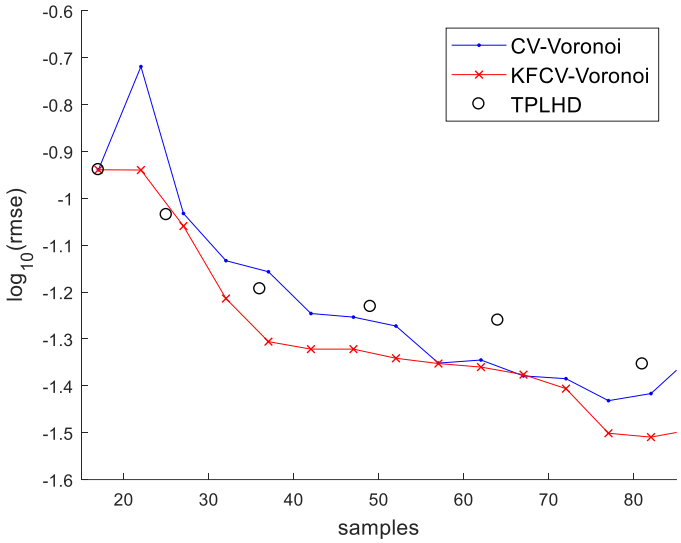


Fig. 19 Convergence history for surrogate models of the lift coefficient of a NACA 0012 airfoil for varying freestream Mach numbers and angles of attack created using CV-and KFCV-based adaptive sampling approaches

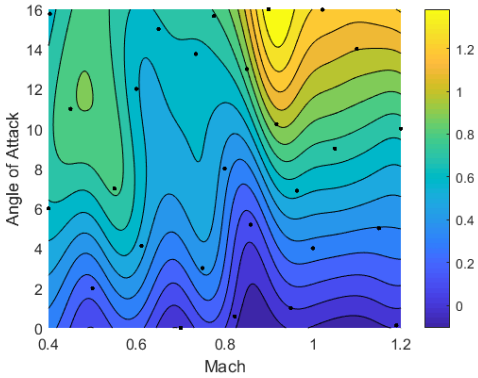
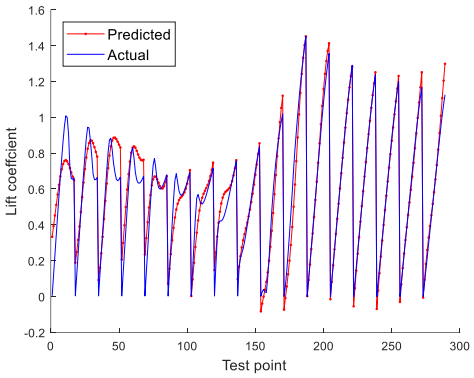
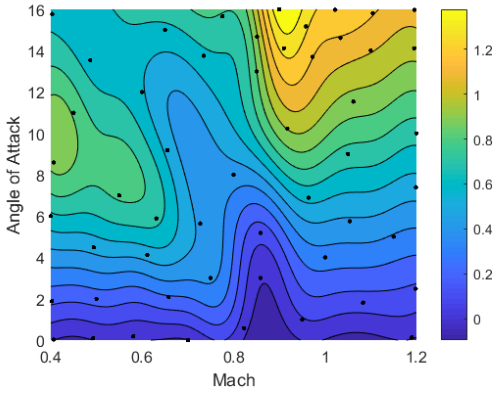
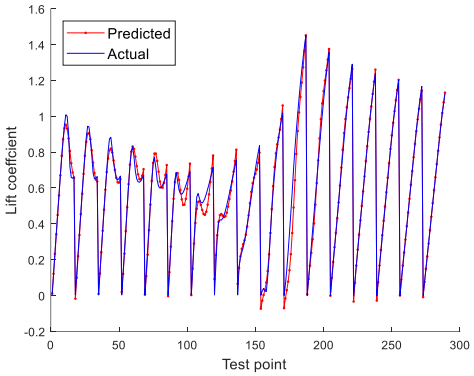
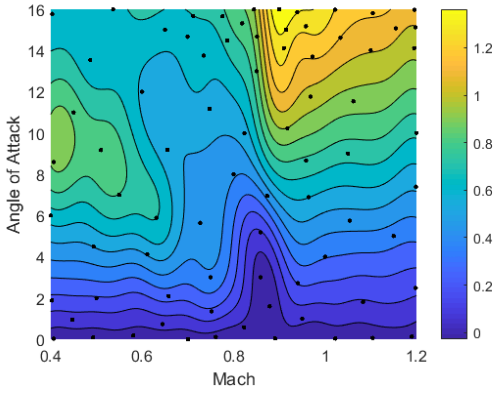
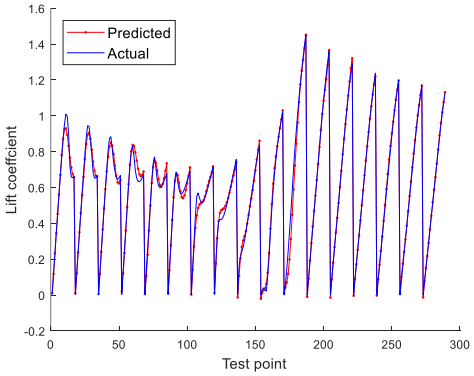
In Fig. 19 , the models built using the adaptive sampling techniques perform better than those sampled in the TPLHD one-shot manner, as the number of samples increases. It can also be seen that the samples selected using the KFCV-Voronoi technique consistently provide more accurate surrogate models. Table 2 presents contours of the surrogate model and comparisons of the predicted and actual values at the test point locations for the surrogate model built using the KFCV-Voronoi adaptive sampling technique at increasing sample quantities.

From the test evaluation and contour of the surrogate model built using 27-samples, we see that initially the model is inaccurate, and the maximum error of the surrogate model is 22% at test point 1, which corresponds to a Mach number of 0.4 and an angle of attack of 0°. For this set of samples test point 1 is an extrapolated value rather than an interpolated value (relative to the 27 sample points), and therefore, added difficulty is expected. Also, for the test points between 0 and 125, which correspond to low Mach numbers, the complex lift coefficient profile is captured very poorly.

As the sampling procedure progresses from 27 to 52 samples, the corners of the parameter space are sampled and samples are concentrated in the region defined by large angles of attack. The model built using 52-samples now includes the complex double peak valley behavior for the low Mach number test slices, but there are still visible discrepancies in the valleys around test point 100. The extrema of the high Mach number values are captured better, which is a result of the adaptive sampling procedure identifying the boundaries as an important sample location. The maximum error of the 52-sample surrogate model is 16% at test point 176, which corresponds to a Mach number of 0.9 and an angle of attack of 5°. The contour plot of the 52-sample model now resembles that constructed by the 287 uniform samples in Fig. 16. It is also interesting to note that the response surface model built using the 52 KFCV-Voronoi samples performs as well as the 81-sample TPLHD model, which corresponds to a sample reduction of 35%.

Advancing from 52 to 77 samples leads to samples concentrated around Mach 0.85 which is the region of the largest gradient and change in gradient. The 77-sample model captures the valleys and the high and low values better. The maximum error is now 10% and occurs in the lift coefficient trough at Mach 0.9 and an angle of attack of 4.5°, and the average prediction error is only 1.9%. Thus, surrogate modeling and adaptive sampling techniques provide models of the aerodynamic loading response of engineering accuracy for a considerably reduced testing cost.

Table 2 Surrogate model evaluation for samples collected using the KFCV-Voronoi adaptive sampling technique

	Surrogate Model Contours	Surrogate Model at Test Points
27 Samples		
52 Samples		
77 Samples		

VI. Conclusions

The goal of the present paper was to reduce the overall cost of surrogate modeling, particularly for high dimension inputs. The adaptive sampling techniques examined here significantly reduce the number of samples required to form accurate surrogate models when compared with one-shot sampling procedures. However, the adaptive sampling procedures add a cost associated with identifying ideal locations for the samples. This cost typically increases with the number of samples, and since high dimension parameter spaces frequently require large sample sets the cost can become intractable. As a result, this work concentrated on reducing cost associated with identifying ideal sample locations with adaptive sampling techniques, through alterations of popular adaptive sampling techniques and implementation of batch sampling.

Alterations were proposed for the LOLA-Voronoi and CV-Voronoi adaptive sampling techniques. For the LOLA-Voronoi technique, the cost associated with calculating the neighborhoods was reduced, which decreases the

computational time associated with sample selection by orders of magnitude. The CV-Voronoi technique was augmented with a preliminary KFCV evaluation, to form a scheme we call the KFCV-Voronoi adaptive sampling technique. The KFCV sweep limits the number response surface models that must be built during evaluation of future sample locations, and typically reduces the computational cost between 50 and 70%. Furthermore, the KFCV sweep acts as a filter that values a region's global impact in addition to its local impact, and has been shown to select samples that lead to improved model accuracy. Finally, a batch sampling procedure is proposed for the Voronoi cell-based adaptive sampling techniques, to prevent over clustering of future samples. The batch procedure does lead to slightly less accurate models, but this is typically more than made up for by maximizing the utilization of testing resources.

The accelerated adaptive sampling techniques were evaluated using benchmark functions of increasing input dimensions, and were shown to offer significant improvement over one-shot techniques. The sampling techniques were also validated for surrogate modeling the aerodynamic loading profile of a NACA 0012 airfoil. The adaptive sampling techniques were shown to dramatically reduce the number of samples required for response surface creation. The KFCV-Voronoi generated samples consistently led to the most accurate surrogate models, and did so with a dramatically lower computational cost.

Acknowledgements

This research is sponsored by US Air Force Test Center under contract FA9302-16-C-0018.

Bibliography

[1] V. Aute, K. Saleh, O. Abdelaziz, S. Azarm and R. Radermacher, "Cross-validation based single response adaptive design of experiments for Kriging metamodeling of deterministic computer simulations," *Structural and Multidisciplinary Optimization*, vol. 48, no. 3, pp. 581-605, 2013.

[2] A. S. VanderWyst, V. Sharma, C. L. Martin and W. A. Silva, "Big Data Challenges in Fluid-Thermal-Structural Interaction Research," in *58th {AIAA/ASCE/AHS/ASC} Structures, Structural Dynamics, and Materials Conference*, Grapevine, TX, 2017.

[3] D. Deschrijver, M. Mrozowski, T. Dhaene and D. De Zutter, "Macromodeling of multiport systems using a fast implementation of the vector fitting method," *IEEE Microwave and Wireless Components Letters*, vol. 18, no. 6, pp. 383-385, 2008.

[4] M. D. McKay, R. J. Beckman and W. J. Conover, "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239-245, 1979.

[5] F. A. Viana, G. Venter and V. Balabanov, "An algorithm for fast optimal Latin hypercube design of experiments," *International journal for numerical methods in engineering*, vol. 82, no. 2, pp. 135-156, 2010.

[6] A. B. Owen, "Orthogonal arrays for computer experiments, integration and visualization," *Statistica Sinica*, pp. 439-452, 1992.

[7] K. Crombecq, D. Gorissen, D. Deschrijver and T. Dhaene, "A novel hybrid sequential design strategy for global surrogate modeling of computer experiments," *SIAM Journal on Scientific Computing*, vol. 33, no. 4, pp. 1948-1974, 2011.

[8] H. Liu, Y.-S. Ong and J. Cai, "A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design," *Structural and Multidisciplinary Optimization*, pp. 1-24, 2017.

[9] T. Mackman and C. Allen, "Investigation of an adaptive sampling method for data interpolation using radial basis functions," *International journal for numerical methods in engineering*, vol. 83, no. 7, pp. 915-938, 2010.

[10] S. M. Clarke, J. H. Griebsch and T. W. Simpson, "Analysis of support vector regression for approximation of complex engineering analyses," *Journal of mechanical design*, vol. 127, no. 6, pp. 1077-1087, 2005.

[11] R. Zimmermann and S. Gortz, "Non-linear reduced order models for steady aerodynamics," *Procedia Computer Science*, vol. 1, no. 1, pp. 165-174, 2010.

[12] M. Guenot, I. Lepot, C. Sainvitu, J. Goblet and R. Filomeno Coelho, "Adaptive sampling strategies for non-intrusive POD-based surrogates," *Engineering computations*, vol. 30, no. 4, pp. 521-547, 2013.

[13] N. Cressie, "Spatial prediction and ordinary kriging," *Mathematical geology*, vol. 20, no. 4, pp. 405-421, 1988.

[14] N. Dyn, D. Levin and S. Rippa, "Numerical procedures for surface fitting of scattered data by radial functions," *SIAM Journal on Scientific and Statistical Computing*, vol. 7, no. 2, pp. 639-659, 1986.

[15] H. Fang and M. F. Horstemeyer, "Global response approximation with radial basis functions," *Engineering Optimization*, vol. 38, no. 04, pp. 407-424, 2006.

[16] F. Douak, F. Melgani, N. Alajlan, E. Pasolli, Y. Bazi and N. Benoudjit, "Active learning for spectroscopic data regression," *Journal of Chemometrics*, vol. 26, no. 7, pp. 374-383, 2012.

[17] W. Hendrickx and T. Dhaene, "Sequential design and rational metamodelling," in *Proceedings of the 37th conference on Winter simulation*, 2005.

[18] R. Jin, W. Chen and A. Sudjianto, "On sequential sampling for global metamodeling in engineering design," in *Proceedings of DETC*, 2002.

[19] J. Sacks, W. J. Welch, T. J. Mitchell and H. P. Wynn, "Design and analysis of computer experiments," *Statistical science*, pp. 409-423, 1989.

[20] M. C. Shewry and H. P. Wynn, "Maximum entropy sampling," *Journal of applied statistics*, vol. 14, no. 2, pp. 165-170, 1987.

[21] D. R. Jones, M. Schonlau and W. J. Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global optimization*, vol. 13, no. 4, pp. 455-492, 1998.

[22] G. Li, V. Aute and S. Azarm, "An accumulative error based adaptive design of experiments for offline metamodeling," *Structural and Multidisciplinary Optimization*, vol. 40, no. 1, pp. 137-155, 2010.

[23] S. Xu, H. Liu, X. Wang and X. Jiang, "A robust error-pursuing sequential sampling approach for global metamodeling based on voronoi diagram and cross validation," *Journal of Mechanical Design*, vol. 136, no. 7, p. 071009, 2014.

[24] G. G. Wang and S. Shan, "Review of metamodeling techniques in support of engineering design optimization," *Journal of Mechanical design*, vol. 129, no. 4, pp. 370-380, 2007.

[25] S. N. Lophaven, H. B. Nielsen and J. Sondergaard, "DACE-A Matlab Kriging toolbox, version 2.0," 2002.

[26] R. T. Clemen, "Combining forecasts: A review and annotated bibliography," *International journal of forecasting*, vol. 5, no. 4, pp. 559-583, 1989.

[27] Z.-H. Han, S. Gortz and R. Zimmermann, "Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function," *Aerospace Science and Technology*, vol. 25, no. 1, pp. 177-189, 2013.

[28] Z.-H. Han, R. Zimmermann and S. Gortz, "A new cokriging method for variable-fidelity surrogate modeling of aerodynamic data," *AIAA Paper*, vol. 1225, p. 2010, 2010.

[29] Z. Han, R. Zimmerman and S. Gortz, "Alternative cokriging method for variable-fidelity surrogate modeling," *AIAA journal*, vol. 50, no. 5, pp. 1205-1210, 2012.

[30] R. P. Liem, C. A. Mader and J. R. Martins, "Surrogate models and mixtures of experts in aerodynamic performance prediction for aircraft mission analysis," *Aerospace Science and Technology*, vol. 43, pp. 126-151, 2015.

[31] R. T. Biedron, J.-R. Carlson, J. M. Derlaga, P. A. Gnoffo, D. P. Hammond, W. T. Jones, B. Kleb, E. M. Lee-Rausch, E. J. Nielsen, M. A. Park and others, "FUN3D Manual: 12.9," 2016.

[32] A. S. VanderWyst, A. B. Shelton, C. L. Martin, L. J. Neergaard and Z. D. Witeof, "Reduced Order Models for Generation of Large, High Speed Aerodynamic Databases with Jet Interactions," in *57th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, 2016.

[33] L. Le Gratiet and C. Cannamela, "Cokriging-based sequential design strategies using fast cross-validation techniques for multi-fidelity computer codes," *Technometrics*, vol. 57, no. 3, pp. 418-427, 2015.