412TW-PA-19305

AN EFFICIENT BATCH K-FOLD CROSS-VALIDATION VORONOI ADAPTIVE SAMPLING TECHNIQUE FOR GLOBAL SURROGATE MODELING

ANDREW L. KAMINSKY
University of Tennessee, Knoxville, TN, 37996, USA
and CFD Research Corporation, Huntsville, AL, 35806, USA

YI WANG
University of South Carolina, Columbia, SC 29208, USA

KAPIL PANT
CFD Research Corporation, Huntsville, AL 35806, USA

WENDY N. HASHII
Edwards Air Force Base, CA 93524, USA

ABRAHAM ATACHBARIAN
Edwards Air Force Base, CA 93524, USA

MAY 2019

412TH TEST WING
EDWARDS AIR FORCE BASE, CALIFORNIA
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

4
1
2
T
W

An Efficient Batch K-Fold Cross-Validation Voronoi Adaptive Sampling Technique for Global Surrogate Modeling

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 widely used when data generation is expensive or time consuming. The accuracy of these models is dependent on the samples used to create them. Therefore, proper selection of samples within the parameter space is of paramount importance. Adaptive sampling procedures have been developed to identify optimal locations for new samples by leveraging response information from existing samples. In this manner, adaptive sampling methods significantly reduce the number of samples required to build a surrogate model of a given accuracy. However, adaptive sampling techniques have a cost associated with determining ideal sample locations, which typically grows with the sample count. The present effort seeks to reduce the cost associated with the adaptive sampling procedure and thereby maximize the efficiency of surrogate model creation. 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 cross-validation (CV)-Voronoi technique. The costs are further reduced through an innovative Voronoi batch sampling technique. The proposed adaptive sampling acceleration techniques are evaluated using benchmark functions with increasing parameter space dimension and aerodynamic loading data.

Nomenclature

β^*	Regression coefficients
e_L	Cross-validation error
e_{KF}	K-fold cross-validation error

[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

F	Design matrix
f	Regression function
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 modeling is used to closely approximate complex functional relationships at considerably reduced computational expense. Consequently, it is becoming increasingly prevalent in engineering analysis, particularly when data collection is time consuming, expensive, or difficult. Within the literature, 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. Once built, a surrogate model of sufficient accuracy can be used to predict the behavior of the untestested interstitial parameter space in place of the full model. However, the accuracy of a surrogate model is dependent upon how well the response behavior is captured by the samples used to create the model. Therefore, 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 grid 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 so that it offers the maximum improvement in model accuracy or certainty. Compared to one-shot techniques, adaptive sampling techniques provide surrogate models of equivalent accuracy with significantly fewer samples [7]. Effective adaptive sampling techniques must balance two contrasting objecitves exploration of the parameter space and exploitation of the response. Exploration is achieved by sampling regions far from existing points, i.e., where the prediction uncertainty is high. Conversely, exploitation is accomplished by leveraging information from response surfaces constructed from a set of initial samples to identify regions of interest (ROI) where additional sampling would provide the most useful information for resolving the complex functional behavior.

Numerous approaches have been developed to identify these ROI, most of which can be classified as gradient-, committee-, variance-, or cross-validation-based methods [8]. Gradient-based adaptive sampling approaches use the gradient and/or the Hessian of the response to identify regions that are difficult to model, i.e. regions with large gradients or Hessians. The Hessian matrix has been employed by the local linear approximation (LOLA)-Voronoi approach [7] and Mackman [9]. Alternatively, the committee strategy uses a combination of different surrogate model approaches, e.g., multivariate polynomials, support vector regression [10], proper orthogonal decomposition [11] [12], Kriging [13], and radial basis functions [14] [15], etc. A committee of one or more surrogate modeling techniques is formed, and a model is built using each method. The region where the models of the committee 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. Conversely, variance-based adaptive sampling techniques use the model prediction variance to identify regions of interest where the model suffers from high uncertainty. These approaches are typically linked to Kriging or Gaussian process-based surrogate models, which provide a prediction variance in addition the surrogate model. The variance or mean square error (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, cross-validation (CV) approaches divide the sample set into training and testing sets, 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 sampling techniques have all been developed to minimize testing costs by ensuring each additional sample provides as much information as possible.

Within the literature the Voronoi type adaptive sampling approaches (e.g. LOLA-Voronoi [7] and CV-Voronoi [23]) have enjoyed considerable popularity. The LOLA-Voronoi approach was one of the first adaptive sampling techniques that could be applied independently of the selected surrogate modeling method. The development of a generic adaptive sampling approach that could be coupled with any adaptive sampling approach led to prevalent adoption. The CV-Voronoi adaptive sampling technique maintained the strength of the LOLA-Voronoi technique, but was demonstrated to provide samples that led to models of improved accuracy. These adaptive sampling approaches have a cost associated with sample selection that increases superlinearly with the number of sample points. The goal of the present work is to reduce the cost associated with identifying future sample locations.

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. Additionally, it is often desirable to select samples in batches rather than one at a time. A novel Voronoi batch sampling method, which significantly outperforms the naïve approach is implemented. Batch sampling further decreases the cost associated with adaptive sampling by reducing the number of calls to the adaptive sampling techniques and enabling testing resources to be used in parallel. This paper is organized as follows: Section II outlines the Kriging method used here for surrogate modeling; Section III introduces the KFCV-Voronoi and Voronoi batch adaptive sampling techniques for reduced cost adaptive sampling; Section IV presents a method to evaluate surrogate model accuracy; and Section V demonstrates the efficacy of the adaptive sampling techniques relative to the current state-of-the-art.

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:

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

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,

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

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]. Though the Kriging model is used excusively to build surrogate models within this work, an advantage of the proposed adaptive sampling method is it has no direct reliance on the Kriging model and can be used with any other surrogate modeling method (like the other Voronoi class adaptive sampling techniques).

III. Proposed Adaptive Sampling Methodology

The accuracy of a surrogate model is dependent upon the information conveyed by the discrete samples of the response used to build it. As a result, the sample location and selection process can have a large impact on the number of samples required to obtain a surrogate model of sufficient accuracy. Samples selected via adaptive sampling techniques have been demonstrated to provide models of a desired accuracy with fewer samples. However, these 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 thereby maximize the cost savings provided by adaptive sampling. A new K-fold cross-validation (KFCV)-Voronoi adaptive sampling technique is presented here to reduce the sample selection costs by adding a global KFCV filter to the cross-validation (CV)-Voronoi technique.

A. Cross-Validation Voronoi

Since the KFCV-Voronoi adaptive sampling technique builds upon the CV-Voronoi technique we begin with a brief summary of the CV-Voronoi adaptive sampling approach. 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. A Voronoi cell corresponds to the region that is closer to a particular sample than any other sample. Fig. 1 presents an example of Voronoi tessellations.

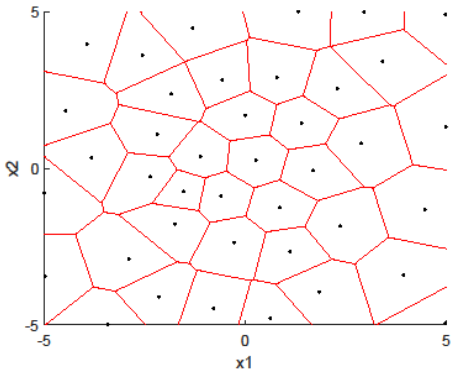


Fig. 1 Voronoi cells consist of the region that is closer to a given point than any other sample

The 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:

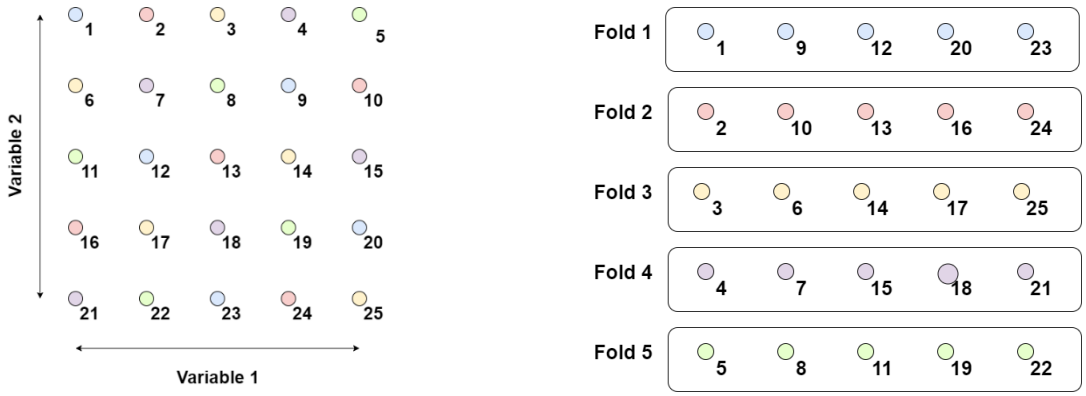
$$e_{LOO}^{i} = \left| y(s_i) - \hat{y}_{S\setminus s_i}(s_i) \right|. \tag{9}$$

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 repeating this process for all off the cells and identifying the Voronoi cell with the largest prediction error. The new sample is then selected as the point within the Voronoi cell with the largest error 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 to evaluate the LOOCV error 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.

B. 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 sample selection cost 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. 2(a). shows this for a two-dimensional parameter space with 25 samples, which is assigned divided into 5 folds (for this example $K = 5$). The fold assignment is portrayed by the sample color, and each fold is shown in Fig. 2(b).



(a) (b)

Fig. 2 (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. 3 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.

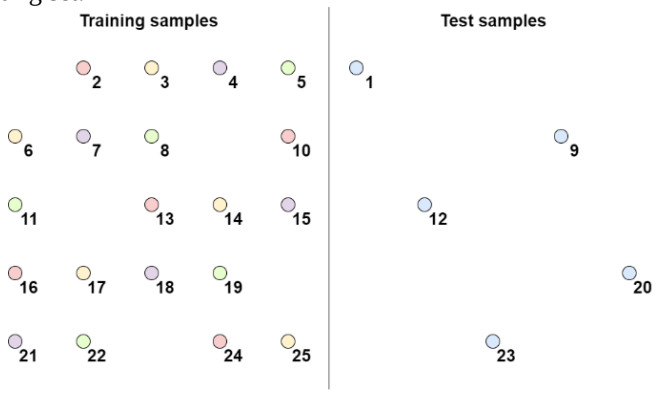


Fig. 3 KFCV utilizes a single fold as a test set for the remaining folds

The error associated with each fold is calculated in a manner similar to the LOOCV error metric, Eq. (9), and it is defined as

$$e_{KF}^{i} = \sum_{j \in kf_i}^{n_K} \left| y(s_j) - \hat{y}_{S\backslash kf_i}(s_j) \right| , \tag{10}$$

where n_K is the number of samples per fold, $y(s_j)$ is the true response at s_j, and $\hat{y}_{S\backslash kf_i}(s_j)$ 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. By implementing KFCV to initially approximate the modeling error the number of surrogate response surfaces required is limited to $K + l\,\frac{n_{samp}}{K}$ instead of n_{samp}.

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

C. Voronoi Batch Sampling

It is often desirable to select samples in batches rather than one at a time. Batches enable testing resources to be used in parallel and reduces the number of calls to the adaptive sampling procedure. However, within the literature the Voronoi type adaptive sampling techniques have only been presented selecting a single point a time. A naïve batch sampling approach would use place a sample within the Voronoi cell of the n cells with the largest CV error that is furthest from the sample (where n is the batch size). However, samples selected in this manner can be overly clustered since cells with large errors are likely to neighbor one another, as shown in Fig. 4. The resulting samples are likely redundant.

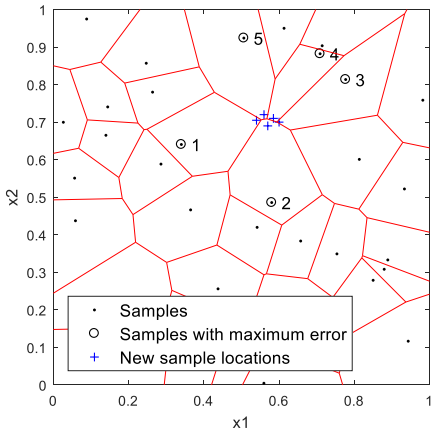


Fig. 4 Future samples from a naïve batch sampling approach.

In this work, a new Voronoi batch sampling approach is presented to overcome the deficiencies of the naïve implementation. The new 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. 5(a). The Voronoi regions are then reassigned, taking the new point into account, as illustrated in Fig. 5(b). The second sample is then selected within the updated Voronoi region associated with the second highest score, as demonstrated in Fig. 5(c). This procedure is then repeated until the number of samples in the batch is reached. Fig. 5(d) presents the resulting 5-sample batch. Comparing the sample distributions of the naïve

approach in Fig. 4 with the new Voronoi batch sampling approach in Fig. 5(d) it is clear that the latter provides a sample set that is less likely to provide redundant information for surrogate model creation.

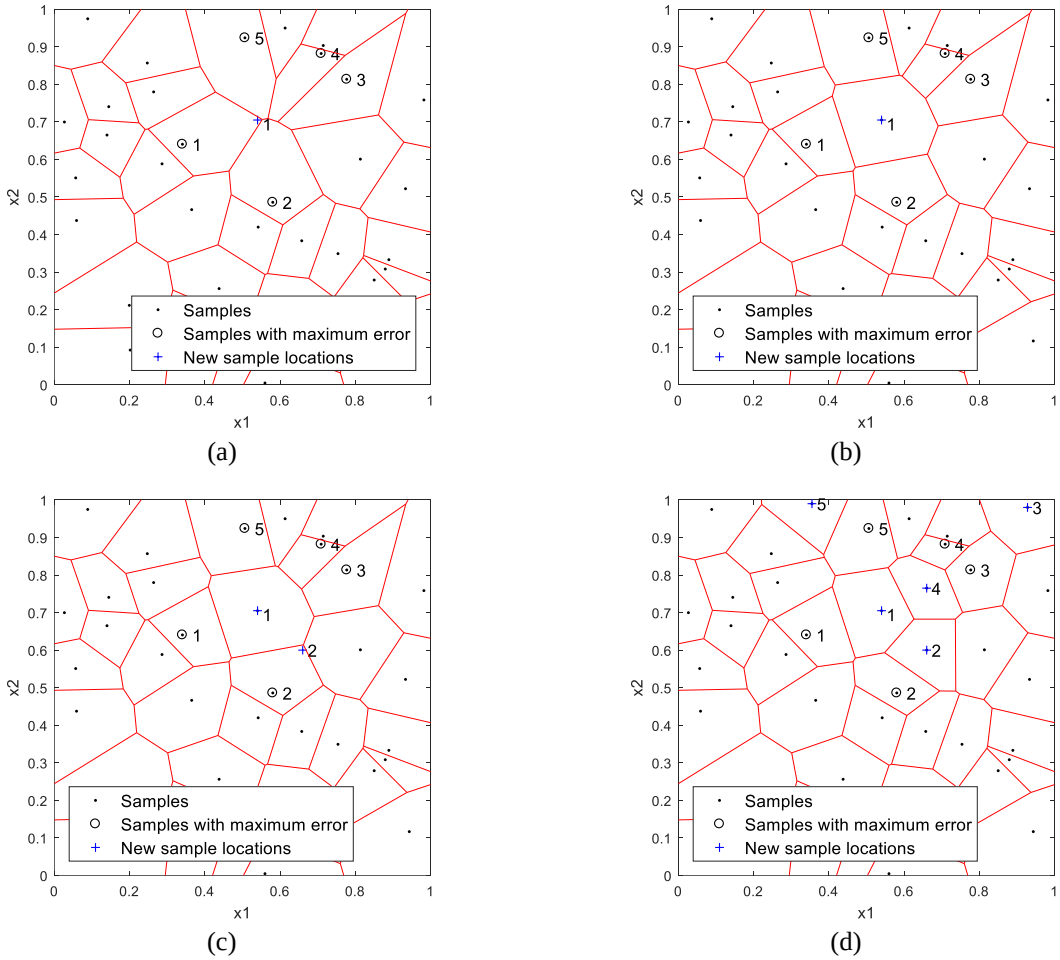


Fig. 5 The Voronoi batch sampling procedure: (a) the first sample is identified, (b) the Voronoi cells are reassigned, (c) the second sample is selected within the new Voronoi cells, (d) the procedure is repeated until the full batch is found.

IV. Response surface accuracy evaluation

The primary goal of an adaptive sampling technique is to choose samples that enable creation of surrogate models of improved quality for a given sample quantity. Therefore, to assess the adaptive sampling techniques we need a metric to asses the quality of a given surrogate model. Many techniques have been proposed to evaluate surrogate model accuracy, but within this work we evaluate the response surfaces using root mean square error (RMSE) or cross-validation error (CVE). RMSE is defined as

$$RMSE = \sqrt{\frac{1}{t}\sum_{i=1}^{t}\left(y(q_i) - \hat{y}(q_i)\right)^2}. \tag{11}$$

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

To demonstrate the efficacy of the KFCV-Voronoi and Voronoi batch sampling methods, we initially compare their performance with popular adaptive sampling methods from the literature for a variety of analytic benchmark functions. Because the benchmark functions can be evaluated so rapidly, the computational time is almost entirely dependent on adtaptive sampling technique. As a result, the benchmark functions enable direct analysis of computational cost savings. We consider benchmark functions of increasing complexity to examine performance of the adaptive sampling techniques over a set of problems. After scrutinizing the performance using these benchmark functions, the adaptive sampling techniques are applied to support modeling of aerodynamic loading coefficients for a NACA 0012 airfoil over a parameter space of interest. The adaptive sampling techniques are also compared with the uniform distributions and translational propagation Latin hyper cube design (TPLHD) [5] one-shot sampling methods.

A. 2D Benchmark functions

1. Peaks function

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

$$y = 3(1-x_1)^2 e^{\left(-x_1^2-(x_2+1)^2\right)} - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right)e^{\left(-x_1^2-x_2^2\right)} - \frac{1}{3}e^{\left(-(x+1)^2-x_2^2\right)} \text{ for } x_{1,2} \in [-5,5]\,, \qquad (12)$$

Fig. 6 presents the peaks function which, features non-linear behavior near the origin but remains constant away from the origin.

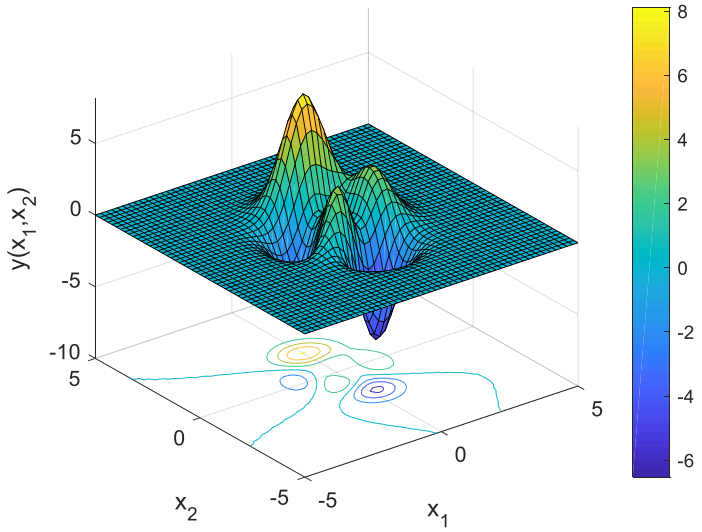


Fig. 6 The peaks function has complex behavior near the origin

The peaks function was chosen because it represents the ideal scenario for demonstrating the value of an adaptive sampling approach. The constant region away from the origin can be correctly modeled using very few samples, and 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 should outperform one-shot sampling techniques by concentrating samples in near the origina where the response is harder to model. Fig. 7 presents the RMSE history of the proposed adaptive sampling techniques applied to the peaks function beginning from 20 samples generated using TPLHD.

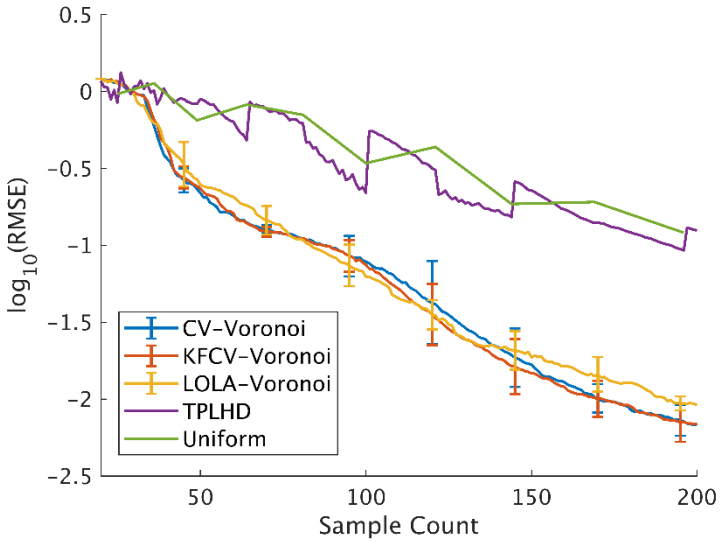
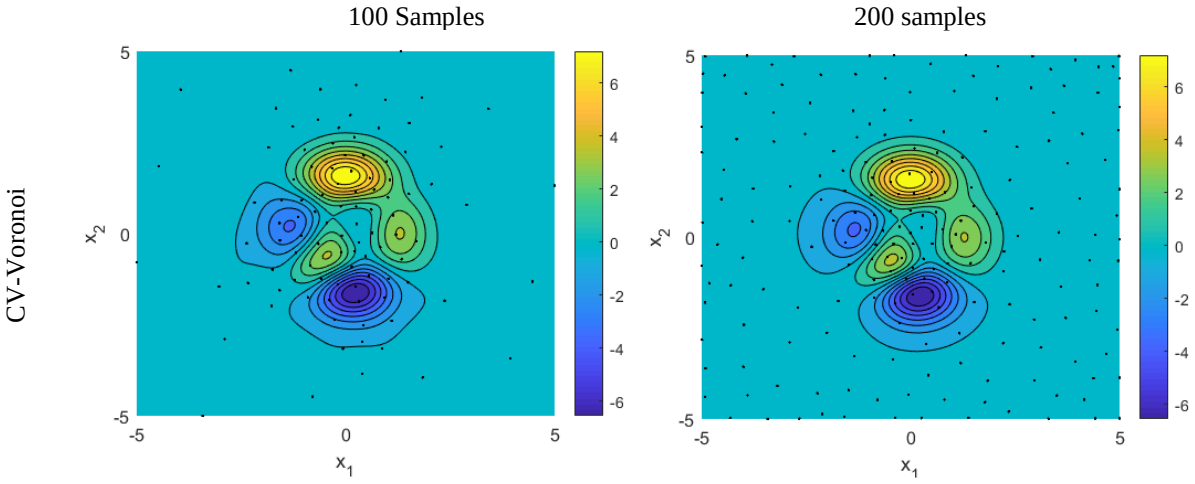
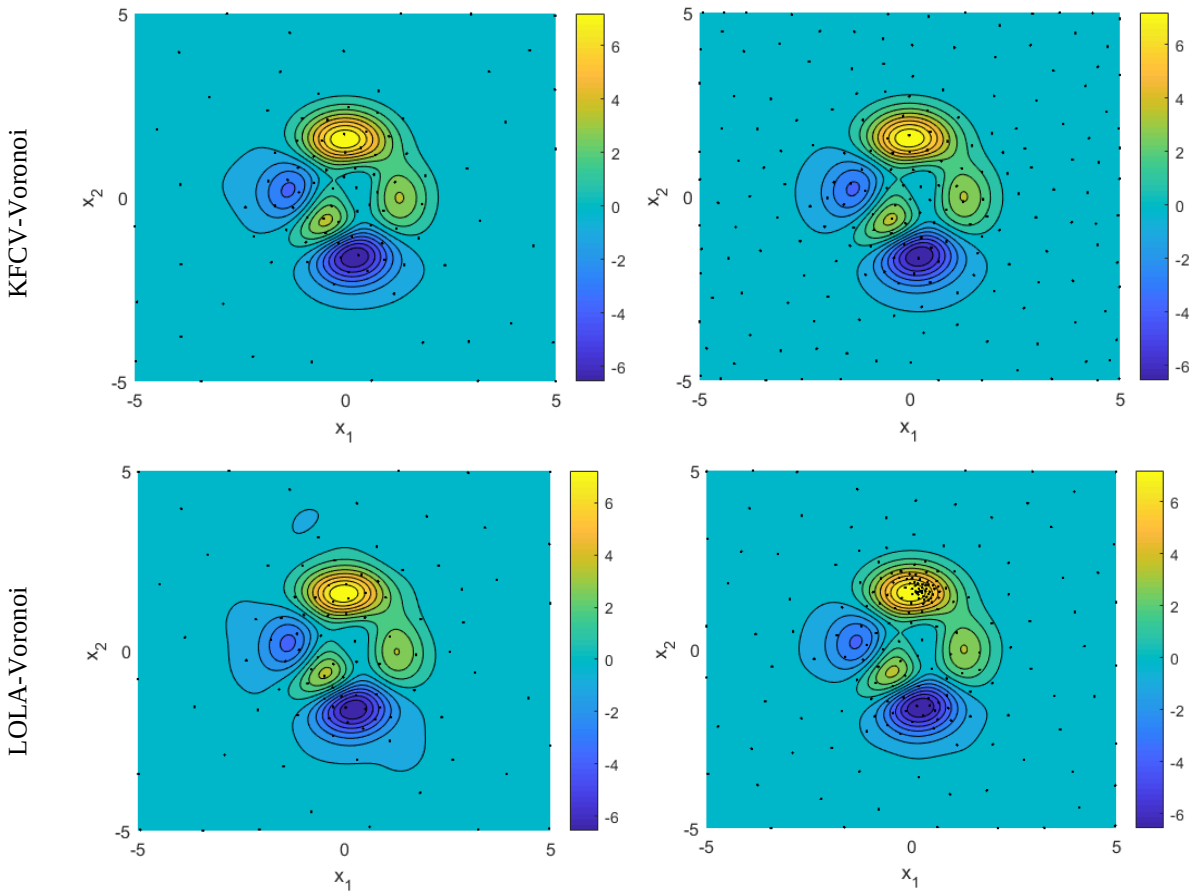


Fig. 7 Convergence history for surrogate response surfaces built using adaptive and one-shot sampling approaches applied to the peaks function.

As expected, the adaptive sampling procedures dramatically outperform the one-shot methods. Initially, the adaptive sampling techniques perform analogously, but after selection of 130 samples the CV- and KFCV-Voronoi approaches outperform the LOLA-Voronoi approach. 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.

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





The KFCV-Voronoi RMSE profile in Fig. 7 was generated by using 10-fold KFCV-Voronoi and the samples of the 3 folds with the highest error were evaluated using LOOCV. Setting the number of folds K to 10, we evaluate the performance of the KFCV-Voronoi technique with LOOCV applied to different fold quantities $l = 1, 3$, and 5. Fig. 8 shows that regardless of the number of folds LOOCV is performed for, the KFCV-Voronoi approach outperforms the CV-Voronoi technique in terms of model quality. Even more interestingly the KFCV-Voronoi does this at a reduced computational cost, as shown in Fig. 8(b). The cost reduction is due to the decrease in the number of surrogate models that must be formed to select the next point.

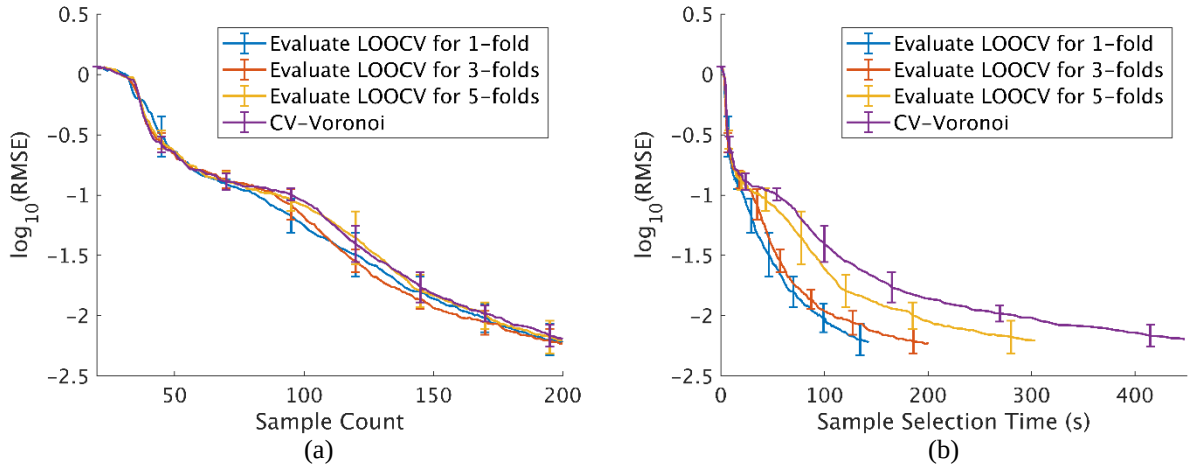


Fig. 8 Convergence history for the surrogate response surfaces built using KFCV-Voronoi applied to the peaks function with different fold quantities (a) compared by sample count and (b) by time.

The KFCV-Voronoi technique was initially implemented purely as a method to reduce the computational cost associated with selecting the next sample location. It was expected that models created using the KFCV-Voronoi adaptive sampling would provide models inferior to those built using CV-Voronoi technique, because the KFCV-Voronoi could potentially misidentify the best Voronoi cell to sample if it was lumped with relatively unimportant points. However, the accuracy of models built using the samples selected through the KFCV-Voronoi adaptive sampling technique is equal to or better than those built from the CV-Voronoi technique. The improved sample selection is potentially due to a global filtering effect that occurs when performing KFCV, since the error in Eq. 10 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.

We also used the peaks function to evaluate samples generated in batches, instead of a single sample at a time. By generating batches of samples, it is possible to evaluate multiple responses in parallel, and make better use of testing resources. Batch sampling also reduces the number of times the adaptive sampling process must be performed, which can reduce the cost associated with sample selection. Fig. 9 presents the RMSE history for the KFCV-Voronoi based adaptive sampling procedure performing LOOCV for different fold quantities and batch sizes.

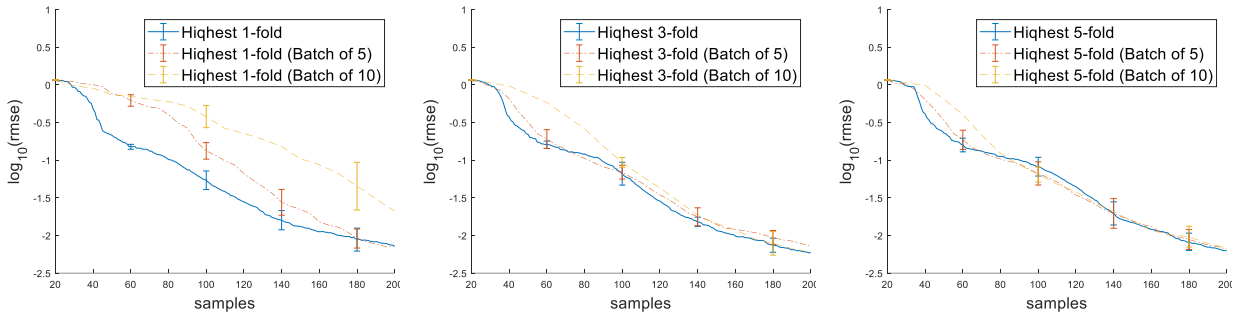


Fig. 9 Convergence history for surrogate response surfaces built using KFCV-Voronoi applied to the peaks function with increasing batch sizes and LOOCV fold counts l : (a) 1-fold, (b) 3-folds, and (c) 5-folds

As the batch size increases, KFCV-Voronoi with LOOCV for only a single fold selects samples poorly, because when LOOCV is performed for only a single fold, 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. 9 (b) and Fig. 9 , 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

$$y = \left(\sum_{i=1}^{5} i\cos\big((i+1)x_1 + i\big)\right) \times \left(\sum_{i=1}^{5} i\cos\big((i+1)x_2 + i\big)\right) \text{ for } x_{1,2} \in [1,3]\,. \tag{13}$$

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. 10, 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.

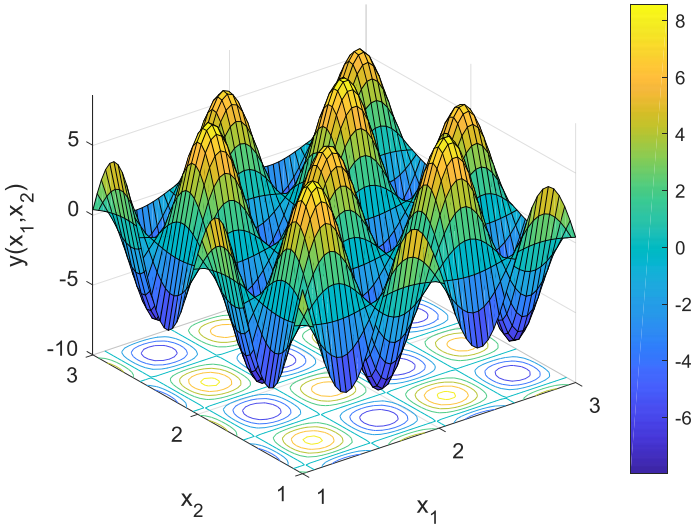


Fig. 10 The Shubert function has a periodic behavior repeated uniformly over the entire domain.

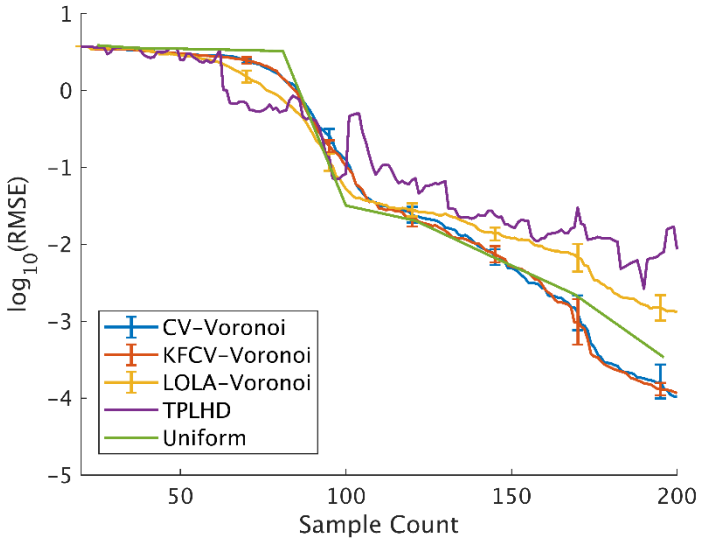


Fig. 11 Convergence history for surrogate response surfaces built using adaptive and one-shot sampling approaches applied to the Shubert function.

The RMSE histories of the adaptive sampling techniques and the one-shot approaches are presented in Fig. 11. The one-shot approaches perform significantly better for this benchmark function, and the uniform sampling approach outperforms the LOLA-Voronoi approach over a significant portion of the sample sizes considered. It is interesting to note that the LOLA-Voronoi technique initially outperforms 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:

$$y = -\sum_{i=1}^{4} c_i e^{\left|-\sum_{j=1}^{3} a_{ij}\left(x_j - p_{ij}\right)^2\right|}, x_j \in [0,1], j = 1,2,3, \tag{14}$$

where

$$\left[a_{ij}\right]_{j=1,\ldots,3} = \begin{bmatrix} 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3.0 & 10 & 30 \\ 0.1 & 10 & 36 \end{bmatrix}, \tag{15}$$

$$c_i = [1 \quad 1.2 \quad 3 \quad 3.2]^T, \tag{16}$$

and

$$\left[p_{ij}\right]_{j=1,\ldots,3} = \begin{bmatrix} 0.3689 & 0.1170 & 0.2673 \\ 0.4699 & 0.4387 & 0.7470 \\ 0.1091 & 0.8732 & 0.5547 \\ 0.0381 & 0.5743 & 0.8828 \end{bmatrix}. \tag{17}$$

The slices of the Hartmann 3 function over the parameter space are included in Fig. 12 .

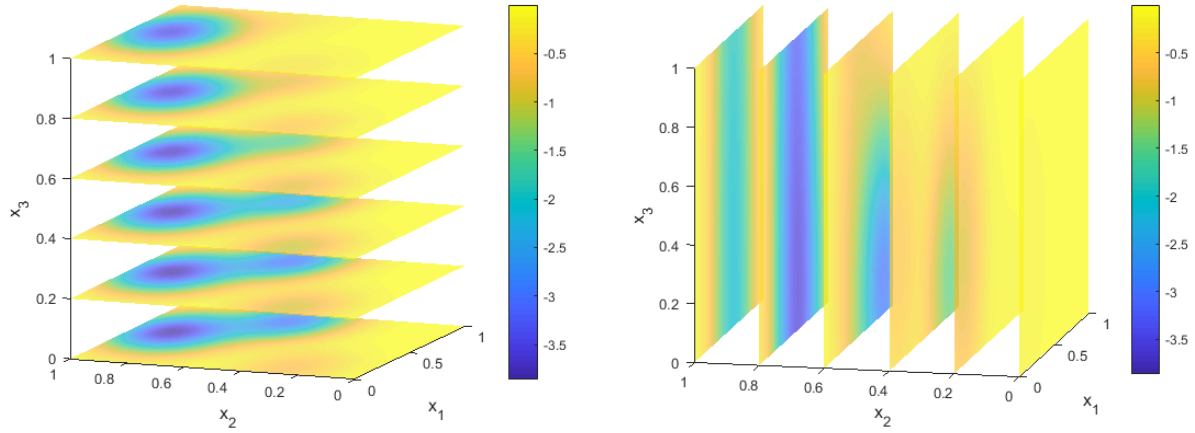


Fig. 12 The Hartmann 3 function is displayed using (a) x_3 slices and (b) x_2 slices.

The RMSE history of the sampling techniques is included in Fig. 13, 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.

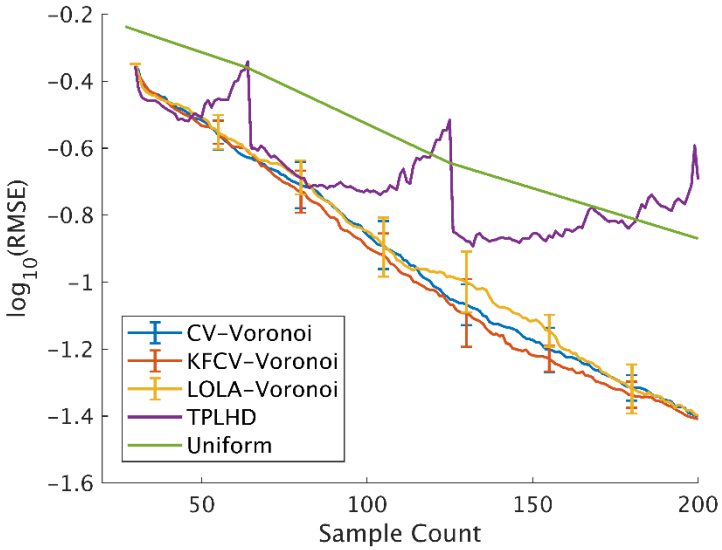


Fig. 13 Convergence history for surrogate response surfaces built using adaptive and one-shot sampling approaches applied to the Hartmann 3 functions.

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. 14. 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.

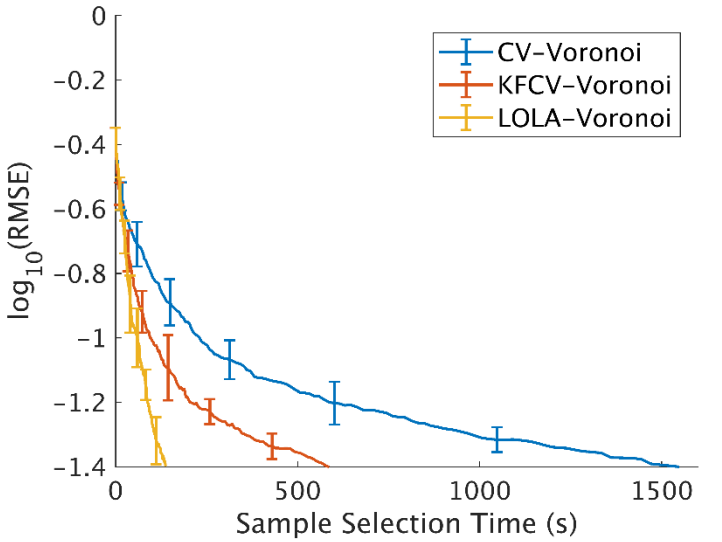


Fig. 14 Convergence history over time for surrogate response surfaces built using adaptive and one-shot sampling approaches applied to the Hartmann 3 function.

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:

$$y = \sum_{i=1}^{4} c_i e^{\left[-\sum_{j=1}^{6} a_{ij}(x_j - p_{ij})^2\right]}, \qquad x_j \in [0,1], j = 1, \ldots, 6\,, \tag{18}$$

where

$$[a_{ij}]_{j=1,\ldots,6} = \begin{bmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{bmatrix}, \tag{19}$$

$$c_i = [1 \quad 1.2 \quad 3 \quad 3.2]^T\,, \tag{20}$$

and

$$[p_{ij}]_{j=1,\ldots,6} = \begin{bmatrix} 0.1312 & 0.1696 & 0.5569 & 0.0124 & 0.8283 & 0.5886 \\ 0.2329 & 0.4139 & 0.8307 & 0.3736 & 0.1004 & 0.9991 \\ 0.2348 & 0.1451 & 0.3522 & 0.2883 & 0.3047 & 0.6550 \\ 0.4047 & 0.8828 & 0.8732 & 0.5743 & 0.1091 & 0.0381 \end{bmatrix}. \tag{21}$$

Only the CV-based Voronoi techniques have been examined in this example, and the RMSE history is included in Fig. 15. Again, the KFCV-Voronoi is competitive with the CV-Voronoi technique with respect to accuracy, at a dramatically reduced computational cost.

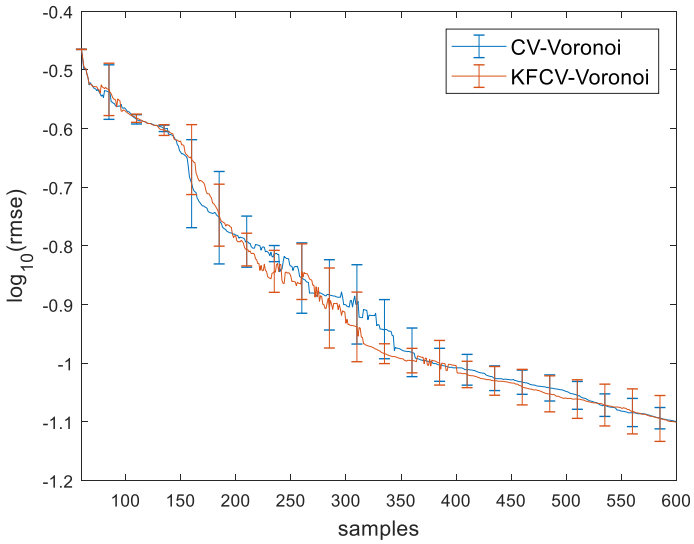


Fig. 15 Convergence history for surrogate response surfaces applied to the six-dimension Hartmann 6 function using the CV- and KFCV-based adaptive sampling approaches

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 [26] 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. 16. 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.

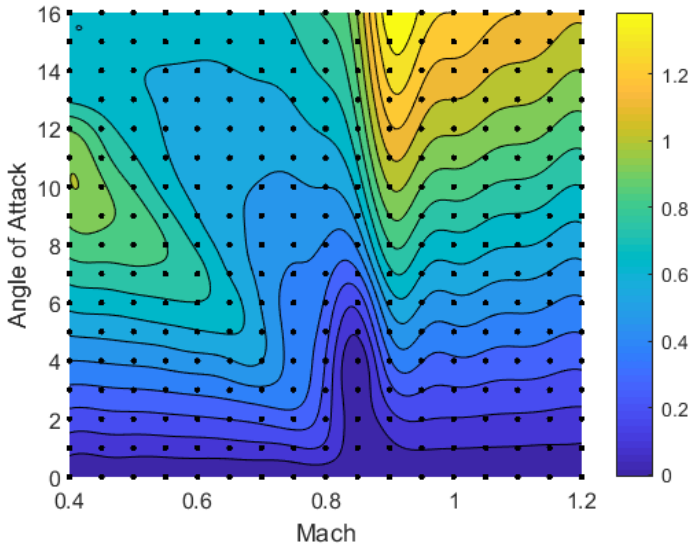


Fig. 16 A Uniform sampling process can be used to approximate the lift coefficient of a NACA 0012 within the parameter space.

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. 17 .

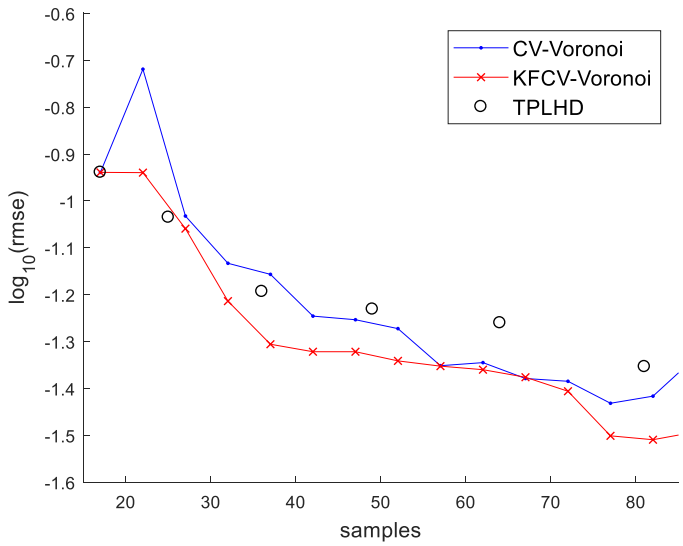


Fig. 17 Convergence history for surrogate models of the lift coefficient of a NACA 0012 airfoil for varying freestream Mach numbers and angles of attack created using CV-and KFCV-based adaptive sampling approaches

In Fig. 17 , 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

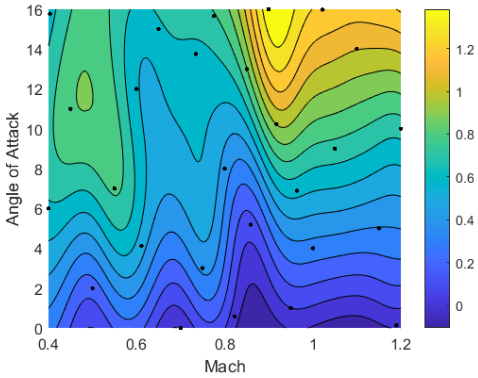
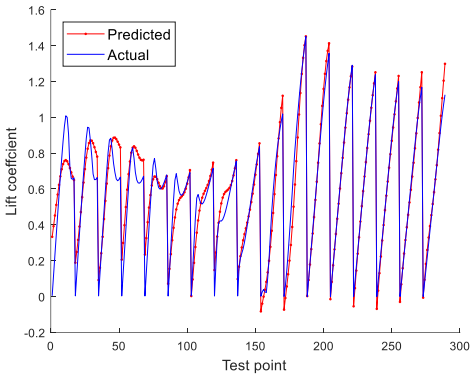
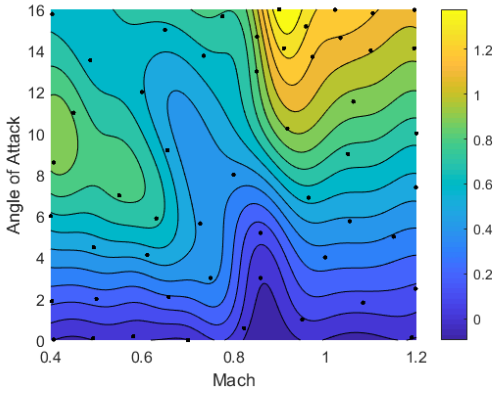
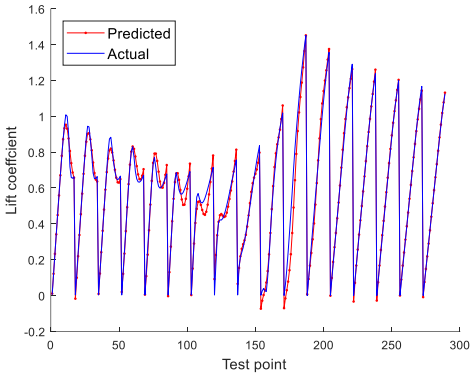
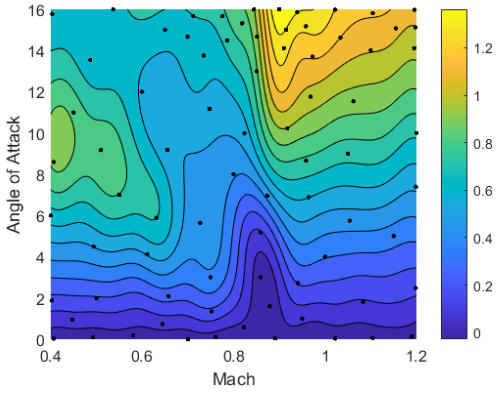
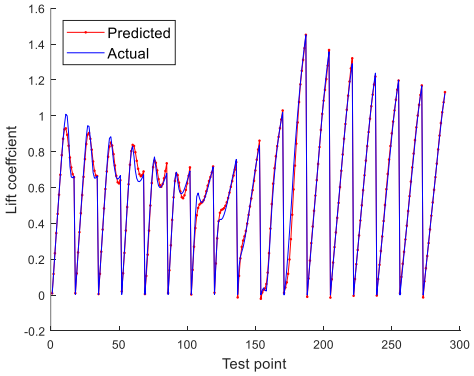
surrogate model and comparisons of the predicted and actual values at the test point locations for the surrogate model built using the KFCV-Voronoi adaptive sampling technique at increasing sample quantities.

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 CV-Voronoi adaptive sampling techniques. 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. 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.

[27] 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.

[28] 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.

[29] 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.

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

[31] 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.

[32] 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.

[33] 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.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
03-06-2019	Technical Paper	N/A

4. TITLE AND SUBTITLE
An Efficient Batch K-Fold Cross-Validation Voronoi Adaptive Sampling Technique for Global Surrogate Modeling

5a. CONTRACT NUMBER
FA9302-16-C-0018

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Andrew L. Kaminsky
Yi Wang
Kapil Pant
Wendy N. Hashii
Abraham Atachbarian

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
CFD Research Corporation, 701 McMillan Way NW, Huntsville, AL 35806
University of South Carolina, 300 Main Street, Columbia, SC 29208

8. PERFORMING ORGANIZATION REPORT NUMBER
412TW-PA-19305

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
412th Test Wing
195 E Popson Ave
Edwards AFB CA 93524

10. SPONSOR/MONITOR'S ACRONYM(S)
N/A

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release A: distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Surrogate models can be used to approximate complex systems at a reduced cost and are widely used when data generation is expensive or time consuming. The accuracy of these models is dependent on the samples used to create them. Therefore, proper selection of samples within the parameter space is of paramount importance. Adaptive sampling procedures have been developed to identify optimal locations for new samples by leveraging response information from existing samples. In this manner, adaptive sampling methods significantly reduce the number of samples required to build a surrogate model of a given accuracy. However, adaptive sampling techniques have a cost associated with determining ideal sample locations, which typically grows with the sample count. The present effort seeks to reduce the cost associated with the adaptive sampling procedure and thereby maximize the efficiency of surrogate model creation. 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 cross-validation (CV)-Voronoi technique. The costs are further reduced through an innovative Voronoi batch sampling technique. The proposed adaptive sampling acceleration techniques are evaluated using benchmark functions with increasing parameter space dimension and aerodynamic loading data.

15. SUBJECT TERMS
Adaptive Sampling, Surrogate modeling, Design of Experiments, Batch Sampling, K-Fold Cross-Validation

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
Unclassified					412 TENG/EN (Tech Pubs)
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	None	21	**19b. TELEPHONE NUMBER** *(include area code)*
Unclassified	Unclassified	Unclassified			661-277-8615