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

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

Yi Wang[2]
University of South Carolina, Columbia, SC, 29208, USA

Kapil Pant
CFD Research Corporation, Huntsville, AL, 35806, USA

And

Wendy N. Hashii and Abraham Atachbarian
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 sample selection within the parameter space is paramount. Numerous design of experiments (DOE) methodologies have been developed with the aim of identifying the optimal sample set to capture the system of interest. Adaptive sampling techniques are a subclass of DOE methods that identify optimal locations for new samples by leveraging response information from existing samples. By exploiting knowledge of the system, adaptive sampling methods have been demonstrated to significantly reduce the number of samples required to build a surrogate model of a given accuracy. However, utilizing the response information of the previous samples adds a computational cost associated with determining the ideal sample locations. Additionally, this cost typically grows with the sample count. This article presents techniques to reduce the cost associated with the adaptive sampling procedure so that the cost savings provided by adaptive sampling are maximized. 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 that is demonstrated to outperform naïve batch sampling. The proposed adaptive sampling acceleration techniques are evaluated using benchmark functions of increasing dimension and aerodynamic loading data.

Nomenclature

β^*	Regression coefficients
e_{LOO}	Leave one out cross-validation error
e_{KF}	K-fold cross-validation error
$\boldsymbol{F}$	Design matrix
f	Regression function
K	Number of folds
l	Number of folds LOO is performed for
p	Number of regression functions
$\boldsymbol{Q}$	Set of test points
$\boldsymbol{R}$	Stochastic-process correlation matrix
$\mathbf{S}$	Set of samples
$\boldsymbol{S}\backslash\boldsymbol{s_r}$	Set of samples without s_r
$\boldsymbol{s}$	Sample
t	Number of test points

[1] Corresponding Author: andrew.kaminsky@cfdrc.com
[2] Corresponding Author: yiwang@cec.sc.edu

Submitted to Journal of Mechanical Design
DISTRIBUTION A. Approved for public release; distribution is unlimited. 412TW-PA-20130.

V	Relative volume
$\boldsymbol{x}$	Input variable vector
$\boldsymbol{Y}$	Set of responses
y	Response
$\hat{y}$	Surrogate response surface
Subscript	
i	Variable number
r	Reference

I. Introduction

Within engineering analysis [1] [2] [3] [4], it is often necessary to investigate output responses of simulations or experiments over a parameter space. Surrogate modeling is a set of techniques used to develop models to approximate complex functional relationships between a response and its explanatory variables. Once built, the surrogate model can be used to predict the response of the untested interstitial parameter space in place of the full model, at considerably reduced computational expense. Surrogate models are constructed from discrete evaluations of responses to inputs within a parameter space of interest. For many real-world engineering problems, collecting each discrete sample from an experiment or simulation can take hours or days to complete, and the accuracy of the surrogate model is dependent upon how well the samples capture the response behavior. As a result, it is desired to identify optimal sample locations which minimize the number of discrete samples required to create a surrogate model of sufficient accuracy.

Design of experiments (DOE) techniques have been developed to identify sampling principles that guide sample collection within the parameter space. In general, these sampling techniques can be classified as either one-shot or adaptive. One-shot sampling approaches, including Latin hypercube design (LHD) [5] [6], orthogonal arrays [7], or uniform grid sampling, distribute the set of samples across the sample space prior to sample collection. The one-shot techniques typically prioritize exploration of the parameter space to ensure the parameter space is evenly covered by distributing samples far away from each other. Conversely, adaptive sampling improves the sampling process by balancing exploration of the parameter space with exploitation of the response information from previously selected samples. Numerous exploitation approaches, which can be classified as gradient-, committee-, variance-, or cross-validation-based methods [8], have been developed to identify regions of interest (ROI) where additional sampling would best resolve the complex functional behavior.

Gradient-based adaptive sampling uses the gradient or Hessian of the response to identify regions of prominent nonlinearity that are difficult to model. The Hessian matrix has been employed by the local linear approximation (LOLA)-Voronoi approach [9] and Mackman [10]. Committee-based strategies [11] [12] use a combination of different surrogate model approaches, e.g., multivariate polynomials, support vector regression [13], kriging [14], radial basis functions [15] [16], etc., to build separate surrogate models and select a new sample where the models disagree the most. Variance-based adaptive sampling techniques employ model prediction variance or mean square error (MSE) to identify ROI where the model suffers from high uncertainty. Examples include sampling procedures based on the maximized mean square error (MMSE) [17]; the integrated mean square error (IMSE) [18]; the maximum entropy (ME) criterion [19]; and the expected improvement (EI) function [20]. Finally, the cross-validation (CV) approaches divide a sample set into training and testing subsets, to respectively build and evaluate a model. Additional samples are then placed in the region with the largest CV error. The CV approach has been used by Aute et al. [1], Li et al. [21], and Xu et al. [22]. By balancing exploration with these exploitation schemes, samples selected via adaptive sampling techniques typically provide surrogate models of improved accuracy with fewer samples [9].

Within the literature, the Voronoi subtype of adaptive sampling approaches (i.e., LOLA-Voronoi [9] and CV-Voronoi [22]) have enjoyed considerable popularity due to their salient sample quality and relative ease of implementation [8] [9] [22] [23]. The Voronoi methods divide the domain into Voronoi cells and then use a metric, (the hybrid score for LOLA-Voronoi and cross-validation error for CV-Voronoi) to select which cell the next sample should be placed in. The sample is then chosen as the point within the selected cell furthest from the existing sample. Both of these approaches have been demonstrated for low-dimension and relatively small data sizes; however, as the dimension and sample set size increase the computational cost associated with selecting the next sample increases dramatically. As a result, cost reduction techniques are needed. For the LOLA-Voronoi technique, the sampling cost increases due to the LOLA operation. Van Der Herten et al. used fuzzy gradient estimation to reduce the cost of the LOLA in their fuzzy-LOLA (FLOLA)-Voronoi algorithm [24] [25]. A similar cost reduction for the CV-Voronoi adaptive sampling technique is highly-desirable, because the CV-Voronoi adaptive sampling technique has been demonstrated to provide samples that led to more accurate surrogate models [22] [26].

The goal of the present work is to maintain the sample quality provided by the CV-Voronoi adaptive sampling technique while reducing the cost associated with identifying future sample locations. To achieve this goal, a new K-fold cross-validation (KFCV)-Voronoi scheme is developed. In contrast to the current CV-Voronoi approach, the novelty of our KFCV-Voronoi approach lies in the addition of a global pre-filter that identifies promising subsets of samples to evaluate with LOOCV (rather than evaluating all samples) to identify the next sample. The KFCV-Voronoi significantly accelerates the adaptive sampling process while providing samples that form surrogate models of accuracy comparable to the CV-Voronoi technique. A second cost reduction method in the form of a novel Voronoi batch sampling method is also developed. The Voronoi batch sampling utilizes a re-tessellation process to tackle the issue of sample clustering within batches, and is demonstrated to outperform the naïve batch sampling approach. This further decreases the computational cost by reducing the number of calls to the adaptive sampling techniques and enabling testing resources to be used in parallel for enhanced utilization. Complex and high-dimension parameter spaces can necessitate the combination of sampling cost reduction techniques.

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 and methods to evaluate the sampling techniques; Section IV demonstrates the efficacy of the adaptive sampling techniques relative to the current state-of-the-art; and Section V summarizes our conclusions.

II. Surrogate Modeling

In this work, surrogate modeling was performed using the kriging technique as implemented in the MATLAB Design and Analysis of Computer Experiments (DACE) toolbox [27]. Kriging is an interpolation method that models a response via a Gaussian process governed by prior covariances. Compared to other interpolation schemes, kriging provides high accuracy from a smaller database size.

Kriging given a set of n sample sites $\boldsymbol{s}_i \in \mathbb{R}^m$:

$$\boldsymbol{S} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_n]^T, \tag{1}$$

and responses:

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

predicts the response at an unsampled point by:

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

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

$$\boldsymbol{F}(\boldsymbol{x}) = \left| f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_p(\boldsymbol{x}) \right|^T. \tag{4}$$

For the set of samples $\boldsymbol{S}$ the $n \times p$ design matrix, $\boldsymbol{F}$, can be built by substituting each sample into $\boldsymbol{f}$, viz., $\boldsymbol{F}_{ij} = \boldsymbol{f}_j(\boldsymbol{s}_i)$. The coefficients $\boldsymbol{\beta}^*$ are determined by solving the generalized least squares solution (with respect to $\boldsymbol{R}$):

$$\boldsymbol{\beta}^* = (\boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{Y} \tag{5}$$

where $\boldsymbol{R}$ is the matrix of stochastic-process correlations between sample points. The second term of Eq. (3) is dependent on the vector of correlations between sampled points $\boldsymbol{S}$ and an unsampled point $\boldsymbol{x}$:

$$\boldsymbol{r}(\boldsymbol{x}) = [\mathcal{R}(\boldsymbol{\theta}, \boldsymbol{s}_1, \boldsymbol{x}), \ldots, \mathcal{R}(\boldsymbol{\theta}, \boldsymbol{s}_n, \boldsymbol{x})]^T. \tag{6}$$

Many choices of correlation forms are possible including exponential, spherical, spline, and Gaussian which was utilized within this work:

$$\mathcal{R}(\boldsymbol{\theta}, \boldsymbol{s}, \boldsymbol{x}) = \prod_{j=1}^{n} \mathcal{R}_j(\boldsymbol{\theta}, \boldsymbol{s}_j - \boldsymbol{x}_j) \tag{7}$$

where

$$\mathcal{R}_j(\boldsymbol{\theta}, \boldsymbol{s}_j - \boldsymbol{x}_j) = \exp\left(-\boldsymbol{\theta}_j (\boldsymbol{s}_j - \boldsymbol{x}_j)^2\right). \tag{8}$$

Through combination of the regression model and random function (stochastic process) the kriging method models the response [27]. Though kriging interpolation was used in this work, the sampling techniques examined here are extendable to other surrogate modeling approaches, and an in-depth examination of prominent surrogate modeling methods can be found in the work of Wang and Shan [4].

III. Adaptive Sampling Methodology

The accuracy of these models is dependent on the samples used to create them, so proper sample selection is key. The focus of this work is to reduce the cost associated with the adaptive selection of optimal samples 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. Additionally, a new Voronoi batch sampling method that reduces sample redundancy is detailed.

A. Cross-Validation Voronoi Adaptive Sampling Method

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. [22], selects samples in regions with the largest prediction error. This is achieved by breaking up the domain into Voronoi tessellations surrounding each sample, where a Voronoi cell corresponds to the region that is closer to a particular sample than any other sample. Figure 1 presents an example of Voronoi tessellations.

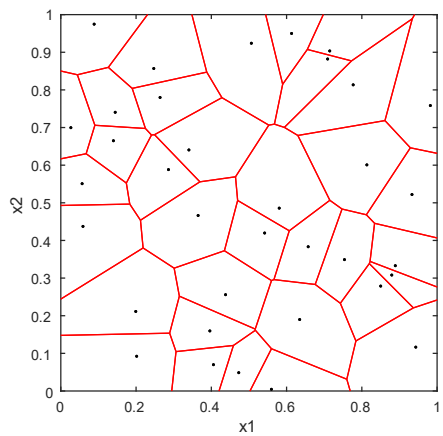


Figure 1. Voronoi cells consist of the region that is closer to a given sample than any other sample

The error associated with each cell is evaluated using the 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:

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

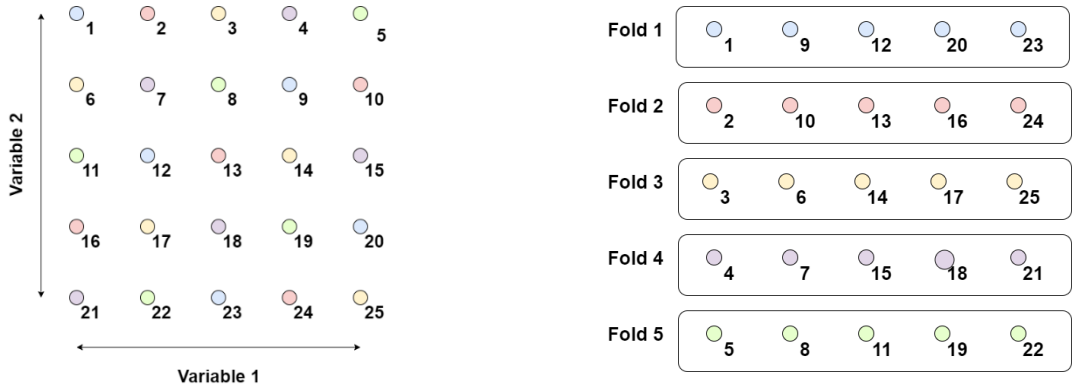
Here, $y(s_i)$ is the measured response at s_i, and $\hat{y}_{S\setminus s_i}(s_i)$ is the response predicted at s_i by the surrogate model constructed without s_i . The new sample location is found by repeating this process for all of the cells and identifying the Voronoi cell with the largest prediction 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 [22] [23] [26]. Unfortunately, the cost of the CV-Voronoi adaptive sampling technique scales superlinearly with number of samples; because, (1) the LOOCV metric requires creation of a separate surrogate model for each sample and (2) the cost of creating a surrogate model for a given data set increases with the number of samples. The kriging metamodeling approach used in this work suffers from a cubic time complexity $O(n^3)$ with the number of samples [28]. Resolution of complex or high dimension parameter spaces typically requires more samples, and sample selection via CV-Voronoi adaptive sampling procedure can become cost prohibitive.

B. K-fold Cross-Validation Voronoi Adaptive Sampling Method

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 surrogate models built during the sampling process. KFCV breaks the sample set into K randomly selected subsets. Figure 2(a) shows this for a two-dimensional parameter space with 25

samples divided into 5 folds (for this example $K = 5$). The fold assignment is portrayed by the sample color, and each fold is shown in Figure 2(b).



(a) (b)

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

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

where n_K is the number of samples per fold, $y(s_j)$ is the true response at s_j, and $\hat{y}_{S \setminus kf_i}(s_j)$ 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. It is posited that the folds with the largest error contain the samples that have the strongest 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 existing 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 reduced 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
3. Select the l folds with the largest K-fold cross-validation errors
4. Use region of interest-leave-one-out cross-validation to evaluate each sample within the l folds
5. Identify the sample with the highest leave-one-out cross-validation error
6. Select the point within this sample's Voronoi cell that is furthest from the existing sample

C. Voronoi Batch Sampling Procedure

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 [9] [22] [23] the Voronoi type adaptive sampling techniques have only been presented selecting a single point a time. A naïve batch sampling approach for the Voronoi methods would place a sample within the n Voronoi cells with the largest error metric (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, and depending upon the shape of the Voronoi cells the resulting samples can be redundant as shown in Figure 3.

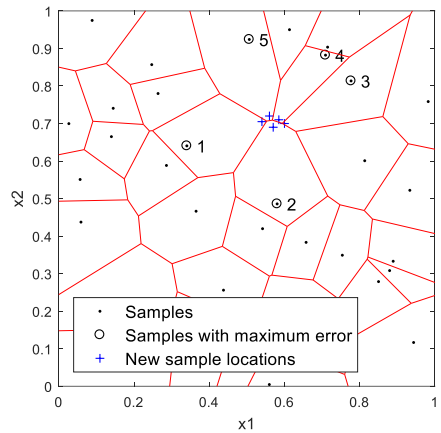


Figure 3. Future samples from a naïve batch sampling approach.

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 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 Figure 4(a). The Voronoi regions are then reassigned, taking the new point into account, as illustrated in Figure 4(b). The second sample is then selected within the updated Voronoi region associated with the second highest score, as demonstrated in Figure 4(c). This procedure is then repeated until the number of samples in the batch is reached. Figure 4(d) presents the resulting 5-sample batch. Comparing the sample distributions of the naïve approach in Figure 3 with the new Voronoi batch sampling approach in Figure 4(d) it is clear that the latter provides a sample set that is less likely to provide redundant information for surrogate model creation. This is because the updated Voronoi tessellations force the new samples to be placed in extremum regions that are as far as possible from existing samples and other samples within the new batch.

The Voronoi batch sampling approach reduces number of calls to the adaptive sampling procedure. This reduces the number of kriging models that must be created during the adaptive sampling process from:

$$N_{Models} = \sum_{j=N_i}^{N_{max}} \left(K + l\,\frac{j}{K}\right) \tag{11}$$

to

$$N_{Models} = \sum_{j=N_i}^{N_{max}} \left(K + l\,\frac{j}{K}\right) \text{ for } j = N_i, N_i + n, N_i + 2n\,, \dots\,, N_{max} \tag{12}$$

where N_i is the initial number samples, N_{max} is the maximum numer of samples, K is the number of folds, l is the number of folds LOOCV is performed for, and n is the batch size. As $N_{max} \to \infty$, the number of kriging models required for adaptive sampling using batches becomes:

$$N_{Models} = \frac{\sum_{j=N_i}^{N_{max}} \left(K + l\,\frac{j}{K}\right)}{n}. \tag{13}$$

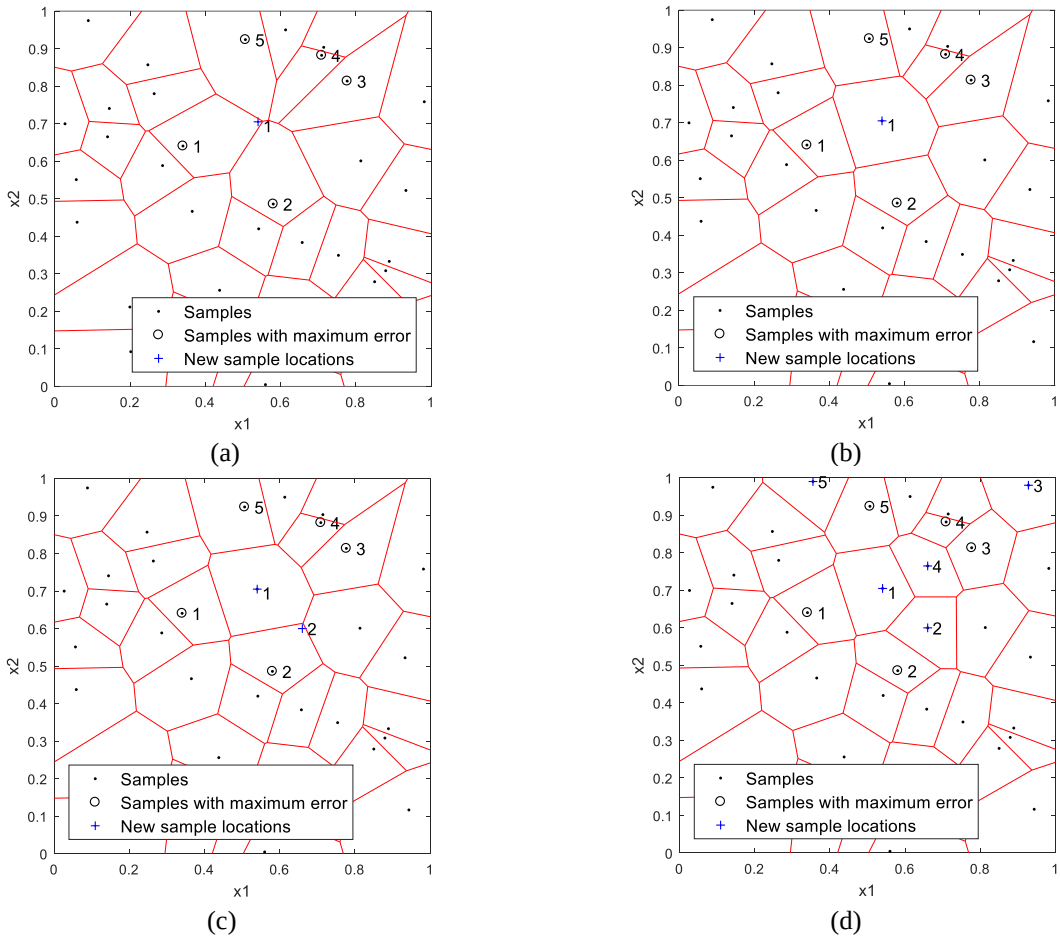


Figure 4. The Voronoi batch sampling procedure: (a) the first sample is identified, (b) the Voronoi cells are reassigned, (c) the second sample is selected within the new Voronoi cells, (d) the procedure is repeated until the full batch is found.

D. Surrogate Model Evaluation Method

In the following section, the KFCV-Voronoi and Voronoi batch sampling methods will be compared with popular design of experiment methods from the literature. The primary goal of the new sampling techniques is to select samples in a cost effective manner that construct surrogate models of improved accuracy for a given sample quantity. To assess the adaptive sampling techniques, we use the normalized root mean square error (NRMSE) [28]:

$$NRMSE = \sqrt{\frac{\sum_{i=1}^{t}\left(y(q_i) - \hat{y}(q_i)\right)^2}{\sum_{i=1}^{t}\left(y(q_i)\right)^2}} \tag{14}$$

to assess the prediction accuracy of a given surrogate model. Here y is the true response, $\hat{y}$ is the response predicted by the surrogate model, and t represents a number of test points, $Q = [q_1,\ q_2, \dots, q_t]$ that have not been used in the surrogate model construction. As t increases, the NRMSE gives an increasingly accurate evaluation of the surrogate response surface. The techniques will also be evaluated based on the sample selection cost.

IV. Evaluation of the KFCV-Voronoi and Voronoi Batch Sampling Method

Evaluation of the proposed KFCV-Voronoi and Voronoi batch sampling for efficient design of experiment to construct global surrogate models is performed in this section. The evaluation is first conducted on four analytical functions [29] of varied dimensionality and complexity. After evaluating the performance for benchmark functions, the methods are applied to resolving aerodynamic loading coefficients of a flight space of interest. To provide a fair

evaluation of the proposed methods a number of steps have been taken: (1) the proposed adaptive sampling techniques are compared with a number of methods from the literature; (2) the kriging method is used for construction of global surrogate models for all sampling approaches; (3) the same initial sample sets are used for the adaptive techniques; (4) the validation points are the same for each sampling method; (5) a large number of validation points ($t = 500{,}000$) are used to evaluate the models; (6) results are based on 10 repeated applications of each adaptive sampling method; (7) randomization is performed using a Mersenne Twister pseudorandom number generator to provide randomization that can be seeded to generate reproducible results; (8) the same seeds are used for evaluation of all methods; and (9) the methods are evaluated for five different response surfaces.

A. Analytical Test Functions

In this section, the four analytical test functions are used to evaluate performance of the adaptive sampling methods, which include

2-D Peaks function:

$$y = 3(1-x_1)^2 e^{\left(-x_1^2-(x_2+1)^2\right)} - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right) e^{\left(-x_1^2-x_2^2\right)} - \frac{1}{3} e^{\left(-(x_1+1)^2-x_2^2\right)} \text{ for } x_{1,2} \in [-5,5], \tag{15}$$

2-D Shubert function:

$$y = \left(\sum_{i=1}^{5} i\cos\big((i+1)x_1 + i\big)\right) \times \left(\sum_{i=1}^{5} i\cos\big((i+1)x_2 + i\big)\right) \text{ for } x_{1,2} \in [1,3]\,, \tag{16}$$

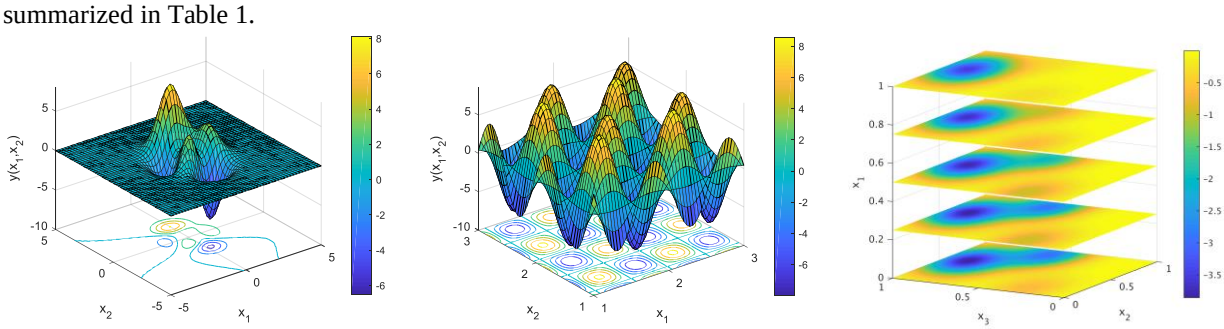
3-D Hartmann function is a highly nonlinear function with three inputs and is given by:

$$y = -\sum_{i=1}^{4} c_i e^{\left|-\sum_{j=1}^{3} a_{ij}\left(x_j - p_{ij}\right)^2\right|}, x_j \in [0,1], j = 1,2,3, \tag{17}$$

6-D Hartmann function, which is a highly nonlinear function with six inputs given by:

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

The response surface of the lower dimension functions are shown in Figure 5, and the behavior of the functions is summarized in Table 1.



(a) Peaks (b) Shubert (c) 3D Hartmann

Figure 5. Response surfaces of the analytical benchmark functions (a) Peaks, (b) Shubert, and (c) 3D Hartmann.

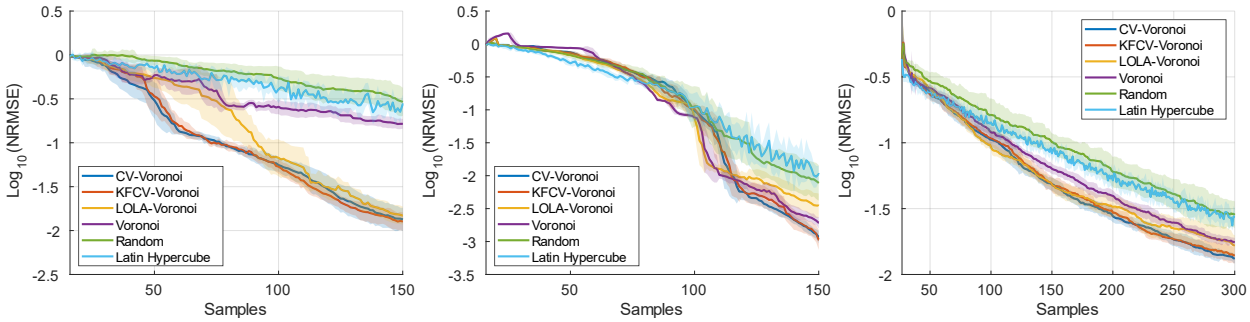
Table 1. Benchmark function behavior description

Function	Range	Description
Peaks	$[-5,5]^2$	Non-linear behavior near the origin but constant away from the origin
Shubert	$[1,3]^2$	Uniform oscillations over parameter space
3D Hartmann	$[0,1]^3$	Three dimensions with significant nonlinearity
6D Hartmann	$[0,1]^6$	Six dimensions with significant nonlinearity

The selected analytic benchmark functions were drawn from the design of experiments literature to cover a variety of behaviors, including varying modality, nonlinearity, irregularity, and dimension.

B. Low-Dimension Performance Assessment

We begin our analysis using the Peaks, Shubert, and 3D Hartmann benchmark functions. For each case we compare the performance of the proposed techniques with samples selected via (1) maximin optimized Latin hypercube sampling, (2) random sampling, (3) Voronoi sampling, (4) LOLA-Voronoi sampling [9], and (5) CV-Voronoi sampling [22]. The Latin hypercube is one of the most commonly used oneshot methods, and the Voronoi sampling approach is a popular sequential sampling approach. Finally, the LOLA-Voronoi and CV-Voronoi methods are popular adaptive sampling techniques that have been demonstrated to perform well for constructing global surrogate models outperforming space-filling cross-validation tradeoff (SFCVT) [1], Monte-Carlo intersite projection (MIPT) [30], mean squared error (MSE) [17], and Accumulate error (ACE) [21], sampling methods within the literature [9] [22]. The adaptive approaches are initialized from a 4 by 4 uniform grid for the 2D functions and a 3 by 3 by 3 uniform grid for the 3D function, and samples are selected one at a time. A comparison of the NRMSE for each sampling method is presented for the three benchmark functions in Figure 6. The solid line represents the mean NRMSE error and the shaded regions signify the standard deviation for each sampling method.



(a) Peaks (b) Shubert (c) 3D Hartmann

Figure 6. NRMSE history as samples are added for the analytical benchmark functions (a) Peaks, (b) Shubert, and (c) 3D Hartmann. (The solid line is the mean and shaded region is the standard deviation)

First examine the NRMSE histories for the Peaks function. The adaptive sampling procedures (LOLA-, CV-, and KFCV-Voronoi) dramatically outperform the sequential (Voronoi and random) and one-shot (Latin hypercube) methods. This is expected, because the Peaks function represents the ideal scenario for demonstrating the exploitative 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, rapidly varying region near the origin is more difficult to model, requiring more samples to accurately approximate this region. Exploitative adaptive sampling techniques should outperform one-shot sampling techniques by concentrating samples near the origin where the response is harder to model. The CV- and KFCV-Voronoi approaches outperform the other adaptive sampling approaches. Though the two methods performed similarly, the CV-Voronoi performed slightly better initially (< 55 samples) and the KFCV-Voronoi approach performed better for larger sample quantities (> 55 samples). It should be noted that the NRMSE is presented on a logarithmic scale, therefore even small differences can represent a significant difference in model accuracy.

The phenomenon of the KFCV-Voronoi technique surpassing the CV-Voronoi technique as the number of samples increases occurs frequently within the results but was actually an unexpected benefit. The KFCV-Voronoi technique was initially devised solely as a method to reduce the computational cost associated with selecting the next sample location while maintaining the quality of the CV-Voronoi selected samples. It was expected that models created using the KFCV-Voronoi adaptive sampling would provide models similar to but inferior to those built using CV-Voronoi technique, because the KFCV-Voronoi could potentially misidentify the best Voronoi cell to sample (i.e., the single one with the worst error) if it was lumped with relatively unimportant points. However, we will see in the results that the accuracy of the models built using the samples selected through the KFCV-Voronoi adaptive sampling technique often surpasses the accuracy of models built from the CV-Voronoi samples. We believe that the improved sample selection is due to the addition of a global filtering effect caused by the KFCV. The KFCV error in Eq. 10 represents the aggregate impact of the samples of each fold at the fold sample locations. This incorporates a distributed (or global) impact of the samples prior to considering the local impact via LOOCV. The KFCV-Voronoi approach is benefitted by taking both the global and local impact into account. However, as we mentioned earlier employing the KFCV

global filter also comes with the chance of a poor fold division where the best sample location is diluted by being lumped with unimportant points. Therefore, the KFCV approach must balance the chance of poor fold division, with the benefit from knowledge of global effects to outperform the CV-Voronoi technique. Fortunately the impact of poor fold division can be limited by randomly partitioning the samples independently during each sampling cycle. Also, the KFCV-Voronoi sampling technique is sequentially identifying points of interest so as the sampling procedure progresses it is statistically more likely that the regions of interest prone to large model errors will be present in all folds and there is a high-probability of capturing the best point for sample refinement.

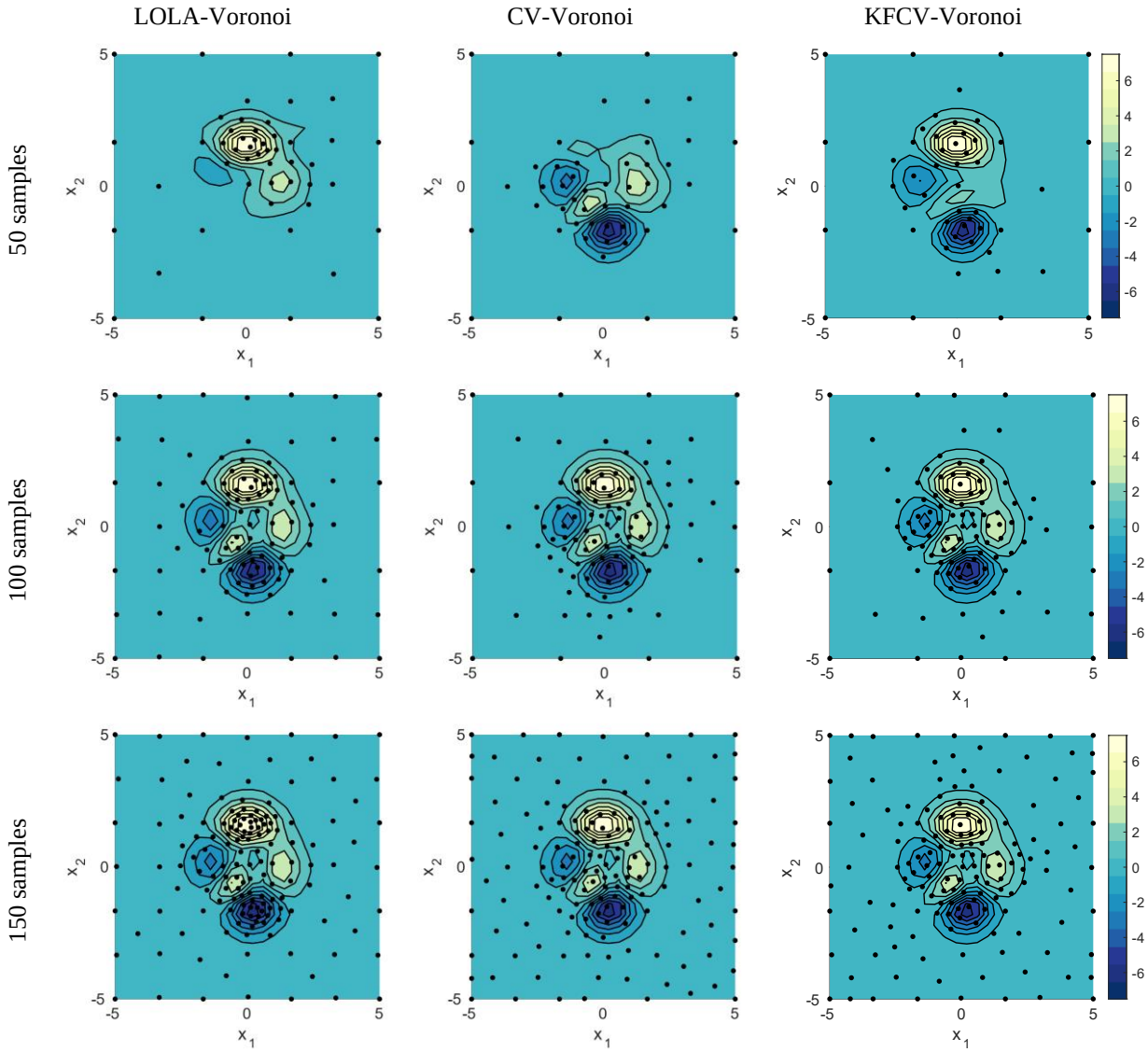


Figure 7. Sample distributions for the adaptive sampling techniques applied to the Peaks function at different sample counts.

Examination of the sample distribution in Figure 7 for each sampling method during the sampling procedure provides insight into why the NRMSE 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 100 and 150 samples. This is remarkable because the KFCV-Voronoi and CV-Voronoi follow different paths to these distributions. 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.

Next, we consider the NRMSE histories for the Shubert function shown in Figure 6(b). The Shubert function features uniform oscillations (shown in Figure 5(b)) which will emphasize the importance of exploratory sampling, because if a single peak or valley is missed, a large NRMSE will be observed. Therefore, in addition to leveraging the exploitative sampling, the adaptive sampling procedure must also feature exploratory sampling. Since exploitative features are less useful for this function, we see that the explorative and one-shot techniques perform more competitively with the adaptive sampling methods. It is interesting to note that the LOLA-Voronoi and Voronoi sampling techniques initially outperform the other sampling techniques. This is likely due to the prioritization of exploration-based behavior capturing all of the peaks and valleys first. LOLA-Voronoi was shown to initially prioritize exploration more in the Peaks function as illustrated in Figure 7. The KFCV- and CV-Voronoi overtake these sampling methods 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.

Finally, the 3D Hartmann function is a highly nonlinear function (shown in Figure 5(c)) with three inputs. Considering the NRMSE histories for the 3D Hartmann function shown in Figure 6(c), we see that the adaptive sampling techniques again outperform the one-shot approaches. The best performance alternates between the KFCV- and CV-Voronoi adaptive sampling techniques. It also should be highlighted that in general as the dimension increases the number of samples required to reach a given accuracy also increases. We again examine the sample distributions of each sampling method at various points during the sampling of the 3D Hartmann function in Figure 8.

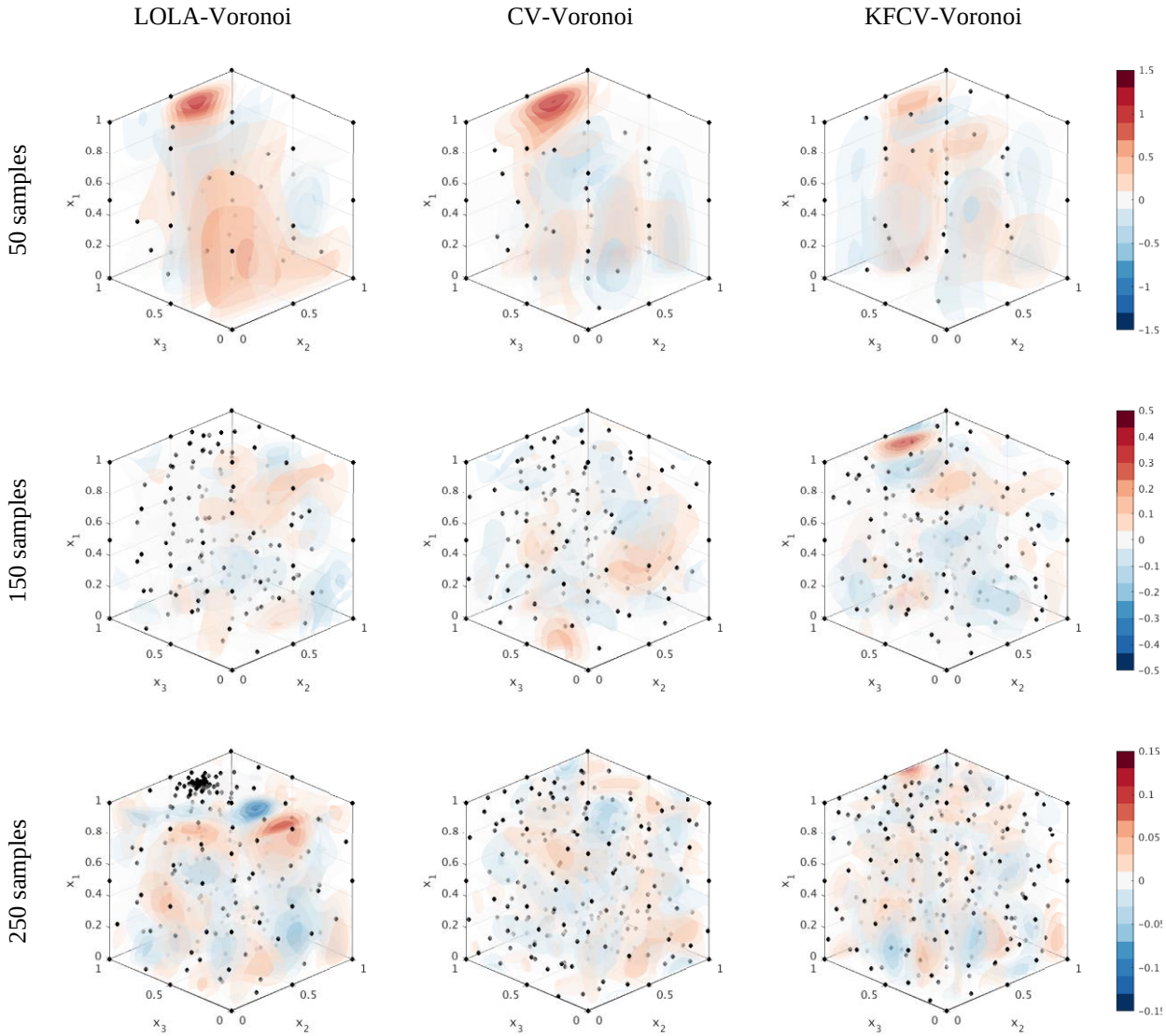
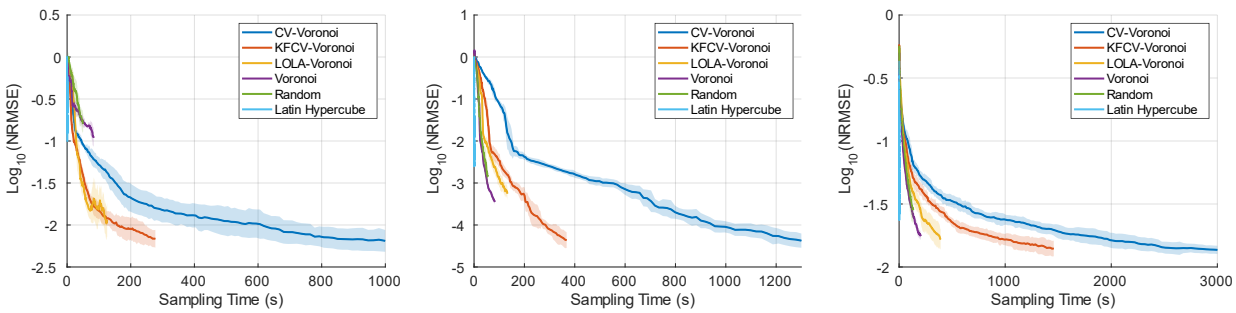


Figure 8. Sample distributions and isosurfaces colored by absolute error for the adaptive sampling techniques at different sample counts for the 3D Hartmann function.

Examination of the sample distribution of each sampling method during the sampling procedure presents trends similar to those found for the Peaks function. The LOLA-Voronoi is more explorative initially, but overly exploitative later, and the KFCV-Voronoi again appears to be slightly more explorative than the CV-Voronoi initially. Examining the error magnitudes, we see that the best model is made from the KFCV-Voronoi samples at 50 samples and from the LOLA-Voronoi samples at 100 samples. At 150 samples the errors are almost identical for the CV-Voronoi and KFCV-Voronoi approaches.

Thus far, the analysis has solely judged the adaptive sampling techniques based on the accuracy of the model created from the selected samples, but the cost of selecting those points is also a factor. The computational cost of evaluating the next sample is not constant between adaptive sampling approaches. As expressed earlier the KFCV-Voronoi technique was originally developed only with computational cost reduction in mind. To examine the relative costs of each sampling method, a comparison of the wall clock time required to achieve a specified NRMSE is included in Figure 9 for the three low-dimension analytic benchmark functions.



(a) Peaks (b) Shubert (c) 3D Hartmann

Figure 9. The NRMSE of each sample selection method and the accumulated time required for selecting those samples on the analytical benchmark functions (a) Peaks, (b) Shubert, and (c) 3D Hartmann. (The solid line is the mean and shaded region is the standard deviation)

It can be seen that KFCV-Voronoi algorithm proposed here maintains the sample quality of the CV-Voronoi technique, but significantly reduces the computational time associated with selecting those samples. The KFCV-Voronoi requires only one third of the computational time when compared to the CV-Voronoi technique. The cost ratio could be further reduced if LOOCV was performed for fewer folds, but we have found performing LOOCV for 3 out of 10 folds has provided good results (i.e. $l = 3$ and $K = 10$). The Latin hypercube method is the cheapest O(1s) because samples only need to be generated once. The Voronoi sequential sampling is only slight more expensive because the sample selection does not consider the response, but only the sample distribution within the parameter space. Next, the accelerated neighborhood LOLA-Voronoi technique (presented in [26]) features a lower sampling costs because only a single kriging model is created during each sampling cycle.

It should be noted that in Figure 9, the time associated with sampling is primary component of the total time because the time associated with evaluating the samples for the benchmark function is small. In most real-world problems, the cost of evaluating the sample is dramatically larger so limiting the number samples required to build a model of a desired accuracy is important. We illustrated in Figure 6 that the CV- and KFCV-Voronoi methods provide samples that create surrogate models of improved accuracy (on a per sample basis). Therefore, though the sampling time is less for the LOLA-Voronoi, Voronoi, LHS, and random sampling is smaller, fewer samples are required to reach $\mathrm{Log}_{10}(\mathrm{NRMSE}) = -1.5$ with the CV- and KFCV-Voronoi as shown in Table 2 for the Hartman 3D function. Therefore, choosing the best adaptive sampling method is dependent upon the relative cost of the sampling process and response evaluation process. KFCV-Voronoi provides high quality samples at low cost.

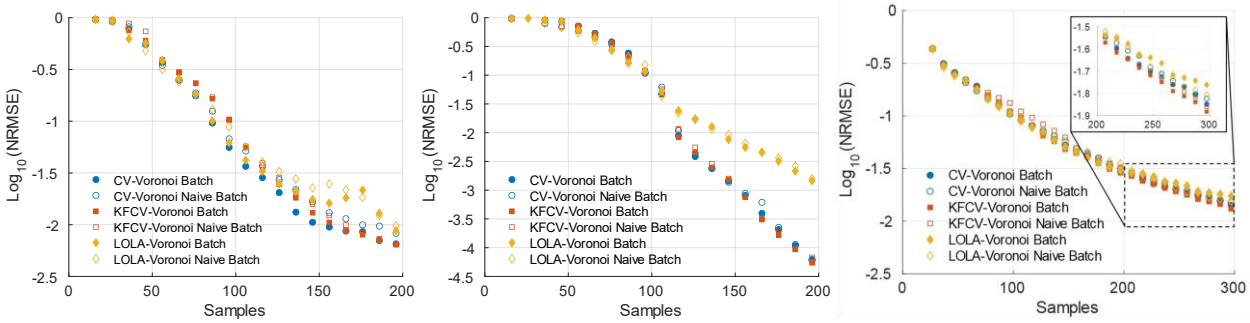
Table 2 Sample counts required to reach $Log_{10}(NRMSE) = -1.5$ for the Hartman 3D function.

Method	CV-Voronoi	KFCV-Voronoi	LOLA-Voronoi	Voronoi	Random	Latin hypercube
Samples	163	167	183	194	253	248

C. Voronoi Batch Sampling Performance Assessment

We also used the Peaks, Shubert, and 3D Hartmann benchmark functions to evaluate selecting samples in batches rather than one at a time. Batch sampling allows us to evaluate multiple responses in parallel, and make better use of

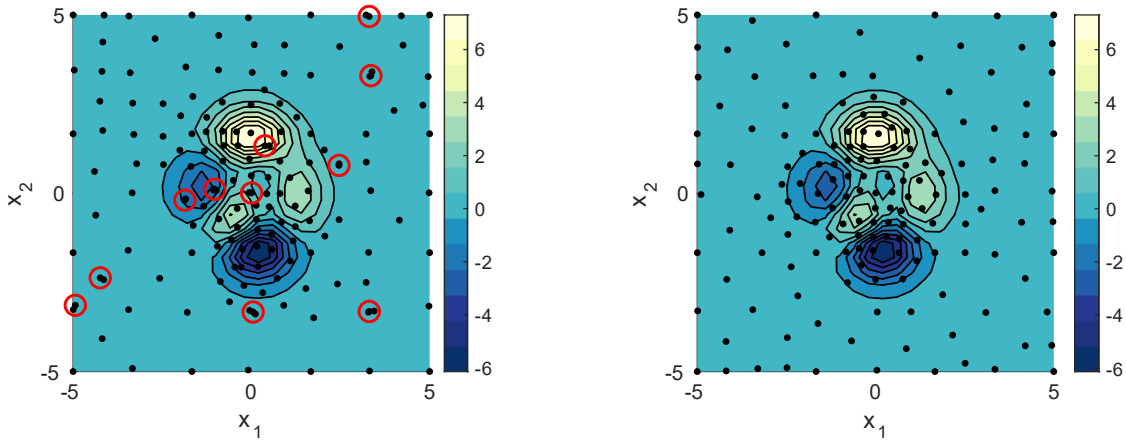
testing resources. It also reduces the number of times the adaptive sampling process must be performed, and as a result the overall sample selection cost. The NRMSE history of the Voronoi batch sampling technique proposed here is compared with the naïve batch sampling technique for the three adaptive sampling techniques applied to the three benchmark functions in Figure 10.



(a) Peaks (b) Shubert (c) 3D Hartmann

Figure 10. The NRMSE history as samples are added in batches of 10 for the analytical benchmark functions (a) Peaks, (b) Shubert, and (c) 3D Hartmann. (Filled symbols signify the proposed Voronoi batch sampling while empty symbols signify the naïve approach.)

It should be noted, the proposed Voronoi batch sampling method has been evaluated using each of the Voronoi adaptive sampling techniques, because it was developed to be generalizable. Examining the figures, it is clear that the new Voronoi batch sampling approach leads to better samples. Improvement is more apparent for the Peaks and 3D Hartmann functions where models of equivalent accuracy are created from CV- and KFCV-Voronoi using one or two less batches. Less improvement is seen for the Shubert function, but this was expected. The Voronoi batch sampling utilizes a re-tessellation process to tackle the issue of sample clustering within batches, but the uniform modality of the Shubert function encourages the selected points to be spread out and explorative. Therefore, sample over-clustering is not as likely. For functions where the sampling is more exploitative the new Voronoi batch approach prevents the overclustering that occurs in the naïve approach. To illustrate this, a comparison of the sample distributions for the Peaks function generated using the KFCV-Voronoi adaptive sampling approach with each of the batch selection approaches is included in Figure 11.



(a) Naïve Batch Approach (b) Proposed Voronoi Batch Approach

Figure 11. Sample distributions for 156 samples selected via (a) the naïve batch approach and (b) the proposed Voronoi batch approach. (Red circles indicate redundant over-clustered samples)

The distribution of samples selected via the naïve batch approach (Figure 11(a)) contains numerous examples of sample over-clustering where redundant information is being acquired. We see that the over-clustering is mitigated within the distribution of samples selected with the proposed Voronoi batch sampling method (Figure 11(b)). This mitigation can lead to considerably improved model accuracy for a given number of samples. This is evidenced by the NRMSE history illustrated in Figure 12 for the two sample sets shown above.

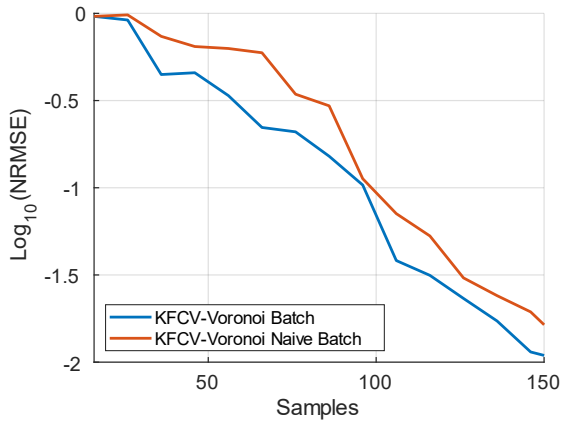
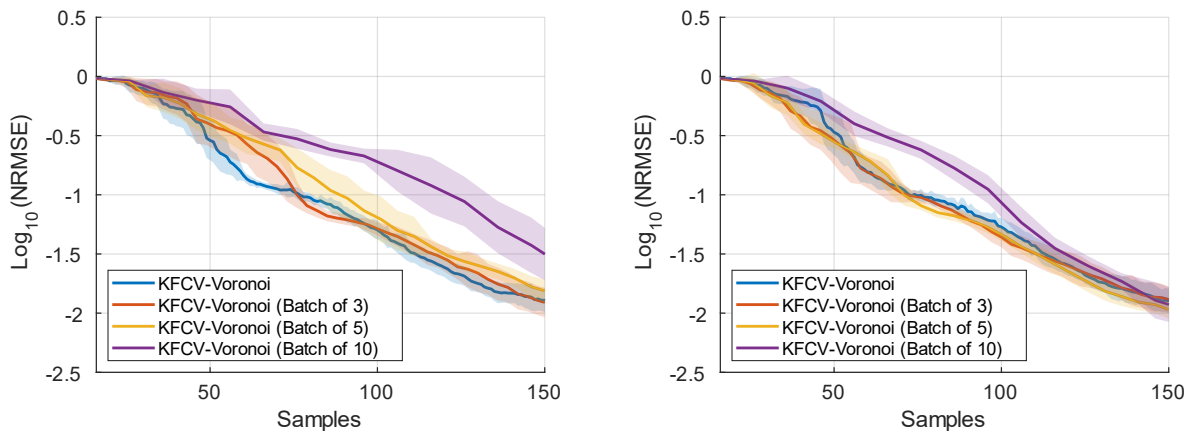


Figure 12. NRMSE for samples selected with the KFCV-Voronoi sampling method with different batch approaches.

We also take this opportunity to highlight proper coupling of the KFCV-Voronoi method and the batch sampling approaches. Figure 13 presents the NRMSE history for the KFCV-Voronoi based adaptive sampling procedure performing LOOCV for different fold quantities (l) to select samples of batch size (n).



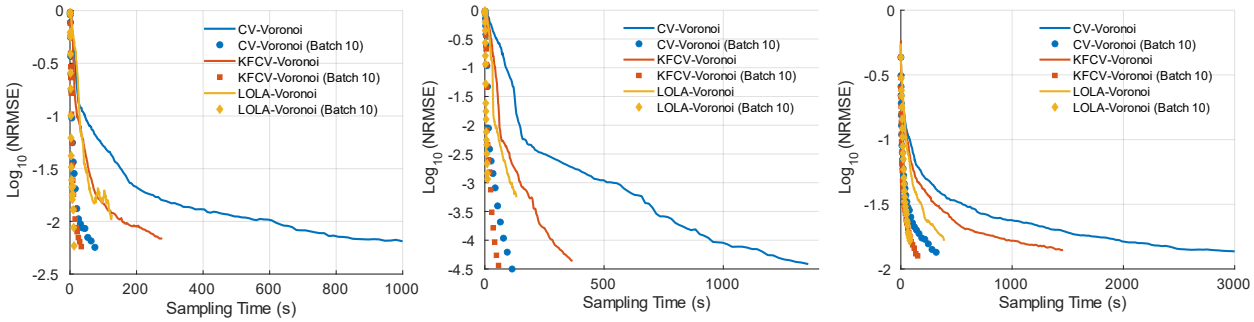
(a) KFCV-Voronoi ($l = 1$) (b) KFCV-Voronoi ($l = 3$)

Figure 13. NRMSE history for surrogate response surfaces built using KFCV-Voronoi applied to the Peaks function with increasing batch sizes and LOOCV fold counts: (a) $l = 1$ fold and (b) $l = 3$ folds.

We can see from Figure 13(a) that as the batch size increases, KFCV-Voronoi with LOOCV for only a single fold (i.e. $K = 10$ and $l = 1$) selects samples poorly. In the KFCV-Voronoi adaptive sampling method KFCV is used to downselect to the l folds with the most promising candidates for individual evaluation. The fold with the largest KFCV error will most likely contain the point with the largest LOOCV error, so this point can be found by performing LOOCV for fewer points. However, when l is small, particularly early in the adaptive sampling process, the number of candidate Voronoi cells is relatively small (N_{samp}/K). As a result, the KFCV-Voronoi method will typically identify the single best location to sample, but additional sample locations may be of diminished quality because it would only be by coincidence that the next best samples would belong to the same fold as the best sample location. KFCV with LOOCV performed for only a single fold therefore is likely to filter out the next best samples for a batch.

This decrease in performance can be improved by performing LOOCV for additional folds, as shown by the 3-fold ($l = 3$) case in Figure 13(b). By increasing the number of folds in which LOOCV is performed, we see that the quality of the samples selected in smaller batch sizes is more comparable to those selected one at a time. Additionally, the larger batch size of 10 samples initially underperforms, but manages to catch up as additional samples are added and the number of samples in each fold increases. Increasing the number of folds LOOCV is performed for increases the cost associated with the sampling method, but this cost is easily justified if the cost to consider the additional folds is less than the cost savings from reducing the number of calls to the adaptive sampling procedure, as can be seen by examining Eqs. 11 and 13. By increasing the batch size to five samples (i.e., $n = 5$), the adaptive sampling procedure

only needs to be called one fifth as frequently as a procedure adding a single point at a time, so as long as l increases by a ratio smaller than n the cost associated with adaptive sampling is reduced. We illustrate that batch sampling can significantly reduce the sampling cost for all the voronoi adaptive sampling approaches in Figure 14.

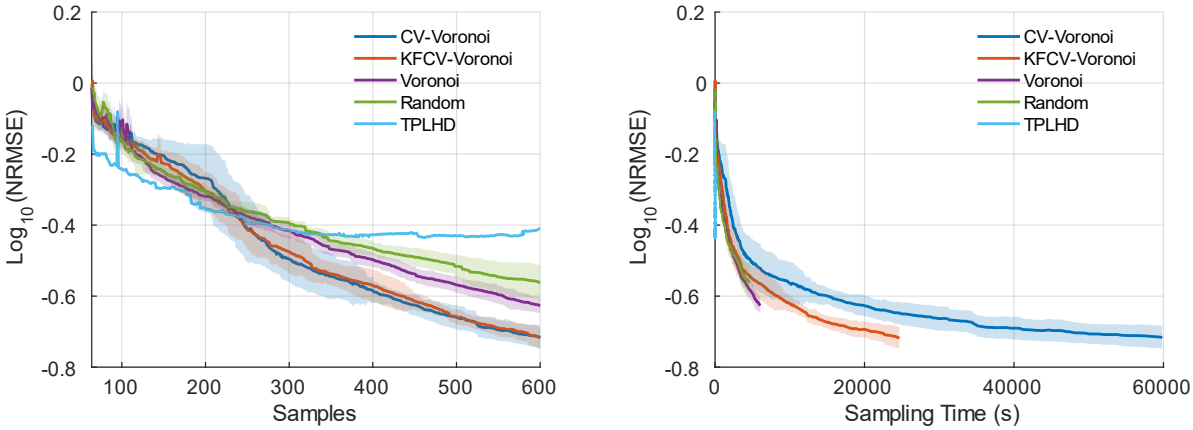


(a) Peaks (b) Shubert (c) 3D Hartmann

Figure 14. The NRMSE of each sample selection method and the accumulated time required for selecting those samples in batches or one at a time on the analytical benchmark functions (a) Peaks, (b) Shubert, and (c) 3D Hartmann.

D. High-Dimension Performance Assessment

In addition to the low-dimension response surfaces, we also evaluated the adaptive sampling for the high-dimension 6D Hartmann analytic function. In the low-dimension cases, the computational cost of the adaptive sampling methods remained affordable, because relatively few points were needed to adequately resolve the response behaviors. Here, the increased dimension of this case requires significantly more sample points to accurately resolve the response. This can be seen by examining the NRMSE histories for each sampling method applied to the 6D Hartmann function in Figure 15(a), where the we began the sampling from 64 points distributed in a 2-point uniform grid of every dimension.



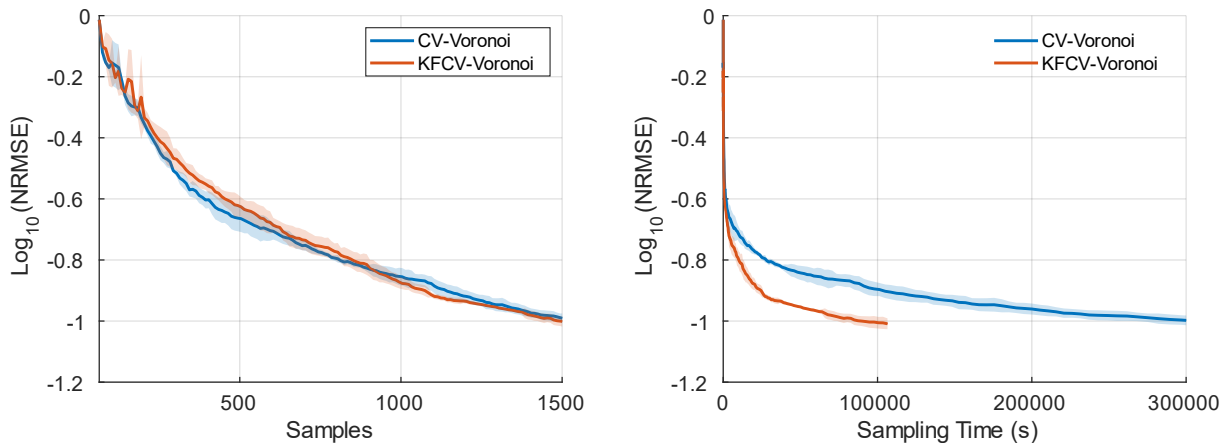
(a) (b)

Figure 15. NRMSE history for the 6D Hartmann analytical benchmark functions examined on (a) a per sample basis and (b) the total sample collection time.

Note the LOLA-Voronoi method was neglected for this case because solving the LOLA operation becomes intractable, and the optimized maximin Latin hypercube sampling was replaced with translational propagation Latin hypercube design (TPLHD) [6] because of its improved performance in high-dimension. Here it can be seen that initially the most accurate models are generated from the TPLHD set, but the model accuracy is surpassed by the models generated via the KFCV- and CV-Voronoi adaptive sampling techniques as additional samples are collected. It is expected that the TPLHD set performs better than the adaptive methods initially; because, TPLHD forms an approximately optimal LHD set, and the adaptive sampling methods are initialized from the 2-point grid which is deliberately suboptimal. It is a testament to the value of the adaptive sampling method that they rapidly overtake the optimal oneshot sets after being initialized from an inferior initial distribution.

It should be noted that there is considerable variation in the computational cost associated with selecting samples using the different approaches. A comparison of the NRMS versus the accumulated computational time required by each sampling method is included in Figure 15(b). It can be seen that the cost associated with adaptive sampling increases with the number of sample points. In this case, the cost is no longer negligible and must be considered when determining the proper sample selection method for the design of experiment. The KFCV-Voronoi provides a great balance between the sample quality and the cost associated with identifying those samples.

We also examined sampling the 6D Hartmann function with batch sampling, because more samples are needed to resolve the response surface and the computational cost of selecting the samples individually was becoming intractable. We again began sampling from an initial 2-point uniform grid of every dimension, and we selected samples in batches of 10. The NRMSE for the batch sampling of the 6D Hartmann function is included in Figure 16(a).



(a) (b)

Figure 16. NRMSE history for the 6D Hartmann analytical benchmark functions examined on (a) a per sample basis and (b) the total sample collection time as samples are added in batches 10.

From the figure, it can be seen that the KFCV- and CV-Voronoi methods perform similarly, although we again see the phenomena of the KFCV-Voronoi technique surpassing the CV-Voronoi technique as the number of samples increases. Selecting samples in batches reduces the associated cost, but the accumulated sampling cost shown in Figure 16(b) is still considerable.

We have also included the CPU time of selecting a batch for increasing sample quantities is included in Figure 17.

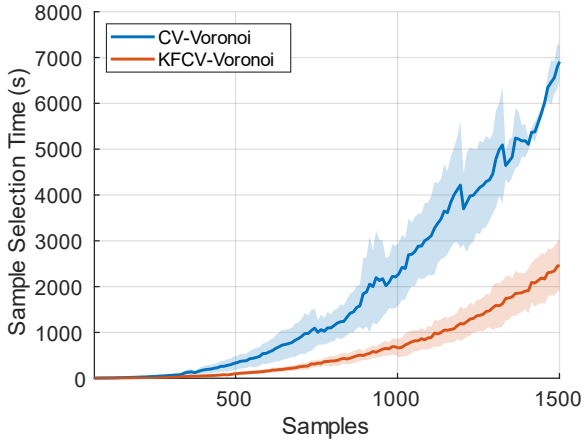


Figure 17. Cost of selected each batch of samples as the number of samples increases.

The sample selection time of the CV-Voronoi is dramically reduced by implementing the KFCV-Voronoi approach. It is important to stress that in real-world problems, the time associated with selecting each batch between experiments or simulations slows down the data collection process which is very undesirable. In some cases, like wind tunnel testing any delay beyond the time required to prepare the next test condition would lead to reduced experimental tests,

since wind tunnel time is often constrained by availability. Parallelization can be used to reduce the wall clock time, but depending upon the facility the user may not have access to HPC resources or be limited to edge level resources. Furthermore, parallelization only reduces the wall clock not the CPU time. Alternatively, the KFCV-Voronoi approach can be used to reduce both CPU and wall clock time to better utilize available resources. If HPC resources are available the wall clock time can be further reduced by combining the KFCV-Voronoi approach with parallelization. The KFCV-Voronoi approach represents a considerable contribution to the design of experiments state of the art.

E. Aerodynamic loading for NACA 00012

In the previous case studies, the capability of each adaptive and batch sampling approach was demonstrated by creating surrogate models of several benchmark functions. Here we extend our consideration to a surrogate model of aerodynamic loading data. An aerodynamic model of this type could be used by a 3DOF or 6DOF simulator for trajectory evaluation. In this case, each sample corresponds to a CFD simulation with a computational cost on the order of 30 minutes with 8 cores. Now that there is a cost associated with collecting each sample there is a true incentive to minimize the number of samples required to resolve the parameter space.

We applied the LOLA-Voronoi, CV-Voronoi, and KFCV-Voronoi adaptive sampling techniques to build surrogate response models for the aerodynamic loading of a NACA 0012 airfoil calculated via CFD simulation. The FUN3D [31] CFD solver was employed to solve the RANS equations for turbulent viscous flow. The lift coefficient was selected as the parameter of interest, and we examined it over a two-dimensional parameter space consisting of freestream Mach numbers from 0.4 to 1.2 and angles of attack between 0° and 16°. To evaluate the adaptive sampling methods the lift coefficient was sampled on a 17×17 uniform grid with spacings of 0.05 in the freestream Mach number and 1° in the angle of attack. These samples are used as the verification data. A surrogate model was built for the 289 uniformly distributed samples to visualize the response over the parameter space in Figure 18.

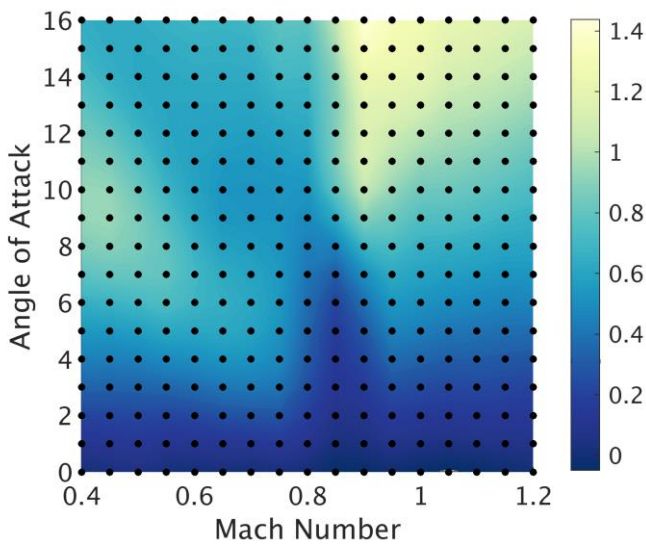


Figure 18. A uniform sampling process can be used to approximate the lift coefficient of a NACA 0012 within the parameter space.

Examining the figure, it is clear that the lift coefficient has regions with steep gradients and changes in gradients, particularly for a freestream Mach number around 0.9. 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.

The LOLA-Voronoi, CV-Voronoi and KFCV-Voronoi adaptive sampling techniques were evaluated for this case and compared with TPLHD oneshot sample sets. The adaptive sampling approaches were initialized from a population of 17 TPLHD samples, and the new Voronoi batch sampling approach was used to collect samples in batches of 5. This process was repeated to reduce the NRMSE of the predicted values at the test points. A plot of the convergence history of the NRMSE for increasing sample quantities is included in Figure 19.

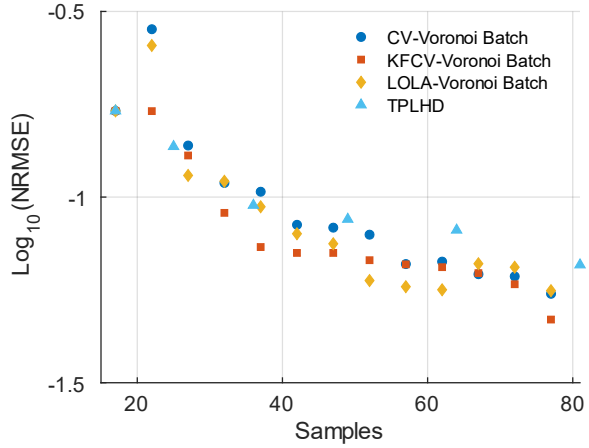
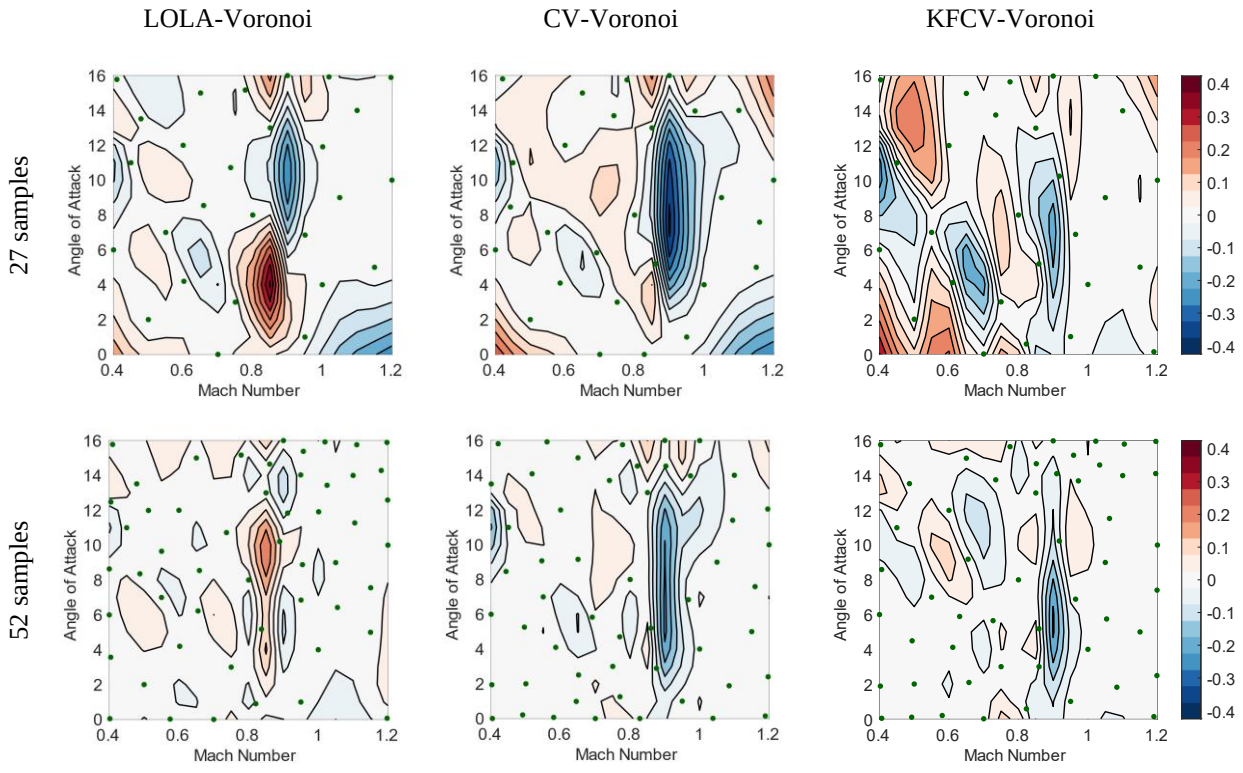


Figure 19. Convergence history for surrogate models of the lift coefficient of a NACA 0012 airfoil for varying freestream Mach numbers and angles of attack created using CV-and KFCV-based adaptive sampling approaches

In the figure, the models built from adaptively selected samples outperform those built using the TPLHD one-shot approach, especially as the number of samples increases. It can also be seen that the KFCV-Voronoi technique provides more accurate surrogate models for the majority of the sample quantities. Only the LOLA-Voronoi selected samples temporarily perform better between sample sets of 52 to 62 samples. In addition to the NRMSE error, the distribution of the error for the surrogate models built using the three adaptive sampling methods is presented in Figure 20. In the figure, the selected samples for each case (shown in green) accompany the contours of the surrogate modeling error.



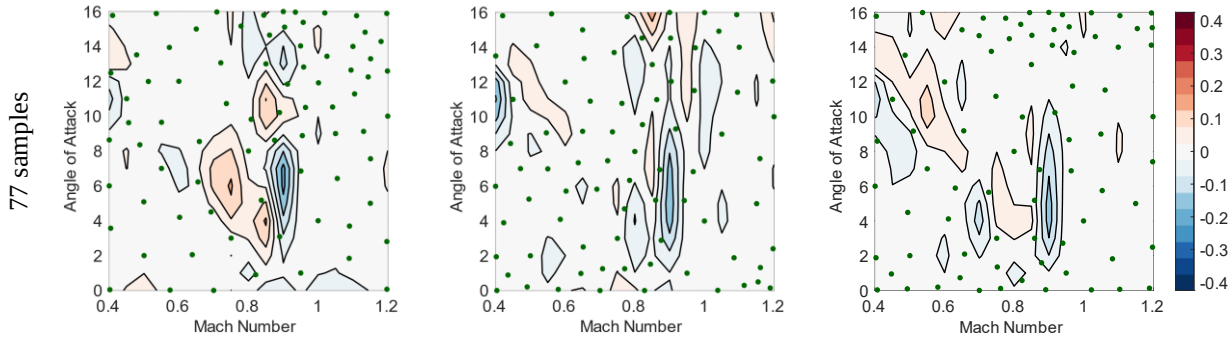


Figure 20. The sample distribution (green dots) of each adaptive sample selection method and contours of the accompanying surrogate model prediction error for the LOLA-Voronoi, CV-Voronoi, and KFCV-Voronoi at different sample quantities.

It can be seen that for all methods the absolute error values get smaller as more samples are added. For the sets of 27 samples, all three sampling approaches create models with large prediction errors in the corners of the design space. This is expected, because until the extremum of the design space is sampled the models must extrapolate rather than interpolate. In the interior, we see that the model from the KFCV-Voronoi samples has smaller magnitude errors in terms of both the positive and negative errors. This is predominantly due to the KFCV-Voronoi approach identifying the area around Mach 0.9 as a region of interest early in the sampling process. In this region, the LOLA and CV-Voronoi approaches have prediction errors that are almost two times larger than those of the KFCV-Voronoi approach.

As more samples are added, we see that the corners of the design space are sampled by all three methods and the errors near the design space boundaries are small for the sets of 52 samples. The maximum errors still occur in the Mach 0.9 region, because this region features nonlinear behavior. In addition, the magnitude of the model errors is now smallest for the LOLA-Voronoi sampling, which was expected from our earlier analysis of the NRMSE. However, as more samples are added the KFCV-Voronoi sampling performs best in the sets of 77 samples. A comparison of the mean and maximum absolute error for the 77 sample sets selected via adaptive sampling methods is included in Table 3.

Table 3. Prediction errors for models created using 77 samples selected by each adaptive sampling method.

Error Metric	LOLA-Voronoi	CV-Voronoi	KFCV-Voronoi
Mean Absolute Error	0.0213	0.0214	0.0189
Max Absolute Error	0.192	0.1818	0.1522

From the table it can be seen that the KFCV-Voronoi sampling performs better in terms of mean and maximum absolute error. Using the lift coefficient range of $[0, 1.45]$ to calculate the percent error (in place of the true values since some are near zero), we find that the KFCV-Voronoi approach reduced the mean percent error from 1.47% to 1.3% (a 13% reduction) and the maximum prediction error from 13.2% to 10.5% (a 25.7% reduction) using the absolute values in the table. It is also important to note that the adaptive sampling methods are able to create models of superior accuracy to the 81-sample TPLHD model with only 52 samples which corresponds to a sample reduction of 35%. Thus, the adaptive sampling techniques provide models of the aerodynamic loading response of engineering accuracy for a considerably reduced testing cost.

V. Conclusions

In this work we presented techniques to maintain the sample quality provided by adaptive sampling techniques while reducing the cost associated with identifying future sample locations. Within the literature the CV-Voronoi adaptive sampling technique has been demonstrated to outperform the LOLA-Voronoi, space-filling cross-validation tradeoff (SFCVT), Monte-Carlo intersite projection (MIPT), mean squared error (MSE), and Accumulate error (ACE) methods in terms of model predictive accuracy for a given number of sample points. However, the cost of the CV-Voronoi technique scales superlinearly with number of samples, because (1) the cross-validation metric requires additional surrogate models to be created and (2) the cost of creating a surrogate model for a given data set increases with the number of samples.

In this work, a new KFCV-Voronoi scheme is developed to address the first issue. The KFCV-Voronoi scheme dramatically reduces the number of surrogate models that must be constructed during evaluation of optimal location

for the next point by adding a K-fold filtering mechanism. By adding the initial K-fold evaluation, the adaptive sampling process can be accelerated by the ratio of folds evaluated via LOOCV, l, to the number of folds K. Within this work $l = 3$ and $K = 10$, therefore the cost was reduced by 70% of the nominal CV-Voronoi approach (i.e., a 3.3× speedup). Despite dramatically reducing the sampling time, the quality of the samples selected via KFCV-Voronoi was maintained, and the accuracy of the models was comparable and often better than those built from CV-Voronoi samples. The improvement was attributed to the addition of the KFCV providing a global filtering affect.

A second cost reduction method in the form of a novel Voronoi batch sampling method was also developed. The Voronoi batch sampling utilizes a re-tessellation process to tackle the issue of sample clustering within batches, and it was demonstrated to outperform the naïve batch sampling approach. Although sampling in batches leads to slightly less accurate models for a given number samples, this typically can be compensated by more efficient utilization of testing resources for data generation. Combination of the KFCV-Voronoi approach with the Voronoi batch sampling approach has been shown to significantly improve the model accuracy and sampling efficiency.

The accelerated adaptive sampling techniques were inspected using benchmark functions of increasing input dimensions, and were shown to offer significant improvement over one-shot techniques. Complex and high-dimension parameter spaces were shown to necessitate the sampling cost reduction techniques. The sampling techniques were also validated for surrogate modeling of the aerodynamic loading 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, while dramatically lowering the associated with sample selection.

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] 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.

[5] 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.

[6] 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.

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

[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] 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.

[10] 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.

[11] 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.

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

[13] 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.

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

[15] 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.

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

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

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

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

[20] 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.

[21] 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.

[22] 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.

[23] H. Liu, S. Xu, X. Wang, S. Yang and J. Meng, "A multi-response adaptive sampling approach for global metamodeling".

[24] J. Van Der Herten, I. Couckuyt, D. Deschrijver and T. Dhaene, "A fuzzy hybrid sequential design strategy for global surrogate modeling of high-dimensional computer experiments," *SIAM Journal on Scientific Computing,* vol. 37, no. 2, pp. A1020-A1039, 2015.

[25] J. Van Der Herten, D. Deschrijver and T. Dhaene, "Fuzzy local linear approximation-based sequential design," in *IEEE SSCI 2014 - 2014 IEEE Symposium Series on Computational Intelligence - CIES 2014: 2014 IEEE Symposium on Computational Intelligence for Engineering Solutions, Proceedings*, 2015.

[26] A. L. Kaminsky, Y. Wang, K. Pant, W. N. Hashii and A. Atachbarian, "Adaptive sampling techniques for surrogate modeling to create high-dimension aerodynamic loading response surfaces," in *2018 Applied Aerodynamics Conference*, Reston, Virginia, 2018.

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

[28] R. Rai and M. I. Campbell, "Q2S2: Qualitative and quantitative sequential sampling a novel approach to exploit qualitative design information," in *International conference on engineering design*, Paris, 2007.

[29] S. a. B. D. Surjanovic, "Virtual Library of Simulation Experiments: Test Functions and Datasets," [Online]. Available: http://www.sfu.ca/~ssurjano. [Accessed 15 July 2019].

[30] K. Crombecq, E. Laermans and T. Dhaene, "Efficient space-filling and non-collapsing sequential design strategies for simulation-based modeling," *European Journal of Operational Research,* vol. 214, no. 3, pp. 683-696, 2011.

[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.