A new method for dynamic sampling of Kriging surrogate models for uncertainty quantification is developed and presented. The criterion for the dynamic adaptive sampling proposed is based on combining the expected uncertainty of the fit and the gradient information resulting from the Kriging predictors, and an error-estimate term (based on the difference in the Kriging predictors with different correlation length scales). The Kriging-based dynamic adaptive sampling method proposed is tested on two-dimensional analytic functions with smoothly and steeply varied responses in the quantities of interest under normal uncertainty distributions. Compared with a classical polynomial chaos expansion method based on the Gauss quadrature rule and a dynamic adaptive sampling method based only on the uncertainty of the Kriging predictor fit, this new method shows superior performance for estimating the statistics of the quantity of interest in terms of both accuracy and robustness, and regardless of either the choice of the initial set of samples or the smoothness of the stochastic space.

**Key Words:** Uncertainty Quantification, Adaptive Sampling, Kriging Surrogate Model

### Nomenclature

- $a, b$: constants in the sandtimer function
- $c$: vector of coefficients in the best linear unbiased predictor
- $D_f$: estimated polynomial error in $f$
- $f$: output solution
- $F$: vector whose $i$-th element is $f(\xi^{(i)})$
- $F$: stochastic process for $f$
- $F_f$: vector whose $i$-th element is $F(\xi^{(i)})$
- $h$: Euclidean distance $|\xi - \xi^{(i)}|$
- $L_n$: likelihood function
- $n$: number of input uncertainties
- $N$: number of sample points
- $P$: number of basis functions
- $r$: vector whose $i$-th element is $\kappa(\xi, \xi^{(i)}|\theta)$
- $R$: matrix whose $(i, j)$ entry is $\kappa(\xi^{(i)}, \xi^{(j)}|\theta)$
- $s$: mean square error (fit uncertainty) in $f$
- $u$: implicit parameter in the sandtimer function
- $v$: explicit function in the sandtimer function
- $x$: vector of input parameters
- $Z$: Kriging local model
- $\alpha$: basis function coefficient
- $\Delta$: distance to the most adjacent sample point
- $\theta$: vector of Kriging hyperparameters
- $\kappa$: Kriging correlation function (kernel)
- $\mu$: Kriging global model
- $\mu_f$: mean of $f$
- $\xi$: vector of input uncertainties in $x$

### Symbols

- $\sigma^2$: Kriging process variance
- $\sigma_f$: standard deviation of $f$
- $\phi$: basis function

### Subscripts

- $i$, $j$, $k$: $i$, $j$, $k$-th input uncertainty
- $l$: $l$-th basis function

### Superscripts

- $(i)$: $i$-th sample point
- $(\hat{i})$: estimated value by polynomial chaos expansion or Kriging model

### 1. Introduction

Numerical simulations of physical systems governed by partial differential equations (PDEs) typically encompass a physical model and a set of well-posed initial and boundary conditions. Often one can exercise multiple options for the choice of the physical model (either through a hierarchy of models of different fidelity or an ensemble of models of similar fidelity) and initial and boundary conditions. In the real world, specific physical models and initial and boundary conditions typically contain uncertainties whose impact on the output quantity of interest must be well understood. For example, physical models calibrated using a fixed number of parameters create significant uncertainties when the values of those parameters are chosen from a limited number of experiments or measurements. Additionally, for simplicity, initial and boundary conditions are typically assumed to be deterministic, while the real-world conditions are not. Therefore, proper consideration of such uncertainties can improve the value of the information produced by numerical simulations of PDEs.

Uncertainty quantification (UQ) is the science of the quan-
titative characterization and reduction of uncertainties in applications of interest that include numerical simulations. A conventional deterministic simulation considers the model that has as an output a unique value of a solution \( f(x) \) at fixed values of input parameters \( x \). On the other hand, a probabilistic simulation considers the model, which assumes that the input parameters \( x \) are subject to a number of uncertainties and are, therefore, represented as random variables \( x(\xi) \). The output of the simulation \( f(x) \) is a distribution itself that depends on the parameter uncertainties \( \xi \), i.e., \( f(x(\xi)) = f(\xi) \). Proper characterization of the relation between input and output uncertainties is a way to extract valuable information from our simulation models.

UQ approaches are typically categorized into intrusive and non-intrusive methods. Non-intrusive methods discussed in this paper only require multiple evaluations of the original (deterministic) model, i.e., users do not need to tamper with the details of the simulation codes. On the other hand, intrusive methods require the formulation and solution of a stochastic version of the original model, i.e., users must spend considerable time and effort modifying the simulation codes. The most popular methods for non-intrusive UQ are sampling methods, such as Monte Carlo (MC) sampling and Latin hypercube sampling (LHS).\(^1\) To estimate the stochastic behavior of an output quantity of interest \( f(\xi) \) resulting from input parameter uncertainties \( \xi \), sampling methods evaluate the values of \( f(\xi) \) at many different locations sampled in the stochastic space of \( \xi \), directly from repeated evaluation of the simulation models. MC samples randomly in the stochastic space, and it requires a large number of samples to ensure the accuracy and convergence of the stochastic estimations. Compared to MC sampling, LHS can reduce the number of sample points that are needed to achieve converged stochastic estimations; LHS samples a point in each equi-probability partition randomly and does not allow overlapping partitions to be sampled for all dimensions of the stochastic space (the so-called orthogonality condition). Therefore, techniques such as LHS can typically span the entire stochastic space with a smaller sample size than MC sampling.

In contrast to such sampling methods which do not rely on any approximation, several methods have been proposed to approximate the stochastic behavior of a solution and use such an approximation (i.e., the so-called surrogate) to predict the values of the solution at additional locations in the stochastic space. Polynomial chaos expansion (PCE)\(^2\) and stochastic collocation (SC)\(^3\) methods are popular UQ approaches that formulate the stochastic behavior of the solution \( f(\xi) \) as a linear combination of basis functions as follows:

\[
f(\xi) \simeq \sum_{i=1}^{P} \alpha_i \phi_i(\xi),
\]

where, \( P \) is the number of basis functions, and \( \phi_i(\xi) \) and \( \alpha_i \) are the \( i \)-th basis function and its scalar coefficient, respectively. For \( \phi_i(\xi) \), PCE employs multivariate orthogonal polynomials (e.g., Legendre, Laguerre, and Jacobi), and SC employs multivariate interpolation polynomials (e.g., Lagrange and Hermite). PCE estimates \( \alpha_i \) for known orthogonal polynomials, while SC forms interpolation functions for known \( \alpha_i \).

Eldred and Burkhart\(^4\) compared the performance of PCE and SC in simple algebraic test problems, and showed that, in principle, their performance is comparable and both approaches result in very fast convergence relative to MC sampling. Hosder et al.\(^5,6\) applied PCE to computational fluid dynamics (CFD) simulations: shock-wave and expansion-wave problems with uncertain geometries and a transonic wing case with uncertain freestream conditions. The results computed using PCE agreed closely with the MC results at a much lower computational cost. Moreover, PCE and SC approaches have been applied and demonstrated in other UQ problems related to CFD simulations (these efforts are well summarized in Bijl et al.\(^7\)).

However, when the stochastic behavior includes a steeply varied response, conventional PCE and SC approaches often suffer from approximation errors and deterioration in accuracy and efficiency due to their reliance on a linear combination of polynomials. For example, Fig. 1 shows the responses of \( f(\xi) = \text{erf}[2(\xi + 1)] \) estimated using PCE (i.e., based on the Hermite polynomials suitable for a normally-distributed uncertainty). There are two types of approximation errors; discretization errors appear around the steeply varied response (\( \xi = -1 \)) due to a lack of sample points,
and polynomial errors appear in the regions near the endpoints ($\xi = -4$ and 4) because of Runge’s phenomenon. Figure 1 indicates that the discretization error can be reduced by additional samples close to the high-gradient region, but the polynomial error is amplified as the number of samples and the polynomial order increase.

To improve the accuracy and efficiency of PCE and SC, multiple researchers have been working on improved PCE and SC approaches. Poëtte et al. proposed a new PCE method based on the entropy variable to reduce the discretization errors in discontinuous responses. This approach bounds the oscillations in the vicinity of steeply varied responses (i.e., discontinuities as the extreme) to a certain range through the entropy of the system and without the use of any adaptive stochastic space discretization. For discontinuous responses, Poëtte and Lucor also proposed an iterative generalized PCE method that recovers the same error bounds in the vicinity of discontinuities as in smoother regions without the need to increase the polynomial order. Lockwood and Mavriplis examined a gradient-based UQ approach to accurately predict the stochastic behaviors in hypersonic flow with a nonlinear response. This approach couples PCE with gradient information evaluated by the adjoint method. Witteveen and Iaccarino proposed the simplex stochastic collocation (SSC) approach to collocate samples in nonhypercubic stochastic spaces for approximating discontinuous responses efficiently. This approach discretizes the stochastic space using a simplex tessellation of the sampling points. Witteveen and Iaccarino then focused on discontinuous/smooth-mixed responses and introduced a stencil selection strategy based on the essentially non-oscillatory (ENO) scheme into the SSC approach for sharper resolution of discontinuities and higher-order approximations in smooth regions.

The approximation errors in PCE and SC are also affected by the choice of sample locations. The Gauss quadrature rule is helpful for the PCE method to mathematically determine where sample points should be collocated in the stochastic space, such that unphysical oscillations will disappear as much as possible. Nevertheless, this rule is sometimes not acceptable for practical use because it may request the location of a sample where numerical simulation is not available (e.g., typical CFD solvers cannot cope with the Mach number of 1 due to its singular flow condition that a shock wave can stand infinitely forward of an object), and it does not allow us to add new samples in a space-filling manner. In addition, Hosder et al. demonstrated that, with a given value of the size of the sample set, PCE based on LHS samples leads to more accurate estimation of statistics than if we were to use brute-force MC samples. Unfortunately, the cost of such approximations exhibits the so-called curse of dimensionality; meaning that the number of samples required for PCE and SC increases exponentially as the number of input parameter uncertainties (i.e., dimensions in the stochastic space) increases.

For these reasons, this paper attempts to adaptively refine the stochastic space using as few samples as possible to approximate the stochastic behavior of a quantity of interest without unphysical local oscillations. In the same spirit, Witteveen and Iaccarino proposed a strategy for the adaptive space-infilling sampling in SSC. It is based on an estimation of the errors derived from hierarchical surpluses, polynomial degree, and local extremum diminishing concepts. This procedure outperformed uniform refinement in terms of accuracy. Palacios et al. utilized the adjoint method and similar ideas in a robust grid adaptation approach to CFD simulations. This approach aims to minimize the numerical discretization error over small variations of the input parameters around a baseline flow state. Similarly, Duraisamy and Chandrashekar proposed a strategy for adaptive sampling in SSC based on an estimation of the error that relied on the use of a stochastic adjoint method.

Unlike polynomial functions used in PCE and SC, the Kriging surrogate model approximates stochastic behaviors in UQ problems as simple algebraic functions based on Bayesian statistics. This approach can be easily adapted to work well with nonlinear functions. It estimates not only the function values, but also an estimate of the uncertainty in the fit that is equivalent to a measure of the approximation error. Although the Kriging model has been developed and utilized for engineering design optimization problems that require many simulations, this paper proposes to use the Kriging model to approximate the stochastic space and determine the effective criteria for both adaptive sampling and detecting steeply varied responses, and thus reducing errors in UQ.

The Kriging-based dynamic adaptive sampling considered here is implemented as shown in Fig. 2. For $N$ given initial samples, the Kriging model is constructed and the location of a new $N + 1$-th sample is determined based on a criterion that is formulated by the Kriging predictors. For $N + 1$ samples updated with the new additional sample, the Kriging model is reconstructed, and thus the next sample is determined and added one-by-one by iterating this process. It is expected to collocate the minimum number of samples necessary to approximate the black-box stochastic space.

Yamazaki has also applied the Kriging model to UQ, but his results indicated that, without adaptive sampling, the performance of the Kriging model is inferior to classical PCE. Dwight and Han proposed a Kriging-based adaptive sampling method that is based on the product of the fit uncertainty in the Kriging predictor and the distributions of the input parameter uncertainties. Bilionis and Zabaras employed a similar adaptive refinement criterion based on the

---

**Fig. 2. Flowchart of Kriging-based dynamic adaptive sampling.**
fit uncertainty predicted in Gaussian process (GP) regression, which is another Bayesian-statistics-based surrogate model similar to the Kriging model. However, as we will show in this paper, such criteria are not effective in speeding up the convergence of the UQ problem when the responses behave steeply.

A highly nonlinear Kriging model enhanced with gradient information \( \partial f(\xi)/\partial \xi \) has recently been investigated for adaptive sampling in UQ. Ulaganathan et al.\(^{26} \) considered nonlinearity appearing in the gradient-enhanced Kriging model as a sampling criterion to detect steeply varied responses. For a similar purpose, Boopathy and Rumpfkeil\(^{27} \) proposed a sampling criterion based on deviating the gradient-enhanced Kriging model from another local surrogate model (i.e., multivariate interpolation and regression). These approaches are promising, but not always practical due to the difficulty of providing the gradient information as the sample data set.

The present study is focused on developing a new dynamic adaptive sampling method based on an ordinary (i.e., gradient-free) Kriging model for effective non-intrusive UQ when the response of an output solution behaves either smoothly or steeply in the stochastic space. We first show that the existing Kriging-based method\(^{23} \) is plagued by persistent errors when the response behaves steeply, and then develop a method that eliminates these errors. The method proposed is then tested on two-dimensional analytic functions with different smoothnesses, each of which models different flow physics related to aeronautics and astronautics. The robustness of the method proposed is also investigated by evaluating statistics using 30 different sets of initial samples generated at random. The Kriging-based method proposed is compared with the existing Kriging-based method as well as a classical PCE method based on the Gauss quadrature rule.

Considering only two dimensions in UQ test cases seems insufficient from the real-world point of view. Nevertheless, this paper targets two-dimensional test cases for fundamental UQ study. The dynamic adaptive sampling method, which is demonstrated in two-dimensional cases here, can also be applied to more-than-two-dimensional cases. In a general sense, however, even this method may be less effective in higher-dimensional cases because it costs more sample points to construct accurate Kriging models as well as the PCE method. Such a bottleneck may disappear if combined with proper methods of dimensionality reduction (e.g., active subspace method\(^{28} \)), which is out of the scope of this paper.

2. Kriging Surrogate Model

2.1. Basic formulation

Consider the approximation of an output function \( f(\xi) \) in terms of a vector of \( n \) input variables \( \xi = [\xi_1, \xi_2, \cdots, \xi_n]^T \), each of which is defined in the range of \( 0 \leq \xi_k \leq 1 \) \((k = 1, 2, \cdots, n)\). Based on Bayesian statistics, the Kriging model starts with a prior distribution over \( f(\xi) \), as illustrated in Fig. 3(a). It treats the deterministic response of \( f(\xi) \) as the realization of a stochastic process \( F(\xi) \), defined as

\[
F(\xi) = \mu + Z(\xi),
\]

where, \( \mu \) is the global model assumed to be constant (i.e., so-called ordinary Kriging) and \( Z(\xi) \) is the local model corresponding to the deviation from \( \mu \) at \( \xi \). \( Z(\xi) \) is assumed to have the stochastic behaviors (i.e., mean \( \mathbb{E}[Z(\xi)] \) and covariance \( \text{Cov}(Z(\xi), Z(\xi')) \)) defined as follows:

\[
\mathbb{E}[Z(\xi)] = 0,
\]

\[
\text{Cov}(Z(\xi), Z(\xi')) = \sigma^2 \kappa(\xi, \xi'|\theta),
\]

where, \( \sigma^2 \) is the process variance and \( \kappa(\xi, \xi'|\theta) \) is the correlation function (or kernel) between any two locations \( \xi \) and \( \xi' \).

\( \kappa(\xi, \xi'|\theta) \) is defined as a function of the Euclidean distance \( h = |\xi - \xi'| \) with a set of constants called hyperparameters \( \theta \).

\( \kappa(\xi, \xi'|\theta) \) satisfies the following conditions:

\[
\lim_{h \to 0} \kappa(\xi, \xi'|\theta) = 1,
\]

\[
\lim_{h \to \infty} \kappa(\xi, \xi'|\theta) = 0,
\]

(a specific choice for \( \kappa(\xi, \xi'|\theta) \) will be given later). This paper assumes \( \kappa(\xi, \xi'|\theta) \) to be stationary, which implies that the hyperparameters \( \theta \) is invariant to \( \xi \). Although a non-stationary correlation function is expected to be more effective for non-linear function approximation,\(^{29-31} \) it is out of scope here and we focus on the stationary correlation function for its simple formulation.

Next, consider that the true value of \( f(\xi) \) is given at \( N \) sample points \( \xi^{(1)}, \xi^{(2)}, \cdots, \xi^{(N)} \). The Kriging model attempts to derive a posterior distribution over \( f(\xi) \), as illustrated in
Fig. 3(b). It supposes that \( F(\xi) \) realizes all \( N \) given samples; that is,

\[
f(\xi) = F(\xi) = \mu + Z(\xi) \quad (i = 1, 2, \ldots, N)
\]

Then, the probability density distribution conditioned on these realizations (i.e., likelihood function) is obtained in logarithmic form as

\[
\log L(\mu, \sigma^2, \theta) = -\frac{N}{2} \log 2\pi - \frac{N}{2} \log \sigma^2 - \frac{1}{2} \ln |R| - \frac{1}{2} \ln |f|,
\]

where \( f = [f(\xi^{(1)}), f(\xi^{(2)}), \ldots, f(\xi^{(N)})]^T \), \( R \) is an \( N \times N \) matrix whose \((i,j)\) entry is \( \kappa(\xi^{(i)}, \xi^{(j)}|\theta) \), and \( \theta \) is \( N \)-dimensional vector with all elements of \( 1 (1 = [1, 1, \ldots, 1]^T) \). The Krigma model needs to determine the values of \( \mu, \sigma^2 \), and \( \theta \) based on the maximum likelihood estimation (MLE). The values of \( \mu \) and \( \sigma^2 \) that maximize \( \log L(\mu, \sigma^2, \theta) \) are solved in closed form as

\[
\hat{\mu} = \frac{1}{\hat{\sigma}^2} f - \frac{1}{\hat{\sigma}^2} R^{-1} f,
\]

\[
\hat{\sigma}^2 = \frac{(f - \hat{\mu})^T \hat{R}^{-1} (f - \hat{\mu})}{N}.
\]

Substituting Eqs. (7a) and (7b) into Eq. (6), the following concentrated likelihood function is obtained:

\[
\log L(\theta) = -\frac{N}{2} \log 2\pi - \frac{N}{2} \log \sigma^2 - \frac{1}{2} \ln |R|,
\]

where \( \theta \) is the set of hyperparameters. This is the function maximized in practice to give the estimates of \( \theta \), and hence an estimate of \( R \). This paper searches for the values of \( \theta \) that maximize \( \log L(\theta) \) using a genetic algorithm (GA). The GA is expected to act as a population-based optimizer to robustly search for the values of \( \theta \) at a global maximum of \( \log L(\theta) \). Then, Eqs. (7a) and (7b) are used to obtain the estimates of \( \hat{\mu} \) and \( \hat{\sigma}^2 \).

Finally, consider the linear predictor \( \hat{f}(\xi) \), which estimates \( f(\xi) \) at the location \( \xi \) where the true value of \( f(\xi) \) is not known, defined as

\[
\hat{f}(\xi) = \hat{c}(\xi) F,
\]

where \( F = [F(\xi^{(1)}), F(\xi^{(2)}), \ldots, F(\xi^{(N)})]^T = f \) from Eq. (5). The Krigma model obtains the best linear unbiased predictor (BLUP) by choosing the \( N \)-dimensional vector \( c(\xi) \) to minimize the following mean squared error (MSE):

\[
\hat{\sigma}^2 = \text{Var}[\hat{f}(\xi) - F(\xi)],
\]

subject to the following unbiased constraint:

\[
E[\hat{f}(\xi)] = E[F(\xi)].
\]

Then, \( c(\xi) \) is solved in closed form as

\[
\hat{c}(\xi) = R^{-1} r(\xi) + \frac{R^{-1} (1 - R^{-1} r(\xi))}{1 + R^{-1} r(\xi)}.
\]

where, \( r(\xi) \) is an \( N \)-dimensional vector, the \( i \)-th element of which is \( \kappa(\xi^{(i)}, \xi^{(j)}|\theta) \). Substituting Eq. (11) into Eq. (9), the BLUP results in final form as

\[
\hat{f}(\xi) = \hat{\mu} + \hat{R}^{-1} (f - \hat{\mu})
\]

This function models the estimate of \( f(\xi) \) at any location \( \xi \) by interpolating the sample points with true values of \( f(\xi) \). Similarly, substituting Eqs. (9) and (11) into Eq. (10a), the MSE results in final form as

\[
\hat{\sigma}^2 = \frac{1}{N} \left[ 1 - R^{-1} (f - \hat{\mu})^T (f - \hat{\mu}) \right].
\]

This function models the fit uncertainty expected in \( \hat{f}(\xi) \) estimated by Eq. (12). It indicates that the accuracy of \( \hat{f}(\xi) \) depends largely on the distance from the given sample points. Intuitively, the closer location \( \xi \) is to the sample points, the less uncertain (i.e., more accurate) the Krigma predictor \( \hat{f}(\xi) \). Note that, at sample points \( \xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(N)} \), the Krigma model satisfies the following conditions:

\[
\hat{f}(\xi) = f(\xi),
\]

\[
\hat{\sigma}(\xi) = 0.
\]

2.2. Correlation function

\[
\kappa(\xi, \xi'|\theta),
\]

which is an \( n \)-variate correlation function between any two locations \( \xi \) and \( \xi' \), is represented as a product of univariate correlation functions for each variable as

\[
\kappa(\xi, \xi'|\theta) = \prod_{k=1}^{n} \kappa(\xi_k, \xi_k'|\theta_k).
\]

This paper uses the \( C^4 \)-smoothness correlation function constructed by Wendland. This is a class of positive definite and compactly supported radial functions consisting of a univariate polynomial. For given smoothness and space dimension, they are of minimal degree and unique up to a constant factor. In two-dimensional space (\( n = 2 \)), for example, this class is constructed with distance \( h_k = |\xi_k - \xi_k'| \) and hyperparameter \( \theta_k > 0 \) as

\[
\kappa(\xi_k, \xi_k'|\theta_k) = \begin{cases} (1 - h_k/\theta_k)^2 & (C^0\text{-smoothness}) \\ (1 - h_k/\theta_k)^4 (4h_k/\theta_k + 1) & (C^2\text{-smoothness}) \\ (1 - h_k/\theta_k)^6 (35h_k^2/\theta_k^2 + 18h_k/\theta_k + 3)/3 & (C^4\text{-smoothness}) \end{cases}
\]

in the range of \( 0 \leq h_k \leq \theta_k \), and

\[
\kappa(\xi_k, \xi_k'|\theta_k) = 0
\]

in the range of \( h_k > \theta_k \). Note that the \( \theta_k \) corresponds to the length scale of correlation along \( \xi_k \); that is, \( \hat{f}(\xi) \) with small/large \( \theta_k \) adapts to given samples locally/globally and yields a quick/smooth response in the \( \xi_k \) direction (i.e., examples will be
shown later). Compared to the Gaussian correlation function:
\[
\kappa(\xi, \xi') = \exp \left( -h \xi / \xi' \right),
\]
which is often used in the Kriging model, the Wendland correlation function with compact support (0 ≤ h ≤ 1) is able to avoid overfitting to given sample points.

3. Dynamic Adaptive Sampling

Using the Kriging model, dynamic adaptive sampling is conducted using the following procedure (see also the flow-chart shown in Fig. 2).

- **Step 1**: Generate N initial samples of input variables \( \xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(N)} \) in the stochastic space and evaluate their output function values \( f(\xi^{(1)}), f(\xi^{(2)}), \ldots, f(\xi^{(N)}) \).

- **Step 2**: Construct Kriging models \( \hat{f}(\xi) \) and \( \sigma(\xi) \) using the N samples.

- **Step 3**: Find the location of a new sample point \( \hat{\xi}^{(N+1)} \), where the value of a criterion (i.e., \( \text{Crit}(\xi) \)) is maximized in the stochastic space, and evaluate its output function value \( f(\hat{\xi}^{(N+1)}) \).

- **Step 4**: Add \( f(\hat{\xi}^{(N+1)}) \) to the N samples, set \( N = N + 1 \), and go back to Step 2.

In Step 1, this study employs LHS\(^1\) to generate initial samples. In Step 3, for simplicity, this study carries out random search, in which \( 10^4 \) samples are generated using MC sampling, to find the location where \( \text{Crit}(\xi) \) is maximum; note that optimizers such as genetic algorithms are also available for better findings. This paper considers the following three criteria to be used in Step 3 for comparing existing sampling methods to investigate effective criteria for UQ.

- **Criterion 1** (existing in Ref. 23)

\[
\text{Crit}(\xi) = \delta(\xi) \text{PDF}(\xi),
\]

where, \( \delta(\xi) \) is the Kriging model fit uncertainty given by Eq. (13), and PDF(\( \xi \)) is the probability density function of input parameter uncertainties \( \xi \). This criterion has been proposed and investigated in a previous study by Dwight and Han.\(^23\) Criterion 1 is expected to sample output solutions evenly in the stochastic space (i.e., from Eq. (14b), new samples will never duplicate existing ones), as well as focus the solutions on locations where there are significant effects on the evaluation of statistics (i.e., solutions where PDF(\( \xi \)) is higher will be sampled more frequently).

- **Criterion 2**

\[
\text{Crit}(\xi) = \left| \frac{\partial \hat{f}(\xi)}{\partial \xi} \right| \delta(\xi) \text{PDF}(\xi),
\]

where,
\[
\frac{\partial \hat{f}(\xi)}{\partial \xi} = \left[ \frac{\partial \hat{f}(\xi)}{\partial \xi_1}, \frac{\partial \hat{f}(\xi)}{\partial \xi_2}, \ldots, \frac{\partial \hat{f}(\xi)}{\partial \xi_N} \right]^T
\]
is a vector of the Kriging predictor gradients \( \hat{f}(\xi) \) given by Eq. (12) with respect to input parameter uncertainties \( \xi \). As the vector \( \mathbf{r}(\xi) \) in \( \hat{f}(\xi) \) consists of differentiable correlation elements, \( \frac{\partial \hat{f}(\xi)}{\partial \xi} \) can be derived analytically as

\[
\frac{\partial \hat{f}(\xi)}{\partial \xi} = \left( \frac{\partial \mathbf{r}(\xi)}{\partial \xi} \right)^T R^{-1} (f - \mathbf{1} \mu).
\]

(20a)

where, for the Wendland correlation with \( \alpha^2 \)-smoothness and two dimensions \( (n = 2) \), \( \kappa(\xi, \xi' | \theta) \) is given by Eq. (16) and \( \kappa'(\xi, \xi' | \theta) \) is defined as

\[
\kappa'(\xi, \xi' | \theta) = -56(1 - h_1 / h_2)^5 (5h_2^2 / \theta_2^2 - h_1 / \theta_1) / \theta_1.
\]

(21)

**Criterion 2** is expected to add new samples where fit uncertainty is large and the gradient is also large. Thus, it works for a balance between sampling evenly in the whole stochastic space and capturing steeply varied responses where output solutions change markedly according to input parameter uncertainties. This is considered to reduce the discretization errors, as seen in Fig. 1.

- **Criterion 3** (proposed in this paper)

Criterion 2 still seems insufficient because new solutions will not be sampled at the locations where \( |\hat{f}(\xi) / \partial \xi| = 0 \) even if \( \delta(\xi) \) is large. This implies that this criterion may not be able to suppress the polynomial errors as seen in the regions near the endpoints in Fig. 1, where the true response is smooth. Therefore, Criterion 2 is now modified into the following Criterion 3, so that the locations where \( |\hat{f}(\xi) / \partial \xi| = 0 \) will still be sampled for new solutions.

\[
\text{Crit}(\xi) = \left| \frac{\partial \hat{f}(\xi)}{\partial \xi} \right| \Delta(\xi) + D(\xi) \delta(\xi) \text{PDF}(\xi)
\]

(22)

|\hat{f}(\xi) / \partial \xi| is now multiplied by \( \Delta(\xi) \) to keep the consistency of units with an extra term \( D(\xi) \). \( \Delta(\xi) \) is set to be the distance from \( \xi \) to the most adjacent sample point; that is,

\[
\Delta(\xi) = \min_{i=1,2,\ldots,N} |\xi - \hat{\xi}^{(i)}|.
\]

(23)

The extra term \( D(\xi) \) in Eq. (22) is assumed to estimate the polynomial errors. This paper formulates \( D(\xi) \) as

\[
D(\xi) = |\hat{f}(\xi | \theta) - \hat{f}(\xi | \theta_0)|.
\]

(24)

where, \( \hat{f}(\xi | \theta) \) is the Kriging predictor modeled with the ordinary values of hyperparameters \( \theta \), which are determined by the MLE, and \( \hat{f}(\xi | \theta_0) \) is the predictor modeled with the doubled values of hyperparameters \( \theta_0 \).

For example, Fig. 4 compares the Kriging predictors with different correlation length scales (\( \hat{f}(\xi | \theta) \), \( \hat{f}(\xi | \theta_0) \), \( \hat{f}(\xi | \theta_40) \), and \( \hat{f}(\xi | 1/40) \) for \( \xi = \text{erf}[2(\xi + 1)] \) (i.e., same example as Fig. 1). The predictor with a smaller correlation length scale (\( \hat{f}(\xi | 1/40) \)) yields an unrealistic and bumpy response that is excessively attracted by the given samples. On the other hand, the predictor with a larger correlation length scale (\( \hat{f}(\xi | 200) \)) slightly fluctuates from the ordinary predictor \( \hat{f}(\xi | \theta_0) \) in regions where the true response is smooth (\(-4 \leq \xi \leq -2 \) and \( 0 \leq \xi \leq 4 \)).
that this fluctuation no longer increases when the correlation length scale is larger than \(4\theta\). This is the reason why this paper employs \(2\theta\) in the formulation of \(D_f(\xi)\) (Eq. (24)).

Thus, \(D_f(\xi)\) often tends to be large in the region where samples are located sparsely even if the true response is smooth. Although there are other possible ways to estimate the errors on surrogate models, as in Mehmani et al.\(^{34}\) Criterion 3 utilizes Eq. (24) for its simple formulation. Therefore, Criterion 3 is proposed in this paper and is expected to improve the quality of adaptation in smooth regions that may induce polynomial errors.

In addition, note that it is important to include the term PDF(\(\xi\)) in all of the criteria. The criteria without PDF(\(\xi\)) achieve slower convergence of statistic estimation compared to that with PDF(\(\xi\)) (i.e., see Shimoyama et al.\(^{35}\) for details). This indicates that the consideration of probabilistic significance enhances the performance of dynamic adaptive sampling.

4. Numerical Tests

This study investigates the capability of the dynamic adaptive sampling criteria in the following two-dimensional UQ test problems.

4.1. Log-Sin function

4.1.1. Problem definition

This function was considered in Hosder et al.\(^{15}\) and Yamazaki,\(^{22}\) and is formulated as

\[
f(\xi_1, \xi_2) = \ln(1 + x_1^2) \sin 5x_2, \quad (25a)
\]

\[x_i = 2.0 + 0.4\xi_i \quad (i = 1, 2). \quad (25b)
\]

The true response of the log-sin function is shown in Fig. 5. This function has a multimodal response, which can be seen in aeroacoustics considering coherent waves, for example. Here, it is assumed that input variables \(\xi_1\) and \(\xi_2\) are uncertain and normally distributed as \(\xi_i \sim \text{Norm}(0, 1^2)\) \((i = 1, 2)\).

4.1.2. Results

The present test starts dynamic adaptive sampling based on either Criteria 1, 2, or 3 from initial samples with \(N = 5\), which are randomly generated within the range of

\[-4 \leq \xi_i \leq 4 \quad (i = 1, 2)\] using LHS. To ensure the robustness of results, the present test averages the results among 30 trials of dynamic adaptive sampling, each of which starts from a different LHS initial sample set.

For the log-sin function, the statistics (i.e., mean \(\mu_f\) and standard deviation \(\sigma_f\)) of the response \(f(\xi_1, \xi_2)\) against the input uncertainties in \(\xi_1\) and \(\xi_2\) are estimated using the Kriging models updated from \(N = 5\) to 100 with Criteria 1–3, respectively. Note that there is an empirical rule suggesting the number of initial samples as \(N = 11n - 1\) for Kriging-surro-
gate-based optimization. This rule gives $N = 21$ to the present test with $n = 2$, which means that five initial samples are quite few and more additional samples are necessary for practical use. Thus, the range from $N = 5$ to 100 used in the current dynamic adaptive sampling is reasonable in an empirical sense.

Figure 6 shows the trial-average errors in $\mu_f$ and $\sigma_f$ estimated by the Kriging-based dynamic adaptive sampling (i.e., 30 trials) compared to higher-fidelity statistics estimated by the MC method with $N = 1 \times 10^6$. Figure 6 also includes the errors in the statistics estimated by pure Kriging without dynamic adaptive sampling (i.e., 30 trials with different LHS sample sets following $\xi_i \sim \text{Norm}(0, 1^2)$ ($i = 1, 2$)), and those estimated by the PCE method based on one trial using the Gauss quadrature rule, which restricts the collocation of sample points in the stochastic space. The PCE results are obtained using “Dakota” open-source software, developed at Sandia National Laboratories.

For $\mu_f$, all three criteria can reduce the error more quickly in the early stage but remain a larger residual, even with $N = 100$, than the PCE method; more detailed data reveals that there is no trial where Kriging-based dynamic adaptive sampling converges to a lower residual than the PCE method. In general, a required level for error reduction depends on user preferences and UQ problems themselves, and thus, it is difficult to make a threshold. What matters here is that the PCE method happens to be better than the Kriging models when $N \geq 49$. This will be discussed further later based on observation from the probability density functions (Fig. 8) and the responses (Fig. 9). Next, for $\sigma_f$, all three of the criteria overwhelm the PCE method, and Criterion 3 is slightly better than Criteria 1 and 2 in terms of the speed of reducing the error. Although pure Kriging can reduce the errors as $N$ increases, Kriging-based dynamic adaptive sampling based on any criterion is still superior in terms of the error reduction speed.

Figure 7 compares the standard deviations of the errors in $f(\xi_1, \xi_2)$ estimated for the log-sin function.
$\mu_f$ and $\sigma_f$ estimated by each method over 30 trials. Compared to the pure Kriging, the Kriging-based dynamic adaptive sampling, especially based on Criterion 3, can quickly reduce the standard deviations of the errors in both $\mu_f$ and $\sigma_f$. It means that the proposed Criterion 3 performs most robustly regardless of the choice of initial samples.

Figure 8 shows the trial-averaged probability density functions of 30 trials estimated using the Kriging models updated by Criteria 1–3 and one trial using the PCE method with $N = 25$, 49, and 100 compared to one trial with $N = 1 \times 10^6$ estimated using the MC sampling method. With $N = 25$ and 49, the PCE method does not agree with the MC sampling method although its error in $\mu_f$ is well reduced with $N = 49$, as seen in Fig. 6(a). This is a proof that
the advantage of the PCE method expected from Fig. 6(a) is not true. On the contrary, Kriging-based dynamic adaptive sampling with \( N = 25 \) can produce estimates similar to the MC method, which indicates a significant advantage of Kriging-based dynamic adaptive sampling to reduce the UQ cost. Criteria 1–3 yield almost the same results because the log-sin function is relatively smooth. This suggests that uniform sampling, which is only followed by a given PDF(\( \xi \)) and independent of the dynamic adaptive sampling criteria, is sufficient to estimate a smooth response.

This paper also investigates how each of Criteria 1–3 collocates sample points and affects the estimated response of a solution in the stochastic space. Figure 9 shows the responses \( \hat{f}(\xi_1, \xi_2) \) estimated using the Kriging models updated by Criteria 1–3, respectively, and the PCE method with \( N = 25, 49, \) and 100. The sample points used are also depicted as white circles in Fig. 9. Compared to the true response (Fig. 5), the PCE method cannot capture the peaks in the \( \xi_2 \) direction when \( N = 25 \) and 49 due to the lack of polynomial order. Particularly, the PCE method when \( N = 25 \) and 49 underestimates the negative peak in \( 0 \leq \xi_2 \leq 1 \), which results in fatal disagreement of the PDF estimated at \( f(\xi_1, \xi_2) \leq 0 \) using MC sampling, as seen in Fig. 8. Criteria 1–3 construct similar Kriging models, which agree quite well with the true response when \( N = 49 \) and 100. Nevertheless, Criterion 3 collocates samples in a wider range, which covers both the smoother and less-smooth regions than Criteria 1 and 2. This indicates that Criterion 3 is ready to mitigate both polynomial errors and discretization errors in any possible case, which will be confirmed in the next test function with a steeper response.

### 4.2. Sandtimer function

#### 4.2.1. Problem definition

This function was considered in Dwight and Han,\(^{23}\) and is defined in the general \( n \)-dimensional form as

\[
\hat{f}(\xi) = \sum_{i=1}^{n-1} u_i^2 + v(\xi_n),
\]

where, \( u_i \) is defined implicitly by

\[
(\xi_n^2 + 1) (u_i + u_i^3) - \xi_i = 0 \quad (i = 1, 2, \ldots, n - 1),
\]

while \( v(\xi_n) \) is explicitly given by

\[
v(\xi_n) = a [x_n^4 + (x_n - 1)^2 - b],
\]

\[
x_n = 0.05\xi_n + 0.25
\]

with \( a = 0.58975451230 \) and \( b = 0.289273423937 \). This paper considers the sandtimer function with \( n = 2 \), as shown in Fig. 10. Compared to the log-sin function, the sandtimer function contains fewer peaks but responds to \( \xi_1 \) and \( \xi_2 \) more steeply. It is worth noting that, at \( \xi_2 \geq 10 \), the value of \( f(\xi_1, \xi_2) \) is not constant and beyond the upper limit of the color range. Such a response can be seen in high-speed aerodynamics, which is comparable with transonic flow around an airfoil involving shock waves; for example, as seen in Kawai and Shimoyama.\(^{38}\) Input uncertainties are assumed as \( \xi_1 \sim \text{Norm}(3, 3^2) \) and \( \xi_2 \sim \text{Norm}(2, 2^2) \).

![Figure 10. True response \( f(\xi_1, \xi_2) \) of the sandtimer function.](image)

![Figure 11. Errors in the statistics (trial averages) estimated for the sandtimer function.](image)

#### 4.2.2. Results

Similar to the previous case for the log-sin function, a total of 30 Kriging-based dynamic adaptive sampling trials with different initial sample sets generated within \( -9 \leq \xi_1 \leq 15 \) and \( -10 \leq \xi_2 \leq 14 \) were implemented to investigate Criteria 1–3 for the sandtimer function. Figures 11 and 12 compare the trial-average errors and their standard deviations, respectively, in \( \mu_f \) and \( \sigma_f \) against the input uncertainties for the sandtimer function, which were estimated using the 30 Kriging-based dynamic adaptive sampling trials, 30 trials us-
ing pure Kriging LHS sampling sets and one trial using the PCE method based on the Gauss quadrature rule. In Fig. 11, Criteria 1–3 seem to successfully reduce the errors in $\mu_f$ and $\sigma_f$ as $N$ increases through the Kriging-based dynamic adaptive sampling. On the other hand, the PCE method seems delayed or stagnant in error reduction even with the Gauss quadrature rule. The pure Kriging method also seems stagnant in error reduction for estimating $\sigma_f$. This indicate a significant advantage of Kriging-based dynamic adaptive sampling for effective UQ. For the dynamic adaptive sampling, Criterion 3 shows the fastest and continuous error reduction until $N = 100$. The error in $\sigma_f$ is almost saturated by Criteria 1 and 2, even with $N > 20$. Hence Criterion 3 seems most promising for efficient dynamic adaptive sampling. In Fig. 12, moreover, Criterion 3 indicates another advantage that performance is less sensitive to the choice of initial samples than Criteria 1 and 2.

To confirm the contribution of the term $D_f(\xi)$ in Criterion 3 (Eq. (22)), Figs. 11 and 12 compare the following Criterion 3’ as well.

**Criterion 3’**

$$\text{Crit}(\xi) = \left| \frac{\partial \hat{f}(\xi)}{\partial \xi} \right| \Delta(\xi) \tilde{d}(\xi) \text{PDF}(\xi).$$  \hspace{1cm} (27)

which eliminates $D_f(\xi)$ from Criterion 3.

Figures 11 and 12 indicate that Criterion 3’ performs well in Kriging-based dynamic adaptive sampling. Nevertheless, Criterion 3 outperforms Criterion 3’ in the later stage (i.e., $N > 35$) to estimate $\sigma_f$. This is because, as explained in Fig. 4, $D_f(\xi)$ in Criterion 3 is capable of detecting small fluctuations caused by different correlation length scales, which affect the accuracy to estimate a smooth response. Hence, consideration of $D_f(\xi)$ is essential to push through the dynamic adaptive sampling toward comprehensive UQ approximation.

Figure 13 compares the trial-average probability density functions using Kriging models updated by 30 trials applying...
Criteria 1–3, respectively, one trial of the PCE method using $N = 25$, 49 and 100, and one trial of the MC sampling method using $N = 1 \times 10^6$. There is still a discrepancy between the PCE method and the MC sampling method even with $N = 100$. On the other hand, the Kriging models with $N = 100$ obtain the probability density functions that almost agree with the MC sampling method. This indicates that Kriging-based dynamic adaptive sampling is effective to reduce the computational cost for UQ. With $N = 100$, Criteria 2 and 3 can catch up with the true peak of probability density around $f(\xi_1, \xi_2) = 0.23$, while Criterion 1 approximates it as two local peaks around $f(\xi_1, \xi_2) = 0.17$ and 0.35. This
can be seen in the enlarged view in Fig. 13(c). Here, we argue that it is very important to predict PDF in an engineering sense. Even a slight change in the number and location of peak(s) in an estimated PDF, which appears in Fig. 13(c), may affect the reliability of engineering design (e.g., Kawai and Shimoyama38) showed that a transonic airfoil involves multiple peaks in the PDF of local surface pressure due to a moving shock wave under an uncertain flight Mach number, and an accurate estimate of these peaks is relevant to verification and validation between CFD and wind tunnel experiments. In addition, with $N = 49$, Criterion 3 can also capture a smooth trend of probability density in the range of $1 \leq f(\xi_1, \xi_2) \leq 1.5$, while Criterion 2 causes slight oscillations, which can be seen in the enlarged view in Fig. 13(b).

It is worth noting that the range of $1 \leq f(\xi_1, \xi_2) \leq 1.5$ corresponds to the smooth region seen in the true response (Fig. 10). This demonstrates that Criterion 3 is effective for reducing the polynomial errors in smooth regions by considering the extra error-reduction term $D_{\xi}(\xi)$ in Eq. (22). This effectiveness is also observed from the sample collocation (Fig. 14). Hence, this suggests that the Kriging-based dynamic adaptive sampling using Criterion 3 is capable of estimating both smoothly and steeply varied stochastic behaviors of a solution.

The responses estimated $\hat{f}(\xi_1, \xi_2)$ using the Kriging models and PCE method are compared in Fig. 14. The PCE method yields a polynomial error due to the oscillations at $\xi_2 \leq -5$, which do not appear in the true response (Fig. 10). For the Kriging model, Criterion 1 produces the samples, which are almost normally distributed corresponding to the given PDF($\xi_1, \xi_2$) regardless of the smoothness of $\hat{f}(\xi_1, \xi_2)$. Hence Criterion 1 does not concentrate on the sandtimer region with a steep gradient around $\xi_2 = 0$ to be captured for accurate estimation of $\mu_f$ and $\sigma_f$. Criterion 2 produces more sample points around $\xi_2 = 0$ than Criterion 1 in the sandtimer region as $N$ increases. However, Criterion 2 seems careless to add sample points in the smooth region far away from $\xi_2 = 0$. Therefore, it cannot reduce the polynomial errors that cause slight oscillations in the probability density functions, as seen at $1 \leq f(\xi_1, \xi_2) \leq 1.5$ in Fig. 13(b). Compared to Criteria 1 and 2, Criterion 3 can keep a balance in sample collocation between the sandtimer region and the other smooth region. Therefore, the present test demonstrates that the proposed criterion, Criterion 3, is the best for dynamic adaptive sampling to make efficient UQ regardless of the choice of initial samples and the smoothness of the response of an output solution in the stochastic space.

5. Conclusions

This paper proposed a new Kriging-based criterion for dynamic adaptive sampling in uncertainty quantification (UQ), and compared its performance in predicting stochastic behaviors of an output solution in two-dimensional test functions (i.e., log-sin function and sandtimer function) under normally distributed uncertainties. The results demonstrated that the criterion named Criterion 2 in this paper, which is based on information about both the fit uncertainty $\delta_{\xi}(\xi)$ and gradient $\partial^2 f(\xi)/\partial \xi^2$ estimated using the Kriging surrogate model, showed better performance for effective adaptive sampling than the existing criterion (i.e., Criterion 1) based only on Kriging model fit uncertainty $\delta(\xi)$. In addition, the polynomial errors appearing near the endpoints in the stochastic space can be reduced by adding an extra term, $D_{\xi}(\xi)$, which represents the difference in Kriging predictors based on different correlation length scales, to the gradient term in the adaptive sampling criterion. Finally, the criterion proposed (i.e., Criterion 3) showed superior performance in terms of accuracy and robustness for estimating the statistics of the output solution in UQ regardless of the choice of initial samples and smoothness of the stochastic space. Furthermore, the Kriging-based dynamic adaptive sampling proposed outperformed a classical polynomial chaos expansion (PCE) method based on the Gauss quadrature rule in both test functions.

This paper focused on analytical test functions that model typical flow physics seen in the fields of aerodynamics and astronautics. Our future work will test the Kriging-based dynamic adaptive sampling proposed by applying it to more realistic UQ problems coupled with CFD to solve Navier-Stokes equations. As a matter of fact, that work is now underway.

Acknowledgments

The authors would like to thank Juan J. Alonso of Stanford University for his valuable comments, which helped to improve the dynamic adaptive sampling method, and the Sandia National Laboratories for Dakota. In addition, the authors were supported in part by the Grant-in-Aid for Scientific Research (B) No. H1503600 and the Grant-in-Aid for Scientific Research (B) No. 18H01620 administered by the Japan Society for the Promotion of Science (JSPS).

References

1) McKay, M. D., Beckman, R. J., and Conover, W. J.: A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, Technometrics, 21, 2 (1979), pp. 239–245.
2) Xiu, D. and Karniadakis, G. E.: The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations, SIAM J. Sci. Comput., 24, 2 (2002), pp. 619–644.
3) Xiu, D. and Hesthaven, J. S.: High-Order Collocation Methods for Differential Equations with Random Inputs, SIAM J. Sci. Comput., 27, 3 (2005), pp. 1118–1139.
4) Eldred, M. S. and Burkardt, J.: Comparison of Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Quantification, 47th AIAA Aerospace Sciences Meeting including The New Horizons Forum and Aerospace Exposition, AIAA Paper 2009-976, AIAA, Reston, VA, January 2009.
5) Hosder, S., Walters, R., and Perez, R.: A Non-Intrusive Polynomial Chaos Method for Uncertainty Propagation in CFD Simulations, 44th AIAA Aerospace Sciences Meeting and Exhibit, AIAA Paper 2006-891, AIAA, Reston, VA, January 2006.
6) Hosder, S., Walters, R. W., and Balch, M.: Point-Collocation Nonintrusive Polynomial Chaos Method for Stochastic Computational Fluid Dynamics, AIAA J., 48, 12 (2010), pp. 409–435.
7) Bigl, H., Lucor, D., Mishra, S., and Schwab, C. (eds.): Uncertainty
Quantification in Computational Fluid Dynamics, Lecture Notes in Computational Science and Engineering, No. 92, Springer, Heidelberg, New York, 2013.
8. Poitee, G., Despré, B., and Lucor, D.: Uncertainty Quantification for Systems of Conservation Laws, J. Comput. Phys., 228, 7 (2009), pp. 4243–4267.
9. Poitee, G. and Lucor, D.: Non Intrusive Iterative Stochastic Spectral Representation with Application to Compressible Gas Dynamics, J. Comput. Phys., 231, 23 (2012), pp. 3587–3609.
10. Lockwood, B. and Moviprilis, D.: Gradient-based Method for Uncertainty Quantification in Hypersonic Flows, Comput. Fluids, 85 (2013), pp. 27–38.
11. Witteveen, J. A. S. and Iaccarino, G.: Simplex Stochastic Collocation with Random Sampling and Extrapolation for Nonhypercube Probability Spaces, SIAM J. Sci. Comput., 34, 2 (2012), pp. A814–A838.
12. Witteveen, J. A. S. and Iaccarino, G.: Simplex Stochastic Collocation with ENO-type Stencil Selection for Robust Uncertainty Quantification, J. Comput. Phys., 239 (2013), pp. 1–21.
13. Stroud, A. H. and Secrest, D.: Gaussian Quadrature Formulas, Prentice-Hall, Englewood Cliffs, NJ, 1966.
14. Stoer, J. and Bulirsch, R.: Introduction to Numerical Analysis, Texts in Applied Mathematics, Vol. 12, Springer, New York, NY, 2002.
15. Hosder, S., Walters, R., and Balch, M.: Efficient Sampling for Non-Intrusive Polynomial Chaos Applications with Multiple Uncertain Input Variables, 9th AIAA Non-Deterministic Approaches Conference, AIAA Paper 2007-1939, AIAA, Reston, VA, April 2007.
16. Witteveen, J. A. S. and Iaccarino, G.: Refined Criteria for Simplex Stochastic Collocation with Local Extremum Diminishing Robustness, SIAM J. Sci. Comput., 34, 3 (2012), pp. A1522–A1543.
17. Palacios, F., Duraisamy, K., Alonso, J. J., and Zuzua, E.: Robust Grid Adaptation for Efficient Uncertainty Quantification, AIAA J., 50, 7 (2012), pp. 1538–1546.
18. Duraisamy, K. and Chandrashekar, P.: Goal-Oriented Uncertainty Propagation Using Stochastic Adjoint, Comput. Fluids, 66 (2012), pp. 10–20.
19. Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P.: Design and Analysis of Computer Experiments, Statist. Sci., 4, 4 (1989), pp. 409–435.
20. Jones, D. R., Schonlau, M., and Welch, W. J.: Efficient Global Optimization of Expensive Black-Box Function, J. Global Optimization, 13 (1998), pp. 455–492.
21. Jeong, S., Minemura, Y., and Obayashi, S.: Optimization of Combustion Chamber for Diesel Engine Using Kriging Model, J. Fluid Sci. Technol., 1, 2 (2006), pp. 138–146.
22. Yamazaki, W.: Stochastic Drag Analysis via Polynomial Chaos Uncertainty Quantification, Trans. Jpn. Soc. Aeronaut. Space Sci., 58 (2015), pp. 89–99.
23. Dwight, R. and Han, Z.-H.: Efficient Uncertainty Quantification Using Gradient-Enhanced Kriging, 11th AIAA Non-Deterministic Approaches Conference, AIAA Paper 2009-2276, AIAA, Reston, VA, June 2009.
24. Bilionis, I. and Zarabas, N.: Multi-Output Local Gaussian Process Regression: Applications to Uncertainty Quantification, J. Comput. Phys., 231, 17 (2012), pp. 5718–5746.
25. Rasmussen, C. E. and Williams, C. K. I.: Gaussian Processes for Machine Learning, The MIT Press, Cambridge, MA, 2006.
26. Ulaganathan, S., Couckuyt, I., Dhaene, T., Degroote, J., and Laermans, E.: High Dimensional Kriging Metamodelling Using Gradient Information, Appl. Math. Modelling, 40, 9–10 (2016), pp. 5260–5270.
27. Boopathy, K. and Rumpfkeil, M. F.: Unified Framework for Training Point Selection and Error Estimation for Surrogate Models, AIAA J., 53, 1 (2015), pp. 215–234.
28. Lukaczyk, T. W., Constantine, P., Palacios, F., and Alonso, J. J.: Active Subspaces for Shape Optimization, 10th AIAA Multidisciplinary Design Optimization Conference, AIAA Paper 2014-1171, AIAA, Reston, VA, January 2014.
29. Xiong, Y., Chen, W., Apley, D., and Ding, X.: A Non-Stationary Covariance-Based Kriging Method for Metamodelling in Engineering Design, Int. J. Numer. Methods Eng., 71, 6 (2007), pp. 733–756.
30. Toal, D. J. J. and Keane, A. J.: Non-Stationary Kriging for Design Optimization, Eng. Optim., 44, 6 (2012), pp. 741–765.
31. Clark, D. L., Jr., Bae, H.-R., Golab, K., and Pennetta, R.: Engineering Design Exploration Using Locally Optimized Covariance Kriging, AIAA J., 54, 10 (2016), pp. 3160–3175.
32. Goldberg, D. E.: Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley Publishing Company, Reading, MA, 1989.
33. Wendland, H.: Piecewise Polynomial, Positive Definite and Compactly Supported Radial Functions of Minimal Degree, Adv. Comput. Math., 4, 1 (1995), pp. 389–396.
34. Mehmani, A., Chowdhury, S., Zhang, J., and Messac, A.: Surrogate-based Design Optimization with Smart Sequential Sampling, 12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, AIAA Paper 2012-5707, AIAA, Reston, VA, September 2012.
35. Shimoyama, K., Kawai, S., and Alonso, J. J.: Dynamic Adaptive Sampling Based on Kriging Surrogate Models for Efficient Uncertainty Quantification, 15th AIAA Non-Deterministic Approaches Conference, AIAA 2013-1470, AIAA, Reston, VA, April 2013.
36. Knowles, J. and Hughes, E. J.: Multiobjective Optimization on a Budget of 250 Evaluations, Coello Coello, C. A., Aguirre, A. H., and Zitzler, E. (eds.), Evolutionary Multi-Criterion Optimization: Third International Conference, EMO 2005, Lecture Notes in Computer Science, Vol. 3410, Springer, Berlin Heidelberg, 2005, pp. 176–190.
37. Dakota Ver: 6.1, Sandia National Laboratories, Albuquerque, NM, USA, November 2014, https://dakota.sandia.gov [cited 17 January 2015].
38. Kawai, S. and Shimoyama, K.: Kriging-Model-Based Uncertainty Quantification in Computational Fluid Dynamics, 32nd AIAA Applied Aerodynamics Conference, AIAA Paper 2014-2737, AIAA, Reston, VA, June 2014.

Nobuyuki Tsuboi
Associate Editor