Efficient Quantification of Aerodynamic Uncertainties Using Gradient-Employing Surrogate Methods

Dishi Liu

Abstract. Uncertainty quantification (UQ) in aerodynamic simulations is hindered by the high computational cost of CFD models. With gradient information obtained efficiently by using an adjoint solver, gradient-employing surrogate methods are promising in speeding up the UQ process. To investigate the efficiency of UQ methods we apply gradient-enhanced radial basis functions, gradient-enhanced point-collocation polynomial chaos, gradient-enhanced Kriging and quasi-Monte Carlo (QMC) quadrature to a test case where the geometry of an RAE2822 airfoil is perturbed by a Gaussian random field parameterized by 10 independent variables. The four methods are compared in their efficiency in estimating some statistics and the probability distribution of the uncertain lift and drag coefficients. The results show that with the same computational effort the gradient-employing surrogate methods achieve better accuracy than the QMC does.

1 Introduction

In aerodynamic simulations it is beneficial to consider uncertainties in the inputs, the formulation and the numerical error of the CFD model. In this work our concern is confined to the uncertainties in the model's input and probabilistic approaches for uncertainty quantification (UQ) for CFD models. The uncertainties in the input propagates to the system response quantities (SRQ) through the model. Minor uncertainties can have an amplified impact in some instances and lead to occurrences of rare catastrophic events. Quantifying the uncertainties associated with the SRQ enhances the reliability of the simulations and enables robust design optimization. Most often this UQ process is done in a probabilistic framework in which the input

Dishi Liu
German Aerospace Center (DLR), Institute für Aerodynamik und Strömungstechnik,
Lilienthalplatz 7, 38108 Braunschweig, Germany
e-mail: dishi.liu@dlr.de

B. Eisfeld et al. (Eds.): *Management & Minimisation of Uncert. & Errors*, NNFM 122, pp. 283–296.
DOI: 10.1007/978-3-642-36185-2_12

uncertainties are represented by random variables, and the consequent uncertainties in the SRQ are quantified by determining its probability distribution or statistical moments.

However, uncertainties in the input, especially those spatially or temporally distributed, like geometric uncertainties, often generate a large number of variables. The "curse of dimensionality" prohibits the use of tensor-product quadratures. In [16] and [21] sparse grid quadratures were employed in aerodynamic UQ problems due to uncertain airfoil geometry. Nevertheless, if the number of variables is larger than 10 even sparse grid methods suffer limitations in applicability [17]. The high computational cost of CFD models also makes the traditional sampling methods such as Monte-Carlo and its variance-reduced variants (e.g. Latin Hypercube method) not efficient due to their slower error convergence rate.

Surrogate methods are gaining more attention in UQ as they provide approximations of the CFD model which are much cheaper to evaluate while maintaining a reasonable accuracy so that the UQ can be performed on the basis of a large number of samples evaluated on the surrogate model. E.g. [12] shows a Kriging surrogate method better than plain Monte Carlo and Latin Hypercube methods in estimating the mean value of a bivariate Rosenbrock function. A comparative study of surrogate methods that are not employing gradients [23] shows Kriging is more accurate than radial basis functions and multivariate polynomial in approximating some 10-variate test functions.

Gradient-employing also give an edge to surrogate methods if the gradients are obtained at a relatively lower cost than that of the SRQ, which is the case when an adjoint CFD solver [5] is used and the number of SRQ is less than the number of variables. It should be noted that the gradient information cannot be effectively utilized by the UQ methods based on direct sampling of the CFD model. A naive augmentation of samples by finite difference brings no benefit because the augmenting samples are not statistically independent of the original ones.

Different sampling schemes are adopted by surrogate methods, majorly of two groups: "on-grid" sampling and scattered sampling. The former is used in some methods based on polynomial approximations, e.g. stochastic collocation methods [2], and affected by the "curse of dimensionality" if the number of variables is large. The latter is more robust since it admits an arbitrary number of samples and arbitrary sample sites. This flexibility not only makes it tolerate sample failures (due to, e.g. poor convergence, as often observed in CFD models), but also makes an incorporation of pre-existing or additional samples possible and enables run-time adaptive sampling.

In this work we apply three gradient-employing surrogate methods, i.e. gradient-enhanced radial basis functions (GERBF), gradient-enhanced point-collocation polynomial chaos (GEPC) and gradient-enhanced Kriging (GEK) [13], and for the purpose of comparison, also the quasi-Monte Carlo quadrature, to a UQ test case where an RAE2822 airfoil is subject to random geometric perturbations, and we compare their efficiency in estimating some statistics and probability distribution of the resulting uncertain lift and drag coefficients. The number of CFD model

evaluations is kept small (≤ 200) in this numerical comparison to make it relevant to large-scale industrial applications.

2 Test Case

The test case we use in this work is a CFD model of the inviscid flow around a 2-dimensional RAE2822 airfoil at a Mach number of 0.73 and an angle of attack of 2.0 degrees. The source of uncertainty is the randomly perturbed airfoil geometry, i.e., the lower and upper surfaces of the airfoil's 2D section (as shown by the solid line in the right part of Figure 1) are each assumed to be subject to a Gaussian random perturbation in the direction normal to the surface. Let $\boldsymbol{p}_l$ and $\boldsymbol{p}_u$ denotes the original lower and upper surface respectively, the perturbed surfaces are

$$\boldsymbol{p}'_l = \boldsymbol{p}_l + \boldsymbol{n} \cdot \theta_l(x)$$

$$\boldsymbol{p}'_u = \boldsymbol{p}_u + \boldsymbol{n} \cdot \theta_u(x)$$

with $x \in [0,1]$. $\boldsymbol{n}$ is the local normal vector of the surface, $\theta(x)$ is a zero-mean Gaussian variable with standard deviations $\sigma(x)$, i.e.

$$\theta_l(x), \theta_u(x) \sim N(0, \sigma(x))$$

in which

$$\sigma(x) = 0.01 \cdot Z_{max} \cdot x(1-x) \cdot \beta(2,2)/1.5$$

with Z_{max} the maximum half-thickness of the airfoil, and β the Beta function. This setting makes the $\sigma(x)$ have its maximum (one percent of Z_{max}) at $x = 0.5$ and being zero at the two ends of the airfoil.

It is assumed that the random deformation is spatially correlated by a Gaussian type correlation function, i.e.

$$\begin{aligned} \mathrm{cov}[\theta_l(x_1), \theta_l(x_2)] &= \sigma(x_1)\sigma(x_2) \exp\left(-\frac{(x_1 - x_2)^2}{\ell^2}\right) \\ &= \mathsf{C}(x_1, x_2) \end{aligned}$$

with $\ell = 0.2$. The same also applies to $\theta_u(x)$.

For the purpose of numerical computation, the correlated random fields $\theta_l(x)$ and $\theta_u(x)$ need to be represented in terms of uncorrelated random variables. This is furnished by Karhunen-Loève expansions (KLE) [1], e.g. for θ_l,

$$\theta_l(x) = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\, \xi_i \Phi_i(x)$$

where ξ_i are independent standard Gaussian variables. λ_i and $\Phi_i(x)$ are the eigenvalues and the eigenfunctions of $\mathsf{C}(x_1, x_2)$, i.e., the solutions of the following integral equation,

$$\int_0^1 \mathsf{C}(x_1,x_2)\Phi_i(x_1)\mathrm{d}x_1 = \lambda_i \Phi_i(x_2) \quad \forall\, i = 1,2,\ldots$$

For practical problems the KLE needs to be truncated so that only a relatively small number of terms is kept, e.g. an approximation with κ terms:

$$\theta_l(x) \approx \sum_{i=1}^{\kappa} \sqrt{\lambda_i}\,\xi_i \Phi_i(x)$$

By taking $\kappa = 5$, $\theta_l(x)$ is parameterized by 5 independent standard Gaussian variables. Applying the same approximation to $\theta_u(x)$,

$$\theta_u(x) \approx \sum_{i=\kappa+1}^{2\kappa} \sqrt{\lambda_i}\,\xi_i \Phi_i(x)$$

we express the randomly perturbed airfoil surface as a function of 10 such variables. This KLE representation is optimal in the sense that it retains the original geometric variance to the maximum degree compared to any other linear-form representation with the same number of variables [1]. Figure 1 shows three examples of random perturbation in the upper and lower surface together with the corresponding perturbed RAE2822 airfoil geometry.

In this test case, the CFD model takes the input variables $\boldsymbol{\xi} = \{\xi_1, \cdots, \xi_{10}\}$ and yields the lift and drag coefficients, C_ℓ and C_d, of the randomly perturbed airfoil. Hereafter, the model is denoted as $f(\boldsymbol{\xi})$ in this paper. We compare the efficiency of the candidate methods in estimating some target statistics, i.e. the means (μ_ℓ, μ_d), the standard deviations (σ_ℓ, σ_d) of C_ℓ and C_d, and the exceedance probabilities $P_{\ell,j} =$

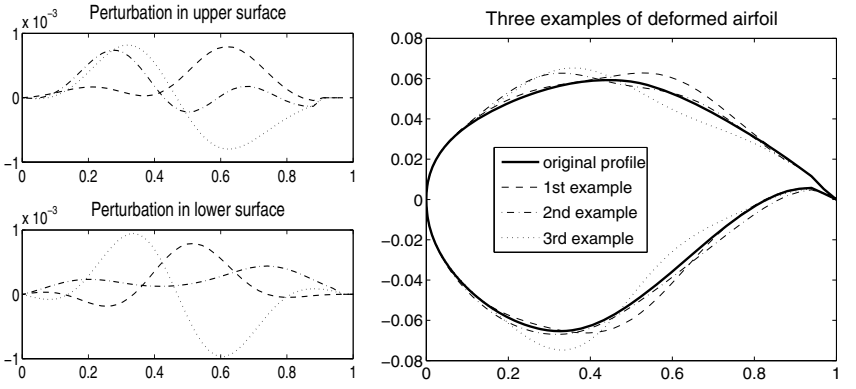

Fig. 1 Three examples of random perturbation in upper and lower surface (left) and three examples of the randomly perturbed airfoil geometry, with the perturbations ten-times exaggerated (right)

$\mathrm{Pro}\{C_\ell < \mu_\ell - j \cdot \sigma_\ell\}$ and $P_{d,j} = \mathrm{Pro}\{C_d > \mu_d + j \cdot \sigma_d\}$ with $j = 2, 3$. The accuracy of the statistics is judged by comparing with reference statistics obtained by a QMC integration with a large sample number ($N = 2 \times 10^5$).

3 Methods

Four methods are applied to the test case. They include three surrogate methods, i.e. gradient-enhanced radial basis functions (GERBF), gradient-enhanced Kriging (GEK) and gradient-enhanced point-collocation polynomial chaos (GEPC), and one direct integration method, i.e. quasi-Monte Carlo (QMC) quadrature. An introduction of them is made in this section.

Since the gradients of the SRQ with respect to all the ten variables are computed by an adjoint solver at an additional cost of approximately one evaluation of the CFD model, to account for this additional cost we introduce the term *elapsed time-penalized sample number M* by making $M = 2N$ for the three gradient-employing methods and $M = N$ for QMC, with N the number of evaluations of the CFD model. Compared to the cost of evaluating the CFD model the computational overhead of constructing surrogates is negligible, so in the efficiency comparison we use M as the measure of computational cost.

In the aspect of design of experiment, the study in [23] shows surrogate models based on samples with relatively high degree of uniformity (using Latin Hyper-cube sampling) are more accurate than those based on samples of lower degree of uniformity (using plain Monte-Carlo sampling). For all the four methods in this work we adopt the QMC sampling scheme [7] because it achieves even higher degree of sample uniformity than Latin Hyper-cube sampling.

We use the DLR-TAU code [10] to solve the CFD model. The geometry perturbation is implemented by using a mesh deformation tool based on radial basis functions incorporated in the DLR-TAU code as described in [14].

3.1 Quasi-Monte Carlo Quadrature

Quasi-Monte Carlo (QMC) quadrature [7] samples at a *low discrepancy* set of points generated by deterministic number-theoretic formulas. The "discrepancy" here is a measurement of how much the distribution of this set of points deviates from the underlying pdf. A low discrepancy set of points achieves a higher degree of uniformity with respect to a given pdf than a pseudo-random set of points does. So QMC is usually much more efficient than a Monte Carlo quadrature. The error bound of QMC is of order $O(N^{-1}(\log N)^d)$ in which d is the number of variables. In many cases this is quite a loose upper bound of the error, i.e. QMC often performs better than that.

A variety of low discrepancy point sets exist, e.g. Van der Corput, Halton, Sobol, Hammersley and Niederreiter point set. The last one is used in this work as it is considered the most efficient when d is large [19]. The point set is generated by the program from [4]. The statistics of the SRQ are directly computed from the samples.

3.2 Gradient-Enhanced Radial Basis Functions

The radial basis function (RBF) method [6] approximates an unknown function by a weighted linear combination of *radial basis functions* each being radially symmetric about a center. An RBF approximation takes the form

$$\hat{f}(\boldsymbol{\xi}) = \sum_{i=1}^{N} w_i \, \phi_i(\|\boldsymbol{\xi} - \boldsymbol{\xi}^{\langle i \rangle}\|)$$

where ϕ_i are radial basis functions, $\|\cdot\|$ denotes the Euclidean norm, and $\boldsymbol{\xi}^{\langle i \rangle}$ are the N sample points each taken as the center of a radial basis function. Making $\hat{f}(\boldsymbol{\xi})$ interpolate the N samples leads to N linear equations. The coefficients w_i are determined by solving this linear system.

Denoting the Euclidean distance from the center as r, popular types of $\phi(r)$ include $\sqrt{r^2+a^2}$ (multiquadric), $1/\sqrt{r^2+a^2}$ (inverse multiquadric), $\exp(-a^2r^2)$ (Gaussian) and $r^2\ln(ar)$ (thin plate spline), in which a is a parameter to be fine-tuned for a particular set of samples. Gradient-employing versions of RBF were proposed in [11, 20] where first-order derivatives of the SRQ are exploited and second-order derivatives of RBF are involved in the system.

In this work we propose a different gradient-employing RBF method that involves only the first-order derivative of RBF, termed gradient-enhanced RBF (GERBF). To accommodate the gradient informations of the SRQ, this method introduces additional RBF that are centered at *non-sampled* points, i.e. an GERBF approximation is

$$\hat{f}(\boldsymbol{\xi}) = \sum_{i=1}^{K} w_i \, \phi_i(\|\boldsymbol{\xi} - \boldsymbol{\xi}^{\langle i \rangle}\|), \quad \text{with} \;\; N < K \leq N(1+d)$$

The $\boldsymbol{\xi}^{\langle i \rangle}$ with $i \leq N$ are sampled points, those with $i > N$ are non-sampled points which can be chosen randomly as long as none of them duplicates the sampled ones. The coefficients $\boldsymbol{w} = \{w_0, w_1, \cdots, w_K\}^T$ are determined by solving the following system,

$$\boldsymbol{\Phi w} = \boldsymbol{f}$$

in which

$$\boldsymbol{\Phi} = \begin{bmatrix} \Phi_1(\boldsymbol{\xi}^{\langle 1\rangle}) & \Phi_2(\boldsymbol{\xi}^{\langle 1\rangle}) & \cdots & \Phi_K(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_1(\boldsymbol{\xi}^{\langle N\rangle}) & \Phi_2(\boldsymbol{\xi}^{\langle N\rangle}) & \cdots & \Phi_K(\boldsymbol{\xi}^{\langle N\rangle}) \\ \Phi_1^{(1)}(\boldsymbol{\xi}^{\langle 1\rangle}) & \Phi_2^{(1)}(\boldsymbol{\xi}^{\langle 1\rangle}) & \cdots & \Phi_K^{(1)}(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_1^{(1)}(\boldsymbol{\xi}^{\langle N\rangle}) & \Phi_2^{(1)}(\boldsymbol{\xi}^{\langle N\rangle}) & \cdots & \Phi_K^{(1)}(\boldsymbol{\xi}^{\langle N\rangle}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_1^{(d)}(\boldsymbol{\xi}^{\langle 1\rangle}) & \Phi_2^{(d)}(\boldsymbol{\xi}^{\langle 1\rangle}) & \cdots & \Phi_K^{(d)}(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_1^{(d)}(\boldsymbol{\xi}^{\langle N\rangle}) & \Phi_2^{(d)}(\boldsymbol{\xi}^{\langle N\rangle}) & \cdots & \Phi_K^{(d)}(\boldsymbol{\xi}^{\langle N\rangle}) \end{bmatrix}, \qquad \boldsymbol{f} = \begin{bmatrix} f(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots \\ f(\boldsymbol{\xi}^{\langle N\rangle}) \\ f^{(1)}(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots \\ f^{(1)}(\boldsymbol{\xi}^{\langle N\rangle}) \\ \vdots \\ f^{(d)}(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots \\ f^{(d)}(\boldsymbol{\xi}^{\langle N\rangle}) \end{bmatrix}$$

with $\Phi_i^{(j)} = \partial\Phi_i/\partial\xi_j$, $f^{(j)} = \partial f/\partial\xi_j$. We chose $K = \frac{N}{2}(1+d)$ in this work, which results in an over-determined system that is solved by a Least Squares method.

A numerical comparison of the accuracy of the aforementioned four types of RBF in approximating this CFD model $f(\boldsymbol{\xi})$ was made by the author. The result favors the inverse multiquadric RBF which is therefore used in this work for the comparison with other UQ methods. The internal parameter a is fine-tuned by a *leave-one-out* error minimizing procedure as in [3].

For this UQ job we first establish a GERBF surrogate model of $f(\boldsymbol{\xi})$ based on QMC samples of the CFD model, and integrate for the target statistics and pdf by a large number (10^5) of QMC samples on the surrogate model.

3.3 Gradient-Enhanced Kriging Method

Kriging [9] approximates $f(\boldsymbol{\xi})$ by a weighted linear combination of samples, i.e.

$$\hat{f}(\boldsymbol{\xi}) = \gamma(\boldsymbol{\xi}) + \sum_{i=1}^{N} w_i(\boldsymbol{\xi}^{\langle i\rangle}) f(\boldsymbol{\xi}^{\langle i\rangle})$$

where $f(\boldsymbol{\xi}^{\langle i\rangle})$ are N samples of the SRQ. γ and w_i are determined by minimizing the variance of the error $\mathfrak{e} = f - \hat{f}$ with the assumptions that the expectation of $\mathfrak{e}$ is zero and that $f(\boldsymbol{\xi})$ honors a spatial correlation model. We use direct gradient-enhanced Kriging (GEK) [8] that incorporates gradient information as secondary samples by an extended spatial correlation model that accommodates gradients. We implement GEK using the *Surrogate-Modeling for Aero-Data Toolbox (SMART)* [13] developed at DLR, opting for ordinary Kriging and a correlation model of spline type which is considered the most efficient in similar situations in [17]. The internal parameters of the correlation model are fine-tuned to fit the sampled data by a maximum likelihood estimation [24].

For this UQ job we first establish a GEK surrogate model of $f(\boldsymbol{\xi})$ based on QMC samples of the CFD model, and integrate for the target statistics and pdf by a large number (10^5) of QMC samples on the surrogate model.

3.4 Gradient-Enhanced Point-Collocation Polynomial Chaos Method

According to Wiener [22], $f(\boldsymbol{\xi})$ can be approximated by a truncated *polynomial chaos expansion* (PCE)

$$\hat{f}(\boldsymbol{\xi}) = \sum_{i=0}^{K} c_i \Psi_i(\boldsymbol{\xi}) \tag{1}$$

where Ψ_i is Hermite polynomial chaos (PC) to which a detailed description can be found in, e.g. [18]. The total number of terms is $K = (p+d)!/(p!d!)$ with p the order of PC.

To determine the coefficients c_i we use a point-collocation method similar to the one used in [15], the difference being that we utilize gradient information. In this gradient-enhanced point-collocation polynomial chaos (GEPC) method the $\boldsymbol{c} = \{c_0, c_1, \cdots, c_K\}^T$ is determined by solving the following system,

$$\boldsymbol{\Psi c} = \boldsymbol{f}$$

$$\text{with } \boldsymbol{\Psi} = \begin{bmatrix} \Psi_0(\boldsymbol{\xi}^{\langle 1\rangle}) & \Psi_1(\boldsymbol{\xi}^{\langle 1\rangle}) & \cdots & \Psi_K(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_0(\boldsymbol{\xi}^{\langle N\rangle}) & \Psi_1(\boldsymbol{\xi}^{\langle N\rangle}) & \cdots & \Psi_K(\boldsymbol{\xi}^{\langle N\rangle}) \\ \Psi_0^{(1)}(\boldsymbol{\xi}^{\langle 1\rangle}) & \Psi_1^{(1)}(\boldsymbol{\xi}^{\langle 1\rangle}) & \cdots & \Psi_K^{(1)}(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_0^{(1)}(\boldsymbol{\xi}^{\langle N\rangle}) & \Psi_1^{(1)}(\boldsymbol{\xi}^{\langle N\rangle}) & \cdots & \Psi_K^{(1)}(\boldsymbol{\xi}^{\langle N\rangle}) \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_0^{(d)}(\boldsymbol{\xi}^{\langle 1\rangle}) & \Psi_1^{(d)}(\boldsymbol{\xi}^{\langle 1\rangle}) & \cdots & \Psi_K^{(d)}(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_0^{(d)}(\boldsymbol{\xi}^{\langle N\rangle}) & \Psi_1^{(d)}(\boldsymbol{\xi}^{\langle N\rangle}) & \cdots & \Psi_K^{(d)}(\boldsymbol{\xi}^{\langle N\rangle}) \end{bmatrix}, \quad \boldsymbol{f} = \begin{bmatrix} f(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots \\ f(\boldsymbol{\xi}^{\langle N\rangle}) \\ f^{(1)}(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots \\ f^{(1)}(\boldsymbol{\xi}^{\langle N\rangle}) \\ \vdots \\ f^{(d)}(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots \\ f^{(d)}(\boldsymbol{\xi}^{\langle N\rangle}) \end{bmatrix}$$

where $\Psi_i^{(k)} = \partial \Psi_i / \partial \xi_k$, $f^{(k)} = \partial f / \partial \xi_k$, and $\boldsymbol{\xi}^{\langle i\rangle} = \{\xi_1, \xi_2, \cdots, \xi_d\}_i$ with $i = 1, \cdots, N$ denote the sample points. The K is chosen to be half of the number of available "conditions", $N(1+d)$, for the best performance according to [15]. This over-determined system is solved by a Least Squares method.

For this UQ job we first establish a GEPC surrogate model of $f(\boldsymbol{\xi})$ based on QMC samples of the CFD model, and compute the mean and the variance of $f(\boldsymbol{\xi})$ directly from the coefficients,

$$\mu = c_0\,, \qquad \sigma^2 = \sum_{i=1}^{K} (c_i)^2 \cdot \mathbb{E}[\Psi_i^2(\boldsymbol{\xi})] \tag{2}$$

The exceedance probabilities and pdf are integrated by a large number (10^5) of QMC samples on the surrogate model.

4 Results and Discussion

The results of the efficiency comparison are shown in Figure 2 to 5. Figure 2 and 3 show the errors of the four methods in estimating the target statistics of C_ℓ and C_d. It is observed there that the three gradient-employing surrogate methods, GEK, GEPC and GERBF are more efficient than the QMC method since the former three reduce their errors faster with an increasing cost measure M. Figure 4 and 5 depict the estimated pdf of C_ℓ and C_d obtained by the four methods, comparing with the reference pdf. There we see that for the same computational cost, the surrogate methods yield much more accurate pdf's. This is consistent with their relative performance in estimating the statistics.

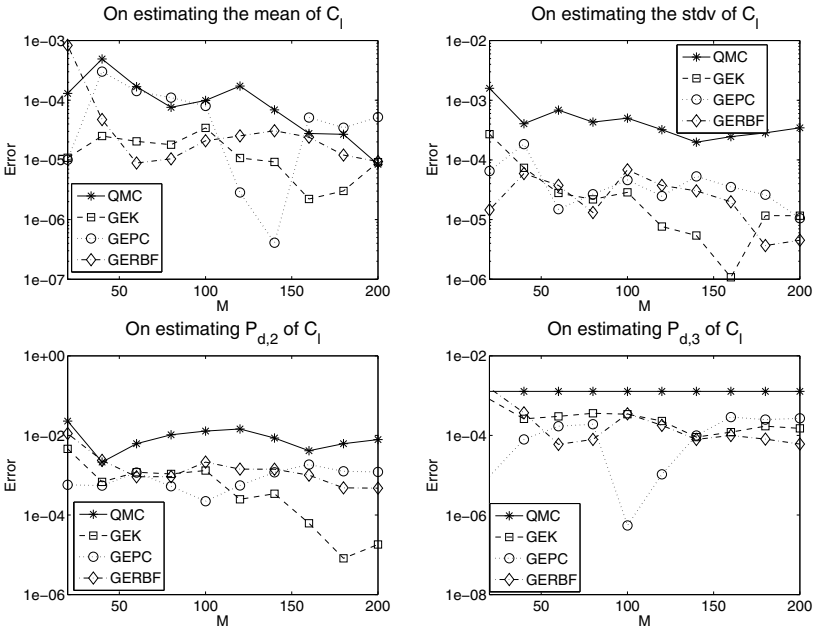

Fig. 2 Error in estimating mean, standard deviation (upper row) and exceedance probabilities (lower row) of C_ℓ

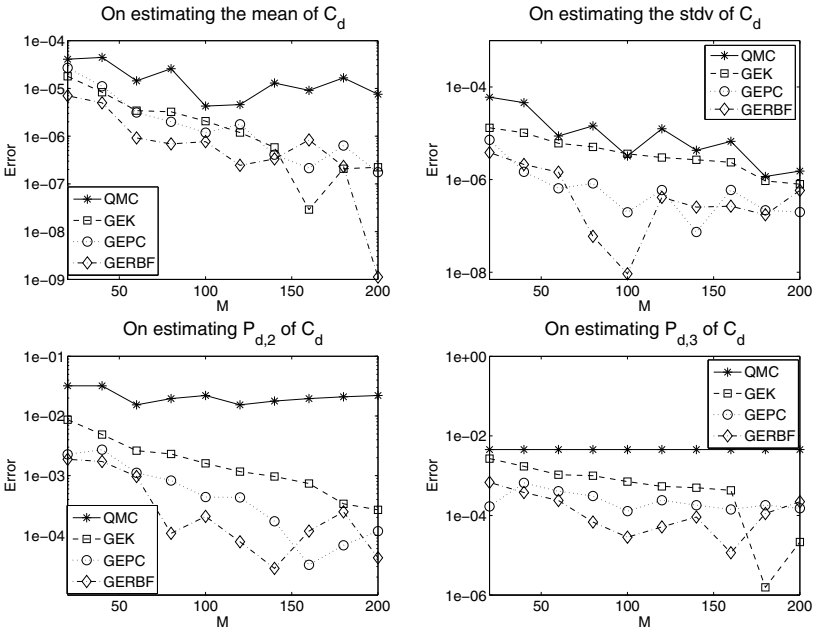

Fig. 3 Error in estimating mean, standard deviation (upper row) and exceedance probabilities (lower row) of C_d

One of the reasons for the relatively better performance of the surrogate methods is that they utilize more information with the same computational cost M, i.e. they use $(1+d)\frac{M}{2}$ conditions (SRQ samples and gradients) while a direct integration method like QMC uses M conditions (SRQ samples only). This advantage is due to the cheaper cost of obtaining gradients by an adjoint solver in the case that the number of SRQ's is smaller than the number of variables d, and is expected to increase with an increasing number of variables, d.

Although it seems that GEK and GERBF perform better than the other surrogate methods in estimating statistics of C_ℓ and C_d respectively, it may not be appropriate to base a general conclusion on that. In Figure 4 and 5 we see the three surrogate methods have similar accuracy in their estimated pdf of C_ℓ and C_d.

The efficiency of GEK or GERBF is sensitive to the choice of the covariance model or the radial basis function and also to the value of the internal parameters, and excelling configurations of them are problem- and data-dependent. In this work, different techniques are used for the optimization of the internal parameters, i.e., maximum likelihood optimization for GEK and leave-one-out error minimization for GERBF. This may also influence their relative efficiency, possibly differently in the C_ℓ and C_d cases. Due to the complex nature of comparative efficiency of

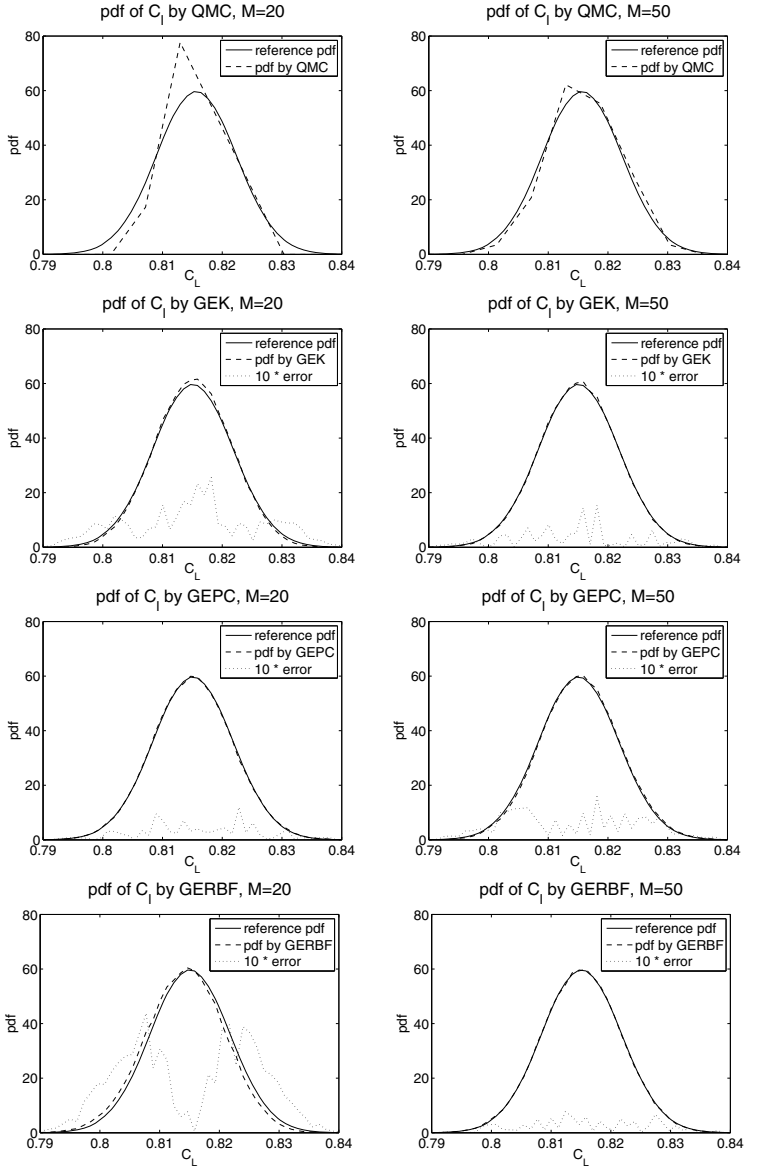

Fig. 4 Estimated pdf (in dash line) of C_ℓ by QMC (1st row), GEK (2nd row), GEPC (3rd row) and GERBF(4th row) at $M = 20$ (left) and $M = 50$ (right), dotted line shows the 10-times scaled up error of the estimated pdf

surrogate methods with different configurations and internal optimization techniques and different target SRQ, here we do not try to draw a conclusion on this issue.

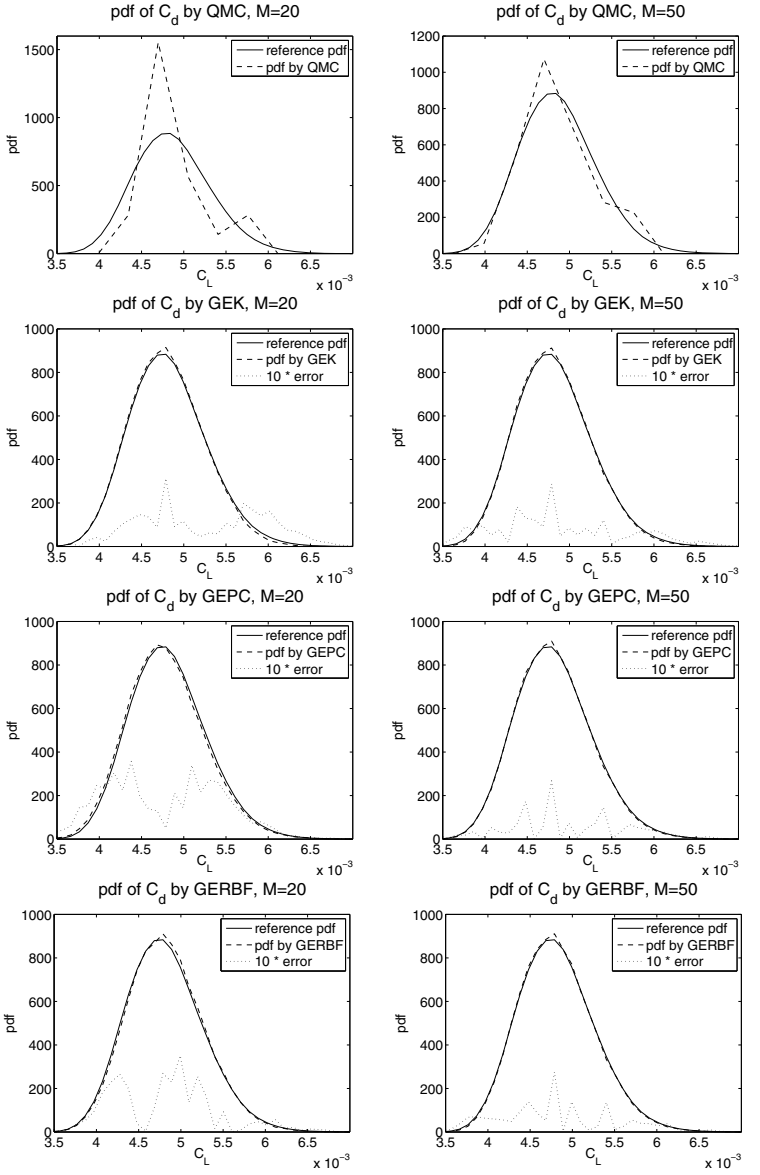

Fig. 5 Estimated pdf (in dash line) of C_d by QMC (1st row), GEK (2nd row), GEPC (3rd row) and GERBF(4th row) at $M = 20$ (left) and $M = 50$ (right), dotted line shows the 10-times scaled up error of the estimated pdf

In the estimation of the statistics of C_ℓ we see GEPC is not always reducing its error with an increasing M. This might be ascribed to the fact that the number of polynomial chaos (PC) terms is not truncated according to the order of PC, but to

an arbitrary number which is half of the number of available conditions. A set of PC terms that is "incomplete" for a particular order might not lead to more accurate approximations than a set with less number of terms but "complete" for a lower order. Nevertheless, GEPC has a favored property that we have no burden of choosing the best-fitting configuration for it.

5 Summary

Gradient-employing surrogate methods have an advantage in handling aerodynamic uncertainty quantification (UQ) problems in the cases that an adjoint solver is used and the number of system response quantities (SRQ) is smaller than the number of variables so that the gradients of SRQ can be obtained at a reduced cost. These methods construct surrogates of the CFD model so that the statistics of an uncertain SRQ can be integrated on the surrogates models.

For investigating the efficiency of the different UQ methods we set up a test case where the geometry of an RAE2822 airfoil is perturbed by a Gaussian random field which is parameterized by 10 independent variables. Three surrogate methods, gradient-enhanced radial basis functions, gradient-enhanced point-collocation polynomial chaos and gradient-enhanced Kriging, together with a direct integration method, quasi-Monte Carlo (QMC) quadrature, are applied to the test case and compared in their efficiency in estimating some statistics and probability distribution of the uncertain lift and drag coefficients. The results show that with the same computational effort the gradient-employing surrogate methods achieve better accuracy than the QMC does.

References

[1] Adler, R.J., Taylor, J.E.: Random Fields and Geometry. Springer, Berlin (2007)
[2] Babuška, I., Nobile, F., Tempone, R.: A stochastic collocation method for elliptic partial differential equations with random input data. SIAM J. Numer. Anal. 45, 1005–1034 (2007)
[3] Bompard, M., Peter, J., Désidéri, J.A.: Surrogate models based on function and derivative values for aerodynamic global optimization. In: V European Conference on Computational Fluid Dynamics, ECCOMAS CFD 2010, Lisbon, Portugal (2010)
[4] Bratley, P., Fox, B.L., Niederreiter, H.: Programs to generate Niederreiter's low-discrepancy sequences. ACM Trans. Math. Softw. 20(4), 494–495 (1994)
[5] Brezillon, J., Dwight, R.: Discrete adjoint of the navier-stokes equations for aerodynamic shape optimization. In: EUROGEN 2005 - Sixth Conference on Evolutionary and Deterministic Methods for Design, Optimization and Control with Applications to Industrial and Societal Problems, Munich, Germany (2005)
[6] Buhmann, M.D.: Radial basis functions. Acta Numerica 9, 1–38 (2000)
[7] Caflisch, R.E.: Monte Carlo and quasi-Monte Carlo methods. Acta Numerica 7, 1–49 (1998)

[8] Chung, H.S., Alonso, J.J.: Using gradients to construct cokriging approximation models for high-dimensional design optimization problems. AIAA paper 2002–0317 (2002)
[9] Cressie, N.: Statistics for Spatial Data. Wiley, New York (1991)
[10] Gerhold, T., Friedrich, O., Evans, J., Galle, M.: Calculation of complex three-dimensional configurations employing the DLR-TAU-Code. AIAA paper 97-0167 (1997)
[11] Giannakoglou, K.C., Papadimitriou, D.I., Kampolis, I.C.: Aerodynamic shape design using evolutionary algorithms and new gradient-assisted metamodels. Computer Methods in Applied Mechanics and Engineering 195(44-47), 6312–6329 (2006)
[12] Giunta, A.A., Eldred, M.S., Castro, J.P.: Uncertainty quantification using response surface approximation. In: 9th ASCE Specialty Conference on Probabolistic Mechanics and Structural Reliability, Albuquerque, New Mexico, USA (2004)
[13] Han, Z.H., Görtz, S., Zimmermann, R.: Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. Journal of Aerospace Science and Technology (2012)
[14] Heinrich, R., Kroll, N., Neumann, J., Nagel, B.: Fluid-structure coupling for aerodynamic analysis and design: a DLR perspective. AIAA paper 2008–561 (2008)
[15] Hosder, S., Walters, R.W., Balch, M.: Efficient sampling for non-intrusive polynomial chaos applications with multiple uncertain input variables. AIAA paper 2007–1939 (2007)
[16] Litvinenko, A., Matthies, H.G.: Sparse data formats and efficient numerical methods for uncertainties quantification in numerical aerodynamics. In: ECCM IV: Solids, Structures and Coupled Problems in Engineering (2010)
[17] Liu, D.S.: A Best Practice Guide: Efficient Quantification of Aerodynamic Uncertainties. Tech. rep., DLR, IB 124-2012/2 (2012)
[18] Matthies, H.G.: Uncertainty quantification with stochastic finite elements. In: Stein, E., de Borst, R., Hughes, T.R.J. (eds.) Encyclopedia of Computational Mechanics. John Wiley & Sons, Chichester (2007)
[19] Niederreiter, H.: Random Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia (1992)
[20] Ong, Y., Lum, K., Nair, P.: Hybrid evolutionary algorithm with hermite radial basis function interpolants for computationally expensive adjoint solvers. Computational Optimization and Applications 39, 97–119 (2008)
[21] Schillings, C., Schulz, V.: On the influence of robustness measures on shape optimization with stochastic uncertainties. Tech. rep., Trierer Forschungsbericht 12-2, Universität Trier (2012)
[22] Wiener, N.: The homogeneous chaos. Amer. J. Math. 60, 897–936 (1938)
[23] Zhao, D., Xue, D.: A comparative study of metamodeling methods considering sample quality merits. Struc. Multidisc. Optim. (42), 923–938 (2010)
[24] Zimmermann, R.: Asymptotic behavior of the likelyhood function of covariance matrices of spatial gaussian processes. J. App. Math., Article ID 494070 (2010)