

Surrogate Model-Based Approaches to UQ and Their Range of Applicability



Daigo Maruyama, Dishi Liu and Stefan Görtz

Introduction

In the context of aerodynamic design under uncertainty, surrogate modeling is considered as one of the suitable approaches to efficiently calculate statistics of the quantity of interest (QoI) under scattered data. The surrogate model-based approaches to UQ here in this chapter are the method that the statistics ideally computed by a large number of data information are obtained by complementary data by an assistance of surrogate models in the uncertainty parameters space. The scattered data as sample points is produced by using Design of Experiments (DOE) and adaptive sampling if necessary in this chapter. The dependency of the statistics of the QoI on the number and distribution of sample points used to build a surrogate model and on the kind of surrogate model is discussed in [1]. In the case of robust design, the statistical value of interest is the sum of the mean and standard deviation of the QoI, or its maximum value. Each of them is considered as the objective function in optimization processes. Note that QoI here is lift coefficient (C_l) or drag coefficient (C_d) evaluated by a CFD computation.

In this chapter as Best Practice Guide it is discussed which kind of methods are the most efficient for computing the statistics of QoI as the objective function in certain tolerances of accuracy compared to the reference, e.g., one drag count ($=10^{-4}$). The errors less than this order can be sometimes regarded as epistemic

D. Maruyama (✉) · D. Liu · S. Görtz

German Aerospace Center (DLR), Institute of Aerodynamics and Flow Technology,
Braunschweig, Germany

e-mail: Daigo.Maruyama@dlr.de

D. Liu

e-mail: Dishi.Liu@dlr.de

S. Görtz

e-mail: Stefan.Goertz@dlr.de

uncertainties due to imperfectness of CFD solvers. To accurately and efficiently compute the statistical values of interest, we focus on the following three aspects:

- Type of surrogate model;
- Number of sample points (used to build the surrogate model); and
- Distribution of sample points (used to build the surrogate model).

The type of surrogate models and a sufficient number of sample points are firstly shown in section “[Selection of Surrogate Models and the Number of Samples](#)”. Then, efficient sampling techniques considering both the number and distribution of sample points are introduced in section “[Sampling Techniques for Different Measures of Robustness](#)” for computing the above-mentioned two kinds of objective functions in the robust design optimization. The CFD solver used to evaluate the aerodynamic coefficients on the sample points is the DLR-TAU-code [2–4]. Fully turbulent computations were performed with the negative Spalart–Allmaras turbulence model [5]. A quasi-two-dimensional hybrid unstructured grid with prisms and tetrahedral elements was used for the RANS simulations.

Selection of Surrogate Models and the Number of Samples

The points to discuss here are which surrogate model is used and how many sample points are selected. The direct integration of quasi-Monte Carlo (QMC) sampling, Kriging, and GEK are compared. The comparison is performed in terms of the accuracy of the statistics for a given number of samples used to build the surrogate model. The influence of different numbers of samples is also studied. The distribution of the sample points is based on the Sobol sequence [6–8], maintaining a high degree of “uniformity” (low-discrepancy) of samples even in high-dimensional cases (≥ 10). Figure 1 shows the distributions of mean and standard deviation of estimated lift coefficient (C_l) [9]. Details of the test case are introduced in [9]. GEK requires the gradients with respect to the input uncertainty parameters, which can be efficiently computed by an adjoint solver. Therefore, $N_c = 2 N$ in case of GEK where N is the number of sample points. Note that the input uncertainty space is 26 dimensions in this test case.

As can be observed in Fig. 1, GEK has comprehensively less errors than the others and converges faster than Kriging along with increase of the number of samples. One can observe that the errors of them when the number of sample points is more than around 15 ($N_c \approx 30$ in Fig. 1) nearly converge to the reference. This is one reason why GEK is recommended.

Another reason to use GEK is further efficiency in high-dimensional cases. When GEK is used, the scattered data information to build a GEK surrogate model is efficiently replenished since the computational cost of an adjoint solver is independent of the dimensionality. This could compensate one of the bottlenecks that the number of sample points should be increased with of the dimensionality

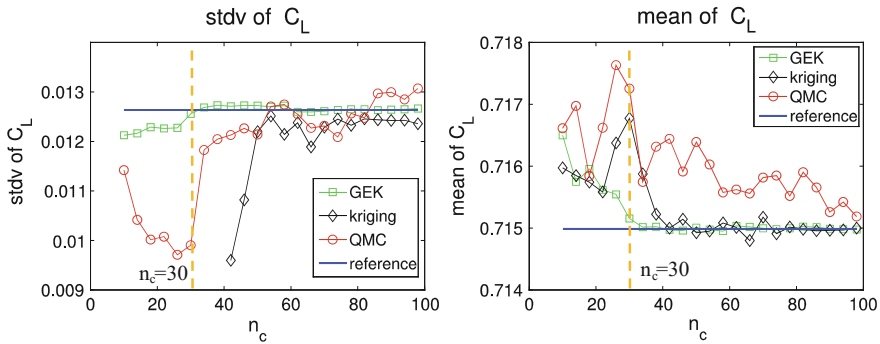


Fig. 1 Convergence of estimate C_l statistics (mean and standard deviation) to the reference statistics by various UQ methods (note that $N_c = 2 N$ in case of GEK because only the gradient of C_l was considered, while $N_c = N$ for Kriging and QMC) [10]

(details can be referred in [10]). We judge that the required number of sample points to satisfy the good accuracy does not change so much even if the dimensionality increases.

Because of the above-mentioned reasons, our conclusion of selecting surrogate models is GEK when the gradients of the QoI are able to be calculated efficiently by an adjoint solver. The number of sample points can be more than around 15. More details on the number of sample points and sampling techniques are introduced in the next section.

Sampling Techniques for Different Measures of Robustness

Criteria to Assess the Accuracy of the Statistics of QoI

Here, we focus also on the accuracy of specific statistical values of the QoI by GEK (also Kriging as comparison) with different distributions of the sample points. Two fixed numbers of sample points (12 and 30) are used to be compared with each other. The QoI considered here is the drag coefficient (C_d). The measures of robustness (objective functions) f considered here are:

$$f \equiv \mu_{C_d} + \sigma_{C_d} \tag{1}$$

$$f \equiv \max_{\mathbf{u}}(C_d(\mathbf{u})) \tag{2}$$

where \mathbf{u} denotes input uncertainty parameters whose dimensionality is 12 in the applications to the UMRIDA BC-02 test case. The optimizations of these stochastic quantities are called “expectation measure with mean-risk approach” and

“worst-case risk measure,” respectively. These statistics expressed by Eqs. (1) and (2) are ideally uniquely determined under fixed probability density functions (pdfs) of the input uncertainty parameters \mathbf{u} . Note that the pdfs of the input uncertainty parameters are assumed to be normally distributed. Details of these equations on how to calculate f by using surrogate models can be seen in [1].

The accuracy is assessed in terms of the following three criteria:

- (1) The expected value (mean) μ_f of f obtained for different distributions of the sample points;
- (2) The dispersion (standard deviation) σ_f of f obtained for different distributions of the sample points; and
- (3) The influence of the above two values μ_f and σ_f on the result of robust design, f_{opt}, χ_{opt} .

The first and second criteria are to investigate the accuracy of f itself. The third criterion is then for examining the accuracy of the optimal solutions in terms of f_{opt}, χ_{opt} in applications to robust design. The closer the mean μ_f is to the reference f_{ref} and the closer the standard deviation σ_f is to 0, the better the accuracy of the estimated f .

The different sets of the sample points are achieved by consecutive rows in the Sobol sequence where “uniformity” (low-discrepancy) is maintained. Three different sets of 30 sample points in two dimensions are shown in Fig. 2i as an example. Each set of sample points was constructed by extracting 30 consecutive rows in the Sobol sequence. These different sets of sample points are transformed by using the cumulative density function (cdf) of uncertain input parameters (e.g., see Fig. 2ii). The uniformity is conserved for each set of sample points. The reason why different sets of sample points are used is as follows.

In robust design, the robustness measures expressed by Eqs. (1) and (2) have to be evaluated at every iteration of the optimization, corresponding to different design variables. In other words, a new surrogate model needs to be built at each iteration and the statistics are evaluated on the surrogate model. Assuming a fixed number of samples and a low-discrepancy distribution of the samples, ideally the statistics should be insensitive to the sample set used to build the surrogate model. The surrogate models we use here allow for different sample sets to be used. This advantage is used below to study the effect of the sample set on the accuracy of the statistics; see Fig. 3.

Note that in this section, Kriging and GEK with a Gaussian kernel (correlation function) were adopted and the hyperparameters were optimized by a global optimizer (a differential evolution algorithm was used) by maximum likelihood estimation (MLE).

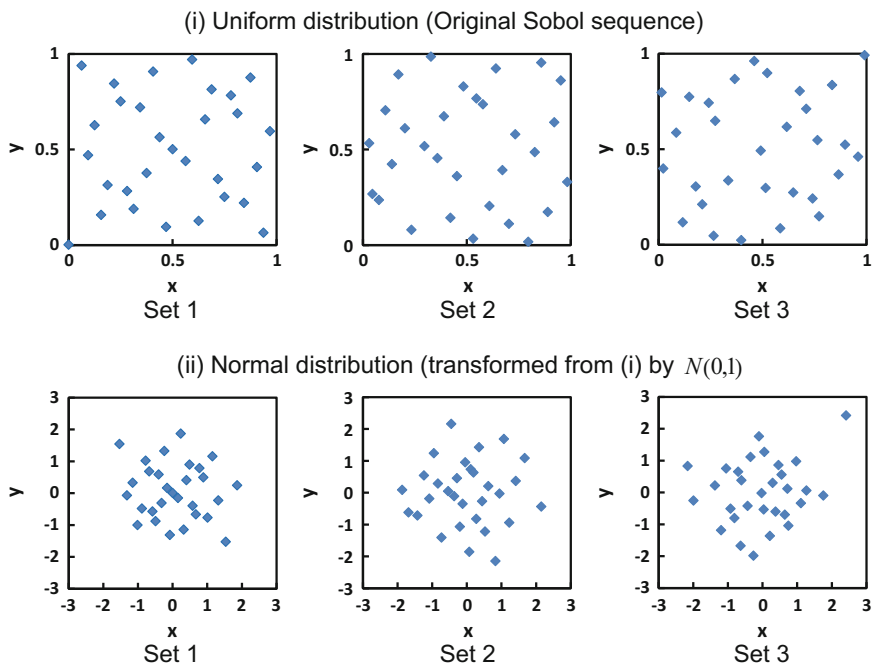


Fig. 2 Three examples of different sets/distributions of sample points (extracted from the Sobol sequence), (i) the number of which is 30 in two-dimensional, and (ii) their transformation into the normal distributions

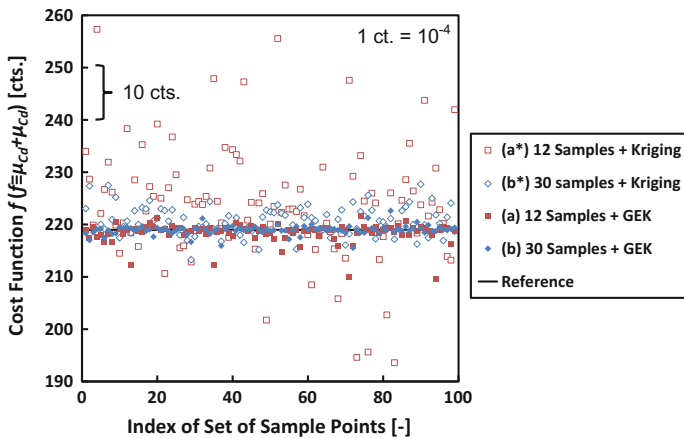


Fig. 3 Cost function $f(f \equiv \mu_{Cd} + \sigma_{Cd})$ distributions evaluated by 100 different sets of sample points for (a) 12 samples with GEK, (b) 30 samples with GEK, (a*) 12 samples with Kriging, (b*) 30 samples with Kriging

Results

Here, two investigations in terms of the criteria (1)–(3) are demonstrated, leading to the best approaches to quantify/optimize the robustness represented by Eqs. (1) and (2).

Investigation of the number of sample points

The first investigation is about influences due to the number of sample points when different sets of sample points of the Sobol sequence are used. The criteria (1)–(3) introduced in the previous sub-section are firstly investigated by using the statistical value f of Eq. (1) as $f \equiv \mu_{cd} + \sigma_{cd}$. The following two numbers of sample points are discussed:

- (a) 12 samples; and
- (b) 30 samples.

Figure 3 shows the cost function f distributions evaluated by 100 different sets of sample points for (a) and (b) extracted from arbitrary consecutive rows of the Sobol sequences. Considering the cases that an adjoint solver is not available, the results obtained by using Kriging are also described as (a*) and (b*) for comparison. The reference value in the figure was evaluated by direct integration of 10^5 Sobol sequence samples. It can be observed that the cost function f evaluated by using GEK has tendency of less dispersions and better agreement with the reference than Kriging. The criteria (1) and (2) in the previous sub-section are discussed here with Fig. 3. Now the mean μ_f and standard deviation σ_f of these distributions can be calculated from the 100 cost functions $f_1 \sim f_{100}$ in Fig. 3.

Figure 4 summarizes μ_f and σ_f , which represent the accuracy of (a) and (b) from the above-mentioned criteria (1) and (2). μ_f is transformed into the absolute error as

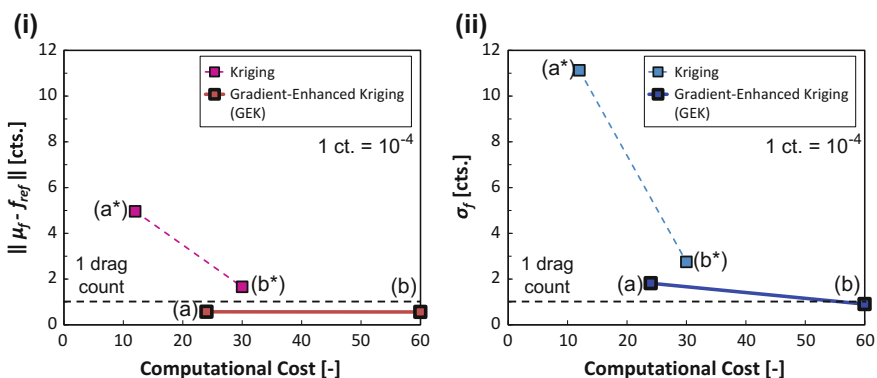


Fig. 4 Comparison between accuracy and computational cost to evaluate a robustness measure, (a) 12 samples with GEK, (b) 30 samples with GEK, (a*) 12 samples with Kriging, (b*) 30 samples with Kriging. The unit of computational cost is one-time state calculation or adjoint calculation in CFD, both of which are assumed to have the same computational time

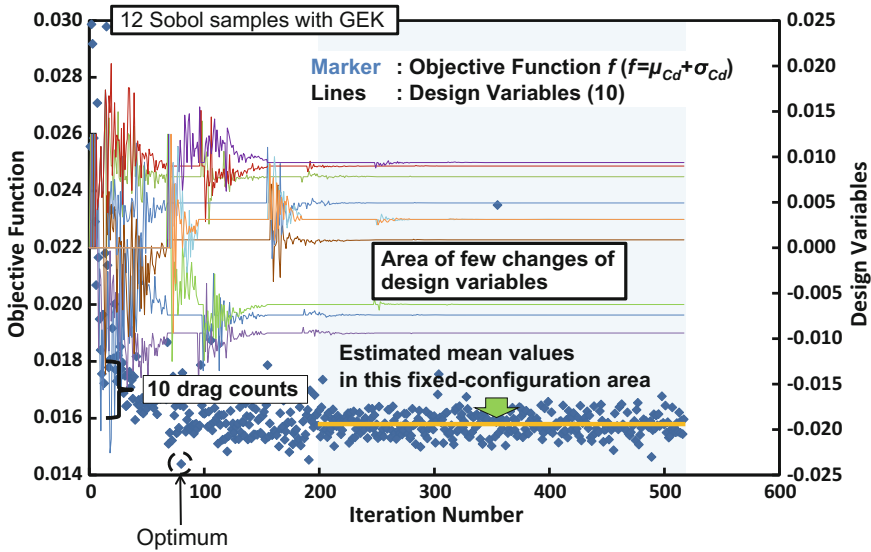


Fig. 5 Histories of objective function evaluated by (a) 12 samples with GEK and design variables in robust design optimization (RDO)

$\|\mu_f - f_{ref}\|$, where f_{ref} is the reference value. The computational cost as the horizontal axis is simply counted by the number of CFD computations with an assumption that the additional computational cost of the adjoint calculation for GEK is identical to that of the state calculation of flow in CFD. Concerning the error tolerances, the errors are discussed with the order of the drag count (where one drag count is 10^{-4} and is denoted as 1 ct.) since the epistemic uncertainties caused by CFD is considered to be not completely negligible in the order of less than 1 ct.

μ_f in (a) and (b) (12 and 30 samples with GEK) have few differences (see Fig. 4i). This fact is in a good agreement with the conclusion in the previous section (see also Fig. 1). On the contrary, σ_f decreases with increase of the number of samples (see Fig. 4ii). The errors of both μ_f and σ_f in (b) are less than 1 ct. How σ_f influences to the optimum solution is introduced next. Note that it can be observed from Fig. 4 that “(a) 12 samples with GEK” is even better than “(b*) 30 samples with Kriging” in terms of both accuracy for μ_f and σ_f , and efficiency.

Now the criterion (3) is discussed with results of applications to robust design optimization. Details of the optimization procedure can be seen in [11]. Figures 5 and 6 show the optimization histories of the cost function $f(f \equiv \mu_{Cd} + \sigma_{Cd})$ and design variables χ . The cost function f of Figs. 5 and 6 at each iteration was computed by (a) and (b), respectively. We can observe that the optimum results f is quite different from each other. This can be confirmed in the areas where the design variables are almost constant; i.e., the configuration is almost fixed. This fact is caused by the difference of σ_f in Fig. 4ii while the mean value μ_f in Fig. 4i is

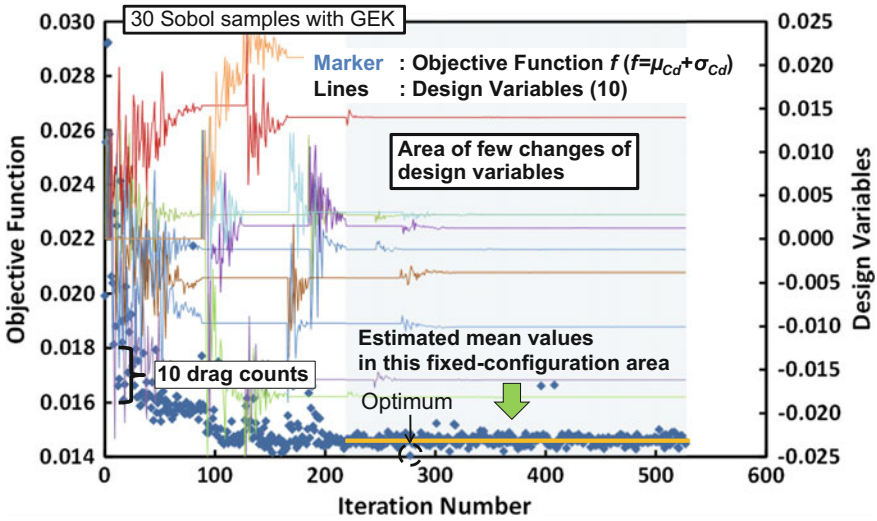


Fig. 6 Histories of objective function evaluated by (b) 30 samples with GEK and design variables in robust design optimization (RDO)

almost constant. Table 1 summarizes the cost function f of the two designed airfoils, which were re-evaluated by the common strategy (b) 30 samples. This table clarifies that more accurate evaluation could lead to an optimal solution with better performance.

Finally as comparison, a fixed set of sample points (see Fig. 2i), i.e. a fixed consecutive row of the Sobol sequence is added as another type of sampling technique for comparison:

(c) 30 samples (by using a fixed set of sample points).

The set of sample points which has the closest f to \hat{f} was picked up from $f_1 \sim f_{100}$ in Fig. 3 and that set of sample points was fixed and used for each iterative process. The histories of f , χ and the re-evaluated f of the optimum configuration are shown in Fig. 7 and Table 1, respectively. There are few differences between the sampling strategies (b) 30 samples and (c) 30 samples (by using a fixed set of sample points). The conclusion here is that σ_f due to different sets of sample points is important here as “another indicator” to determine the number of sample points.

Table 1 Values of objective function $f(f \equiv \mu_{Cd} + \sigma_{Cd})$ of optimized airfoils obtained by different strategies (a) 12 samples, (b) 30 samples, (c) 30 samples (by using a fixed set of sample points)

	(a) (cts.)	(b) (cts.)	(c) (cts.)
μ_{Cd}	144.4	135.9	136.2
σ_{Cd}	14.1	9.9	10.8
$f(\equiv \mu_{Cd} + \sigma_{Cd})$	158.5	145.8	147.0

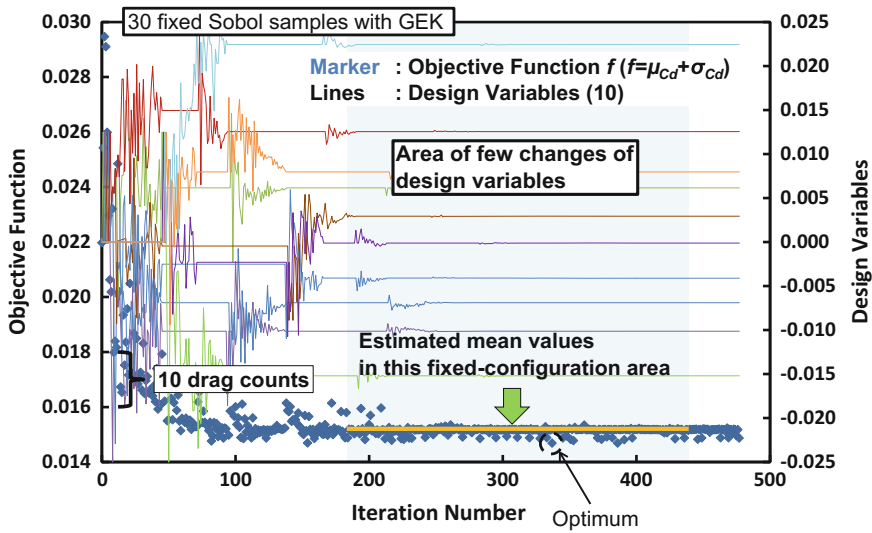


Fig. 7 Histories of objective function evaluated by (c) 30 samples (by using a fixed set of sample points) with GEK and design variables in robust design optimization (RDO)

Investigation of distribution of sample points

In the second investigation, more details of influences of distribution of sample points are demonstrated. That is, the distributions of the sample points are not only by the original Sobol sequence but the one transformed into input pdf (normal distributions here, see Fig. 2ii) and/or the one with dynamic infilled sample points. The number of samples and the surrogate model are fixed at 30 and GEK, respectively. The sampling techniques used are:

- (a) input pdf (normal distributions);
- (b) uniform distributions;
- (c) uniform distributions and an adaptive sampling.

Suitable sample techniques for different measures of robustness (statistical values of QoI) are introduced in [1]. The results obtained here are straightforward as follows:

- For evaluating mean and standard deviation of QoI (expectation measure with mean-risk approach) represented by Eq. (1), the same distributions as the pdf of the input uncertainty parameters (normal distributions are often used) can be applicable.
- For evaluating maximum or minimum value of QoI (worst-case risk measure) represented by Eq. (2), an adaptive sampling technique in the uniform distributions leads to good accuracy.

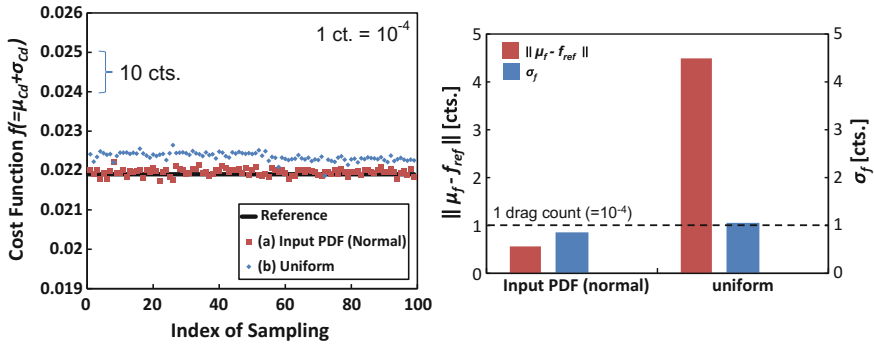


Fig. 8 Cost function $f(f \equiv \mu_{Cd} + \sigma_{Cd})$ distributions evaluated by 100 different distributions of Sobol sample points with (a) input pdf (normal distributions), (b) uniform distributions

Qualitative substantiation of them is demonstrated here. Figures 8 and 9 show cost function $f(f \equiv \mu_{Cd} + \sigma_{Cd}$ and $f \equiv \max_u(C_d(\mathbf{u}))$, respectively) distributions evaluated by 100 different sets of Sobol sample points by (a) input pdf (normal distributions), (b) uniform distributions, (c) uniform distributions with adaptive sampling (only for Fig. 9), and μ_f and σ_f , as with Figs. 3 and 4 for the first approach. The adaptive sampling technique here is an Expected-Improvement (EI)-based approach to search for the maximum or minimum value of QoI on the surrogate model. The initial sample points are the Sobol sequence with 24 points. The surrogate model is updated in stages by an imposed sample point until the total number of sample points reaches to 30. Details of the adaptive sampling technique can be found in [1].

In Fig. 8, for the cost function $f \equiv \mu_{Cd} + \sigma_{Cd}$ by Eq. (1), μ_f and σ_f by the input pdf (normal distributions) are lower than the uniform distributions and are lower than 1 ct, respectively. Note that μ_f and σ_f by the normal distributions correspond to (b) 30 samples in Figs. 2 and 3, and also its optimization result can be seen in Fig. 5.

On the other hand, for the cost function $f \equiv \max_u(C_d(\mathbf{u}))$ by Eq. (2), the expected value μ_f by the input pdf (normal distributions) is lower than 1 ct., whereas σ_f is quite large as can be also confirmed in Fig. 9(a) that f varies widely with different sets of the Sobol sample points. On the contrary, σ_f by the uniform distributions is low but μ_f is overestimated with around 3 cts. as also can be seen in Fig. 9(a). The uniform distributions with an adaptive sampling technique bring the same accuracy as the input pdf for $f \equiv \mu_{Cd} + \sigma_{Cd}$ in terms of both μ_f and σ_f , which are less than 1 ct.

In this chapter, the types of surrogate models, the number and distribution of sample points were discussed. In the worst-case risk measure, dynamic adaptive sampling techniques are necessary to keep the same order of accuracy as the expectation measure with mean-risk approach. Further improvement of the accuracy of evaluating the cost functions could be led by a variety of adaptive sampling techniques to enhance the quality of the Kriging-based surrogate models [12, 13].

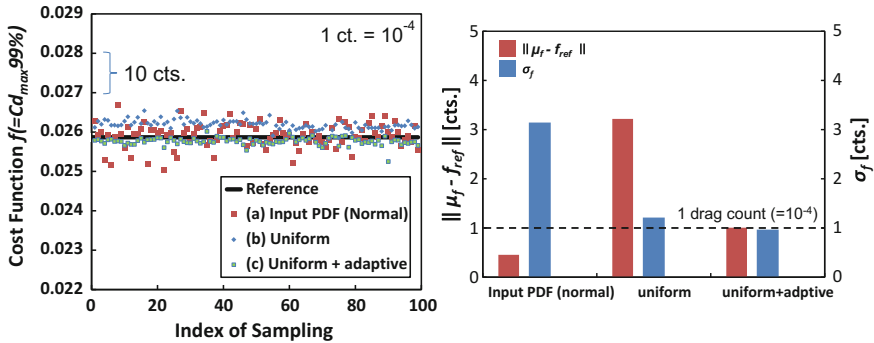


Fig. 9 Cost function $f(f \equiv C_{d99\%max})$ distributions evaluated by 100 different distributions of Sobol sample points with (a) input pdf (normal distributions), (b) uniform distributions, (c) uniform distributions with adaptive sampling

Summary

The accuracy and efficiency of surrogate model-based approaches to UQ and their application to robust design were demonstrated for the UMRIDA BC-02 test case. Twelve uncertain parameters, yielding a 12-dimensional input parameter space for surrogate model construction, were considered. Both Kriging and gradient-enhanced Kriging (GEK) were investigated. GEK was shown to lead to a good agreement of the statistical values such as the mean and standard deviation of the aerodynamic coefficients with reference values when the number of samples is more than around 12. GEK is the best choice when an adjoint solver is available. The accuracy of the statistics was also investigated from the point of view of how the sampling influences the surrogate model used in robust design. It was confirmed that the error dispersions of the statistical cost function is a function of the number of samples, the distribution of the samples. Sampling techniques in accordance with statistics to be evaluated are required to reduce the error dispersion and to achieve good robust design solutions. Different robustness measures can be evaluated accurately to within one drag count by using 30 sample points.

References

1. Maruyama, D., Liu, D., Görtz, S.: Comparing surrogates for estimating aerodynamic uncertainties of airfoils. In: Hirsch, C. et al. (eds.) *Uncertainty Management for Robust Industrial Design in Aeronautics*, Chap. 13, pp. xx–xx. (2017)
2. Galle, M., Gerhold, T., Evans, J.: Parallel computation of turbulent flows around complex geometries on hybrid grids with the DLR-TAU code. In: Ecer, A., Emerson, D.R. (eds.) *In: Proceedings 11th Parallel CFD Conference*, Williamsburg, VA, North-Holland, 23–26 May 1999

3. Gerhold, T., Hannemann, V., Schwaborn, D.: On the validation of the DLR-TAU code. In: Nitsche, W., Heinemann, H.J., Hilbig, R. (eds.) *New Results in Numerical and Experimental Fluid Mechanics, Notes on Numerical Fluid Mechanics*, vol. 72, pp. 426–433. Vieweg (1999). ISBN 3-528-03122-0
4. Schwaborn, D., Gerhold, T., Heinrich, R.: The DLR TAU-code: recent applications in research and industry, invited lecture. In: Wesseling, P., Oate, E., Piaux, J. (eds.) *Proceedings of the European Conference on Computational Fluid Dynamics (ECCOMAS CFD 2006)*, The Netherlands (2006)
5. Allmaras, S.R., Johnson, F.T., Spalart, P.R.: Modifications and clarifications for the implementation of the Spalart-Allmaras turbulence model. In: *Seventh International Conference on Computational Fluid Dynamics (ICCFD7)*, ICCFD7-1902, Hawaii, July 2012
6. Sobol, I.M.: Distribution of points in a cube and approximate evaluation of integrals. *Zh. Vychisl. Mat. Mat. Fiz.* **7**(4), 784–802 (1967)
7. Joe, S., Kuo, F.Y.: Remark on algorithm 659: Implementing Sobol’s quasirandom sequence generator. *ACM Trans. Math. Softw.* **29**, 49–57 (2003)
8. Joe, S., Kuo, F.Y.: Constructing Sobol sequences with better two-dimensional projections. *SIAM J. Sci. Comput.* **30**, 2635–2654 (2008)
9. Liu, D., Litvinenko, A., Schillings, C., Schulz, V.: Quantification of airfoil geometry-induced aerodynamic uncertainties—comparison of approaches. *SIAM/ASA J. Uncertainty Quant.* (2016)
10. Liu, D., Maruyama, D., Görtz, S.: Geometrical uncertainties—accuracy of parametrization and its influence on UQ and RDO. In: Hirsch, C. et al. (eds.) *Uncertainty Management for Robust Industrial Design in Aeronautics*, Chap. 51, pp. xx–xx. (2017)
11. Maruyama, D., Görtz, S., Liu, D.: Robust design measures for airfoil shape optimization. In: Hirsch, C. et al. (eds.) *Uncertainty Management for Robust Industrial Design in Aeronautics*, Chap. 32, pp. xx–xx. (2017)
12. Dwight, R., Han, Z.H.: Efficient uncertainty quantification using gradient-enhanced Kriging. *AIAA Paper 2009–2276* (2009)
13. Shimoyama, K., Kawai, S., Alonso, J.J.: Dynamic adaptive sampling based on Kriging surrogate models for efficient uncertainty quantification. In: *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. AIAA paper 2013-1470 (2013)