Uncertainty Analysis of Structural Output with Closed-form Expression Based on Surrogate Model

Yuan-Lv Chen [a], Yan Shi [a, b, c*], Hong-Zhong Huang [a, b], Dong Sun [a, b], Michael Beer [c, d, e]

[a] *School of Mechanical and Electrical Engineering, University Electronic Science and Technology of China, Chengdu 611731, China*
[b] *Center for System Reliability and Safety, University Electronic Science and Technology of China, Chengdu 611731, China*
[c] *Institute for Risk and Reliability, Leibniz Universität Hannover, Hannover 30167, Germany*
[d] *Institute for Risk and Uncertainty, University of Liverpool, Liverpool L69 7ZF, United Kingdom*
[e] *International Joint Research Center for Resilient Infrastructure & International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, Shanghai 200092, China*

Abstract: Uncertainty analysis (UA) is the process that quantitatively identifies and characterizes the output uncertainty and has a crucial implication in engineering applications. The research of efficient estimation of structural output moments in probability space plays an important part in the UA and has great engineering significance. Given this point, a new UA method based on the Kriging surrogate model related to closed-form expressions for the perception of the estimation of mean and variance is proposed in this paper. The new proposed method is proven effective because of its direct reflection on the prediction uncertainty of the output moments of metamodel to quantify the accuracy level. The estimation can be completed by directly using the redefined closed-form expressions of the model's output mean and variance to avoid excess post-processing computational costs and errors. Furthermore, a novel framework of adaptive Kriging estimating mean (AKEM) is demonstrated for more efficiently reducing uncertainty in the estimation of output moment. In the adaptive strategy of AKEM, a new learning function based on the closed-form expression is proposed. Based on the closed-form expression which modifies the computational error caused by the metamodeling uncertainty, the proposed learning function enables the updating of metamodel to reduce prediction uncertainty efficiently and realize the decrease in computational costs. Several applications are introduced to prove the effectiveness and efficiency of the AKEM compared with a universal adaptive Kriging method. Through the good performance of AKEM, its potential in engineering applications can be spotted.

Keywords: Uncertainty analysis; Adaptive procedure; Kriging surrogate model; Closed-form expression; Epistemic uncertainty

* Corresponding author
E-mail: yanshi@uestc.edu.cn

1. Introduction

Uncertainty widely exists in many fields of science and engineering, which brings great challenges in the analysis or optimization problems. Refs. [1, 2] discussed the uncertainty description and definition, established the basic principle of uncertainty analysis, and pointed out that the modeling method directly affects the uncertainty degree. The epistemic uncertainty that we are trying to research is considered to be caused by a lack of knowledge, which can be reduced if more information is obtained. What's more, the greatest importance of uncertainty analysis is to ensure the accuracy of the output results, thus consideration should be paid to the effects of uncertainty. With the development of engineering applications, uncertainties with increased complexity have been emphasized in recent decades. Most of the researches about uncertainty can be summarized by uncertainty quantification (UQ). The important steps of UQ contain uncertainty propagation (UP) and uncertainty analysis (UA) [3, 4]. Finding an efficient and practical UQ method is attractive to researchers in many fields. The study of the characteristics and consequences of uncertainties, as well as their mathematical modeling in reliability analysis, has been proven useful for reliability evaluation and decision-making [5-9].

UP is aimed at the process of propagation from a set of uncertain inputs to the distribution of uncertain output. UP plays an important part in research areas such as reliability analysis, reliability design, optimization problems, and so on. Former researches have tested the practicability of UP in the analysis of mechanical properties of materials [10], risk and resilience analysis [11], or reliability optimization [12]. Some proposed UP methods, for example, the edge detection for multi-dimensional UP [13], UP based on the direct probability integral method and exponential convex model [14] or Bayesian probabilistic integration [15], and UP applied in the construction of response surfaces [16] or topological structures [17] are available for reference. UA, on the other hand, identifies and characterizes the variability of output due to uncertain input of a system or model. Traditional Monte Carlo (MC) and further developed Quasi-Monte Carlo (QMC) methods with satisfying robustness, are easy to understand and be applied to the UQ of complex structures [18-21]. QMC focuses on an efficient sampling approach, allowing a reliable estimation of the accuracy and the ability to facilitate the sequential addition of samples [22]. For parametric UA, there is a development based on probability boxes [23-27]. Some findings of research like non-intrusive reduced-order modeling were proposed to overcome the unaffordable computational burden in the analysis of high dimensional situation [28]. Other methods combining UA with neural networks and utilizing deep learning techniques are also provided [29, 30].

The accurate estimation of moments is greatly influential in the research about uncertainty. Except being preliminary for UP, predicting mean and variance is a fundamental task of UA and also crucial to the robust design optimization (RDO), a representative paragon for engineering design under uncertainty [31-33]. Although previous UA methods contributed to the evaluation of prediction and are proven to be suitable for engineering applications, it is noticed that the majority of existing methods to complete estimation require a number of simulations to ensure accuracy [34-36]. When it comes to complex models, huge computational costs are needed [37] since finite element (FE) simulation is commonly used to evaluate the output corresponding to a single input. These numerical simulation models can be computationally very expensive and may spend a few hours to days or even months to simulate a set of inputs [38, 39]. What's more, paying attention to model characteristics, model dimension, and distribution of the input variables, and performing evaluation for higher-order moments like skewness and kurtosis, seems to be a consensus [40]. Therefore, the improvement of efficiency in the estimation of moments exists as a big challenge.

At present, surrogate model or metamodel methods have been widely used in engineering modeling [41], which is also appropriate for the mentioned requirement. These methods have been developed for the excessive computational costs existing in the simulations with high complexity. Relevant research and applications can be found in optimization problems [42], sensitivity or reliability analysis [43, 44], and so on [45, 46]. High dimensional model representation [47], state dependent parameter [48], polynomial chaos expansion [49], support vector regression [50], and Gaussian process regression [51] are surrogate models commonly used. As one kind of Gaussian process that uses a spatial covariance function as kernel, Kriging is another widely used surrogate model, which can be easily accomplished by the computer [52-54]. After taking several input samples of training sets with their corresponding output values to calculate the mapping relation, the Kriging model is able to provide output prediction for any possible inputs. Its advantage was explained by its tendency to find the best linear unbiased predictor while minimizing the mean square error of the prediction [55]. The rapid development of Kriging-based methods started from an active learning reliability analysis method combining Kriging and MCS [56], which makes full use of good convergence of the Kriging model and cooperation with adaptive strategies. After this new research area was explored, more strategies were proposed to eliminate the shortages of existing methods, such as AK-MCSi [57], AK-IS [58], AK-SS [59] and eAK-MCS [60]. As a versatile analysis tool, Kriging surrogate model is also used to approximate the dynamic system or estimate the failure probability border of imprecise probability models [61, 62]. However, the majority of research on Kriging theory focuses on reliability analysis [63, 64], more concerned about the determination of the limit state and failure detection. Hence, the potential of the Kriging model in the task of moments estimation is underestimated. A few of former researches revealed the practicability of this work. In an adaptive Kriging method proposed by Song et al. [65], the concept of Kriging prediction covariance is introduced to describe the prediction responses of two points and further display the correspondence of the prediction uncertainty. Inspired by a similar design for probabilistic integration of the Gaussian process regression model with basic kernel function [66], it is noticed that the Kriging covariance can be expanded to the estimation with further closed-form expression, for the prediction covariance exposes the effect of prediction uncertainty on the output. What's more, previous conclusions indicated that following a specific probability distribution, the moments of the first four orders can be calculated by using the parameters of Kriging [67].

In this paper, the closed-form expressions for the evaluation of Kriging output mean and variance are newly established with the metamodeling uncertainty additionally embedded inside. The prediction variance, or the so-called posterior variance in metamodeling, is also derived in closed-form expressions to identify the prediction uncertainty of the output mean, in the form of the integral of the Kriging covariance in the probability space. Considering the accuracy of output mean, the moment prediction variance can directly reflect the goodness level of metamodeling under this specific requirement, which has not been emphasized before in Kriging-based applications. What's more, the effect of any certain point in the probability space on the metamodeling uncertainty is discovered in this process. Through the practicable analysis, the identified uncertainty can be exploited to develop estimation method. We propose a novel framework of adaptive Kriging estimating mean (AKEM) for the efficient estimation of the structural output mean and variance. With the new-proposed adaptive strategy, the method of AKEM successfully reaches the achievement of better efficiency of estimation and economization of computational costs in the application of adaptive Kriging to fill the gap in research. In AKEM, the estimation of outputs is completed by directly using the established closed-form expressions to approach output, instead of combining sampling methods with surrogate models, where the latter leads to excess post-processing computational costs and error. A new learning function is proposed as the core of the adaptive training framework of AKEM, which quantifies the contribution of possible inputs to the identified prediction uncertainty. The probability density and prediction variance of a single point are also considered to test sampled points in the learning function to ensure better stability of AKEM. The adaptive strategy is able to better update the training set to reduce the uncertainty of estimation iteratively, and the improvement in the estimation efficiency and accuracy is achieved. Stopping criteria are set to make sure the iteration should end with satisfying computational accuracy. For the performance of AKEM, the considered uncertainty in constructed learning function is aimed at the mean, so this output turns out to approach real value efficiently. Furthermore, the output variance also converges well as a result of an influential advantage of the proposed strategy. For the dependability of AKEM is tested in several cases, as well as its numerical accuracy and computational efficiency are proven, AKEM is recognized as able to reach advanced prediction accuracy and operation efficiency on the prediction of the original function's mean and variance value. In the future research, this method is considered eligible to be further developed and applied in RDO or more applications of engineering problems.

The rest of this paper consists of the following parts: Section 2 reviews previous calculation methods of output mean and variance and introduces a brief explanation of the Kriging theory. Section 3 proposes novel closed-form expressions. Section 4 introduces the construction of adaptive strategy algorithms. Example tests are shown in Section 5. Conclusions are given in Section 6.

2. Moments and Kriging model

2.1. Mean and variance

Assume that $\mathbf{X}=\left[X_1, X_2, ..., X_n\right]^{\mathrm{T}}$ is a vector of random variables with a probability density function (PDF) $f(\mathbf{X})$. $y=g(\mathbf{X})$ is a function of $\mathbf{X}$. The mean and variance of y are defined as Eq.(1) and Eq.(2) respectively.

$$E(y)=\int g(\mathbf{X}) f(\mathbf{X})\mathrm{d}\mathbf{X} \tag{1}$$

$$V(y)=\int\left(g(\mathbf{X})-E(g(\mathbf{X}))\right)^2 f(\mathbf{X})\mathrm{d}\mathbf{X} \tag{2}$$

When the random vectors are one-dimensional, the integral is unquestioned. However, it should be noticed that $\mathbf{X}$ can be multidimensional, which means $\int f(\mathbf{X})\mathrm{d}\mathbf{X}$ can be multiple integrals of each component over the probability space. In this paper, the single integral symbol is used to represent possible multiple integrations of a certain vector for convenience. Variance $V(y)$ is generally the concept representing the dispersion degree of y. However, another characteristic of variance is worth attention in this paper: it can be used to describe the error involved with the prediction of metamodel as a measure of uncertainty, which is known as the prediction variance.

These two moments are concerned in this paper. Based on the assumption of engineering problems, there exists a functional relationship $y = g(\mathbf{X})$ between input and output, where $g(\mathbf{X})$ is the real performance function in application. However, direct integration for $E(y)$ and $V(y)$ is usually impossible for the probability density function $f(\mathbf{X})$ is frequently hard to obtain or non-integrable. Generally, MC or QMC methods can be used to approximate the numerical integration. An origin data set $S = \left\{(\mathbf{X}^{(i)}, y^{(i)})\right\}_{i=1}^{N}$ of size N is first considered. $\mathbf{X}$ is a vector of input variables and y is the corresponding output. In this term, samples with size N is generated obeying the distribution of each component. Then the values of y corresponding to each sample are calculated to obtain the mean and variance of output. These methods have good robustness, and the result is supposed to be accurate when N is big enough due to the law of large numbers and its lemma. However, the disadvantage is also obvious, since a complex performance function makes N sets of calculations unaffordable. Consequently, numerous alternative methods have been proposed to address the problem of excess computational costs, in which surrogate model method is included.

2.2. Kriging surrogate model

The surrogate model, also known as metamodel, can be understood as the model of model. The surrogate model method is used in engineering to improve the efficiency of calculation by emulating the behavior of the original simulation model whose exact output generally requires high computational costs to obtain. In this paper, we introduce the Kriging method to construct a surrogate model of the origin structural performance function $y = g(\mathbf{X})$, and the surrogate model is written as $\hat{g}(\mathbf{X})$. According to Kriging theory [52], the ordinary Kriging is comprised of a constant part and a Gaussian process part as

$$\hat{g}(\mathbf{X}) = \beta + Z(\mathbf{X}) \tag{3}$$

, where β is the regression parameter vector and $Z(\mathbf{X})$ is a steady-state Gaussian process with zero as mean and $\mathrm{cov}(\mathbf{X}^{(i)}, \mathbf{X}^{(j)}) = \sigma^2 R(\mathbf{X}^{(i)}, \mathbf{X}^{(j)})$ as covariance, where $\mathbf{X}^{(i)}, \mathbf{X}^{(j)}$ are two input samples of $\mathbf{X}$ and σ^2 is the process variance. R is the correlation function given as:

$$R(\mathbf{X}^{(i)}, \mathbf{X}^{(j)}) = \exp\left\{-\sum_{k=1}^{n} \theta_k (X_k^{(i)} - X_k^{(j)})^2\right\} \tag{4}$$

in which θ_k is the k-th correlation scale parameter, which will be mentioned below with β and σ^2.

$X_k^{(i)}, X_k^{(j)}$ are the k -th components of $\mathbf{X}^{(i)}, \mathbf{X}^{(j)}$ respectively. n is the dimension of inputs.

Assuming a training set with $\mathbf{X}^* = (\mathbf{X}^{*(1)}, \mathbf{X}^{*(2)}, ...\mathbf{X}^{*(m)})$ and $\mathbf{y} = (y^{(1)}, y^{(2)}, ...y^{(m)})^{\mathrm{T}}$ as inputs and outputs individually, for regression matrix $\boldsymbol{F}$ and correlation matrix $\boldsymbol{R}$. $\boldsymbol{F}$ is a $m \times 1$ vector whose elements all equal to 1 and $R_{ij} = R(\mathbf{X}^{*(i)}, \mathbf{X}^{*(j)})(i, j = 1, 2, ...m)$ in $\boldsymbol{R}$. The regression parameter β and the variance of the Gaussian process σ^2 can be obtained as $\beta = (\boldsymbol{F}^{\mathrm{T}} \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{F}^{\mathrm{T}} \boldsymbol{R}^{-1} \mathbf{y}^{\mathrm{T}}$, and $\sigma^2 = \dfrac{1}{m}(\mathbf{y} - \boldsymbol{F}\beta)^{\mathrm{T}} \boldsymbol{R}^{-1} (\mathbf{y} - \boldsymbol{F}\beta)$. The correlation scale parameters $\boldsymbol{\theta} = \left[\theta_1, \theta_2, ...\theta_n\right]$ could be obtained by maximum likelihood estimation, e.g., $\boldsymbol{\theta} = \arg\max\limits_{\boldsymbol{\theta}} \left(-\dfrac{m}{2} \ln \sigma^2 - \dfrac{1}{2} \ln |\boldsymbol{R}| \right)$.

For an unknown point $\mathbf{X}$, the mean and variance of predicted response at point are obtained as:

$$\mu_{\hat{g}}(\mathbf{X}) = \beta + \boldsymbol{r}^{\mathrm{T}}(\mathbf{X}) \boldsymbol{R}^{-1} (\mathbf{y} - \boldsymbol{F}\beta) \tag{5}$$

$$\sigma_{\hat{g}}^2(\mathbf{X}) = \sigma^2 \left[1 - \boldsymbol{r}^{\mathrm{T}}(\mathbf{X}) \boldsymbol{R}^{-1} \boldsymbol{r}(\mathbf{X}) + \boldsymbol{u}^{\mathrm{T}}(\mathbf{X}) (\boldsymbol{F}^{\mathrm{T}} \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{u}(\mathbf{X}) \right] \tag{6}$$

in which $\boldsymbol{r}(\mathbf{X}) = \left[R(\mathbf{X}, \mathbf{X}^{*(1)}), R(\mathbf{X}, \mathbf{X}^{*(2)}), ...R(\mathbf{X}, \mathbf{X}^{*(m)}) \right]^{\mathrm{T}}$, and $\boldsymbol{u}(\mathbf{X}) = \boldsymbol{F}^{\mathrm{T}} \boldsymbol{R}^{-1} \boldsymbol{r}(\mathbf{X}) - 1$. Eq.(6) can be expanded to the covariance of the predicted response at two points $\mathbf{X}$ and $\mathbf{X}'$ as follows:

$$\mathrm{Cov}(\hat{g}(\mathbf{X}), \hat{g}(\mathbf{X}')) = \sigma^2 \left[R(\mathbf{X}, \mathbf{X}') - \boldsymbol{r}^{\mathrm{T}}(\mathbf{X}) \boldsymbol{R}^{-1} \boldsymbol{r}(\mathbf{X}') + \boldsymbol{u}^{\mathrm{T}}(\mathbf{X}) (\boldsymbol{F}^{\mathrm{T}} \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{u}(\mathbf{X}') \right] \tag{7}$$

In the next section, the above parameters, matrixes, and equations based on Kriging surrogate model will be employed to establish closed-form expressions of structural output.

3. Uncertainty analysis with closed-form expression

3.1. Closed-form expression of output mean

After the Kriging surrogate model of structural performance function is built, the mean of structural output can be obtained by numerical integration of combining sampling methods. However, this process leads to post-processing computational costs and error. An existing and efficient way of dealing with this issue is to establish the closed-form expression of output mean by directly using the information of the Kriging surrogate model. The closed-form expression of the output mean is expressed as follows [68]:

$$E(\mu_{\hat{g}}(\mathbf{X})) = \beta + \boldsymbol{\gamma} \boldsymbol{R}^{-1} (\mathbf{y} - \boldsymbol{F}\beta) = d \tag{8}$$

in which $\boldsymbol{\gamma} = \left(|A|^{\frac{1}{2}} S^{(1)}, |A|^{\frac{1}{2}} S^{(2)}, ...|A|^{\frac{1}{2}} S^{(m)} \right)$, $A = \mathrm{diag}\left(\dfrac{1}{2\theta_1}, \dfrac{1}{2\theta_2}, ...\dfrac{1}{2\theta_n} \right)$,

$S^{(i)} = \left|A+\Sigma_X\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X}^{*(i)}-\boldsymbol{\mu}_X)^{\mathrm{T}}(A+\Sigma_X)^{-1}(\mathbf{X}^{*(i)}-\boldsymbol{\mu}_X)\right\}$. The premise of this conclusion is that the variables are normally distributed with mean vector $\boldsymbol{\mu}_X$ and diagonal covariance matrix Σ_X. Accordingly, it should be noticed that the preprocessing of data such as probability integral transformation is required.

Eq.(8) outputs in constant, but the new proposed uncertainty analysis provides an additional explanation. For every definite input $\mathbf{X}$, we consider the corresponding kriging prediction of structure output $\hat{g}(\mathbf{X}) = y$ as a random variable following normal distribution $y \sim N(\mu_{\hat{g}}(\mathbf{X}), \sigma_{\hat{g}}^2(\mathbf{X}))$. The prediction value and uncertainty are combined in the form of another normal probability space, differing from traditional methods which treat them individually. Consider the probability integration y in the space of the distribution of $\mathbf{X}$:

$$\bar{y} = E_{\mathbf{X}}(\hat{g}(\mathbf{X})) = \int \hat{g}(\mathbf{X}) f(\mathbf{X}) \mathrm{d}\mathbf{X} \tag{9}$$

As the integral of random variable, $\bar{y}$ is still a normally distributed random variable. To obtain the numerical expectation as a constant, a further step should be made. Different from Eq.(9), E_g represents the probability integral for the PDF of $\hat{g}(\mathbf{X})$ in the probability space of prediction:

$$E_g(\hat{g}(\mathbf{X})) = \int \hat{g}(\mathbf{X}) \pi(g) \mathrm{d}g = \mu_{\hat{g}}(\mathbf{X}) \tag{10}$$

where $\pi(g)$ is the PDF of $\hat{g}(\mathbf{X})$, which corresponds to a normal distribution $N(\mu_{\hat{g}}(\mathbf{X}), \sigma_{\hat{g}}^2(\mathbf{X}))$ that $\hat{g}(\mathbf{X})$ follows. Do one more integration, there is:

$$E(\bar{y}) = E_g(E_{\mathbf{X}}(\hat{g}(\mathbf{X}))) = E_{\mathbf{X}}(E_g(\hat{g}(\mathbf{X}))) = E_{\mathbf{X}}(\mu_{\hat{g}}(\mathbf{X})) = d \tag{11}$$

Therefore, the output mean from $\bar{y}$ is the same as Eq.(8).

The prediction variance of $\bar{y}$ which characterizes the prediction uncertainty in the estimation of the output value of $E(\bar{y})$ is also available. Through $\mathrm{Cov}(\hat{g}(\mathbf{X}), \hat{g}(\mathbf{X}'))$ mentioned in Eq.(7), the prediction variance $V(\bar{y})$ can be obtained by double integrals of $\mathrm{Cov}(\hat{g}(\mathbf{X}), \hat{g}(\mathbf{X}'))$ over $\mathbf{X}$ and $\mathbf{X}'$:

$$\begin{aligned} V(\bar{y}) &= \int\int \mathrm{Cov}(\hat{g}(\mathbf{X}), \hat{g}(\mathbf{X}')) f(\mathbf{X}) f(\mathbf{X}') \mathrm{d}\mathbf{X}\mathrm{d}\mathbf{X}' \\ &= \sigma^2 \int\int \left[R(\mathbf{X}, \mathbf{X}') - \boldsymbol{r}^{\mathrm{T}}(\mathbf{X}) \boldsymbol{R}^{-1} \boldsymbol{r}(\mathbf{X}') + \boldsymbol{u}^{\mathrm{T}}(\mathbf{X}) (\boldsymbol{F}^{\mathrm{T}} \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{u}(\mathbf{X}') \right] f(\mathbf{X}) f(\mathbf{X}') \mathrm{d}\mathbf{X}\mathrm{d}\mathbf{X}' \end{aligned} \tag{12}$$

To continue from Eq.(12), an already proved theorem is necessary to be introduced: the product of two normal PDFs is equal to a new normal PDF multiplied by a constant [68]. Supposing that $f_i(\mathbf{X})$ is a normal PDF of n-dimensional random variables $\mathbf{X}$ with mean vector $\boldsymbol{\mu}_i$ and covariance matrix Σ_i shown below,

$$f_i(\mathbf{X}) = (2\pi)^{-\frac{n}{2}} \left|\Sigma_i\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu}_i)^{\mathrm{T}} \Sigma_i^{-1} (\mathbf{X}-\boldsymbol{\mu}_i)\right\} \tag{13}$$

Then,

$$f_i(\mathbf{X}) f_j(\mathbf{X}) = (2\pi)^{-\frac{n}{2}} \left|\Sigma_i + \Sigma_j\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^{\mathrm{T}} (\Sigma_i + \Sigma_j)^{-1} (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)\right\} f_{ij}(\mathbf{X}) \tag{14}$$

$f_{ij}(\mathbf{X})$ is normal PDF with mean vector $\boldsymbol{\mu}_{ij} = \Sigma_{ij}(\Sigma_i^{-1}\boldsymbol{\mu}_i + \Sigma_j^{-1}\boldsymbol{\mu}_j)$ and covariance matrix $\Sigma_{ij} = (\Sigma_i^{-1} + \Sigma_j^{-1})^{-1}$.

Notice $R(\mathbf{X},\mathbf{X}')$ in Eq.(12), a transformation can be made:

$$\begin{aligned} R(\mathbf{X},\mathbf{X}') &= \exp\left\{-\sum_{k=1}^{n} \theta_k (X_k - X_k')^2\right\} \\ &= (2\pi)^{\frac{n}{2}} \left|A\right|^{\frac{1}{2}} (2\pi)^{-\frac{n}{2}} \left|A\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X}-\mathbf{X}')^{\mathrm{T}} A^{-1} (\mathbf{X}-\mathbf{X}')\right\} \\ &= (2\pi)^{\frac{n}{2}} \left|A\right|^{\frac{1}{2}} m(\mathbf{X}) \end{aligned} \tag{15}$$

in which $m(\mathbf{X})$ is the PDF of a normal distribution with mean vector $\mathbf{X}'$ and covariance matrix A.

In Eq.(12), the following integration is first considered:

$$\begin{aligned} &\int R(\mathbf{X},\mathbf{X}') f(\mathbf{X}) \mathrm{d}\mathbf{X} = \\ &= (2\pi)^{\frac{n}{2}} \left|A\right|^{\frac{1}{2}} \int (2\pi)^{-\frac{n}{2}} \left|A+\Sigma_X\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X}'-\boldsymbol{\mu}_X)^{\mathrm{T}} (A+\Sigma_X)^{-1} (\mathbf{X}'-\boldsymbol{\mu}_X)\right\} m_f(\mathbf{X}) \mathrm{d}\mathbf{X} \\ &= (2\pi)^{\frac{n}{2}} \left|A\right|^{\frac{1}{2}} (2\pi)^{-\frac{n}{2}} \left|A+\Sigma_X\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X}'-\boldsymbol{\mu}_X)^{\mathrm{T}} (A+\Sigma_X)^{-1} (\mathbf{X}'-\boldsymbol{\mu}_X)\right\} = (2\pi)^{\frac{n}{2}} \left|A\right|^{\frac{1}{2}} m'(\mathbf{X}') \end{aligned} \tag{16}$$

$m'(\mathbf{X}') = (2\pi)^{-\frac{n}{2}} \left|A+\Sigma_X\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X}'-\boldsymbol{\mu}_X)^{\mathrm{T}} (A+\Sigma_X)^{-1} (\mathbf{X}'-\boldsymbol{\mu}_X)\right\}$ is a normal PDF following normal distribution $N(\boldsymbol{\mu}_X, A+\Sigma_X)$. $m_f(\mathbf{X})$ is also a normal PDF and $\int m_f(\mathbf{X})\mathrm{d}\mathbf{X} = 1$. Then,

$$\begin{aligned} &\int m'(\mathbf{X}') f(\mathbf{X}') \mathrm{d}\mathbf{X}' \\ &= \int (2\pi)^{-\frac{n}{2}} \left|A+2\Sigma_X\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{\mu}_X - \boldsymbol{\mu}_X)^{\mathrm{T}} (A+\Sigma_X)^{-1} (\boldsymbol{\mu}_X - \boldsymbol{\mu}_X)\right\} m_f'(\mathbf{X}') \mathrm{d}\mathbf{X}' \\ &= (2\pi)^{-\frac{n}{2}} \left|A+2\Sigma_X\right|^{-\frac{1}{2}} \end{aligned} \tag{17}$$

Thus,

$$
\begin{aligned}
&\iint R(\mathbf{X},\mathbf{X}') f(\mathbf{X}) f(\mathbf{X}')\mathrm{d}\mathbf{X}\mathrm{d}\mathbf{X}' \\
&= (2\pi)^{\frac{n}{2}} |A|^{\frac{1}{2}} \int m'(\mathbf{X}') f(\mathbf{X}')\mathrm{d}\mathbf{X}' \\
&= (2\pi)^{\frac{n}{2}} |A^{-1}|^{-\frac{1}{2}} (2\pi)^{-\frac{n}{2}} |A+2\Sigma_X|^{-\frac{1}{2}} = |2A^{-1}\Sigma_X + I|^{-\frac{1}{2}}
\end{aligned} \tag{18}
$$

Similarly, for $R(\mathbf{X},\mathbf{X}^{*(i)})$ in $\boldsymbol{r}(\mathbf{X})$ can be transformed into $(2\pi)^{\frac{n}{2}} |A|^{\frac{1}{2}} m_i(\mathbf{X})$, where $m_i(\mathbf{X})$ is the PDF of a normal distribution with mean vector $\mathbf{X}^{*(i)}$ and covariance matrix A,

$$
\int R(\mathbf{X},\mathbf{X}^{*(i)}) f(\mathbf{X})\mathrm{d}\mathbf{X} = |2A^{-1}\Sigma_X + I|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X}^{*(i)}-\boldsymbol{\mu}_X)^{\mathrm{T}}(A+\Sigma_X)^{-1}(\mathbf{X}^{*(i)}-\boldsymbol{\mu}_X)\right\} \tag{19}
$$

Therefore, considering $\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})$, a row vector whose i-th element is exactly $R(\mathbf{X},\mathbf{X}^{*(i)})$,

$$
\int \boldsymbol{r}^{\mathrm{T}}(\mathbf{X}) f(\mathbf{X})\mathrm{d}\mathbf{X} = \boldsymbol{\gamma} \tag{20}
$$

$\boldsymbol{\gamma}$ is the same as in Eq.(8).

$\mathbf{X}$ and $\mathbf{X}'$ are obviously symmetrical in Eq.(12), so based on Eq.(19) and Eq.(20) , $\int \boldsymbol{r}(\mathbf{X}') f(\mathbf{X}')\mathrm{d}\mathbf{X}' = \boldsymbol{\gamma}^{\mathrm{T}}$. In addition, notice $\boldsymbol{u}(\mathbf{X}) = \boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X}) - 1$,

$$
\int \boldsymbol{u}^{\mathrm{T}}(\mathbf{X}) f(\mathbf{X})\mathrm{d}\mathbf{X} = \int \boldsymbol{u}(\mathbf{X}') f(\mathbf{X}')\mathrm{d}\mathbf{X}' = \boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{\gamma} - 1 \tag{21}
$$

Combining all results, the following closed-form expression is obtained:

$$
V(\bar{y}) = \sigma^2\left[|2A^{-1}\Sigma_X + I|^{-\frac{1}{2}} - \boldsymbol{\gamma}R^{-1}\boldsymbol{\gamma}^{\mathrm{T}} + (\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{\gamma}-1)^{\mathrm{T}}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F})^{-1}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{\gamma}-1)\right] \tag{22}
$$

$V(\bar{y})$, as the moment prediction variance of $\bar{y}$, is the statistical concept in surrogate model method that measures the uncertainty when $E(\bar{y})$ is predicted. As $V(\bar{y})$ decreases and converges closely to 0 in the iterations of adaptive Kriging model, the estimated output mean is considered more credible.

3.2. Closed-form expression of output variance

Similar to the last section, a new closed-form expression of output variance is proposed. Considering $\hat{g}(\mathbf{X})$ as a random variable, the variance can be calculated by:

$$
D = E_{\mathbf{X}}((\hat{g}(\mathbf{X})-\bar{y})^2) = \int (\hat{g}(\mathbf{X})-\bar{y})^2 f(\mathbf{X})\mathrm{d}\mathbf{X} \tag{23}
$$

Similarly being a normal random variable, D in Eq.(23) can be transformed into:

$$
E_{\mathbf{X}}((\hat{g}(\mathbf{X})-\bar{y})^2) = E_{\mathbf{X}}(\hat{g}(\mathbf{X})^2 - 2\hat{g}(\mathbf{X})\cdot\bar{y} + \bar{y}^2) = E_{\mathbf{X}}(\hat{g}(\mathbf{X})^2) - 2E_{\mathbf{X}}(\hat{g}(\mathbf{X})\cdot\bar{y}) + E_{\mathbf{X}}(\bar{y}^2) \tag{24}
$$

, in which $E_{\mathbf{X}}(\bar{y}) = \bar{y}$ because $\bar{y}$ itself is already an integral in the probability space of $\mathbf{X}$. And,

$$D = E_{\mathbf{X}}(\hat{g}(\mathbf{X})^2) - 2E_{\mathbf{X}}(\hat{g}(\mathbf{X})) \cdot \bar{y} + E_{\mathbf{X}}(\bar{y}^2) = E_{\mathbf{X}}(\hat{g}(\mathbf{X})^2) - \bar{y}^2 \tag{25}$$

$E_{\mathbf{X}}$ and E_g are having different meanings as before. Because D is a random variable similar to $\bar{y}$, the variance of structural output should be further derived and established. Notice that $E_{\mathbf{X}}(\hat{g}(\mathbf{X})^2) = \int \hat{g}(\mathbf{X})^2 f(\mathbf{X})\mathrm{d}\mathbf{X}$ and $E_g(\hat{g}(\mathbf{X})^2) = \mu_{\hat{g}}(\mathbf{X})^2 + \sigma_{\hat{g}}^2(\mathbf{X})$, then

$$E(D) = E_g(E_{\mathbf{X}}(\hat{g}(\mathbf{X})^2) - \bar{y}^2) = E_{\mathbf{X}}(E_g(\hat{g}(\mathbf{X})^2)) - E_g(\bar{y}^2) \tag{26}$$

in which d and $V(\bar{y})$ have been introduced before and $E_g(\bar{y}^2) = d^2 + V(\bar{y})$ because of the definition of variance. So,

$$E(D) = E_{\mathbf{X}}(\mu_{\hat{g}}(\mathbf{X})^2 + \sigma_{\hat{g}}^2(\mathbf{X})) - (d^2 + V(\bar{y})) = E_{\mathbf{X}}(\mu_{\hat{g}}(\mathbf{X})^2) + E_{\mathbf{X}}(\sigma_{\hat{g}}^2(\mathbf{X})) - d^2 - V(\bar{y}) \tag{27}$$

Based on Eq.(5), firstly there is:

$$E_{\mathbf{X}}(\mu_{\hat{g}}(\mathbf{X})^2) = \int \left(\beta + \boldsymbol{r}^{\mathrm{T}}(\mathbf{X})\boldsymbol{R}^{-1}(\mathbf{y} - \boldsymbol{F}\beta)\right)^2 f(\mathbf{X})\mathrm{d}\mathbf{X} \tag{28}$$

Let the $m \times 1$ vector $\boldsymbol{R}^{-1}(\mathbf{y} - \boldsymbol{F}\beta) = B$. Because $\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})B$ is 1×1 parameter, $\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})B = B^{\mathrm{T}}\boldsymbol{r}(\mathbf{X})$. It is already known that $E(\mu_{\hat{g}}(\mathbf{X})) = \int \left(\beta + \boldsymbol{r}^{\mathrm{T}}(\mathbf{X})B\right) f(\mathbf{X})\mathrm{d}\mathbf{X} = \beta + \boldsymbol{\gamma}B = d$, so

$$\int \boldsymbol{r}^{\mathrm{T}}(\mathbf{X})B \cdot f(\mathbf{X})\mathrm{d}\mathbf{X} = \boldsymbol{\gamma}B = d - \beta \tag{29}$$

In addition, $(\beta + \boldsymbol{r}^{\mathrm{T}}(\mathbf{X})B)^2 = \beta^2 + 2\beta\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})B + B^{\mathrm{T}}\boldsymbol{r}(\mathbf{X})\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})B$, so

$$\begin{aligned} E_{\mathbf{X}}(\mu_{\hat{g}}(\mathbf{X})^2) &= \int \left(\beta^2 + 2\beta\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})B + B^{\mathrm{T}}\boldsymbol{r}(\mathbf{X})\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})B\right) f(\mathbf{X})\mathrm{d}\mathbf{X} \\ &= \beta^2 + 2\beta(d - \beta) + B^{\mathrm{T}} \cdot \int \boldsymbol{r}(\mathbf{X})\boldsymbol{r}^{\mathrm{T}}(\mathbf{X}) f(\mathbf{X})\mathrm{d}\mathbf{X} \cdot B \end{aligned} \tag{30}$$

in which $\boldsymbol{r}(\mathbf{X})\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})$ is a $m \times m$ matrix with whose element in i-th row and j-th column being $R(\mathbf{X}, \mathbf{X}^{*(i)}) \cdot R(\mathbf{X}, \mathbf{X}^{*(j)})$.

We have concluded that $R(\mathbf{X}, \mathbf{X}^{*(i)})$ can be transformed into $(2\pi)^{\frac{n}{2}} |A|^{\frac{1}{2}} m_i(\mathbf{X})$, where $m_i(\mathbf{X})$ is the PDF of a normal distribution with mean vector $\mathbf{X}^{*(i)}$ and covariance matrix A, so

$$\begin{aligned} & R(\mathbf{X}, \mathbf{X}^{*(i)}) \cdot f(\mathbf{X}) = (2\pi)^{\frac{n}{2}} |A|^{\frac{1}{2}} m_i(\mathbf{X}) f(\mathbf{X}) \\ & = \left|A^{-1}\Sigma_X + I\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X}^{*(i)} - \boldsymbol{\mu}_X)^{\mathrm{T}}(A + \Sigma_X)^{-1}(\mathbf{X}^{*(i)} - \boldsymbol{\mu}_X)\right\} f_i(\mathbf{X}) \\ & = \gamma_i f_i(\mathbf{X}) \end{aligned} \tag{31}$$

γ_i is the i-th element of $\boldsymbol{\gamma}$. $f_i(\mathbf{X})$ is a normal PDF with mean vector $\boldsymbol{\mu}_i = \Sigma_i(A^{-1}\mathbf{X}^{*(i)} + \Sigma_X^{-1}\boldsymbol{\mu}_X)$ and covariance matrix $\Sigma_i = (A^{-1} + \Sigma_X^{-1})^{-1}$. Using the conclusion of Eq.(14), it can be obtained that:

$$R(\mathbf{X},\mathbf{X}^{*(i)})\cdot R(\mathbf{X},\mathbf{X}^{*(j)})\cdot f(\mathbf{X})=\left(R(\mathbf{X},\mathbf{X}^{*(i)})\cdot f(\mathbf{X})\right)\cdot R(\mathbf{X},\mathbf{X}^{*(j)})=\gamma_i f_i(\mathbf{X})R(\mathbf{X},\mathbf{X}^{*(j)}) \tag{32}$$

Because $R(\mathbf{X},\mathbf{X}^{*(j)})=(2\pi)^{\frac{n}{2}}\left|A\right|^{\frac{1}{2}}m_j(\mathbf{X})$,

$$\begin{aligned}&\gamma_i f_i(\mathbf{X})R(\mathbf{X},\mathbf{X}^{*(j)})=(2\pi)^{\frac{n}{2}}\left|A\right|^{\frac{1}{2}}\gamma_i m_j(\mathbf{X})f_i(\mathbf{X})\\&=\gamma_i\left|A^{-1}\Sigma_i+I\right|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\mathbf{X}^{*(j)}-\boldsymbol{\mu}_i)^{\mathrm{T}}(A+\Sigma_i)^{-1}(\mathbf{X}^{*(j)}-\boldsymbol{\mu}_i)\right\}f_{ij}(\mathbf{X})\\&=G_{ij}f_{ij}(\mathbf{X})\end{aligned} \tag{33}$$

$f_{ij}(\mathbf{X})$ is a normal PDF with mean vector $\boldsymbol{\mu}_{ij}=\Sigma_{ij}(A^{-1}\mathbf{X}^{*(i)}+\Sigma_i^{-1}\boldsymbol{\mu}_i)$ and covariance matrix $\Sigma_{ij}=(A^{-1}+\Sigma_i^{-1})^{-1}$. $G_{ij}=\gamma_i\left|A^{-1}\Sigma_i+I\right|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\mathbf{X}^{*(j)}-\boldsymbol{\mu}_i)^{\mathrm{T}}(A+\Sigma_i)^{-1}(\mathbf{X}^{*(j)}-\boldsymbol{\mu}_i)\right\}$ is the element in i-th row and j-th column of $\boldsymbol{G}$. Therefore,

$$\int R(\mathbf{X},\mathbf{X}^{*(i)})\cdot R(\mathbf{X},\mathbf{X}^{*(j)})f(\mathbf{X})\mathrm{d}\mathbf{X}=\int G_{ij}f_{ij}(\mathbf{X})\mathrm{d}\mathbf{X}=G_{ij} \tag{34}$$

$$\int \boldsymbol{r}(\mathbf{X})\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})f(\mathbf{X})\mathrm{d}\mathbf{X}=\boldsymbol{G} \tag{35}$$

Then,

$$E_{\mathbf{X}}(\mu_{\hat{g}}(\mathbf{X})^2)=\beta^2+2\beta(d-\beta)+B^{\mathrm{T}}\boldsymbol{G}B \tag{36}$$

For $E_{\mathbf{X}}(\sigma_{\hat{g}}^2(\mathbf{X}))$ in Eq.(27), the following deformation is firstly given:

$$\begin{aligned}E_{\mathbf{X}}(\sigma_{\hat{g}}^2(\mathbf{X}))&=\int\sigma^2\left(1-\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})+\boldsymbol{u}^{\mathrm{T}}(\mathbf{X})(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F})^{-1}\boldsymbol{u}(\mathbf{X})\right)f(\mathbf{X})\mathrm{d}\mathbf{X}\\&=\sigma^2-\sigma^2\int\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})f(\mathbf{X})\mathrm{d}\mathbf{X}+\frac{\sigma^2}{\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F}}\int\left(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})-1\right)^2f(\mathbf{X})\mathrm{d}\mathbf{X}\end{aligned} \tag{37}$$

It is noticed that $\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})$ is the sum of a $m\times m$ matrix whose element in i-th row and j-th column being $R(\mathbf{X},\mathbf{X}^{*(i)})\cdot R(\mathbf{X},\mathbf{X}^{*(j)})\cdot(\boldsymbol{R}^{-1})_{ij}$. Therefore,

$$\int R(\mathbf{X},\mathbf{X}^{*(i)})\cdot R(\mathbf{X},\mathbf{X}^{*(j)})\cdot(\boldsymbol{R}^{-1})_{ij}f(\mathbf{X})\mathrm{d}\mathbf{X}=G_{ij}\cdot(\boldsymbol{R}^{-1})_{ij} \tag{38}$$

Then

$$s_2=\int\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})f(\mathbf{X})\mathrm{d}\mathbf{X}=\sum_{i=1}^{m}\sum_{j=1}^{m}G_{ij}\cdot(\boldsymbol{R}^{-1})_{ij} \tag{39}$$

For the last term in Eq.(37), for $\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})$ is 1×1 parameter,

$$\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})=(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X}))^{\mathrm{T}}=\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})(\boldsymbol{R}^{-1})^{\mathrm{T}}\boldsymbol{F} \tag{40}$$

Then,

$$
\begin{aligned}
&\int (\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})-1)^2 f(\mathbf{X})\mathrm{d}\mathbf{X} \\
&=\int \left(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})\boldsymbol{r}^{\mathrm{T}}(\mathbf{X})(\boldsymbol{R}^{-1})^{\mathrm{T}}\boldsymbol{F}-2\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})+1\right) f(\mathbf{X})\mathrm{d}\mathbf{X} \\
&=\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{G}(\boldsymbol{R}^{-1})^{\mathrm{T}}\boldsymbol{F}-2\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{\gamma}^{\mathrm{T}}+1
\end{aligned}
\tag{41}
$$

Thus,

$$
s_3=\frac{\int (\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X})-1)^2 f(\mathbf{X})\mathrm{d}\mathbf{X}}{\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F}}=\frac{\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{G}(\boldsymbol{R}^{-1})^{\mathrm{T}}\boldsymbol{F}-2\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{\gamma}^{\mathrm{T}}+1}{\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F}}
\tag{42}
$$

Three parts of Eq.(37) are all derived, so $E_{\mathbf{X}}(\sigma_{\hat{g}}^2(\mathbf{X}))$ can be calculated by using the following closed-form expression.

$$
E_{\mathbf{X}}(\sigma_{\hat{g}}^2(\mathbf{X}))=\sigma^2(1-s_2+s_3)
\tag{43}
$$

Referring to Eqs.(27)-(43), the closed-form expression of output variance $E(D)=E_{\mathbf{X}}(\mu_{\hat{g}}(\mathbf{X})^2)+E_{\mathbf{X}}(\sigma_{\hat{g}}^2(\mathbf{X}))-d^2-V(\bar{y})$ established in AKEM can be obtained as

$$
E(D)=\beta^2+2\beta(d-\beta)+B^{\mathrm{T}}\boldsymbol{G}B+\sigma^2(1-s_2+s_3)-d^2-V(\bar{y})
\tag{44}
$$

For comparison, previous output variance is expressed as:

$$
\begin{aligned}
V_t&=E_{\mathbf{X}}((\mu_{\hat{g}}(\mathbf{X})-d)^2)=E_{\mathbf{X}}(\mu_{\hat{g}}(\mathbf{X})^2)-d^2 \\
&=\beta^2+2\beta(d-\beta)+B^{\mathrm{T}}\boldsymbol{G}B-d^2
\end{aligned}
\tag{45}
$$

Obviously, the probability integral in the prediction space is not included in Eq.(45). The output variance of $E(D)$ has extra terms $E_{\mathbf{X}}(\sigma_{\hat{g}}^2(\mathbf{X}))-V(\bar{y})$. The extra terms exist as a complement of the uncertainty which is not concerned in V_t. The result of V_t turns out to be almost the same as the result by combing the surrogate model with sampling method. However, a differential exists between $E(D)$ and V_t, unlike $E(\bar{y})=d$. It is found that as the prediction accuracy improves, both $E(D)$ and V_t will approach the true value of the structural output. Tests prove that $E(D)$ is usually closer to the true value than V_t in Kriging output. For this reason, the advance of using $E(D)$ as output instead of V_t is credible.

3.3. Adaptive strategy

Sec 3.1 and 3.2 have introduced two mainly concerned indicators. In the design of AKEM, the adaptive Kriging framework must be aiming at reducing the estimation error efficiently to ensure the accuracy of the indicators. For a Kriging surrogate model, each point in the probability space can be analyzed as contributing to the prediction uncertainty $V(\bar{y})$ and making effect on output inaccuracy.

Eq.(12) clarifies the expression of $V(\bar{y})$, which is the integral of $\mathbf{X}$ and $\mathbf{X}'$ over the entire probability space. Considering a definite point $\mathbf{X}_t$ replacing $\mathbf{X}$ in $\mathrm{Cov}(\hat{g}(\mathbf{X}),\hat{g}(\mathbf{X}'))$, $H(\mathbf{X}_t)$ represents its contribution to $V(\bar{y})$:

$$\begin{aligned} H(\mathbf{X}_t) &= \int \mathrm{Cov}(\hat{g}(\mathbf{X}_t),\hat{g}(\mathbf{X}'))f(\mathbf{X}')\mathrm{d}\mathbf{X}' \\ &= \sigma^2 \int \left(R(\mathbf{X}_t,\mathbf{X}') - \boldsymbol{r}^{\mathrm{T}}(\mathbf{x}_t)\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X}') + \boldsymbol{u}^{\mathrm{T}}(\mathbf{x}_t)(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F})^{-1}\boldsymbol{u}(\mathbf{X}') \right) f(\mathbf{X}')\mathrm{d}\mathbf{X}' \end{aligned} \tag{46}$$

Since $\int R(\mathbf{X}_t,\mathbf{X}')f(\mathbf{X}')\mathrm{d}\mathbf{X}' = \left|A^{-1}\Sigma_X + I\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X}_t-\boldsymbol{\mu}_X)^{\mathrm{T}}(A+\Sigma_X)^{-1}(\mathbf{X}_t-\boldsymbol{\mu}_X)\right\}$ and $\int \boldsymbol{r}^{\mathrm{T}}(\mathbf{X})f(\mathbf{X})\mathrm{d}\mathbf{X} = \boldsymbol{\gamma}$ are already proven, $H(\mathbf{X}_t)$ can be expressed by

$$H(\mathbf{X}_t) = \sigma^2(H_1(\mathbf{X}_t) - H_2(\mathbf{X}_t) + H_3(\mathbf{X}_t)) \tag{47}$$

in which

$$H_1(\mathbf{X}_t) = \left|A^{-1}\Sigma_X + I\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X}_t-\boldsymbol{\mu}_X)^{\mathrm{T}}(A+\Sigma_X)^{-1}(\mathbf{X}_t-\boldsymbol{\mu}_X)\right\} \tag{48}$$

$$H_2(\mathbf{X}_t) = \boldsymbol{r}^{\mathrm{T}}(\mathbf{X}_t)\boldsymbol{R}^{-1}\boldsymbol{\gamma}^{\mathrm{T}} \tag{49}$$

$$H_3(\mathbf{X}_t) = (\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\mathbf{X}_t)-1)^{\mathrm{T}}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F})^{-1}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{\gamma}^{\mathrm{T}}-1) \tag{50}$$

The larger value of $H(\mathbf{X}_t)$ identifies the more important contribution of uncertainty in $\mathbf{X}_t$ that eventually affects the posterior variance of the estimated output mean. On the other hand, a sample point with a large value of $H(\mathbf{X}_t)$ is well worth adding to the design set, and the accuracy of model can be increased as well as the uncertainty of prediction being reduced efficiently in this way. In addition, other features of training points should also be included. In this work, two representative features, point prediction variance $\sigma_{\hat{g}}^2(\mathbf{X}_t)$ and the PDF $f(\mathbf{X}_t)$ of samples are considered. These two indexes are introduced to guarantee the general reliability of AKEM.

Therefore, we define the learning function in AKEM as follows:

$$L(\mathbf{X}_t) = H(\mathbf{X}_t)\sigma_{\hat{g}}^2(\mathbf{X}_t)f(\mathbf{X}_t) \tag{51}$$

$L(\mathbf{X}_t)$ theoretically contains both uncertainty and density information of sample points. During the updating of adaptive Kriging, the sample point with the maximum value of $L(\mathbf{X}_t)$ is added to the training set in one certain iteration. Therefore, the best point to be added in the design point is chosen by:

$$\mathbf{X}_{new}^{*(m+1)} = \arg\max_{i=1}^{N} L(\mathbf{X}^i) \tag{52}$$

4. Implementation of the AKEM framework

Based on the established closed-form expressions and learning function proposed above, the algorithm framework of AKEM is introduced in this section. In the program, a sample set S_{MC} consisting of n_{MC} DoE points is first generated. Generally, though the n_{MC} can be defined flexibly, it is recommended to be large enough (depending on the complexity of the original function) to make sure the inaccurate points of small probability density are included. To start the iteration, an original training set S_{DoE} of n_0 points is obtained to build the initial model. Points of S_{DoE} are sampled following uniform distribution, because they are expected to spread out as much as possible.

The stopping criteria consists of two sub-conditions separately in order to guarantee the accuracy of both output mean and variance to their true values. Two thresholds t_1 and t_2 are established. The first sub-condition is aiming at $\bar{y}$, based on the coefficient of variation (C.O.V) that $C.O.V_{\bar{y}} = \left|\frac{\sqrt{V(\bar{y})}}{E(\bar{y})}\right|$, as $C.O.V_{\bar{y}} < t_1$. The second sub-condition is aiming at $E(D)$. The changing rate is defined as $\Delta = \left|\frac{E(D)-E'(D)}{E(D)}\right|$, in which $E(D)$ and $E'(D)$ respectively represent two consecutive outputs of previous and subsequent iterations. $\Delta < t_2$ is defined as the second one. Sometimes the model mistakenly stops and provides a definite output which deviates from the true value obviously. To avoid this situation, delayed judgment is introduced. The stopping condition is passed only if two sub-conditions are both satisfied in 2 times of consecutive iterations. Therefore, the stopping criteria can be summarized into the two sub-conditions of the main condition and the delayed judgment. Only if $C.O.V_{\bar{y}} < t_1$ and $\Delta < t_2$ are both true in two consecutive iterations, the program passes the main stopping criteria and makes output. Basically, the design is made for perfecting the model with fewer original function callings. However, the relatively complex situation makes the program iterate too much even if the output has been acceptable. According to test conditions, t_1 and t_2 will be renewed as the doubled ones between the 100^{th} and 150^{th} iterations, or tripled ones in the iterations after the 150^{th}. The specific process of the program can refer to **Fig. 1.**

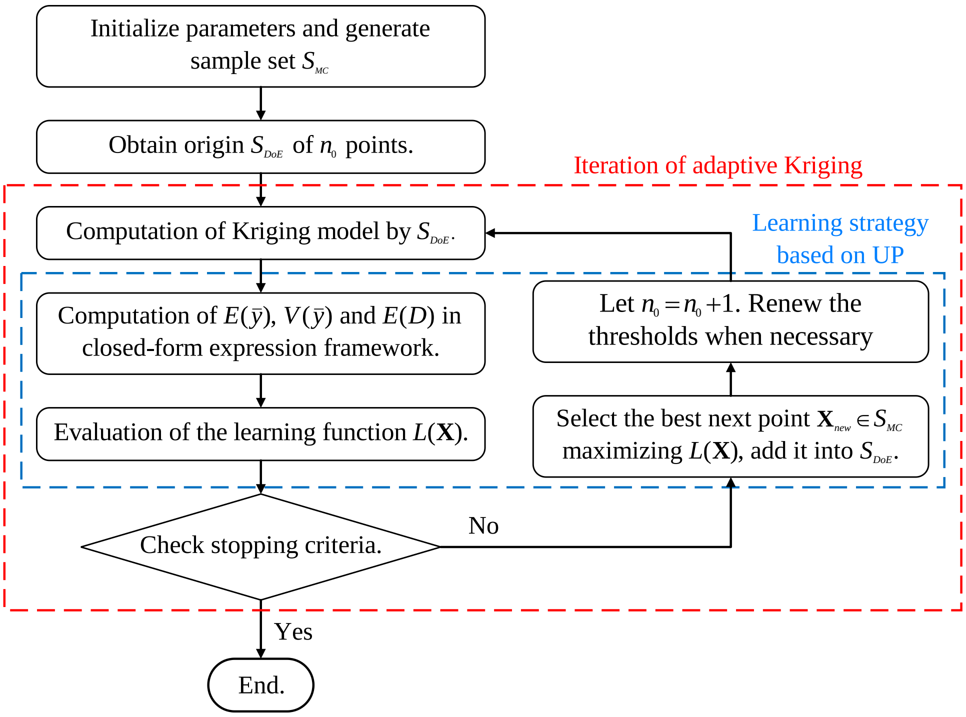


Fig. 1. Flowchart of AKEM framework

Each iteration is done along with a new point being added into S_{data}, and the original function is called once. The eventual output n_0 in this algorithm records the calling times of original function and the difference between initial n_0 and the output n_0 is the times of iteration. The output mean and variance is directly calculated based on the closed-form expressions in Sec 3.1 and 3.2.

It should be noted that some of the parameters in this algorithm are customizable to adapt to different situations, such as n_0, t_1, t_2, and the times of delayed judgment or iterations limit. In the later application examples, it is defined as $n_0 = \min\{2n+1, 12\}$ for a n-dimensional performance function and $t_1 = 0.01$, $t_2 = 0.001$.

5. Applications

5.1. *A highly-nonlinear function*

Ishigami function [69] is widely used to test uncertainty analysis methods. This function has a highly nonlinear characteristic. Respond is formulated as:

$$g(\mathbf{X}) = \sin(X_1) + a\sin(X_2)^2 + bX_3^4\sin(X_1) \tag{53}$$

where a and b are constant that can be freely defined. In this example, they are set as $a = 7$ and

$b = 0.25$. X_1, X_2 and X_3 are components of $\boldsymbol{X}$ and independently follow the same uniform distribution $U(-\pi, \pi)$. $L(\mathbf{X}_t) = H(\mathbf{X}_t)\sigma_{\hat{g}}^2(\mathbf{X}_t) f(\mathbf{X}_t)$

For reference to AKEM, one basic function should be introduced.

$$L_t(\mathbf{X}_t) = \sigma_{\hat{g}}^2(\mathbf{X}_t) \tag{54}$$

For reference, $L_t(\mathbf{X}_t)$ is used to replace the learning function of AKEM to construct a comparative adaptive Kriging framework to compare with AKEM. In the following examples, this reference method will be demonstrated in the following figures marked as Simple Kriging. The reference value of output is obtained by an extremely large number of calls of the original function and displayed in the figures as the true value.

Two graphs in **Fig. 2** visualize the convergence progress of output mean and variance estimated by the AKEM method. **Fig. 2** (a) shows that the output mean estimated by AKEM converges quickly and settles in an interval around the true value. Additionally, the confidence interval $\left[E(\bar{y})-1.96V(\bar{y}), E(\bar{y})+1.96V(\bar{y})\right]$ is also presented as a solid color area. **Fig. 2** (b) shows the change of output variance estimated by AKEM. From **Fig. 2** (b), one can see that the advantage of AKEM rather than the previous expression of output variance is quite clear, for its proximity to reference. It can also be seen that Simple Kriging takes far more iterations to converge, which means the computational costs are significantly reduced through AKEM. The learning function $L(\mathbf{X}_t)$ have been defined in Sec 3.3 to test the points of sample set. The maximum value of learning function is considered as $\mathrm{L}_{\max}$. **Fig. 3** confirms the view that the changing trend of $\mathrm{L}_{\max}$ with iteration theoretically should be decreasing as the result of the improvement of model accuracy. 136 times of original function calling are used for AKEM in total. Other simulation methods including simple MC, the Latin hypercube sampling (LHS) and the Sobol sequence are tested under the limit of the same number of function calls. These methods provide close results sometimes, but relatively inaccurate ones within a wide interval more often, see **Fig. 4**. The three methods are individually simulated 100 times. Each point in **Fig. 4** corresponds to a single result of simulation. For Simple Kriging obviously provides worse output than AKEM as shown in **Fig. 2**, it is not listed in the scatter plots. The probability of the result is expressed by the transparency of color. The distribution of these points is very dispersed, which means there is a higher probability for the output of MC or QMC lying in a more inaccurate interval of value. This deficiency exists at both output mean and output variance. This situation illustrates the instability of sampling methods under the limit of computational costs. These three sampling methods perform differently but are all unable to overcome this disadvantage. However, the only result provided by AKEM is a relatively more reliable output.

Table 1 shows the outputs of different methods to make a straight comparison with AKEM, with mean and variance's relative error written as δ_M and δ_V additionally. The truth can be acquired from δ_M and δ_V that the prediction of AKEM is significantly more accurate than the others. All indications point out that AKEM is a significantly advanced method because of its stability and accuracy with evidently cheap computational costs.

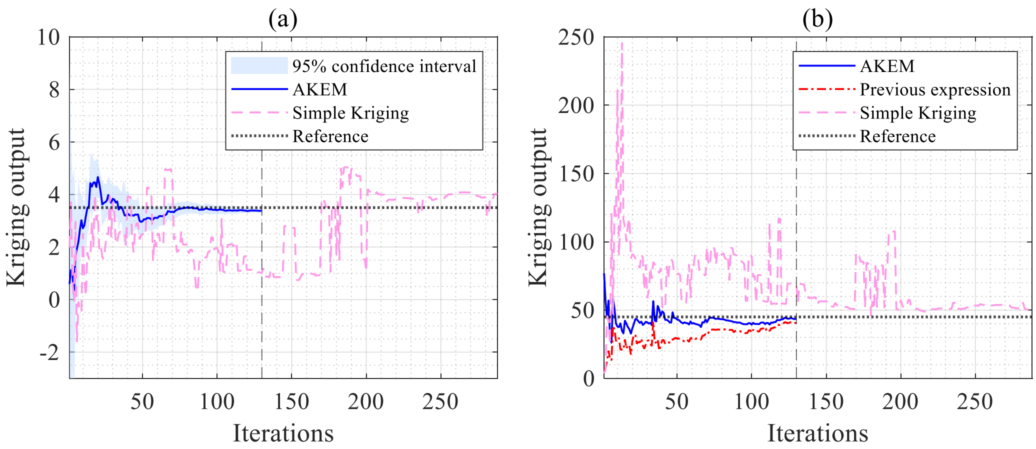


Fig. 2. Output mean and variance for example 5.1: (a) Output mean; (b) Output variance

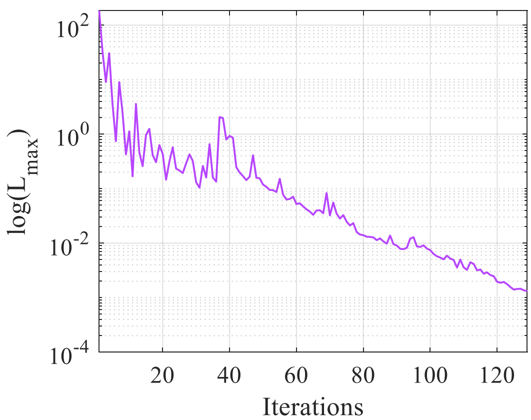


Fig. 3. $log(\mathrm{L}_{max})$ through iterations for example 5.1

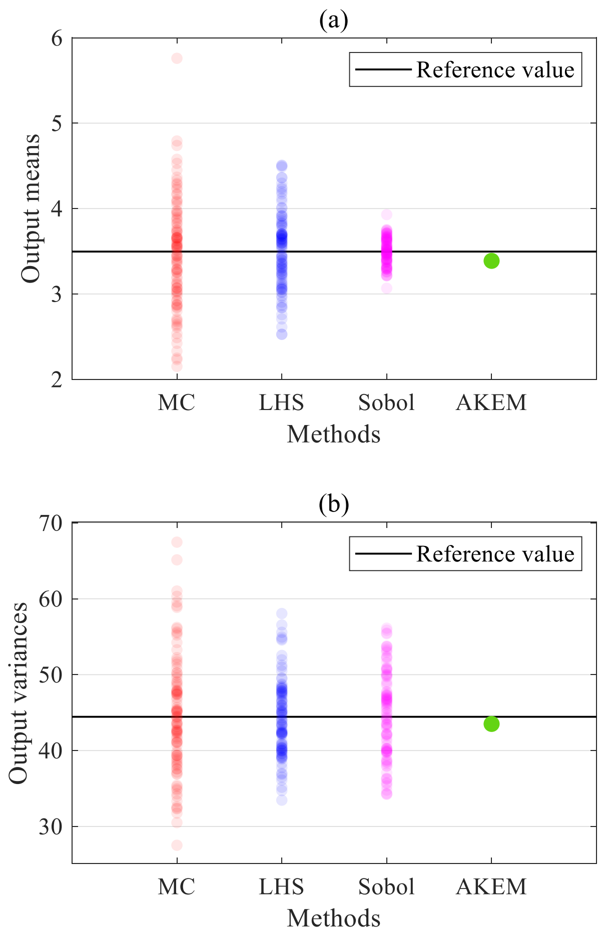


Fig. 4. Output variations of different methods for example 5.1: (a) Output mean; (b) Output variance

Table 1 Estimations of Ishigami function

Methods	Means	δ_M	Variance	δ_V	n_0
AKEM	3.39	3.21%	43.51	2.22%	136
MCS	3.29	5.95%	47.72	7.23%	
LHS	3.36	4.09%	51.96	16.75%	
Sobol	3.41	2.67%	47.01	5.65%	
Simple Kriging	4.03	15.08%	50.07	11.27%	294
Reference	3.5	0.00%	44.5	0.00%	10^6

5.2. Dynamic response of a non-linear oscillator

This example introduces a common undamped single-degree-of-freedom oscillating system shown in **Fig. 5**, which is widely used in the test of metamodel [56]. The performance function of the oscillator is defined as:

$$g(c_1, c_2, m, r, F_1, t_1) = 3r - \left| \frac{2F_1}{m\omega_0^2} \sin\left(\frac{\omega_0 t_1}{2}\right) \right| \tag{55}$$

where $\omega_0 = \sqrt{\dfrac{c_1 + c_2}{m}}$. The distribution parameters of these variables are listed in **Table 2**. The 6-dimensional application is tested and compared with other methods the same as the last section.

In this example, AKEM completes iteration by calling the original function only 36 times in total. The convergence process of outputs and comparison of Simple Kriging can be seen in **Fig. 6**. Decreasing of $\mathrm{L}_{\max}$ is presented in **Fig. 7**. **Fig. 8** contains two scatterplots of different methods' outputs. The original function calling for MC and QMC is still limited. Numeral outputs are shown in **Table 3**. Corresponding outputs are obviously highly variable, but the only output that AKEM provides is close to the reference value. Therefore, AKEM has been convinced with excellent stability and ability of efficient estimation in this example.

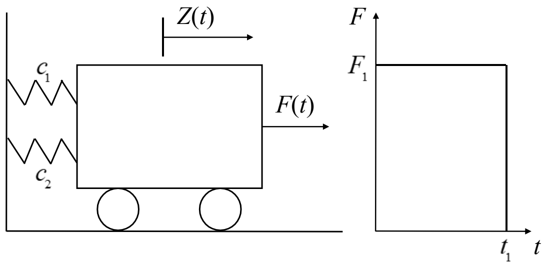


Fig. 5. A non-linear oscillator

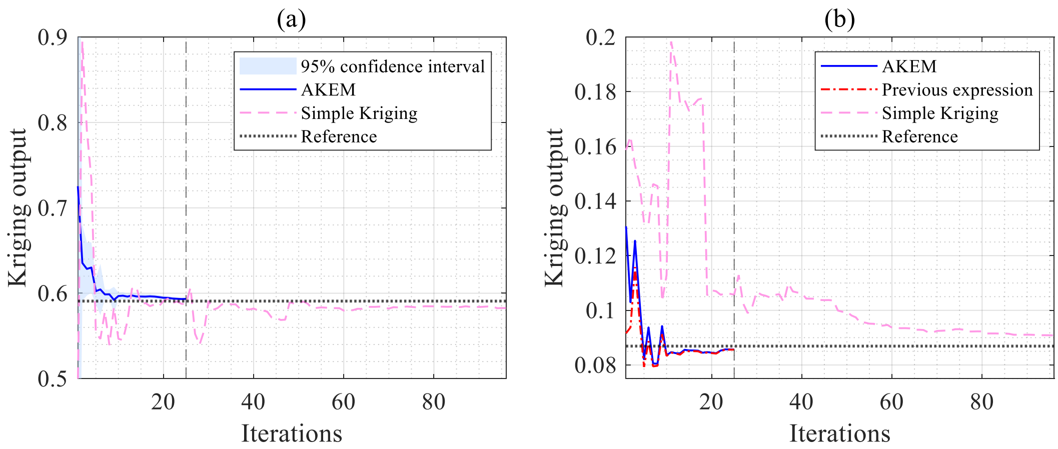


Fig. 6. Output mean and variance for example 5.2: (a) Output mean; (b) Output variance

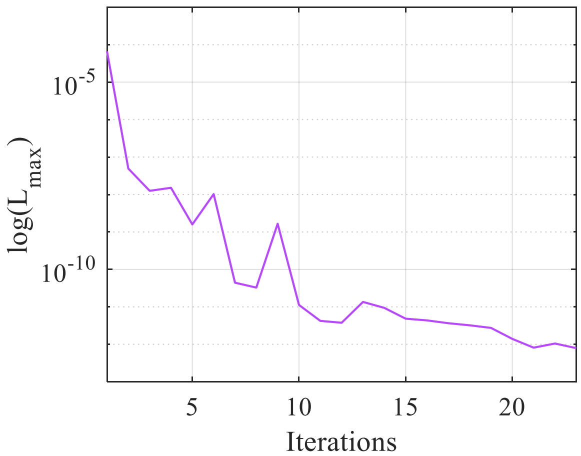


Fig. 7. $log(\mathrm{L}_{max})$ through iterations for example 5.2

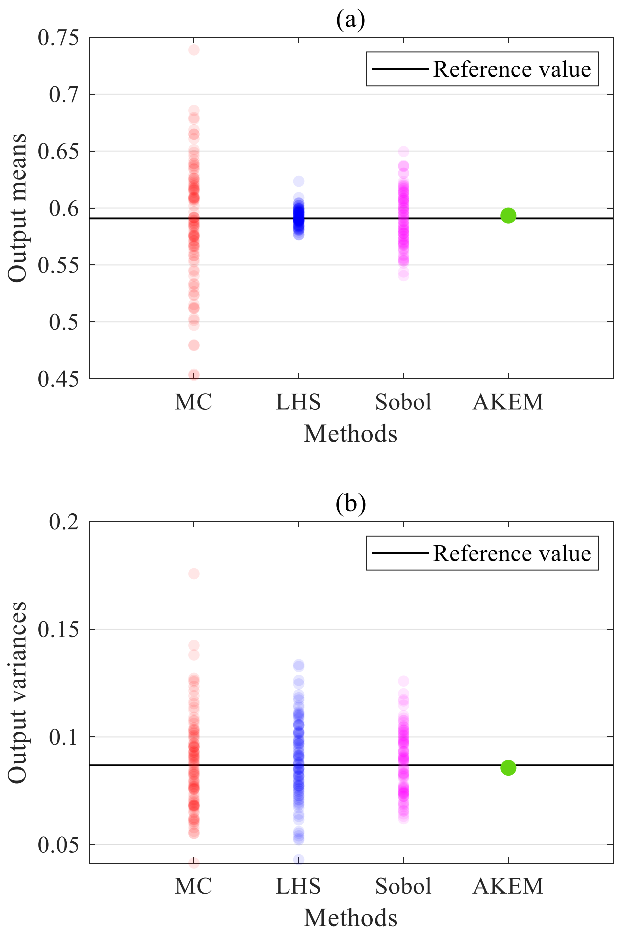


Fig. 8. Output variations of different methods for example 5.2: (a) Output mean; (b) Output variance

Table 2 Distribution of variables of non-linear oscillator

Variable	Distribution	Mean	Standard deviation
c_1	Normal	1	0.1
c_2	Normal	0.1	0.01
m	Normal	1	0.05
r	Normal	0.5	0.05
F_1	Normal	1	0.2
t_1	Normal	1	0.2

Table 3 Estimations of non-linear oscillator's response

Methods	Means	δ_M	Variance	δ_V	n_0
AKEM	0.5934	0.42%	0.0857	1.38%	
MCS	0.4970	15.89%	0.1007	15.88%	36
LHS	0.5845	1.08%	0.0929	6.90%	
Sobol	0.5725	3.11%	0.0973	11.97%	
Simple Kriging	0.5827	1.39%	0.0908	4.49%	107
Reference	0.5909	0.00%	0.0869	0.00%	10^6

5.3. Borehole function

This function describes the water flow through a borehole in one year (m^3/year) [45]:

$$g(\mathbf{X})=\frac{2\pi T_u(H_u-H_l)}{\ln(\frac{r}{r_w})(1+\frac{T_u}{T_l}+\frac{2LT_u}{\ln(\frac{r}{r_w})r_w^2K_w})} \tag{56}$$

$\mathbf{X}=(r_w,r,T_u,H_u,T_l,H_l,L,K_w)$ is input. Inputs $r_w,r,T_u,H_u,T_l,H_l,L,K_w$ respectively represent radius of the borehole, radius of influence, the transmissivity of the upper aquifer, the potentiometric head of the upper aquifer, the transmissivity of the lower aquifer, the potentiometric head of the lower aquifer, the length of the borehole and the hydraulic conductivity of the soil. Their distributions are listed in **Table 4**.

The convergence process of AKEM and comparison are displayed in **Fig. 9**. $\mathrm{L}_{\max}$ in this case can be found in **Fig. 10** and seem to have larger value than previous ones, as a result of the larger original function value. Relatively, L_{max} is still enough to guarantee the confidence of prediction. What's more, the change of output indicators reveals a law of the adaptive learning strategy of Kriging: the value of Kriging output fluctuates in a few iterations, but it eventually ends up being close to the true value in multiple iterations. The dispersion of sampling methods' results is also plotted in **Fig. 11** and details of outputs are given in **Table 5**. The prediction of AKEM is apparently more accurate than the others. In summary, the results of this example illustrate the advantages of AKEM well.

Table 4 Distribution of variables in borehole-function

Variable (unit)	Distribution	Parameter 1	Parameter 2
r_w (m)	Uniform	0.05	0.15
r (m)	Lognormal	7.71	1.0056
T_u (m^2/year)	Uniform	63070	115600
H_u (m)	Uniform	990	1110
T_l (m^2/year)	Uniform	63.1	116
H_l (m)	Uniform	700	820
L (m)	Uniform	1120	1680
K_w (m^2/year)	Uniform	9855	12045

Note: Parameter 1 and parameter 2 are minimum and maximum for uniform distribution, mean and standard deviation for the natural logarithm for lognormal distribution. Variables are independent to each other.

Table 5 Estimations of Borehole function

Methods	Means	δ_M	Variance	δ_V	n_0
AKEM	78.05	0.03%	2023	4.10%	76
MCS	74.94	3.95%	1872	11.28%	
LHS	76.96	1.36%	1846	12.51%	
Sobol	80.11	2.68%	2394	13.48%	
Simple Kriging	76.44	2.02%	2231	5.73%	169
Reference	78.02	0.00%	2110	0.00%	10^6

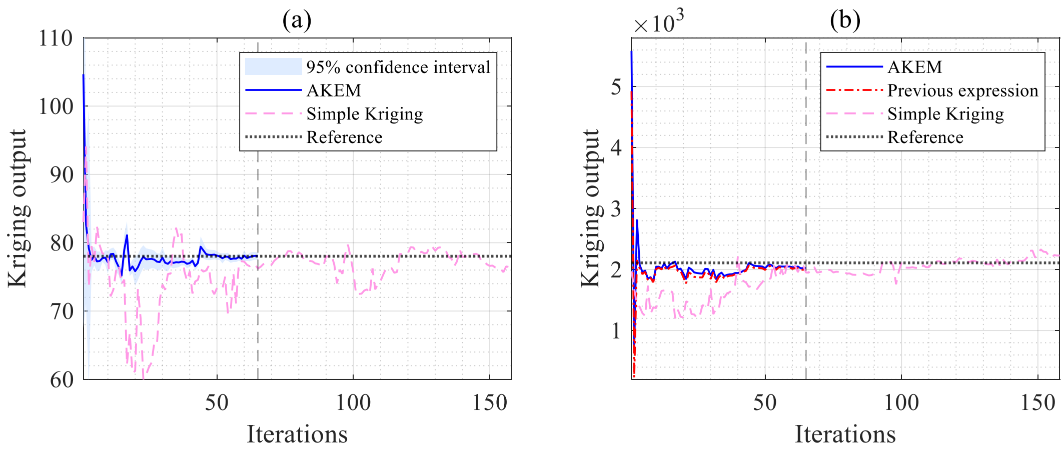


Fig. 9. Output mean and variance for example 5.3: (a) Output mean; (b) Output variance

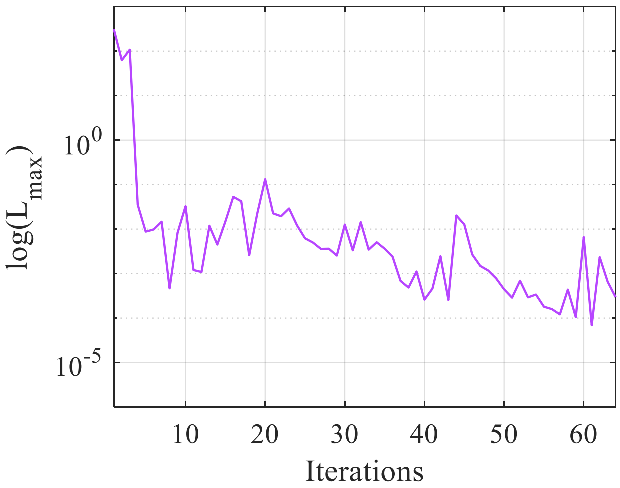
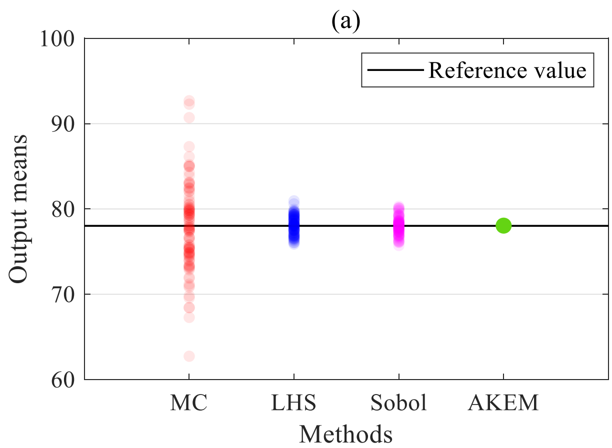


Fig. 10. $log(L_{max})$ through iterations for example 5.3



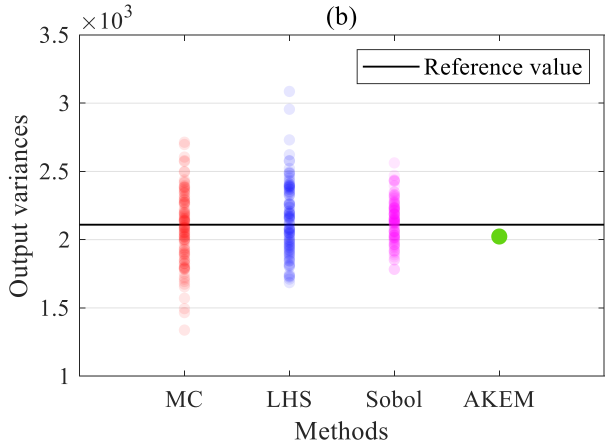


Fig. 11. Output variations of different methods for example 5.3: (a) Output mean; (b) Output variance

5.4. A front wing reinforcing rib

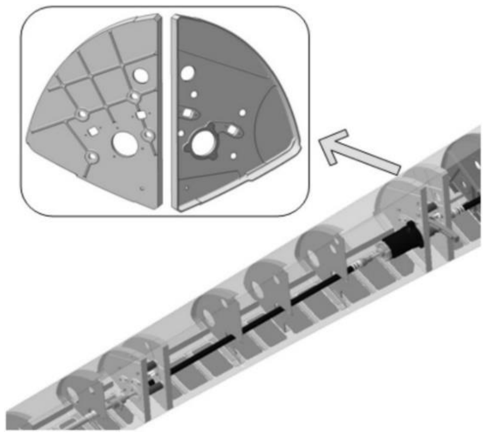


Fig. 12. Front wing reinforcing rib used for civil aircraft

(a) (b)

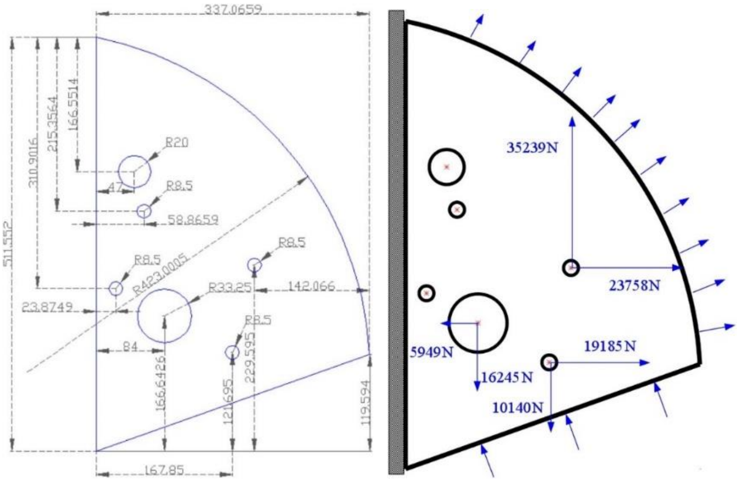


Fig. 13. (a): Structure diagram of the reinforcing rib and (b): Force diagram of the reinforcing rib

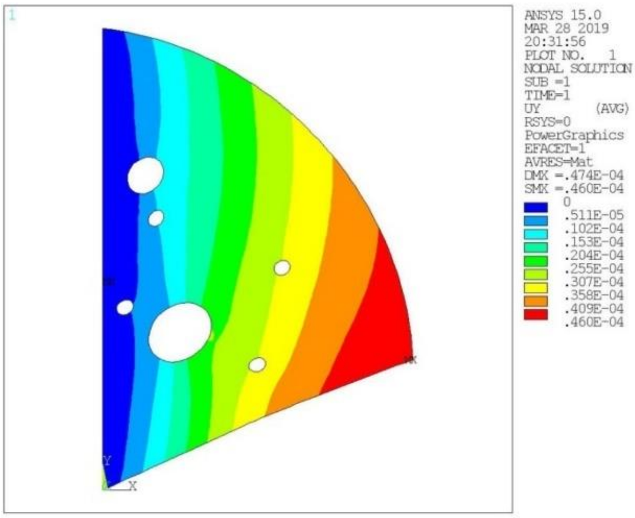


Fig. 14. Finite element model of the reinforcing rib

Fig. 12 shows one kind of front wing reinforcing rib from Ref. [69], which will be used to test AKEM. **Fig. 13** (a) is the simplified structure diagram of the reinforcing rib, where the upper edge and lower edge are respectively approximated as arc segment and line segment. Six round holes are set in the middle of the rib. The largest one of them is employed to fix the engine that generates the torque to retract the slat, the one at the top is used to insert pipes and cables, and the four remaining are for supporting the slide rails for retracting the slat. The material of this reinforcing rib is aluminum alloy 7075-T7451 whose Poisson ratio is $\nu = 0.3$. It refers to **Fig. 13** (b) for the loads uploading on the reinforcing rib. Ten random inputs are included. They are web thickness d , elastic modulus E , aerodynamic loads $P_1 = P_2 = 5000\text{GPa}$, and concentrated loads F_1 to F_6 . These random variables can be referred from **Table 6**. The structure fails when the maximum longitudinal displacement exceeds 0.068 mm. Therefore, denoting the displacement as Δ , the performance function of the reinforcing rib is concluded by

$$g(\mathbf{X}) = 0.068 - |\Delta| \tag{57}$$

Δ of the reinforcing rib can be simulated by finite element model of Ansys 15.0 as **Fig. 14**. Due to the requirement of accuracy in this example, the first threshold is set as $t_1 = 0.0025$. Applying AKEM, the original model is used 167 times. According to AKEM output, there are mean $E(\bar{y}) = 0.0093$, moment prediction variance of mean $V(\bar{y}) = 2.3247 \times 10^{-9}$ and output variance $E(D) = 1.2419 \times 10^{-5}$. It can be concluded that the expectation of displacement $E\left(|\Delta|\right) = 0.0587$. Output mean and variance converge as **Fig. 15**. In addition, 500 points are randomly sampled to make a reference to its output values can also be seen in the figure. **Fig. 16** expresses the value of $\mathrm{L}_{\max}$ in iterations. The visualized iterative process has the same trend of convergence as the three examples above. AKEM provides a result close to the reference one but spends less computational costs. In practical applications, relative parameters can be set according to the real situation, in order to further refine the model and guarantee better accuracy.

Table 6 Distribution of variables in front wing reinforcing rib

Variable (unit)	Distribution	Mean	Variation coefficient
d (mm)	Lognormal	0.05	0.05
E (GPa)	Lognormal	100	
P_1 (GPa)	Lognormal	5000	
P_2 (GPa)	Lognormal	5000	
F_1 (N)	Lognormal	32539	
F_2 (N)	Lognormal	23758	
F_3 (N)	Lognormal	5949	
F_4 (N)	Lognormal	16245	
F_5 (N)	Lognormal	10140	
F_6 (N)	Lognormal	19185	

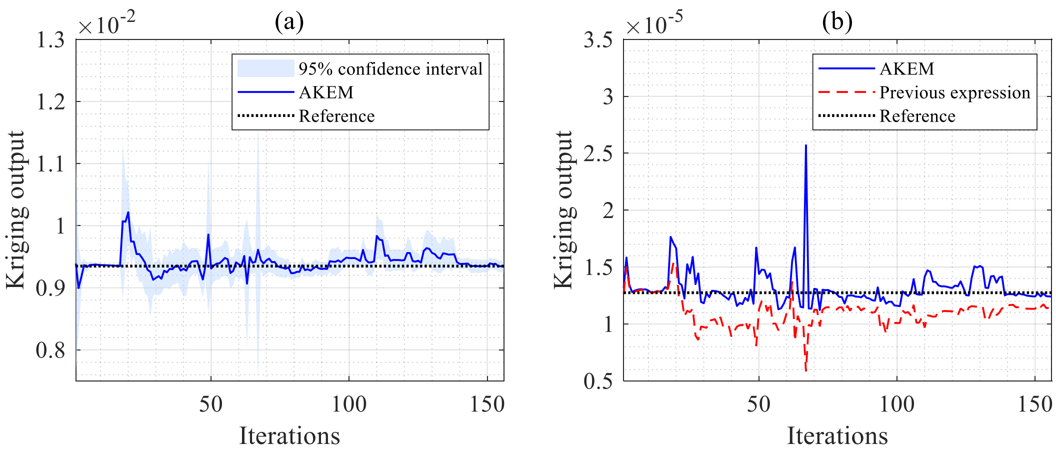


Fig. 15. Output mean and variance for example 5.4: (a) output mean; (b) output variance

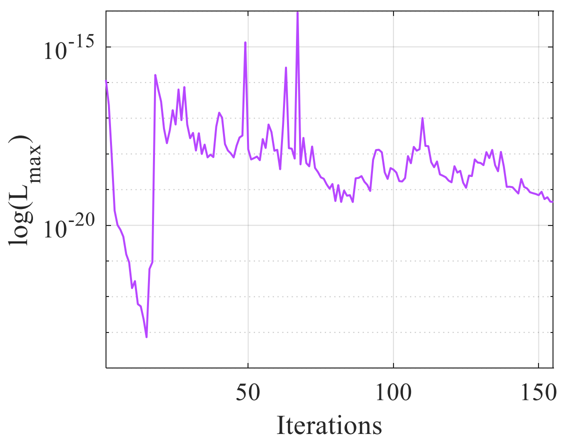


Fig. 16. $log(\mathrm{L}_{\max})$ through iterations for example 5.4

6. Conclusions

In the numerical analysis of complex systems or functions in probability space, uncertainty always brings challenges to our work. Surrogate model methods are popularly researched and used in reliability analysis, optimization, or other fields to make a simpler and sufficiently accurate substitute for the simulations difficult to complete. This paper focuses on the UA of Kriging, one kind of surrogate model. Our work is based on the characteristics of Kriging models, such as the ability to be easily combined with adaptive learning strategies to design algorithms. We solve the problem when estimating two main indicators: the mean and variance of a function under uncertainty. For statistical significance, closed-form expressions of outputs consisting of uncertainty are redefined. Traditionally, the Kriging prediction is only analyzed at a certain point as a value with prediction variance signifying confidence level. In the proposed UA, the prediction is considered as a random variable rather than a simple value, by combining the predicted value with the prediction variance. It is also noticed that uncertainty interacts in the whole space, which between two points is expressed as prediction covariance when expanded to the probability space rather than a certain point. Therefore, the prediction variance of mean, which enables the direct reflection of the uncertainty of the concerned estimation, whose effect on the accuracy of output prediction was more deeply discovered and analyzed.

Based on the above inference, the adaptive framework of AKEM is designed in order to best efficiently reduce the uncertainty in prediction and minimize computational costs for computer simulation. The key to the adaptive Kriging algorithm is a learning function based on the evaluation of uncertainty. The contribution to posterior variance can be quantified as a certain value for each generated point. Subsequently, the point with more potential to improve accuracy can be selected to update the Kriging model. Four examples are used to test AKEM. Results reveal its accuracy of output and stability of working, as well as its advantage over MC, QMC and the Simple Kriging methods. By presenting visualized data in the figures, the adaptive strategy is proved to be feasible and effective for the improvement of adaptive Kriging model. Therefore, this method is considered qualified to be applied to a variety of problems with different situations. The biggest advantage of AKEM is its efficient prediction of output mean and variance. Thus, the prospect of applying AKEM in other applications like RDO is well worth attention. In addition, it has been noticed that this method can further improve the accuracy by virtue of results in some cases, and the exploration of the prediction uncertainty of variance estimation can be pursued. Related researches will be carried out in the future.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant 52205252), the Natural Science Foundation of Sichuan Province (Grant 2023NSFSC0876), and the China Postdoctoral Science Foundation (Grant 2022M710613). The corresponding author would also thanks for the support of the Alexander von Humboldt Foundation of Germany.

References

[1] A. D. Kiureghian, O. Ditlevsen, Aleatory or epistemic? Does it matter? Structural Safety. 31, 105–112 (2009).

[2] H. B. Liu, C. Jiang, Z. Xiao, Efficient uncertainty propagation for parameterized p-box using sparse-decomposition-based polynomial chaos expansion. Mechanical Systems and Signal Processing. 138, 106589 (2020).

[3] A. Gel, R. Garg, C. Tong, M. Shahnam, C. Guenther, Applying uncertainty quantification to multiphase flow computational fluid dynamics. Powder Technology. 242, 27–39 (2013).

[4] C. Dang, P. F. Wei, M. G. Faes, M. A. Valdebenito, M. Beer, Interval uncertainty propagation by a parallel Bayesian global optimization method. Applied Mathematical Modelling. 108, 220–235 (2022).

[5] J. C. Helton, J. D. Johnson, W. L. Oberkampf, C. J. Sallaberry, Representation of analysis results involving aleatory and epistemic uncertainty. International Journal of General Systems. 39, 605–646 (2010).

[6] M. Valdebenito, H. Jensen, H. Hernández, L. Mehrez, Sensitivity estimation of failure probability applying line sampling. Reliability Engineering & System Safety. 171, 99–111 (2018).

[7] A. Kalinina, M. Spada, D. F. Vetsch, S. Marelli, C. Whealton, P. Burgherr, B. Sudret, Metamodeling for Uncertainty Quantification of a Flood Wave Model for Concrete Dam Breaks. Energies. 13, 3685 (2020).

[8] M. Moustapha, B. Sudret. Surrogate-assisted reliability-based design optimization: a survey and a unified modular framework. Structural and Multidisciplinary Optimization, 60, 2157-2176 (2019).

[9] J. Mi, N. Lu, Y.-F. Li, H.-Z. Huang, L. Bai, An evidential network-based hierarchical method for system reliability analysis with common cause failures and mixed uncertainties. Reliability Engineering & System Safety. 220, 108295 (2022).

[10] J. C. García-Merino, C. Calvo-Jurado, E. García-Macías, Polynomial chaos expansion for uncertainty propagation analysis in numerical homogenization of 2D/3D periodic composite microstructures. Composite Structures. 300, 116130 (2022).

[11] A. Tabandeh, N. Sharma, P. Gardoni, Uncertainty propagation in risk and resilience analysis of hierarchical systems. Reliability Engineering & System Safety. 219, 108208 (2022).

[12] T. Sun, R. Jiang, H. Sun, D. Liu, Z. Pan, Multiscale uncertainty propagation analysis and reliability optimization of the CFRP crossbeam of the twist beam axle. International Journal of Mechanical Sciences. 242, 108022 (2023).

[13] S. Kawai, W. Yamazaki, A. Oyama, Gegenbauer reconstruction method with edge detection for multi-dimensional uncertainty propagation. Journal of Computational Physics. 468, 111505 (2022).

[14] Z. Meng, J. Zhao, G. Chen, D. Yang, Hybrid uncertainty propagation and reliability analysis using direct probability integral method and exponential convex model. Reliability Engineering & System Safety. 228, 108803 (2022).

[15] F. Liu, P. He, Y. Dai, A new Bayesian probabilistic integration framework for hybrid uncertainty propagation. Applied Mathematical Modelling. 117, 296–315 (2023).

[16] V. J. Romero, L. P. Swiler, A. A. Giunta, Construction of response surfaces based on progressive-lattice-sampling experimental designs with application to uncertainty propagation. Structural Safety. 26, 201–219 (2004).

[17] Z. Wang, M. Daeipour, H. Xu, Quantification and propagation of Aleatoric uncertainties in topological structures. Reliability Engineering & System Safety. 233, 109122 (2023).

[18] M. Shinozuka, Monte Carlo solution of structural dynamics. Computers & Structures. 2, 855–874 (1972).

[19] I. Sobol', On the distribution of points in a cube and the approximate evaluation of integrals. USSR Computational Mathematics and Mathematical Physics. 7, 86–112 (1967).

[20] C. P. Xing, H. Niederreiter, A construction of low-discrepancy sequences using global function fields. Acta Arithmetica. 73, 87–102 (1995).

[21] R. Cools, F. Y. Kuo, D. Nuyens, Constructing Embedded Lattice Rules for Multivariate Integration. SIAM Journal on Scientific Computing. 28, 2162–2188 (2006).

[22] T. F. Hou, D. Nuyens, S. Roels, H. Janssen, Quasi-Monte Carlo based uncertainty analysis: Sampling efficiency and error estimation in engineering applications. Reliability Engineering & System Safety. 191, 106549 (2019).

[23] S. Ferson, L. R. Ginzburg, Different methods are needed to propagate ignorance and variability. Reliability Engineering & System Safety. 54, 133–144 (1996).

[24] S. Ferson, What Monte Carlo methods cannot do. Human and Ecological Risk Assessment: An International Journal. 2, 990–1007 (1996).

[25] W. L. Oberkampf, C. J. Roy, Verification and validation in scientific computing. Cambridge University Press (2010).

[26] M. Beer, S. Ferson, V. Kreinovich, Imprecise probabilities in engineering analyses. Mechanical Systems and Signal Processing. 37, 4–29 (2013).

[27] H. B. Liu, C. Jiang, X. Y. Jia, X. Y. Long, Z. Zhang, F. J. Guan, A new uncertainty propagation method for problems with parameterized probability-boxes. Reliability Engineering & System Safety. 172, 64–73 (2018).

[28] A. Abdedou, A. Soulaïmani, A non-intrusive reduced-order modeling for uncertainty propagation of time-dependent problems using a B-splines Bézier elements-based method and proper orthogonal decomposition: Application to dam-break flows. Computers & Mathematics with Applications. 102, 187–205 (2021).

[29] B. Zhang, Y. C. Shin, An adaptive Gaussian mixture method for nonlinear uncertainty propagation in neural networks. Neurocomputing. 458, 170–183 (2021).

[30] T. Q. D. Pham, T. V. Hoang, X. V. Tran, S. Fetni, L. Duchêne, H. S. Tran, A. M. Habraken, Characterization, propagation, and sensitivity analysis of uncertainties in the directed energy deposition process using a deep learning-based surrogate model. Probabilistic Engineering Mechanics. 69, 103297 (2022).

[31] D. Lee, R. Jahanbin, S. Rahman, Robust design optimization by spline dimensional decomposition. Probabilistic Engineering Mechanics. 68, 103218 (2022).

[32] F. Pellizzari, G. C. Marano, A. Palmeri, R. Greco, M. Domaneschi, Robust optimization of MTMD systems for the control of vibrations. Probabilistic Engineering Mechanics. 70, 103347 (2022).

[33] S. Chakraborty, S. Das, S. Tesfamariam, Robust design optimization of nonlinear energy sink under random system parameters. Probabilistic Engineering Mechanics. 65, 103139 (2021).

[34] S. Mohammadi, S. Cremaschi, Efficiency of Uncertainty Propagation Methods for Estimating Output Moments. Computer Aided Chemical Engineering, 487–492 (2019).

[35] J. C. Helton, F. J. Davis, Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliability Engineering & System Safety. 81, 23–69 (2003).

[36] T. Hou, D. Nuyens, S. Roels, H. Janssen, Quasi-Monte Carlo based uncertainty analysis: Sampling efficiency and error estimation in engineering applications. Reliability Engineering & System Safety. 191, 106549 (2019).

[37] Y. Q. Liu, H. V. Gupta, Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. Water Resources Research. 43, W07401 (2007).

[38] Z. Wang, C. Jiang, P. W. Liu, W. H. Yang, Y. Zhao, M. F. Horstemeyer, L. Q. Chen, Z. Hu, L. Chen, Uncertainty quantification and reduction in metal additive manufacturing. npj Computational Materials. 6, 175 (2020).

[39] M. Arnst, J. P. Ponthot, An overview of nonintrusive characterization propagation, and sensitivity analysis of uncertainties in computational mechanics. International Journal for Uncertainty Quantification. 4, 387–421 (2014).

[40] S. Mohammadi, S. Cremaschi, Efficiency of uncertainty propagation methods for moment estimation of uncertain model outputs. Computers & Chemical Engineering. 166, 107954 (2022).

[41] M. Gibanica, T. J. Abrahamsson, Data-driven modal surrogate model for frequency response uncertainty propagation. Probabilistic Engineering Mechanics. 66, 103142 (2021).
[42] D. R. Jones, M. Schonlau, W. J. Welch, Efficient Global Optimization of Expensive Black-Box Functions. Journal of Global Optimization. 13, 455–492 (1998).
[43] Q. Guo, H. Zhai, B. Suo, W. Zhao, Y. Liu, Time-variant reliability global sensitivity analysis with single-loop Kriging model combined with importance sampling. Probabilistic Engineering Mechanics. 72, 103441 (2023).
[44] A. Thapa, A. Roy, S. Chakraborty, Reliability analysis of underground tunnel by a novel adaptive Kriging based metamodeling approach. Probabilistic Engineering Mechanics. 70, 103351 (2022).
[45] J. An, A. Owen, Quasi-regression. Journal of Complexity. 17, 588–607 (2001).
[46] P. F. Wei, F. C. Liu, C. H. Tang, Reliability and reliability-based importance analysis of structural systems using multiple response Gaussian process model. Reliability Engineering & System Safety. 175, 183–195 (2018).
[47] L. J. Gomez, A. C. Yucel, L. Hernandez-Garcia, S. F. Taylor, E. Michielssen, Uncertainty Quantification in Transcranial Magnetic Stimulation via High-Dimensional Model Representation. IEEE Transactions on Biomedical Engineering. 62, 361–372 (2015).
[48] M. Ratto, A. Pagano, P. Young, State Dependent Parameter metamodelling and sensitivity analysis. Computer Physics Communications. 177, 863–876 (2007).
[49] M. M. Rajabi, B. Ataie-Ashtiani, C. T. Simmons, Polynomial chaos expansions for uncertainty propagation and moment independent sensitivity analysis of seawater intrusion simulations. Journal of Hydrology. 520, 101–122 (2015).
[50] K. Cheng, Z. Z. Lu, Y. C. Zhou, Y. Shi, Y. H. Wei, Global sensitivity analysis using support vector regression. Applied Mathematical Modelling. 49, 587–598 (2017).
[51] R. Tripathy, I. Bilionis, M. Gonzalez, Gaussian processes with built-in dimensionality reduction: Applications to high-dimensional uncertainty propagation. Journal of Computational Physics. 321, 191–223 (2016).
[52] S. N. Lophaven, H. B. Nielsen, J. Søndergaard, DACE: a Matlab kriging toolbox (Vol. 2). IMM, Informatics and Mathematical Modelling, The Technical University of Denmark (2002).
[53] D. Rochman, W. Zwermann, S. C. V. D. Marck, A. J. Koning, H. Sjöstrand, P. Helgesson, B. Krzykacz-Hausmann, Efficient Use of Monte Carlo: Uncertainty Propagation. Nuclear Science and Engineering. 177, 337–349 (2014).
[54] Ning-Cong Xiao, Kai Yuan, Hongyou Zhan. System reliability analysis based on dependent Kriging predictions and parallel learning strategy. Reliability Engineering and System Safety. 218, 108083 (2022).
[55] Y. S. Liu, L. Y. Li, S. H. Zhao, S. F. Song, A global surrogate model technique based on principal component analysis and Kriging for uncertainty propagation of dynamic systems. Reliability Engineering & System Safety. 207, 107365 (2021).
[56] B. Echard, N. Gayton, M. Lemaire, AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. Structural Safety. 33, 145–154 (2011).
[57] N. Lelièvre, P. Beaurepaire, C. Mattrand, N. Gayton, AK-MCSi: A Kriging-based method to deal

with small failure probabilities and time-consuming models. Structural Safety. 73, 1–11 (2018).

[58] B. Echard, N. Gayton, M. Lemaire, N. Relun, A combined Importance Sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. Reliability Engineering & System Safety. 111, 232–240 (2013).

[59] X. X. Huang, J. Q. Chen, H. P. Zhu, Assessing small failure probabilities by AK–SS: An active learning method combining Kriging and Subset Simulation. Structural Safety. 59, 86–95 (2016).

[60] N. Razaaly, P. M. Congedo, Extension of AK-MCS for the efficient computation of very small failure probabilities. Reliability Engineering & System Safety. 203, 107084 (2020).

[61] P. F. Wei, J. W. Song, S. F. Bi, M. Broggi, M. Beer, Z. Z. Lu, Z. F. Yue, Non-intrusive stochastic analysis with parameterized imprecise probability models: I. Performance estimation. Mechanical Systems and Signal Processing. 124, 349–368 (2019).

[62] P. F. Wei, J. W. Song, S. F. Bi, M. Broggi, M. Beer, Z. Z. Lu, Z. Z. Yue, Non-intrusive stochastic analysis with parameterized imprecise probability models: II. Reliability and rare events analysis. Mechanical Systems and Signal Processing. 126, 227–247 (2019).

[63] X. Peng, T. Ye, W. Hu, J. Li, Z. Liu, S. Jiang, Construction of adaptive Kriging metamodel for failure probability estimation considering the uncertainties of distribution parameters. Probabilistic Engineering Mechanics. 70, 103353 (2022).

[64] W. Yun, Z. Lu, L. Wang, K. Feng, P. He, Y. Dai, Error-based stopping criterion for the combined adaptive Kriging and importance sampling method for reliability analysis. Probabilistic Engineering Mechanics. 65, 103131 (2021).

[65] Z. Z. Song, H. Y. Zhang, L. Zhang, Z. Liu, P. Zhu, An estimation variance reduction-guided adaptive Kriging method for efficient time-variant structural reliability analysis. Mechanical Systems and Signal Processing. 178, 109322 (2022).

[66] P. F. Wei, X. Zhang, M. Beer, Adaptive experiment design for probabilistic integration. Computer Methods in Applied Mechanics and Engineering. 365, 113035 (2020).

[67] Y. Shi, Z. Z. Lu, S. Y. Chen, L. Y. Xu, A reliability analysis method based on analytical expressions of the first four moments of the surrogate model of the performance function. Mechanical Systems and Signal Processing. 111, 47–67 (2018).

[68] T. Ishigami, T. Homma, An importance quantification technique in uncertainty analysis for computer models. Proceedings. First International Symposium on Uncertainty Modeling and Analysis. IEEE, 398–403 (1990).

[69] Y. Shi, Z. Lu, R. He, Y. Zhou, S. Chen, A novel learning function based on Kriging for reliability analysis. Reliability Engineering & System Safety. 198, 106857 (2020).