SVR enhanced Kriging for optimization with noisy evaluations

Youquan Du[1], Keshi Zhang[1*], Peixia Lu[1] and Zhonghua Han[1]

[1] School of Aeronautics, Northwestern Polytechnical University, Xi'an, 710072, P.R. China
zhangkeshi@nwpu.edu.cn

Abstract. Numerical noise is an unavoidable by-product of Computational Fluid Dynamics (CFD) simulations, which bring challenges to optimizations. In the former work, we have proposed the ε-kriging model that can adaptively filter the numerical noise in the sample data by adding the insensitive factor (ε) of a support vector regression (SVR) model to the diagonal of the correlation matrix of a kriging model. Here we aim to develop the surrogate optimization method based on it for tackling the problems with noisy evaluations. The infilling criterion is developed to guide global optimization. It is compared with the classical kriging based optimization for couples of benchmark problems varying nonlinearity and dimension, with noise of low, medium and high intensity. The results show that our method successfully converged to the global optimums no matter how strong the numerical noise is. Drag minimization of NACA0012 airfoil also obtained satisfactory results. The results indicate that our method is effective and robust for optimizations affected by noise.

Keywords: Surrogate based optimization, Numerical noise, ε-kriging, SVR

1 Introduction

During the past two decades, the efficient global optimization method (EGO) [1-4] based on kriging model and expected improvement (EI) has gained a great attention in engineering design optimizations [5,6]. Though widely applied, numerical noise [7-10] still a hindrance, so that the researchers put much effort on developing an EGO method to identify the global optimum in the optimizations with noisy evaluations [11-16].

Due to the numerical noise, Giunta et al. [17] encountered convergence problems in aerodynamic design of a high-speed civil transport aircraft (HSCT), in which the optimization process was trapped into local minima. Shy et al. [18] and Burman et al. [19] had to cope with numerical noise in their researches too. Forrester et al. [20] explored the source of numerical noise in aerodynamic design of airfoils and figured out that discretization error, incomplete convergence and inaccurate applications of boundary conditions were the main reasons. Gilkeson et al. [21,22] conducted series of CFD simulations of a generic road vehicle and then demonstrated that the numerical noise is a result of interaction between the turbulence model and the grid type. Our research group [23] explored the physical mechanism and observed that numerical noise was generated in airfoil and wing optimizations, then concluded that mesh-quality discrepancy of

different geometries might lead to small truncation error shown as low-level numerical noise, while incomplete convergence is most likely encountered in the complex flow condition such as transonic buffet and drag divergence, which incurs high-level numerical noise that must have critical impact on optimizations. We also found that low-boom design of a supersonic transport was probably interfered by the numerical noise [24].

It is common to account for "noise" in the data by adding a regularization factor to the main diagonal of the correlation matrix of an ordinary kriging model, which is called "nugget-effect" kriging model that is the term from geo-statistics [25], and this regularization factor is called nugget. In the case of noisy data, nugget should be valued proportional to the noise variance. Cressie [25] valued nugget as the ratio of the noise variance to the process variance. Yin et al. [26-28] extended this method to handle the heteroscedastic variance case by varying the nugget with the sample location. Forrester [20] proposed to optimize nugget together with the hyperparameters of a kriging model and applied it to an airfoil aerodynamic optimization problem interfered by numerical noise. Sakata et al. [29,30] added the nugget into the response of the samples, then derived a new noise-filtering kriging model, ns-kriging. Chen et al. [31] proposed to value the nugget with the variance of the support vectors obtained by SVR. Bostanabad and Chen et al. [32] analyzed impact of the nugget on the log-likelihood function and proposed adaptively adjusting nugget in the modeling process to simultaneously filter noise and improve prediction performance. Though some research about noisy optimization has been studied based on "nugget-effect" kriging model, there is still room for improvement on construction of the model and optimization method. The objective of this article is to develop a surrogate optimization method based on ε-kriging [33] model which could adaptively filter the noise, for the optimizations with noisy evaluations.

The paper is organized as follows: Section 2 briefly introduces the theory of the ε-kriging model; Section 3 proposed a primary optimization framework based on it; Section 4 show the validation results via analytical test functions and an airfoil aerodynamic design case.

2 Background formulation of ε-Kriging model

Given a training data set $\boldsymbol{D} = (\boldsymbol{X}_S, \boldsymbol{Y}_S) = \{(\boldsymbol{x}^{(i)}, y^{(i)}) | i = 1,2,\cdots,n\}$, where $\boldsymbol{x}^{(i)} \in \mathbb{R}^m$ denotes the input vector and $y^{(i)}$ is its corresponding response ($y^{(i)} = f(\boldsymbol{x}^{(i)}) + s_i$). It is assumed that the noise in the sample data is homoscedastic and $s_i \sim N(0, \sigma_s^2)$.

Assume a random process $Y(\boldsymbol{x})$ corresponding to the unknown function.

$$Y(\boldsymbol{x}) = \beta_0 + Z(\boldsymbol{x}) \tag{1}$$

where β_0 is an unknown constant that depicts the global trend, and $Z(\boldsymbol{x})$ is defined as a stationary random process.

Inspired by the ε-tube of SVR which is favored due to good capability of filtering numerical noise, we proposed to value the nugget factor [25] by the insensitive factor (ε) and call it ε-kriging model [33]. The linear regression function $\hat{f}(\boldsymbol{x}) = \langle \boldsymbol{w} \cdot \boldsymbol{x} \rangle + b$ of ε-SVR is constructed to deviate least from the training set according to Vapnik's ε-

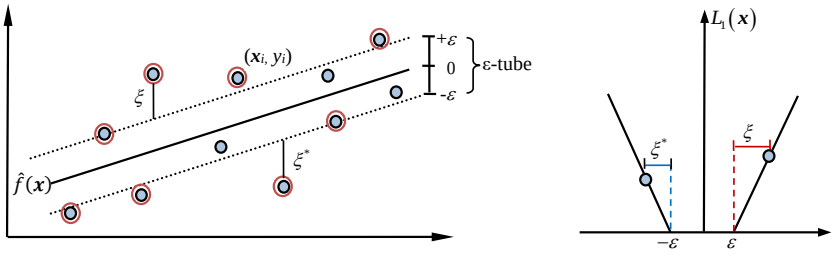
insensitive loss function [34]

$$L_1(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \left|\hat{f}(\boldsymbol{x}) - y\right| \le \varepsilon \\ \left|\hat{f}(\boldsymbol{x}) - y\right| - \varepsilon & \text{otherwise} \end{cases} \tag{2}$$

while at the same time is as "flat" as possible (i.e., $\|\boldsymbol{w}\|$ is as small as possible). Mathematically, this means

$$\begin{aligned} \min \quad & \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \\ \text{s.t.} \quad & y_i - \left\langle \boldsymbol{w} \bullet \boldsymbol{x}_i \right\rangle - b \le \varepsilon + \xi_i \\ & \left\langle \boldsymbol{w} \bullet \boldsymbol{x}_i \right\rangle + b - y_i \le \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \ge 0 \end{aligned} \tag{3}$$

where (ξ_i, ξ_i^*) indicates that the model allows some exceptions and the constant $C(C > 0)$ controls the tradeoff between complexity of the model and deviations larger than ε. A key assumption of this formulation is that there exists a function $f(\boldsymbol{x})$ that can approximate all pairs of $(\boldsymbol{x}_i, y_i)$ in ε precision by a so-called "ε-tube" or "ε-insensitive zone". In a linear case, the ε-tube and the insensitive loss function is shown in Fig. 1. By mapping $\boldsymbol{x}$ to $\boldsymbol{z} = \psi(\boldsymbol{x})$ in a feature space via a nonlinear map ψ (kernel function), any nonlinear function can be modeled.



a) ε-tube in the sample space b) ε-insensitive loss function

Fig. 1. ε-tube and error allowance of the ε-SVR (samples in a red circle: support vectors)

The parameter ε controls width of the ε-tube. The ε-tube covers most of the samples with "normal" noise ($\boldsymbol{x} \sim N(0, \sigma_s^2)$) (inside or on the border of the ε-tube), while allowing for some exceptions outside it (like the "bad points" occasionally happened in the numerical simulations). ε-Kriging takes ε as the nugget factor to be added to the diagonal of the correlation matrix of the kriging model. The optimal ε is obtained adaptively via our in-house SVR code [35], and the details of it would be found in [33]

Then the ε-kriging model is built based on data set $\boldsymbol{D}$, and the prediction of the function $y(\boldsymbol{x})$ at any untried point $\boldsymbol{x}$ can be written as

$$\hat{f}(\boldsymbol{x}) = \beta_{\varepsilon 0} + \boldsymbol{r}^{\mathrm{T}}(\boldsymbol{x}) \underbrace{\boldsymbol{R}_{\varepsilon}^{-1}(\boldsymbol{Y}_{\mathrm{S}} - \beta_{\varepsilon 0}\boldsymbol{F})}_{\boldsymbol{V}_{\varepsilon\ \mathrm{Krig}}} \tag{4}$$

where $\beta_{\varepsilon 0} = (\boldsymbol{F}^{\mathrm{T}}{\boldsymbol{R}_{\varepsilon}}^{-1}\boldsymbol{F})^{-1}\boldsymbol{F}^{\mathrm{T}}{\boldsymbol{R}_{\varepsilon}}^{-1}\boldsymbol{Y}_{\mathrm{S}}$, and the column vector $V_{\varepsilon\ \mathrm{Krig}}$ is a product of the given samples, so it can be stored for the subsequent predictions; $\boldsymbol{F} \in \mathbb{R}^{\boldsymbol{n}}$ is the column vector of each element being 1, $\boldsymbol{R}_{\boldsymbol{\varepsilon}}$ is the correlation matrix in which each element is the correlation function product of two samples. where $r(\cdot)$ is the spatial correlation function, which is only dependent on the spatial distance between two samples.

$$\boldsymbol{R}_{\varepsilon} = \begin{bmatrix} 1+\varepsilon & r\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}\right) & \cdots & r\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(n)}\right) \\ r\left(\boldsymbol{x}^{(2)}, \boldsymbol{x}^{(1)}\right) & 1+\varepsilon & \cdots & r\left(\boldsymbol{x}^{(2)}, \boldsymbol{x}^{(n)}\right) \\ \vdots & \vdots & \vdots & \vdots \\ r\left(\boldsymbol{x}^{(n)}, \boldsymbol{x}^{(1)}\right) & r\left(\boldsymbol{x}^{(n)}, \boldsymbol{x}^{(2)}\right) & \cdots & 1+\varepsilon \end{bmatrix} \in \mathbb{R}^{n\times n} \quad \boldsymbol{r}(\boldsymbol{x}) = \begin{bmatrix} r(\boldsymbol{x}^{(1)}, \boldsymbol{x}) \\ r(\boldsymbol{x}^{(2)}, \boldsymbol{x}) \\ \vdots \\ r(\boldsymbol{x}^{(n)}, \boldsymbol{x}) \end{bmatrix} \tag{5}$$

3 Optimization based on ε-Kriging and EI method

Based on the ε-kriging method, this work aims to priliminarily develop an optimization framework to solve the optimization problem with noisy responses below.

$$\begin{aligned} &\min y(\boldsymbol{x}) + N(0, \sigma_s^2) \\ &s.t. g_i(\boldsymbol{x}) \le 0, i = 1, 2 \ldots N_C \\ &w.r.t. \boldsymbol{x}_{low} \le \boldsymbol{x} \le \boldsymbol{x}_{up} \end{aligned} \tag{6}$$

where $y(\boldsymbol{x})$ and $g_i(\boldsymbol{x})$ denote the objective and constraint functions, N_c is the number of constraint functions, and the test functions of this paper are unconstrainted; $\boldsymbol{x}_{low}$ and $\boldsymbol{x}_{up}$ are the lower and upper bounds of the design variables $\boldsymbol{x}$.

Any function in Eq.(6) that is probably affected by noise would be modeled by the ε-kriging method. Then the EI infilling criterion is used to add the new samples to improve the modeling accuracy as well as guide to the global optimum, and the ε-kriging models are repetitively updated until the convergence criteria are reached. The optimization framework is shown in Fig. 2. The main steps are described as follows:

1) The initial samples are generated by the DoE method (such as the Latin hypercube sampling (LHS)) and their responses are are solved via the high-fidelity analysis tools such as Renolds-averaged Navier-Stokes (RANS) equations.

2) The optimal hyperparameter ε is obtained via training the SVR model in which its hyperparameters are optimized by minimizing the generalization error of the model. The generalization error can be estimated by the cross validation (CV) method or the leave-one-out (LOO) bound method. The optimizers including Genetic Algorithms (GA), Covariance Matrix Adaptation Evolution Strategy (CMA-ES) and Bayesian Optimization (BO) are optional for the hyperparameter optimization.

3) The optimal hyperparameters of SVR (ε) is added to the diagonal of the correlation matrix, subsequently the kriging modeling module of the in-house SurroOpt [36] code will be called to build the ε-kriging model.

4) The locations of the new samples are generated by maximizing the EI function via the built-in optimization methods (GA, Hooke&Jeeves, BFGS or the combined method of them), and the resoponse(s) are evaluated.

5) The new samples are infilled to the data set "***D***" and then the ε-kriging models are re-trained. The steps 2)–5) are repeated until the termination condition is satisfied.

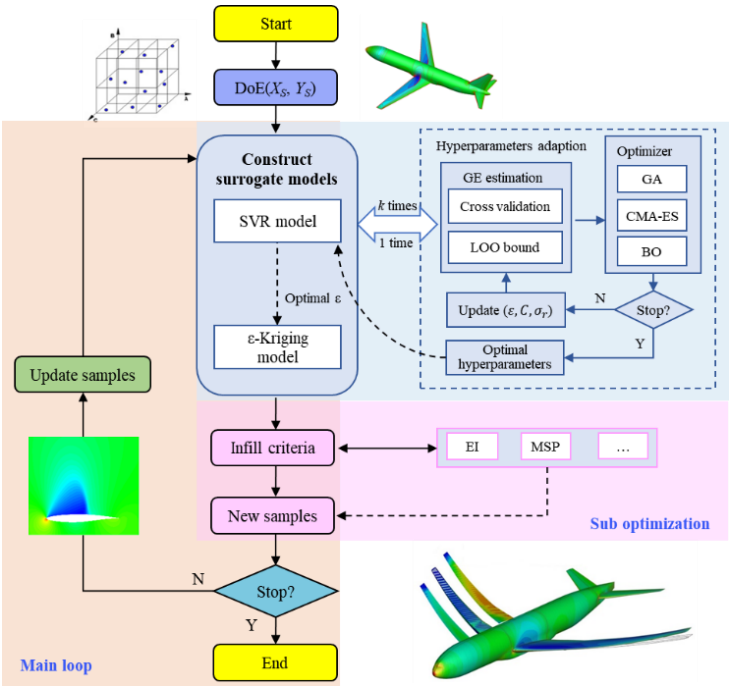


Fig. 2. Flowchart of optimization based on developed ε-kriging and EI method

The infilling criterion has critical impact on optimal efficiency and results. Here the EI criterion [1] is investigated to see what happens to it when the noise exists. Assume that the prediction of ε-kriging model at any untried site $\boldsymbol{x}$ obeys a normal distribution $\hat{Y}(\boldsymbol{x}) \sim N[\hat{f}(\boldsymbol{x}), s^2(\boldsymbol{x})]$, where the mean is the surrogate prediction $\hat{f}(\boldsymbol{x})$ and the variance $s^2(\boldsymbol{x})$ is the corresponding MSE [15]. Then the statistical improvement at any untried location observed so far y_{min} is defined as:

$$I(\boldsymbol{x}) = \max\left(y_{\min} - \hat{Y}(\boldsymbol{x}), 0\right) \tag{7}$$

Then the function of EI can be written as

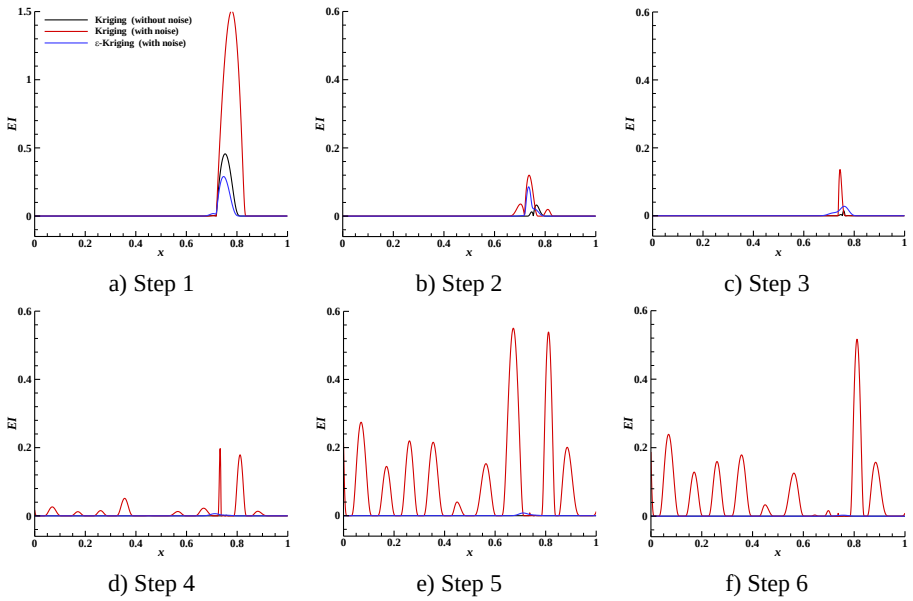
$$E\left(I\left(\boldsymbol{x}\right)\right)=\begin{cases}\left(y_{\min}-\hat{f}\left(\boldsymbol{x}\right)\right)\Phi\left(\dfrac{y_{\min}-\hat{f}\left(\boldsymbol{x}\right)}{s\left(\boldsymbol{x}\right)}\right)+s\left(\boldsymbol{x}\right)\phi\left(\dfrac{y_{\min}-\hat{f}\left(\boldsymbol{x}\right)}{s\left(\boldsymbol{x}\right)}\right), & s\left(\boldsymbol{x}\right)>0\\ 0 & , s\left(\boldsymbol{x}\right)=0\end{cases} \tag{8}$$

where Φ and ϕ are the cumulative distribution function and probability density function of standard normal distribution, respectively. Then a new sample point is obtained by solving the following sub-optimization problem:

$$\boldsymbol{x}_{new}=\arg\max_{\boldsymbol{x}_l\le\boldsymbol{x}\le\boldsymbol{x}_u} EI(\boldsymbol{x}) \tag{9}$$

The noise of $N(0,1.2^2)$ is added to the function in Eq.(10) to investigate the effect of noise on the EI function. The noise standard variance is 25% of the function standard variance. 10 samples are generated using the LHS method to establish the initial model, then the EI criterion is adopted to infill the new samples. Fig. 3 shows the EI function changes in the first 6 iteration steps. It is shown that, the new sample is exactly infilled to the maximum-EI location, and the maximum EI-value decreases to nearly zero in only a few steps. The proposed method based on the noisy samples converges as well as the ordinary kriging-based optimization based on the corresponding noise-free samples, which indicates its effectiveness.

$$f(x)=(6x-2)^2\sin(12x-4) \tag{10}$$



a) Step 1 b) Step 2 c) Step 3

d) Step 4 e) Step 5 f) Step 6

Fig. 3. The influence of noise on the EI criterion

The ε-kriging model and the aforementioned EI method have been integrated to an in-house optimization code called SurroOpt [36]. In next section, SurroOpt will be employed to run the numerical examples whose responses are interfered by noise. Then we will demonstrate the developed method with benchmark aerodynamic shape optimization which has noisy evaluation. In all of these examples, two different optimization methods will be compared: kriging based optimization; ε-kriging based optimization. To conduct a fair comparison, all the settings excluding the choice of surrogate models will be kept the same.

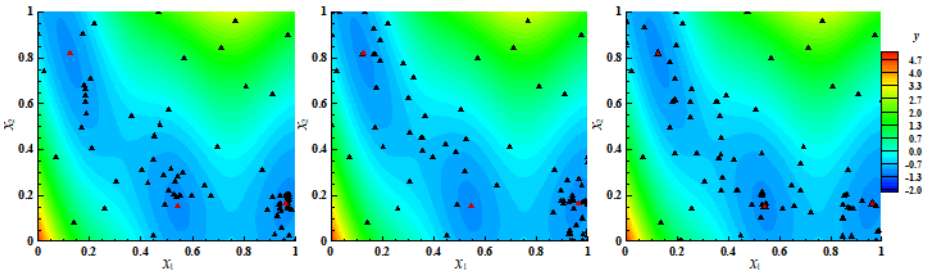
4 Numerical examples

4.1 Analytical test functions

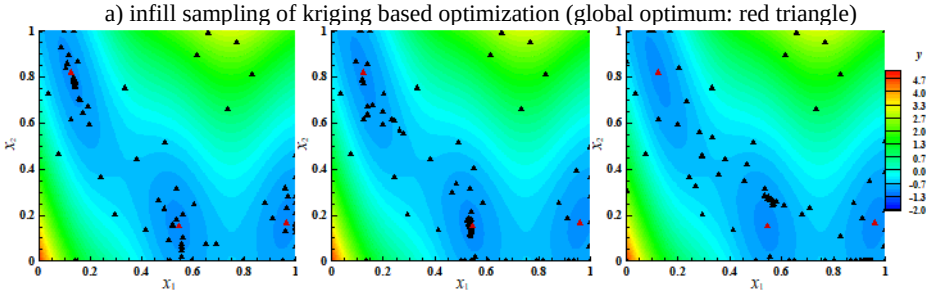
The function in Eq.(11) is used to demonstrate the peformance of ε-kriging based optimization. Fig. 4 shows the infill sampling process, Fig. 5 shows the convergence history of optimization of kriging and ε-kriging based optimization method under different noise levels. It can be seen that ε-kriging based optimization can efficiently find the global optimal region in the noisy design space, and the process of infill sampling more accurate.

$$y(x)=\frac{1}{51.95}\left[\left(\bar{x}_2-\frac{5.1\bar{x}_1^2}{4\pi^2}+\frac{5\bar{x}_1}{\pi}-6\right)^2+\left(10-\frac{10}{8\pi}\right)\cos(\bar{x}_1)-44.81\right] \tag{11}$$

where $\bar{x}_1=15x_1-5, \bar{x}_2=15x_2$

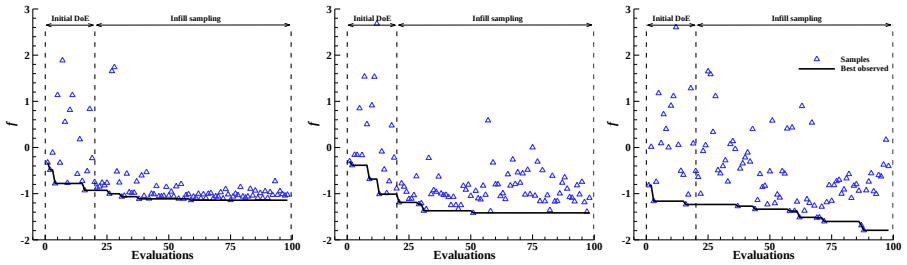


a) infill sampling of kriging based optimization (global optimum: red triangle)

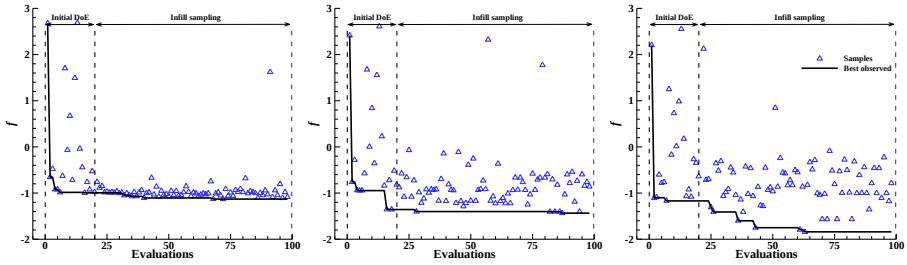


b) infill sampling of ε-kriging based optimization (global optimum: red triangle)

Fig. 4. Infill sampling of surrogate model based optimization under different noise levels (noise level from left to right: 5%, 25%, 50% of the function standard variance)



a) convergence history of kriging based optimization



b) convergence history of ε-kriging based optimization

Fig. 5. Convergence history of surrogate model based optimization under different noise levels (noise level from left to right: 5%, 25%, 50% of the function standard variance)

Besides, the peformance of ε-kriging based optimization have been researched on some numerical examples, whose response have the numerical nosie. We used some benchmark problems in order to test the performance of the algorithm (in Table 1). This set could include many different kinds of problems such as unimodal, multimodal, regular, irregular, separable, non-separable and multi-dimensional. The definition of function characteristics can be found in the literature [37]. In some functions, global minima is very small when compared to whole search space (Michalewicz) and so on. In Table 1 characteristics of each function are given under the column titled C. Noting the original functions have been rescaled to map their search space to $[0, 1]^d$, and measure optimization results impartial.

Table 1. Formulation of the benchmark functions

No	Range	C	Function	Formulation
1	[0,1]	MS	Forrester	$y(x) = (6x-2)^2 \sin(12x-4)$
2	$[0,1]^2$	MS	Branin-Hoo	$y(x) = \frac{1}{51.95}\left[\left(\bar{x}_2 - \frac{5.1\bar{x}_1^2}{4\pi^2} + \frac{5\bar{x}_1}{\pi} - 6\right)^2 + \left(10 - \frac{10}{8\pi}\right)\cos(\bar{x}_1) - 44.81\right]$ *where* $\bar{x}_1 = 15x_1 - 5, \bar{x}_2 = 15x_2$
3	$[0,1]^2$	MN	Six Hump Camel	$y(x) = (4 - 2.1\bar{x}_1^2 + \frac{\bar{x}_1^4}{3})\bar{x}_1^2 + \bar{x}_1\bar{x}_2 + (-4 + 4\bar{x}_2^2)\bar{x}_2^2$ *where* $\bar{x}_1 = 6x_1 - 3, \bar{x}_2 = 4x_2 - 2$
4	$[0,1]^2$	MN	GoldStein-Price	$y(x) = \frac{1}{2.427}\left[\log\left(\left[1 + (\bar{x}_1 + \bar{x}_2 + 1)^2(19 - 14\bar{x}_1 + 3\bar{x}_1^2 - 14\bar{x}_2 + 6\bar{x}_1\bar{x}_2 + 3\bar{x}_2^2)\right]\right.\right.$ $\left.\left.\left[30 + (2\bar{x}_1 - 3\bar{x}_2)^2(18 - 32\bar{x}_1 + 12\bar{x}_1^2 + 48\bar{x}_2 - 36\bar{x}_1\bar{x}_2 + 27\bar{x}_2^2)\right]\right) - 8.693\right]$ *where* $\bar{x}_i = 4x_i - 2$, *for all* $i = 1,2$
5	$[0,1]^2$	MS	Michalewicz	$y(x) = -\sum_{i=1}^{d} \sin(\bar{x}_i)\sin^{2m}(\frac{i\bar{x}_i^2}{\pi})$ *where* $\bar{x}_i = 3x_i$ *for all* $i = 1,...,d = 2$
6	$[0,1]^4$	MN	Hartmann4	$y(x) = \frac{-1}{1.94}\left[2.58 + \sum_{i=1}^{4} C_i \exp\left(-\sum_{j=1}^{4} a_{ji}(x_j - p_{ji})^2\right)\right]$ *where* $\boldsymbol{C} = [1.0 \;\; 1.2 \;\; 3.0 \;\; 3.2]$ $\boldsymbol{a} = \begin{bmatrix} 10.00 & 0.05 & 3.00 & 17.00 \\ 3.00 & 10.00 & 3.50 & 8.00 \\ 17.00 & 17.00 & 1.70 & 0.05 \\ 3.50 & 0.10 & 10.00 & 10.00 \\ 1.70 & 8.00 & 17.00 & 0.10 \\ 8.00 & 14.00 & 8.00 & 14.00 \end{bmatrix}$ $\boldsymbol{p} = \begin{bmatrix} 0.1312 & 0.2329 & 0.2348 & 0.4047 \\ 0.1696 & 0.4135 & 0.1451 & 0.8828 \\ 0.5569 & 0.8307 & 0.3522 & 0.8732 \\ 0.0124 & 0.3736 & 0.2883 & 0.5743 \\ 0.8283 & 0.1004 & 0.3047 & 0.1091 \\ 0.5886 & 0.9991 & 0.6650 & 0.0381 \end{bmatrix}$
7	$[0,1]^6$	MN	Hartmann6	$y(x) = \frac{-1}{1.94}\left[2.58 + \sum_{i=1}^{4} C_i \exp\left(-\sum_{j=1}^{6} a_{ji}(x_j - p_{ji})^2\right)\right]$

#M: means that the function is multimodal

#U: means that the function is unimodal

#S: means that the function is separable

#N: means that the function is non-separable

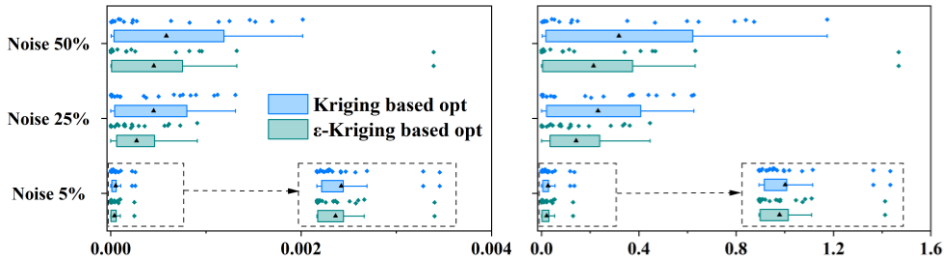
To get a better ε, 10d (d is the dimension) initial samples are chosen by the LHS method to build an initial ε-kriging model for the objective function. Then the EI infilling criterion is used to repetitively generate new samples to update the ε-kriging model, until the global optimum is found.

For reducing the impact of randomness on optimization, every function has been optimized 20 times. We use the assessment EX (gap of optimal location and optimized location) and EY (gap of optimal value and optimized value) [38] to messure the optimization quality, from two aspects reflcting the peformance of optimization. The box plots of the EX and EY are shown in Fig. 6, in which the median is the mean value of all of the 20 results. The results show that,

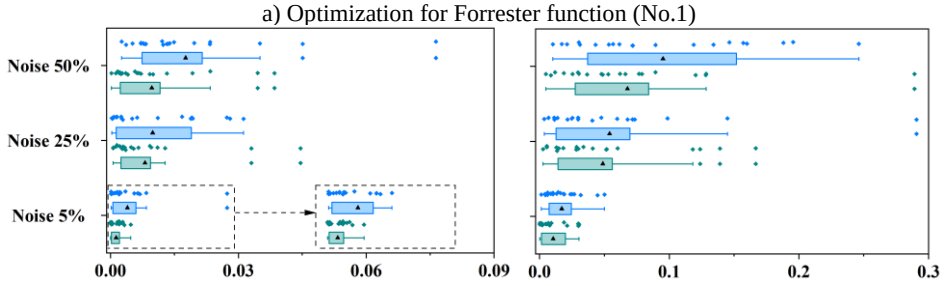
1) The average EX, EY(solid triangle in the box plots) and robustness of ε-kriging based optimization method is better than the kriging based optimization. It is indicated that the ε-kriging based optimization could filter the nosie in the process of optimization on different noise levels, to a certain extent.

2) With 5 % noise all the average results of both optimization methods are very close to the actual optimum, because the existence of assessment EX and EY equal to zero in this case. It is showed that low intensity noise almost no impact on the surrogate model based optimization. With 25% and 50% noise the ε-kriging based optimization method could identify the minima of function with precision, better than the kriging based optimization method.

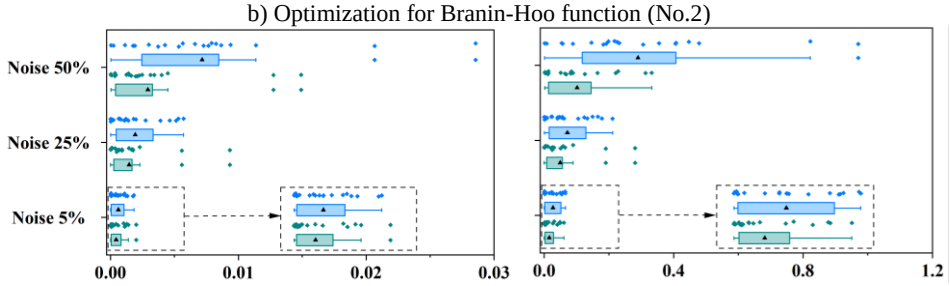
3) There are some functions fail to optimize because the value of EX and EY is very high(such as Goldstein-Price, indeed the basin of the global minimum is relatively small and was not found for several runs). Besides, with the intensity noise increacing, the performance of kriging based optimization method reduce dramatically, but the ε-kriging based optimization method could alleviate the fall trend, especially in the function Michalewicz and Six Hump Camel.



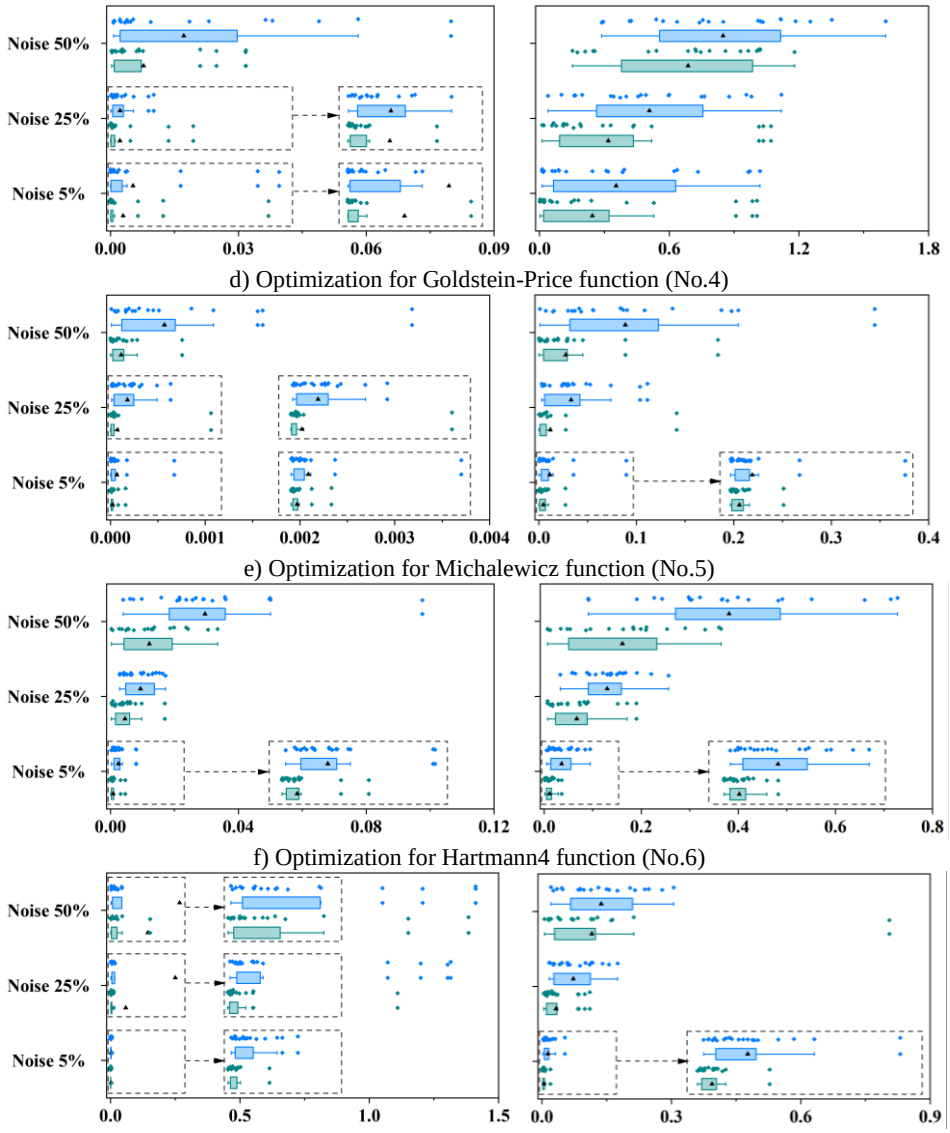
a) Optimization for Forrester function (No.1)



b) Optimization for Branin-Hoo function (No.2)



c) Optimization for Six Hump Camel function (No.3)



d) Optimization for Goldstein-Price function (No.4)

e) Optimization for Michalewicz function (No.5)

f) Optimization for Hartmann4 function (No.6)

g) Optimization for Hartmann6 function (No.7)

Fig. 6. The EX and EY box-plots of different test functions (left: the assessment EX, right: the assessment EY)

4.2 Drag minimization of NACA0012 airfoil in transonic flow

One representative example will be used to demonstrate the application of the developed method to aerodynamic shape optimizations. The benchmark problem defined by the AIAA aerodynamic design optimization discussion group (ADODG) [39], which

is the drag minimization of a NACA0012 airfoil, subject to a full thickness constraint. The airfoil is optimized at a freestream Mach number 0.85 and zero angle of attack in inviscid flow. It can be written as an optimization problem:

$$\begin{aligned} &\min \quad C_d \\ &s.t. \quad C_l = 0 \\ &\qquad y \geq y_{baseline} \forall x \in [0,1] \end{aligned} \tag{12}$$

where $y \geq y_{baseline}$ means that the local thickness must always larger than one of the baseline, along the airfoil from the leading edge to the trailing. The modified airfoil features zeros thickenss at trailing edge:

$$y_{baseline} = \pm 0.6(0.2969\sqrt{x} - 0.1260x - 0.3516x^2 + 0.2843x^3 - 0.1036x^4) \tag{13}$$

The parameterization method of the airfoil is sixteen-order CST [40] method, so the number of design variables is 17 in total. We force the airfoil is always symmetric and the lift coefficient is zero at zero angle of attack. The design space is defined by expanding the initial CST coefficients by 1.5 times and narrowing it by half (in Fig. 7):

$$\boldsymbol{x} \in [0.5\boldsymbol{x}_{base}, 1.5\boldsymbol{x}_{base}] \tag{14}$$

where $\boldsymbol{x}_{base}$ denotes the baseline shape.

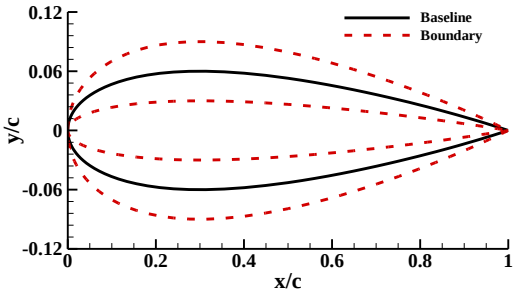


Fig. 7. Baseline airfoil and definition of design sapce

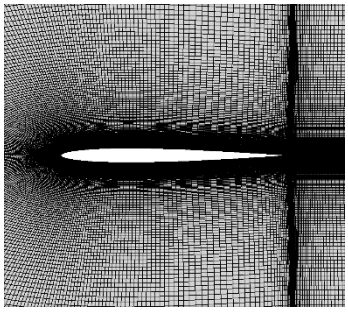


Fig. 8. Sketch of C-grid for airfoil optimization

A special technique called conformal transformation is employed to generate the C-type grid of airfoil to guarantee good uniformity and orthogonality, and the structural grid used for optimization has 512 points in the stream-wise direction and 256 points in the direction normal to the airfoil surface, as Fig. 8 sketches. Here, an in-house flow solver PMNS2D [39] is employed to perform flow simulation. To start optimization process, LHS method is used to generate 50 samples to construct initial ε-kriging and kriging model, then infilling criterion is employed to refine the model, and 4 samples are infilled per cycle. The optimization are terminated when the total number of CFD evaluations reaches 200. In the optimization, it should be noted that there is numerical error in the artificial dissipation which is the source of the noise[20].

The convergence histories are shown in Fig. 9, and the ε-kriging based method perform better than the kriging based method optimization. The comparison of the baseline and optimized airfoils' aerodynamic force coefficients are shown in Table 2, and the aerodynamic force coefficients of optimial airfoil are evaluated with noiseless method. It can be observed that the drag of ε-kriging based method is reduced by 43.85%, better than kriging based method, the constraints are strictly satisfied, and the time consuming almost identical.

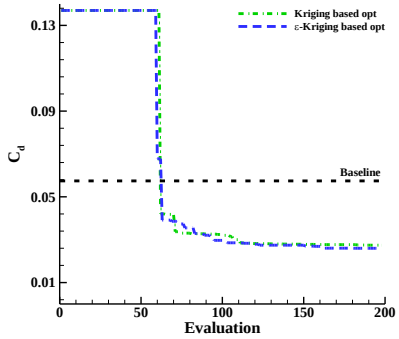


Fig. 9. The convergence histories of optimizing NACA0012 airfoil using kriging and ε-kriging based methods

Table 2. Optimization results for NACA0012

	C_d(counts)	Reduction of C_d (%)	Time(h)
Baseline airfoil	473.64	/	/
kriging based method	281.51	40.56	25.77
ε-kriging based method	265.96	43.85	25.96

The comparison of geometric shapes, pressure coefficient distributions, and pressure contour of the baseline and the optimal airfoils are shown in Fig. 10 and Fig. 11 respectively. It could be seen that both the leading and trailing edges of optimal airfoils are much thicker than the baseline airfoil, and the shock waves are largely reduced. The final design result still has a strong shock wave, because the shape of leading and trailing edges has reached the boundary of the design and better design may be located outside the design space.

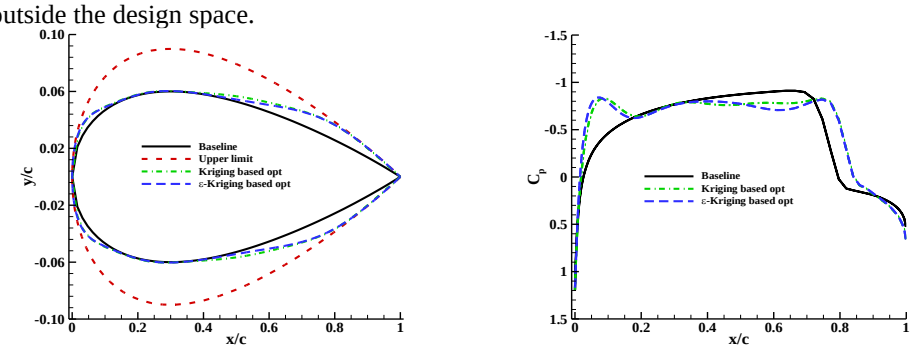
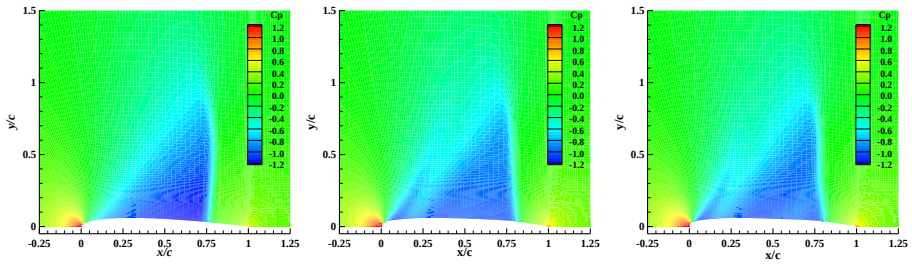


Fig. 10. Comparison of the shapes (left) and pressure distributions (right) for baseline and optimized airfoils



a) baseline b) kriging based method c) ε-kriging based method

Fig. 11. Comparison of pressure coefficient contours for baseline and optimized airfoils

5 Conclusion

In this paper a method termed as ε-kriging based optimization is developed and preliminarily validated to filter out the noise in the response during identifying the optimum in the optimization. The core of the developed method is to introduce the ε-insensitive loss of a SVR model into the diagonal of kriging correlation matrix. This method is verified by a numerical example and applied to an aerodynamic shape optimization of minimizing the drag of NACA0012 airfoil in transonic flow. The results indicate that the optimization accuracy and robustness is dramatically improved compared to classic kriging, which confirms that the developed method has a great potential for applications on the problem which the evaluation is interfered by noise.

Acknowledgment. This research was sponsored by the Practice and Innovation Funds for Graduate Students of Northwestern Polytechnical University PF2023039 and the National Science Foundation under grant U20B2007.

References

1. Jones, D.R., Schonlau, M. & Welch, W.J.: Efficient Global Optimization of Expensive Black-Box Functions. Journal of Global Optimization **13**(4), pp. 455–492 (1998).
2. Forrester, A. I. J., Keane. A. J.: Recent advances in surrogate-based optimization. Progress in aerospace sciences, **45**(1-3), pp. 50-79 (2009).
3. Han, Z. H., Zhang, K. S.: Surrogate-based optimization. Real-world applications of genetic algorithms, **343**, pp. 343-362 (2012).
4. Zhang, Q.: Simulation-driven design optimization and modeling for microwave engineering. Imperial College Press, London (2013).
5. Sobester, A., Forrester, A., Keane, A.: Engineering design via surrogate modelling: a practical guide. John Wiley & Sons, New York (2008).
6. Liu, J., Song, WP., Han, ZH. et al.: Efficient aerodynamic shape optimization of transonic wings using a parallel infilling strategy and surrogate models. Struct Multidisc Optim **55**(3), pp. 925–943 (2017).

7. Viana, F. A. C., Simpson, T. W., Balabanov. V., et al.: Special section on multidisciplinary design optimization: metamodeling in multidisciplinary design optimization: how far have we really come. AIAA journal, **52**(4), pp. 670-690 (2014).
8. Sakata, S., Ashida, F., and Zako, M.: Microstructural design of composite materials using fixed-grid modeling and noise-resistant smoothed Kriging-based approximate optimization. Structural and Multidisciplinary Optimization, **36**(3) pp. 273-287 (2008).
9. Zhang, Z., Yu, F., Wang, Q., et al.: Thermal-electric coupling systems analysis of electronic cabinets with consideration of numerical noise. Journal of Vibration and Shock, **36**(13), pp. 214-222 (2017).
10. Qiu, N., Gao, Y., Fang, J., et al.: Crashworthiness optimization with uncertainty from surrogate model and numerical error. Thin-Walled Structures, **129**, pp. 457-472 (2018).
11. Wang, Q., Nakashima, T., Lai, C., et al.: An improved system for efficient shape optimization of vehicle aerodynamics with "noisy" computations. Structural and Multidisciplinary Optimization, **65**(8), pp. 215 (2022).
12. Picheny, V., Ginsbourger, D.: Noisy kriging-based optimization methods: a unified implementation within the DiceOptim package. Computational Statistics & Data Analysis, **71**, pp. 1035-1053 (2014).
13. Jalali, H., Van, Nieuwenhuyse, I., Picheny, V.: Comparison of kriging-based algorithms for simulation optimization with heterogeneous noise. European Journal of Operational Research, **261**(1), pp. 279-301 (2017).
14. Sambakhé, D., Rouan, L., Bacro, J, N., et al.: Conditional optimization of a noisy function using a kriging metamodel. Journal of Global Optimization, **73**(3), pp. 615-636 (2019).
15. Letham, B., Karrer, B., Ottoni, G., et al.: Constrained Bayesian optimization with noisy experiments. Bayesian Analysis, **14**(2), pp. 495-519 (2019).
16. Picheny, V., Ginsbourger, D., Richet Y., et al.: Quantile-based optimization of noisy computer experiments with tunable precision. Technometrics, **55**(1), pp. 2-13 (2013).
17. Giunta, A. A., Dudley, J. M., Narducci, R., et al.: Noisy aerodynamic response and smooth approximations in HSCT design. In: Proceedings of 5th AIAA/USAF/NASA/ISSMO Symp on Multidisciplinary and Structural Optimization, pp. 1117-1128, Panama (1994).
18. Shyy, W., Papila, N., Vaidyanathan, R., et al.: Global design optimization for aerodynamics and rocket propulsion components. Progress in Aerospace Sciences, **37**(1), pp. 59-118 (2001).
19. Burman, J., Gebart, B. R.: Influence from numerical noise in the objective function for flow design optimization. International Journal of Numerical Methods for Heat & Fluid Flow, **11**(1), pp. 6-19 (2001).
20. Forrester, A. I., Keane, A. J., Bressloff, N. W.: Design and analysis of Noisy computer experiments. AIAA journal, **40**(1), pp. 2331-2339 (2006).
21. Gilkeson, C. A., Toropov, V. V., Thompson, H. M., et al.: The Curse of Numerical Noise and Implications for CFD-Based Design Optimization. In Proceedings of Proceedings of the 10th ASMO UK Conference Engineering Design Optimization, Delft (2014).
22. Gilkeson, C. A., Toropov, V. V., Thompson, H. M., et al.: Dealing with numerical noise in CFD-based design optimization. Computers & Fluids, **94**, pp. 84-97 (2014).
23. Ling, S. B.: Researches on wing aero/structural optimization with numerical noises, School of Aeronautics. Northwestern Polytechnical University, Xi'an, MS (2021).
24. Qiao, J. L.: On sonic boom prediction and low-boom design optimization for supersonic transports, School of Aeronautics., Northwestern Polytechnical University, Xi'an, MS (2019).
25. Cressie, N.: Statistics for spatial data, revised edition, Wiley, New York (1993).

26. Yin, J., Ng, S. H., Ng, K. M.: Kriging model with modified nugget effect for random simulation with heterogeneous variances. In: Proceedings of the 2008 IEEE International Conference on Industrial Engineering and Engineering Management, pp. 1714-1718, IEEE, Singapore (2008).
27. Yin, J., Ng, S. H., Ng, K. M.: A study on the effects of parameter estimation on Kriging model's prediction error in stochastic simulations. In: Proceedings of the 2009 Winter Simulation Conference (WSC), pp. 674-685, IEEE, Austin (2009).
28. Yin, J., Ng, S. H., and Ng, K. M.: Kriging metamodel with modified nugget-effect: The heteroscedastic variance case. Computers & Industrial Engineering, **61**(3), pp. 760-777 (2011).
29. Sakata, S., Ashida, F., Zako, M.: On applying Kriging-based approximate optimization to inaccurate data. Computer methods in applied mechanics and engineering, **196**(13-16), pp. 2055-2069 (2007).
30. Sakata. S., Ashida, F.: Ns-kriging based microstructural optimization applied to minimizing stochastic variation of homogenized elasticity of fiber reinforced composites. Structural and Multidisciplinary Optimization, **38**(5), pp. 443-453 (2009).
31. Chen, L., Qiu, H., Jiang, C., et al.: Support Vector enhanced Kriging for metamodeling with noisy data. Structural and Multidisciplinary Optimization, **57**(4), pp. 1611-1623 (2018).
32. Bostanabad, R., Kearney, T., Tao, S., et al.: Leveraging the nugget parameter for efficient Gaussian process modeling. International journal for numerical methods in engineering, **114**(5), pp. 501-516 (2018).
33. Zhang, K. S., Du, Y. Q., Lu, P. X., et al.: ε-Kriging: A Novel Surrogate Modeling Method for Noisy Aerodynamic Data AIAA journal (to be published)
34. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, Berlin (1999).
35. Wang, P. H.: Investigations on SVR hyper-parameters Adaptive method and its application in aerodynamic Optimization. School of Aeronautics. Northwestern Polytechnical University, Xi'an, MS (2021).
36. Han, Z. H.: SurroOpt: a generic surrogate-based optimization code for aerodynamic and multidisciplinary design. In: Proceedings of the 30th Congress of the International Council of the Aeronautical Sciences, pp. 2016-0281, ICAS Secretariat, Daejeon (2016).
37. Karaboga, D., Akay, B.: A comparative study of Artificial Bee Colony algorithm. Applied Mathematics and Computation, **214**(1), pp. 108-132 (2009).
38. Wackers, J., Visonneau, M., Ficini, S., et al.: Adaptive N-fidelity metamodels for noisy CFD data. In: Proceedings of AIAA AVIATION 2020 FORUM, pp. 3161 (2020).
39. Zhang, Y., Han, Z. H., Shi, L., et al.: Multi-round surrogate-based optimization for benchmark aerodynamic design problems. In: Proceedings of 54th AIAA Aerospace Sciences Meeting, pp. 2016-1545. AIAA paper, SanDiego (2016).
40. Kulfan, B. M.: Universal parametric geometry representation method. Journal of aircraft, **45**(1), pp. 142-158 (2008).