

A cooperative radial basis function method for variable-fidelity surrogate modeling

Xu Li¹ · Wenkun Gao¹ · Liangxian Gu¹ · Chunlin Gong¹ · Zhao Jing² · Hua Su¹

Received: 21 July 2016 / Revised: 9 March 2017 / Accepted: 24 April 2017 / Published online: 15 May 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract By coupling the low-fidelity (LF) model with the high-fidelity (HF) samples, the variable-fidelity model (VFM) offers an efficient way to overcome the expensive computing challenge in multidisciplinary design optimization (MDO). In this paper, a cooperative radial basis function (Co-RBF) method for the VFM is proposed by modifying the basis function of RBF. The RBF method is constructed on the HF samples, while the Co-RBF method incorporates the entire information of the LF model with the HF samples. In Co-RBF, the LF model is regarded as a basis function of Co-RBF and the HF samples are utilized to compute the Co-RBF model coefficients. Two numerical functions and three engineering problems are adopted to verify the proposed Co-RBF method. The predictive results of Co-RBF are compared with those of RBF and Co-Kriging, which show that the Co-RBF method improves the efficiency, accuracy and robustness of the existing VFMs.

Keywords VFM · Surrogate model · RBF · Co-RBF · Co-Kriging

1 Introduction

Multidisciplinary design optimization (MDO) is a field of engineering that focuses on the use of numerical optimization for the design of systems that involve several disciplines. However, MDO is computationally expensive for the multidisciplinary analysis of complex models (Sobieszczanski-Sobieski and Haftka 1997; Breitkopf and Coelho 2010). To overcome the computing challenge, the surrogate model (also known as metamodel, response surface method or approximation model) is adopted to approximate the HF model. The general surrogate model is constructed with HF samples under certain assumption (Booker et al. 1998; Leary et al. 2003; Simpson et al. 2004; Gorissen et al. 2010). However, with the increase of model nonlinearity, the surrogate model requires a large number of expensive HF samples to satisfy the predefined accuracy. Therefore, the information of fast-cheap LF model is an alternative choice to reduce the computational cost (Rodriguez 2001; Keane and Nair 2005; Marduel et al. 2006; Zahir and Gao 2012; Courier et al. 2016). Whereas, the LF model often results in precision loss. Thus, the advantages of HF model and LF model should be coupled to construct a surrogate model to improve the computational efficiency and accuracy.

Variable-fidelity (also known as multi-fidelity) model (VFM) couples the HF samples with the LF model information in a surrogate model, which offers a trade-off solution between the computing cost and the accuracy. The advantage of the VFM is that it reduces the expensive HF samples while guarantees necessary accuracy of the surrogate model (Rodriguez 2001; Zheng et al. 2012; Liem et al. 2015). For instance, the flight vehicle CFD model of 2 million meshes takes several hours with Navier-Stocks code while several

✉ Xu Li
lukebest@mail.nwpu.edu.cn

Zhao Jing
jingzhao.behunters@gmail.com

¹ School of Astronautics, Northwestern Polytechnical University, Xi'an, People's Republic of China

² School of Civil Engineering and Mechanics, Huazhong University of Science and Technology, Hubei, People's Republic of China

minutes with Euler inviscid flow method, and even faster with engineering empirical method. Thus, the VFM model makes it possible to optimize the time-consuming design with good accuracy (Kuya et al. 2011; Leifsson and Koziel 2015).

Zero-order scaling is the simplest approach for VFM, involving a scale factor based on the value of the HF and LF models at a single point \mathbf{x}_0 . This method is also referred as the local-global approximation strategy (Haftka 1991). A straightforward extension of the method (Haftka 1991; Eldred et al. 2004) can be formulated by Taylor expansions of the scaling factors. It generates first-order scaling and second-order scaling respectively. However, when the value of LF model is zero, the scaling factor tends to infinity, which is a failure case for the method. To avoid the problem, Lewis and Nash (2005) developed an additive bridge function to replace the scaling factor. Whereas, Gano et al. (2005, 2006) showed that the additive bridge functions did not always have better prediction accuracy than the multiplicative scaling method. Hence, an adaptive hybrid method that combines the multiplicative and additive methods was developed (Gano et al. 2005, 2006). The method builds a correlation (or bridge) function between the HF and LF models, which has good local prediction accuracy during the optimization process. However, it doesn't work well when a global surrogate model is required (Han 2013; Zheng 2013).

Co-Kriging is another powerful VFM constructed under the assumption that the HF and LF models are almost linear dependent (Kennedy and O'Hagan 2000; Forrester et al. 2007, 2008; Laurenceau and Sagaut 2015). It is a kind of Kriging method that correlates multiple sets of data. Thus, Co-Kriging performs a higher approximation accuracy than the surrogate model with the only HF samples. However, a difficult and time-consuming multivariable optimization is required for the estimation of the parameters. Moreover, Co-Kriging uses numbered LF samples other than the entire LF model. Therefore, the parameter estimation process should be simplified and more LF model information should be included into the VFM.

Han et al. (2012, 2013) improved Co-Kriging via gradient-enhanced Kriging, which takes more HF information into consideration. The improved method performs much better than the original Co-Kriging when gradient information is available. Nevertheless, the gradient information is difficult to obtain in most practical cases, which limits the application of the improved Co-Kriging.

This paper develops a cooperative radial basis function (Co-RBF) method for variable fidelity surrogate modeling under the assumption that the HF and LF models are almost linear dependent (Forrester et al. 2008). Under this assumption, the LF model is regarded as a basis function of Co-RBF. Then the surrogate model is built with the HF

samples. Finally, the Co-RBF model shape parameter is optimized with the leave-one out cross validation (LOOCV) error (Hastie et al. 2000). In the Co-RBF method, the LF model basis function captures most of the model features, and the other basis functions of original RBF captures the left model features. Therefore, Co-RBF gets better prediction accuracy with less HF samples.

The remainder of this paper is structured as follows. In Section 2, the RBF, Co-Kriging surrogate models and the model validation methods are introduced, and then the Co-RBF is constructed as well as the features are discussed. Subsequently, in Section 3 two analytical functions and three engineering problems are used to compare Co-RBF with RBF and Co-Kriging. Finally, the conclusions and suggestions are given in Section 4.

2 Cooperative radial basis function method

In practical engineering problems, different simulation models are often used in different design stages. LF models are computationally cheap (eg. Euler equations in aerodynamic), while HF models are expensive (eg. Navier-Stocks equations). In order to describe the relationship, the cooperative radial basis function method proposed in this paper utilizes the idea of radial basis functions and the basic hypothesis of Co-Kriging, by combining a small number of HF sample points with a large number of LF sample points. In this section, the basic theories of RBF and Co-Kriging are introduced in Sections 2.1 and 2.2 respectively. Thereafter, in Sections 2.3–2.5, the theory and implementation of Co-RBF are introduced in detail.

2.1 Radial basis function method

A surrogate model is a prediction model for the unknown points according to the given samples, and an interpolation or a regression model mathematically (Simpson et al. 2004). It is also a branch of data mining and machine learning (Hastie et al. 2000). The common used surrogate models include polynomial response surface method (PRSM), radial basis functions (RBF), Kriging, support vector machine (SVM) and artificial neural net (ANN). (Gutmann 2000; Forrester et al. 2008; Boopathy and Rumpfkeil 2015). Since RBF works well in nonlinear cases and is simple to build, this paper modifies RBF with HF samples and LF model. For a given data set $S = \{(\mathbf{x}_i, y_i) | i = 1, 2, \dots, n\}$, written in a form of matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T$, $\mathbf{y} = [y_1, y_2, \dots, y_n]^T$, where n is the number of samples; $\mathbf{X} \in \mathbb{R}^{n \times d}$ is the input samples matrix; $\mathbf{y} \in \mathbb{R}^{n \times 1}$ is the output response vector; n is the number of samples; d is the number of design variables. The prediction value at \mathbf{x} is $\hat{y}(\mathbf{x})$.

RBF is performed through linear combination of nonlinear basis functions. The formulation is presented as follows,

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^n \beta_i f(\|\mathbf{x} - \mathbf{x}_i\|) = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} \tag{1}$$

where β_i is the i th coefficient for RBF; $f(\|\mathbf{x} - \mathbf{x}_i\|)$ is the i th radial basis function. The conventional form of the basis function is presented in Table 1. Substituting the n samples into (1)

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} f(\|\mathbf{x}_1 - \mathbf{x}_1\|) & \cdots & f(\|\mathbf{x}_1 - \mathbf{x}_n\|) \\ f(\|\mathbf{x}_2 - \mathbf{x}_1\|) & \cdots & f(\|\mathbf{x}_2 - \mathbf{x}_n\|) \\ \vdots & \ddots & \vdots \\ f(\|\mathbf{x}_n - \mathbf{x}_1\|) & \cdots & f(\|\mathbf{x}_n - \mathbf{x}_n\|) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$$

and rewrite it in a matrix form

$$\mathbf{y} = \mathbf{F}\boldsymbol{\beta} \tag{2}$$

The equation has a unique solution $\boldsymbol{\beta} = \mathbf{F}^{-1}\mathbf{y}$ since $\text{rank}(\mathbf{F})$ is full rank, say $\mathbf{F} \in \mathbb{R}^{n \times n}$, according to the linear algebra theory. Substituting the solution into (1), the prediction model is given by

$$\hat{y}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \mathbf{F}^{-1}\mathbf{y} \tag{3}$$

In (3), $\mathbf{f}(\mathbf{x})^T$ is determined by the prediction point \mathbf{x} and the matrix of samples \mathbf{X} . $\mathbf{F}^{-1}\mathbf{y}$ is just relevant to \mathbf{X} and \mathbf{y} , and is calculated only once. For a given prediction point \mathbf{x} , the prediction value $\hat{y}(\mathbf{x})$ can be calculated after evaluation of $\mathbf{f}(\mathbf{x})$. It should be pointed out that the shape parameter c makes significant influence on the prediction performance. Thus, the additional optimization criterion is required to determine the parameter c .

2.2 Co-Kriging

When additional LF information is available, the accuracy of a surrogate of the HF model improves. To make use of the

Table 1 Radial basis functions

Function type	Function form $f(\ \mathbf{x} - \mathbf{x}_i\)$
Gaussian	$\exp(-cr^2)$
Multiquadric	$\sqrt{r^2 + c^2}$
Inverse Multiquadric	$\frac{1}{\sqrt{r^2 + c^2}}$
Thin plate spline	$r^2 \log(cr^2 + 1)$

LF samples, some form of correction process which combines the differences between the LF and HF models must be adopted. The expensive HF data have values \mathbf{y}_H at points \mathbf{X}_H and the cheap LF data have values \mathbf{y}_L at points \mathbf{X}_L . To simplify the formulation of a correction process, we assume the HF sample locations coincide with a subset of the LF ones ($\mathbf{X}_H \subset \mathbf{X}_L$). Co-Kriging is a representative variable-fidelity method (Kennedy and O’Hagan 2000; Forrester et al. 2008; Kuya et al. 2011). The correction process takes the form

$$y_H(\mathbf{x}) = \rho y_L(\mathbf{x}) + Z_d(\mathbf{x}) \tag{4}$$

where $y_H(\mathbf{x})$ and $y_L(\mathbf{x})$ represent the features of the HF and LF models respectively. In the correction process, the HF model is approximated as the LF model multiplied by a constant scaling factor ρ plus a Gaussian process $Z_d(\mathbf{x})$ which represents the difference between $\rho y_L(\mathbf{x})$ and $y_H(\mathbf{x})$. The Co-Kriging prediction model is given by

$$\hat{y}_H(\mathbf{x}) = \hat{\mu} + \mathbf{c}^T \mathbf{C}^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu}) \tag{5}$$

where

$$\hat{\mu} = \mathbf{1}^T \mathbf{C}(\mathbf{X}, \mathbf{X})^{-1} \mathbf{y} / \mathbf{1}^T \mathbf{C}(\mathbf{X}_H, \mathbf{X}_H)^{-1} \mathbf{1} \tag{6}$$

and \mathbf{c} is a column vector of the covariance between \mathbf{X} and \mathbf{x} , \mathbf{C} is the covariance matrix. For more details about the derivation, see Forrester et al. (2008).

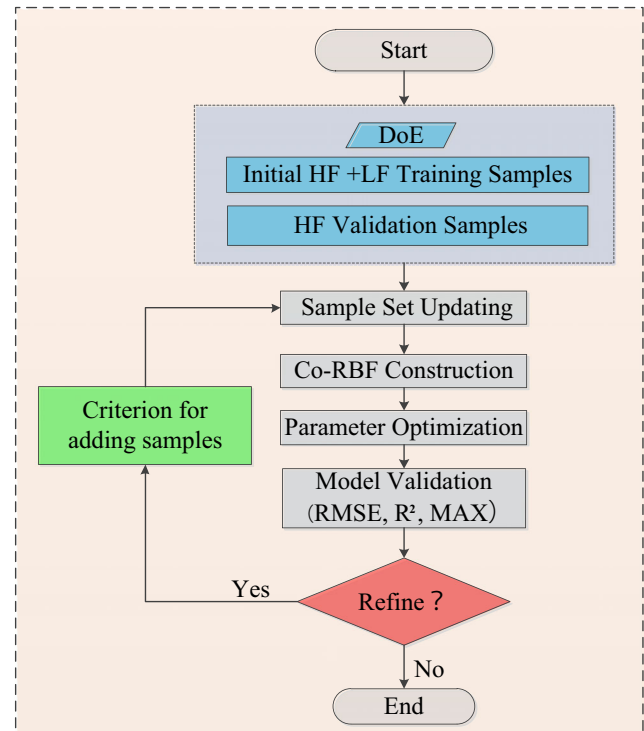


Fig. 1 The refinement flow chart of Co-RBF

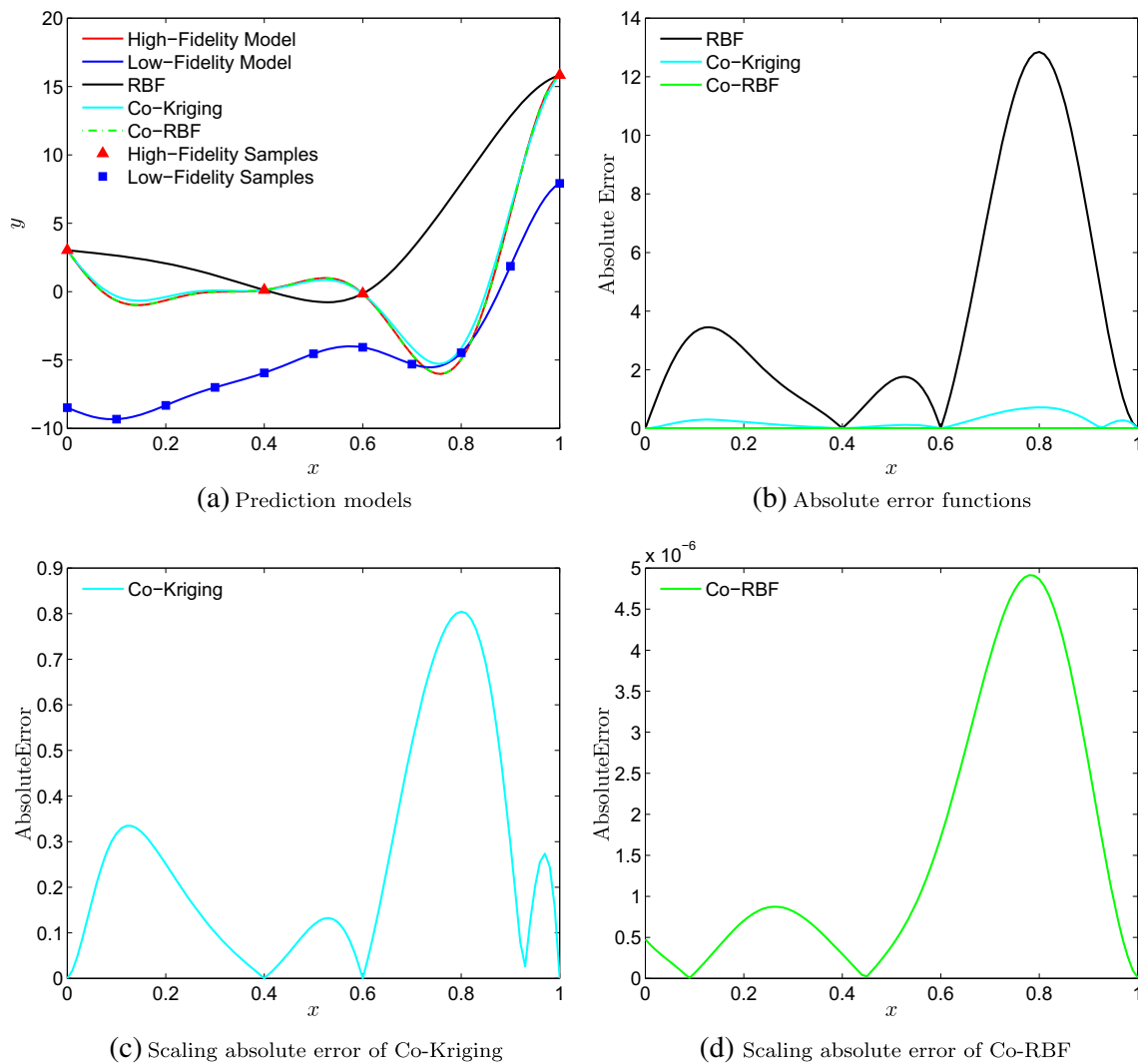


Fig. 2 Prediction models and the absolute errors

2.3 Validation of the surrogate model

The generalization performance of a surrogate model relates to its prediction capability on independent test data. Assessment of this performance is extremely important in practice, since it guides the choice of the model or the parameter, and it gives us a measure of the quality of the ultimate model (Hastie et al. 2000).

Three assessments are adopted to validate the accuracy of the surrogate models over the training samples: the root-mean-square error (RMSE), the maximum absolute error (MAX) and the square of the coefficient of multiple correlation R^2 (Qian et al. 2006). They are defined as

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}(\mathbf{x}_i; S))^2} \tag{7}$$

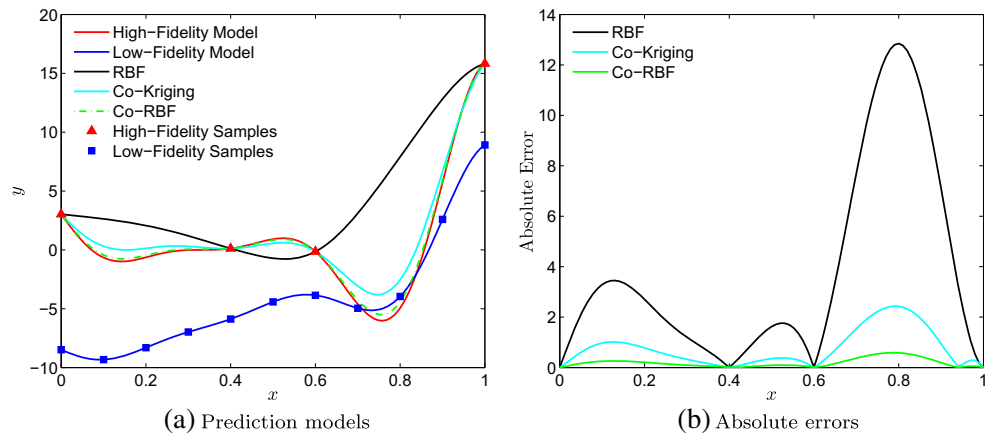
$$MAX = \max_{i=1,2,\dots,n} |y_i - \hat{y}(\mathbf{x}_i; S)| \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}(\mathbf{x}_i; S))^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{9}$$

Table 2 Performance comparison of 1D example

Models	R^2	RMSE	MAX
LF	-0.4845	6.5762	11.5136
RBF	0.0730	5.1969	12.8466
Co-Kriging	0.9921	0.4798	1.1734
Co-RBF	1.0000	3.1442×10^{-6}	5.4126×10^{-6}

Fig. 3 Predictions and absolute errors with more complex correction function $\delta(\mathbf{x})$



where n denotes the total number of validation samples, \bar{y} denotes the mean value of observed value y_i and \hat{y}_i is the predicted value. RMSE is utilized to measure the global accuracy of the surrogate model; MAX is employed to evaluate the local accuracy; R^2 is computed to reflect the linear dependence between the predicted and the actual values. R^2 takes values below 1; a higher value indicates a better predictability of the HF model from LF model, with a value of 1 indicating that the predictions are exactly correct.

However, when the prediction surrogate model goes through the samples, $RMSE = 0$, $MAX = 0$ and $R^2 = 1$, additional validation samples are required to measure the prediction performance of the surrogate model. To avoid the problem, the cross validation method is adopted. The samples are divided into K roughly equal-sized parts. For the k th part ($k = 1, 2, \dots, K$), the model is fitted with the other $K - 1$ parts of the samples, while the k th part of samples are used to estimate the prediction error. When $K = n$, the method is named leave-one-out cross validation (LOOCV) error (Hastie et al. 2000; Fasshauer and Zhang 2007; Forrester et al. 2008). Thus the RBF shape parameter c can be estimated through the following sub optimization problem

$$\min_c LOOCV = \frac{1}{n} \sum_{i=1}^n [y_i - \hat{y}_i(\mathbf{x}_i; S - \{(\mathbf{x}_i, y_i)\})]^2 \quad (10)$$

where, for a given c the surrogate model is calculated for n times. LOOCV doesn't require additional samples. It is an ideal method to describe the match degree between samples and the model. Moreover, the LOOCV based parameter optimization avoids the potential over fitting problem (Tetko et al. 1995). Since the surrogate model and parameter is determined, the VFM can be constructed.

2.4 Construction of the cooperative radial basis function model

In engineering, the LF model information is often computationally cheaper than the HF samples. For instance, the

LF model can be calculated via empirical equations while the HF samples are evaluated with finite element analysis or computational fluid dynamics. Co-RBF makes the best of the HF samples and LF model, by regarding LF model as a basis of RBF. In this paper, a similar correction process to Co-Kriging (See Section 2.2) is used, which is originally from the auto-regression model of Kennedy and O'Hagan (2000), assuming $cov\{y_H(\mathbf{x}_i), y_L(\mathbf{x})|y_L(\mathbf{x}_i)\} = 0, \forall \mathbf{x} \neq \mathbf{x}_i$, which means that no more information can be learnt about $y_H(\mathbf{x}_i)$ from the LF model if the value of the HF model at \mathbf{x}_i is known (this is known as a Markov property which, in essence, says we assume that the HF model is correct and any inaccuracies lie wholly in the LF model). Using the auto-regressive model we essentially approximate the HF model as the LF model multiplied by a constant scaling factor β_L plus a correction function. Actually, the HF and LF models have some similarities most of which can be captured by the scaling LF model. Moreover, the correction function is adopted to capture the remaining nonlinear information. Given the HF samples $S = \{(\mathbf{x}_i, y_H(\mathbf{x}_i)) | i = 1, 2, \dots, n\}$. The formulation is presented as

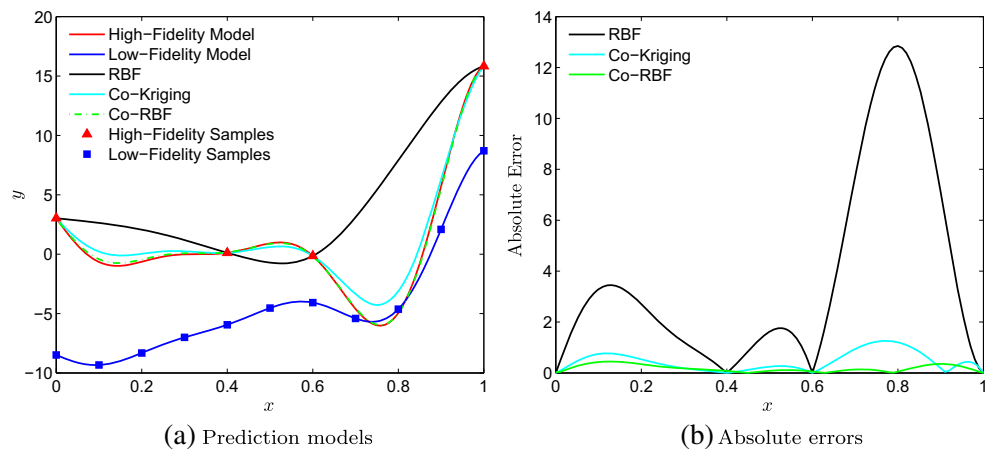
$$y_H(\mathbf{x}) = \beta_L y_L(\mathbf{x}) + \delta(\mathbf{x}) + \epsilon(\mathbf{x}) \quad (11)$$

Where $y_H(\mathbf{x})$ and $y_L(\mathbf{x})$ are the HF and LF model respectively; β_L is an unknown constant coefficient; $\delta(\mathbf{x})$ is the correction function; $\epsilon(\mathbf{x})$ is the error function. Furthermore, the correction function $\delta(\mathbf{x})$ can be replaced by radial basis functions, defined as $\delta(\mathbf{x}) = \mathbf{f}_\delta(\mathbf{x})^T \boldsymbol{\beta}_\delta$; $\epsilon(\mathbf{x})$ is the Gauss

Table 3 Performance comparison of 1D example with the effect of $\delta(\mathbf{x})$

Models	R^2	RMSE	MAX
LF	-0.4101	6.4095	11.5136
RBF	0.0730	5.1969	12.8466
Co-Kriging	0.9659	0.9961	2.4235
Co-RBF	0.9980	0.2400	0.5810

Fig. 4 Predictions and absolute errors with nonlinear coefficient β_L



stochastic process, $\epsilon(\mathbf{x}) \sim \mathbb{N}(0, \sigma^2)$, σ is an unknown constant. Substituting the HF samples into (11)

$$\begin{bmatrix} y_H(\mathbf{x}_1) \\ y_H(\mathbf{x}_2) \\ \vdots \\ y_H(\mathbf{x}_n) \end{bmatrix} = \begin{bmatrix} y_L(\mathbf{x}_1) & f_\delta(\|\mathbf{x}_1 - \tilde{\mathbf{x}}_1\|) & \cdots & f_\delta(\|\mathbf{x}_1 - \tilde{\mathbf{x}}_m\|) \\ y_L(\mathbf{x}_2) & f_\delta(\|\mathbf{x}_2 - \tilde{\mathbf{x}}_1\|) & \cdots & f_\delta(\|\mathbf{x}_2 - \tilde{\mathbf{x}}_m\|) \\ \vdots & \vdots & \ddots & \vdots \\ y_L(\mathbf{x}_n) & f_\delta(\|\mathbf{x}_n - \tilde{\mathbf{x}}_1\|) & \cdots & f_\delta(\|\mathbf{x}_n - \tilde{\mathbf{x}}_m\|) \end{bmatrix} \begin{bmatrix} \beta_L \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \epsilon \\ \epsilon \\ \vdots \\ \epsilon \end{bmatrix} \quad (12)$$

Where $f_\delta(\cdot)$ takes the similar form as the one in (1); $\tilde{\mathbf{x}}_i$ is the given center point of the i th radial basis function determined by design of experiment (DOE), and the point doesn't need to be the same as the HF input samples. However, here we choose HF samples as the center points of the radial basis functions. n is the number of samples; m is the number of radial basis function center points; Eq. 12 can be written in a simple form of matrix

$$\mathbf{y} = \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (13)$$

where the response samples vector becomes a multivariable Gaussian distribution $\mathbf{y} \sim \mathbb{N}(\mathbf{F}\boldsymbol{\beta}, \boldsymbol{\Sigma})$; the covariance matrix $\boldsymbol{\Sigma} = \text{cov}(\boldsymbol{\epsilon}, \boldsymbol{\epsilon}) = \sigma^2[r(\|\mathbf{x}_i - \mathbf{x}_j\|)]_{n \times n} = \sigma^2\mathbf{R}$, $r(\cdot)$ is the correlation function and \mathbf{R} is the correlation matrix. The probability distribution function (PDF) is formulated as

$$L(\mathbf{y}|\mathbf{X}; \mathbf{F}\boldsymbol{\beta}, \sigma^2\mathbf{R}) = \frac{1}{\sqrt{|2\pi\sigma^2\mathbf{R}|}} \exp[-\frac{1}{2}(\mathbf{y} - \mathbf{F}\boldsymbol{\beta})^T(\sigma^2\mathbf{R})^{-1}(\mathbf{y} - \mathbf{F}\boldsymbol{\beta})] \quad (14)$$

For convenience, transforming the PDF into logarithm form

$$\ln L = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2} \ln(|\mathbf{R}|) - \frac{1}{2}(\mathbf{y} - \mathbf{F}\boldsymbol{\beta})^T(\sigma^2\mathbf{R})^{-1}(\mathbf{y} - \mathbf{F}\boldsymbol{\beta}) \quad (15)$$

According to the maximum likelihood estimation (MLE)

$$\begin{cases} \frac{\partial \ln L}{\partial \boldsymbol{\beta}} = \mathbf{0} \\ \frac{\partial \ln L}{\partial \sigma^2} = 0 \end{cases} \quad (16)$$

Furthermore, from the first component of (16), setting $\tilde{\mathbf{F}} = \mathbf{R}^{-1}\mathbf{F}$, then

$$\tilde{\mathbf{F}}\boldsymbol{\beta} = \tilde{\mathbf{y}} \quad (17)$$

Where $\tilde{\mathbf{F}} \in \mathbb{R}^{n \times (m+1)}$, $\tilde{\mathbf{y}} \in \mathbb{R}^{n \times 1}$, according to the matrix theory (Kaare and Michael 2012), if $n = m + 1$, say $\text{rank}(\tilde{\mathbf{F}}^T) = n$, there exists unique solution $\boldsymbol{\beta} = \tilde{\mathbf{F}}^{-1}\tilde{\mathbf{y}}$; if $n > m + 1$, say $\text{rank}(\tilde{\mathbf{F}}^T) = m + 1$, there exists unique least square solution $\boldsymbol{\beta} = (\tilde{\mathbf{F}}^T\tilde{\mathbf{F}})^{-1}\tilde{\mathbf{F}}^T\tilde{\mathbf{y}}$, which means the surrogate doesn't strictly go through the HF samples; if $n < m + 1$, say $\text{rank}(\tilde{\mathbf{F}}^T) = n$, there exists unique least norm solution $\boldsymbol{\beta} = \tilde{\mathbf{F}}^T(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T)^{-1}\tilde{\mathbf{y}}$, which means the surrogate strictly go through the HF samples. Assuming $\tilde{\mathbf{F}}^+$ to be the pseudo-inverse (or Moore-Penrose inverse) of $\tilde{\mathbf{F}}$, then the parameter estimation can be obtained in an unified form

$$\begin{cases} \boldsymbol{\beta} = \tilde{\mathbf{F}}^+\tilde{\mathbf{y}} \\ \sigma^2 = \frac{1}{n}(\mathbf{y} - \mathbf{F}\tilde{\mathbf{F}}^+\tilde{\mathbf{y}})^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\tilde{\mathbf{F}}^+\tilde{\mathbf{y}}) \end{cases} \quad (18)$$

Substituting the estimations (18) into (11), we get access to the prediction model

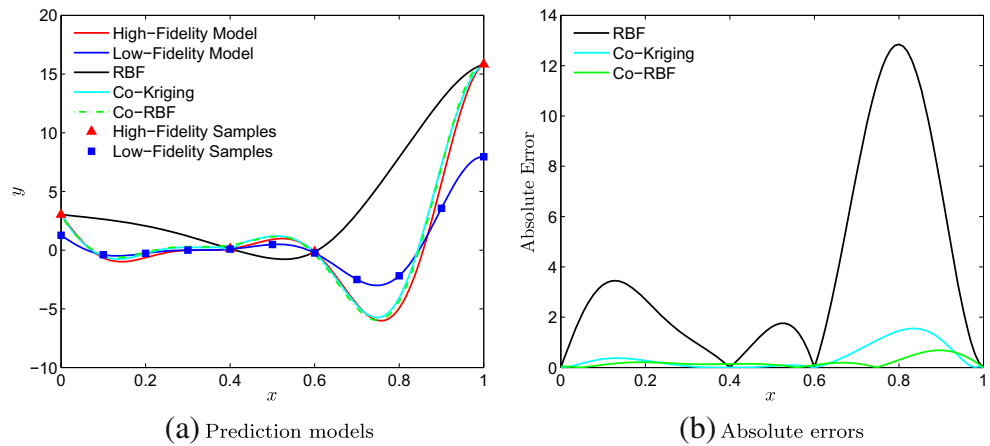
$$\hat{y}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T\tilde{\mathbf{F}}^+\tilde{\mathbf{y}} \quad (19)$$

Intuitively, the formulation of the prediction model is similar to that of RBF in (3). However, the basis functions are different. $\tilde{\mathbf{F}}^+$ includes both the HF and LF information

Table 4 Performance comparison of 1D example with the effect of β_L

Models	R ²	RMSE	MAX
LF	-0.4036	6.3946	11.5136
RBF	0.0730	5.1969	12.8466
Co-Kriging	0.9897	0.5482	1.2019
Co-RBF	0.9984	0.2128	0.4303

Fig. 5 Predictions and absolute errors with the variable disturbance



and $\tilde{\mathbf{y}}$ contains HF information. It is worth to mention that the shape parameter c in $\tilde{\mathbf{F}} \in \mathbb{R}^{n \times (m+1)}$ has a significant effect on the performance of the surrogate model. The shape parameter c can be determined with the LOOCV method (See Section 2.3). Since there is only one parameter, the sub optimization of parameter is easier than the multivariable optimization of Co-Kriging (Forrester et al. 2008). The correlation matrix \mathbf{R} illustrates the correlation between the HF samples, which means that the larger the distance is, the weaker the relation is. Here we assume the samples are totally independent, so \mathbf{R} becomes an identity matrix, and $\tilde{\mathbf{F}} = \mathbf{F}$. Thus,

$$\hat{\mathbf{y}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \mathbf{F}^+ \mathbf{y} \tag{20}$$

After the prediction model is obtained, the values of the base functions of Co-RBF are obtained by substituting the prediction point into the formula for the new prediction point. This is a vector-valued function, in which the first term is the value of the LF model, the left are the values of the traditional radial basis functions. However, the model coefficient $\beta = \mathbf{F}^+ \mathbf{y}$ only needs the initial HF and LF samples. It is worth noting that, in order to achieve better prediction accuracy, we usually need to optimize shape parameter c . Finally, the new prediction point can be obtained by the inner product of the prediction base vector function $\mathbf{f}(\mathbf{x})$ and the model coefficient vector β .

The Co-RBF method takes the similar correction process of Co-Kriging with a constant scaling factor (See Section 2.2) between the HF and LF models. The LF model includes most information of the HF model, and the remainder radial basis functions make up for the components that can't be captured by the LF model. When predicting the new points, the values of LF at the new locations are utilized, while Co-Kriging just uses the initial HF and LF samples. Since the LF samples at the prediction locations contains some information of the HF model, Co-RBF gets better prediction accuracy. However, when the LF values at the prediction locations are not available we have to

calculate the values with the LF samples by extra interpolation technique. Moreover, whether the linear dependent assumption of the constant scaling factor holds or not is unknown in practice and the assumption is difficult to validate directly. Instead, in the community of surrogate model (Forrester et al. 2007, 2008; Qian et al. 2006), different prediction error estimations (See Section 2.3) are adopted to verify the model assumption indirectly. In what follows, several examples are used to demonstrate the performance of Co-RBF.

2.5 Refinement of the surrogate model

In the practical engineering problems, the VFM constructed through the initial samples may not directly satisfy the required precision, therefore, it is necessary to refine the surrogate model by adding samples in the important region which we are interested in. Since the uncertainty of the prediction model increases with the HF samples sparsity (Bischl et al. 2014; Haftka et al. 2016) and the differences between the HF and LF models, we improve the global accuracy of the surrogate model by adding points where the HF samples is spare and the LF model differs greatly from the HF samples with large uncertainty. The detailed process is shown in Fig. 1. The process does not depend on the constructed surrogate model, but only on the HF sample density and the differences between the LF model and the HF samples, which is actually a multi-objective optimization problem. We transform it into a single objective optimization problem by multiplying the two objectives. The criterion for adding points is given by

$$\max_{\mathbf{x}} \text{dist}(\mathbf{x}, \mathbf{X}_H) \cdot \text{dist}([\mathbf{x}; \hat{\mathbf{y}}_L(\mathbf{x})], [\mathbf{X}_H, \mathbf{y}_H]) \tag{21}$$

where $\hat{\mathbf{y}}_L(\mathbf{x})$ is the surrogate model of the LF model constructed with LF samples; $[\mathbf{x}; \hat{\mathbf{y}}_L(\mathbf{x})]$ is a $1 \times (m + 1)$ column vector, $[\mathbf{X}_H, \mathbf{y}_H]$ is a $n \times (m + 1)$ matrix; $\text{dist}(\mathbf{x}, \cdot)$ is the smallest standardized Euclidean distance between \mathbf{x} and the

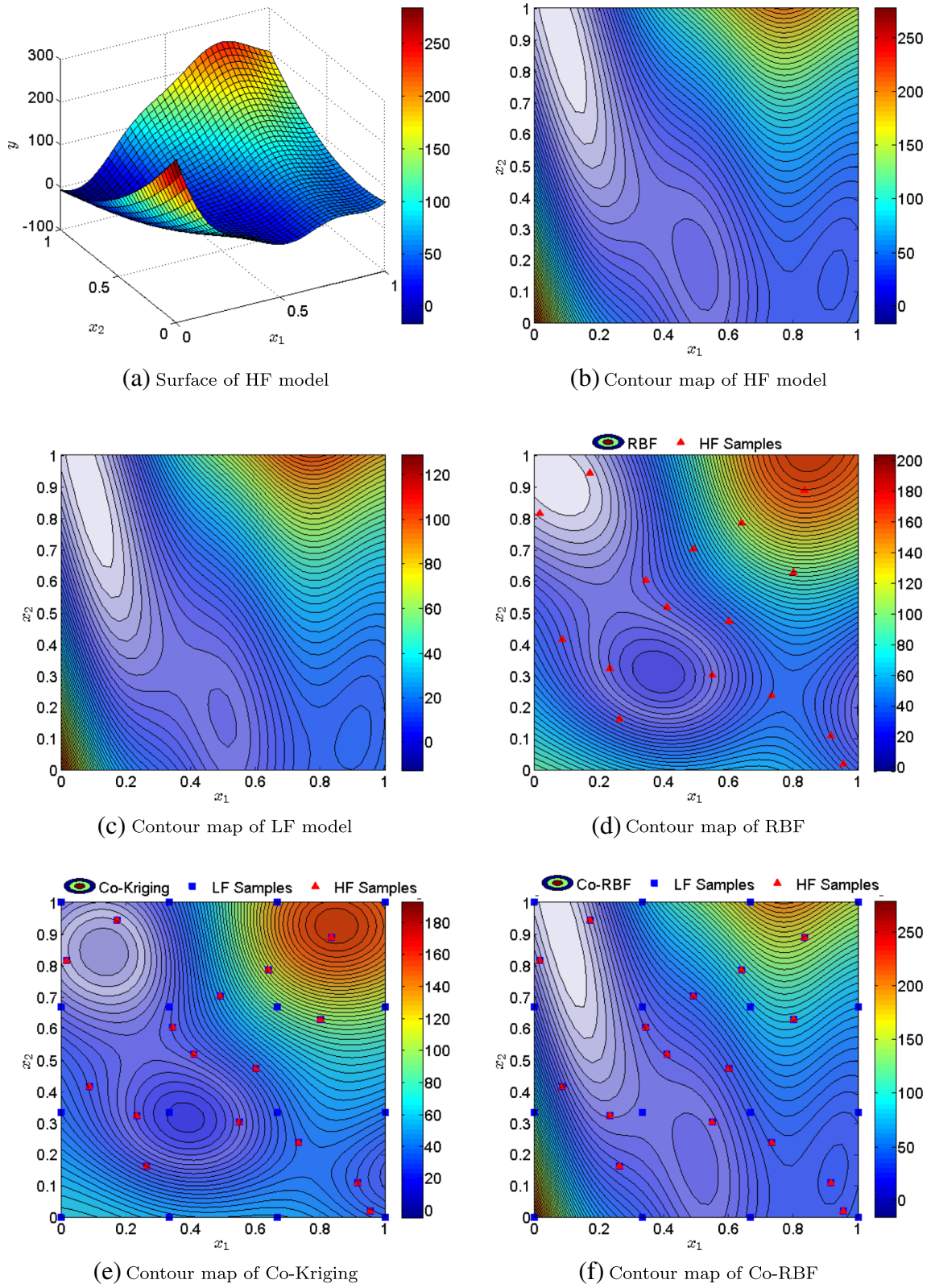


Fig. 6 Different prediction models

Table 5 Performance comparison of 1D example with the effect of the variable disturbance

Models	R ²	RMSE	MAX
LF	-0.4865	6.5808	11.6374
RBF	0.0730	5.1969	12.8466
Co-Kriging	0.9881	0.5880	1.4170
Co-RBF	0.9978	0.2528	0.6835

sample matrix. Given by $n \times p$ sample matrix \mathbf{X} , which is combined with n ($1 \times p$) row vectors $\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_n^T$, the distance is defined as

$$\text{dist}(\mathbf{x}, \mathbf{X}) = \min_{\mathbf{x}_i} \sqrt{(\mathbf{x} - \mathbf{x}_i)^T \mathbf{D}^{-1} (\mathbf{x} - \mathbf{x}_i)} \quad (22)$$

where \mathbf{D} is the $p \times p$ diagonal matrix whose j th ($j = 1, 2, \dots, p$) diagonal element is the variance of the j th column of the data matrix \mathbf{X} .

Table 6 Performance comparison of example 2

Models	R ²	RMSE	MAX
LF Model	0.3808	47.4268	151.5645
RBF	0.8049	26.6253	162.4648
Co-Kriging	0.7554	29.8114	190.9640
Co-RBF	1.0000	9.5152×10^{-7}	5.4481×10^{-6}

3 Numerical examples

In this section, 2 numerical and 3 engineering examples are used to demonstrate the performance of Co-RBF, compared with RBF and Co-Kriging. The 1D and 2D examples demonstrate how Co-RBF works in the simple cases, while the engineering examples illustrate the effectiveness of Co-RBF in practice. Co-Kriging is realized by Forrester’s MATLAB toolbox (Forrester et al. 2008) which can be found

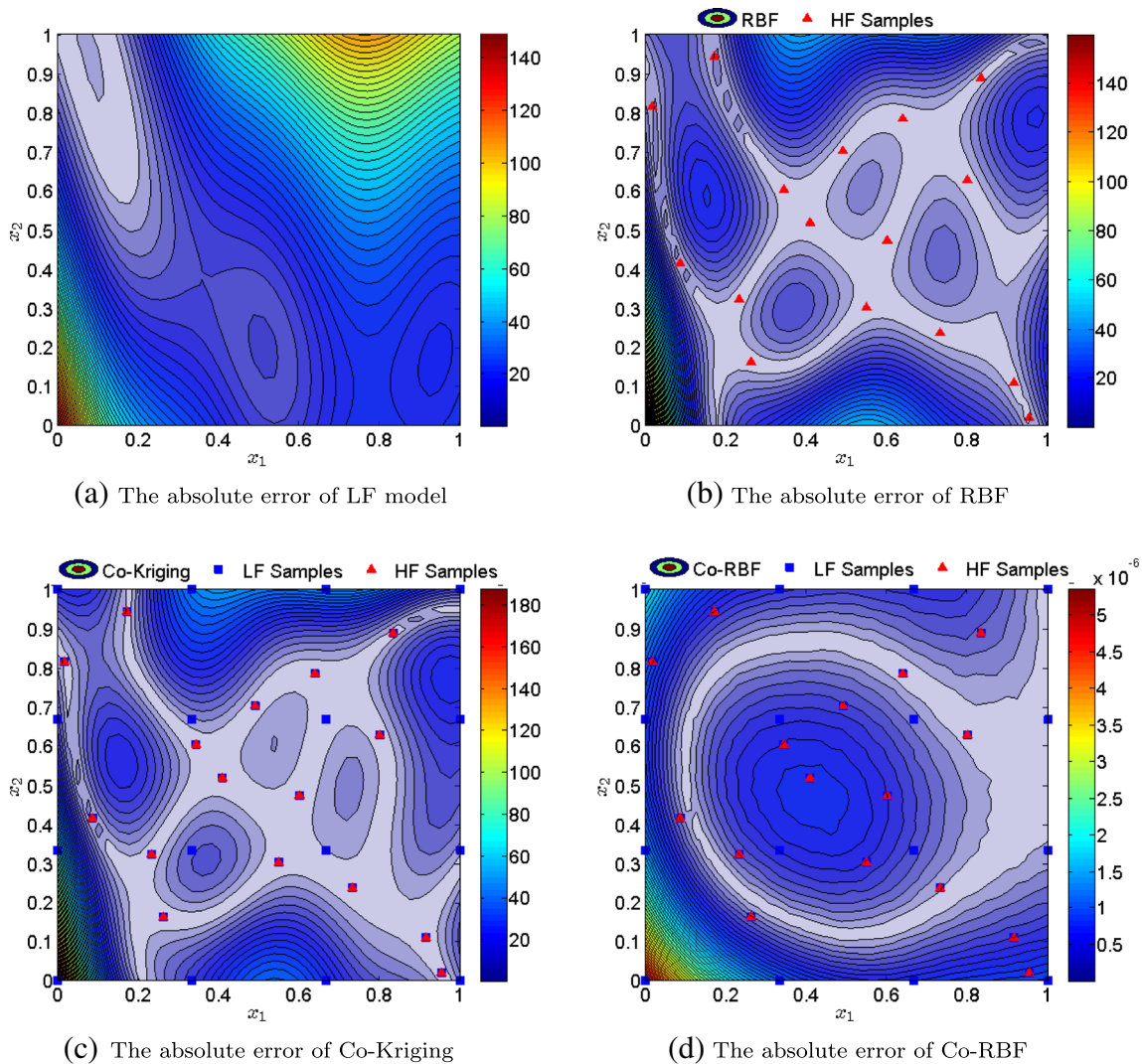
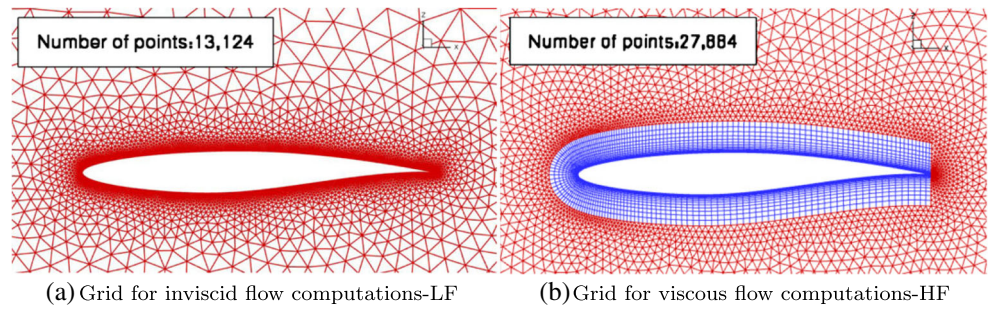


Fig. 7 Contour map of the prediction absolute errors

Fig. 8 Grids for RAE2822 (Han 2013)



on the book website at <http://www.wiley.com/go/forrester>. RBF and Co-RBF are realized by the authors' in-house MATLAB toolbox.

3.1 Example 1: 1D function

Consider the following 1D numerical model in which the LF model is almost linear dependent with the HF one

(Forrester et al. 2007, 2008). The coefficient β_L is constant and the correction function is $10(x - 0.5) - 5$. HF Model:

$$y_H(x) = (6x - 2)^2 \sin(12x - 4), x \in [0, 1]$$

LF Model:

$$y_L(x) = 0.5(6x - 2)^2 \sin(12x - 4) + 10(x - 0.5) - 5$$

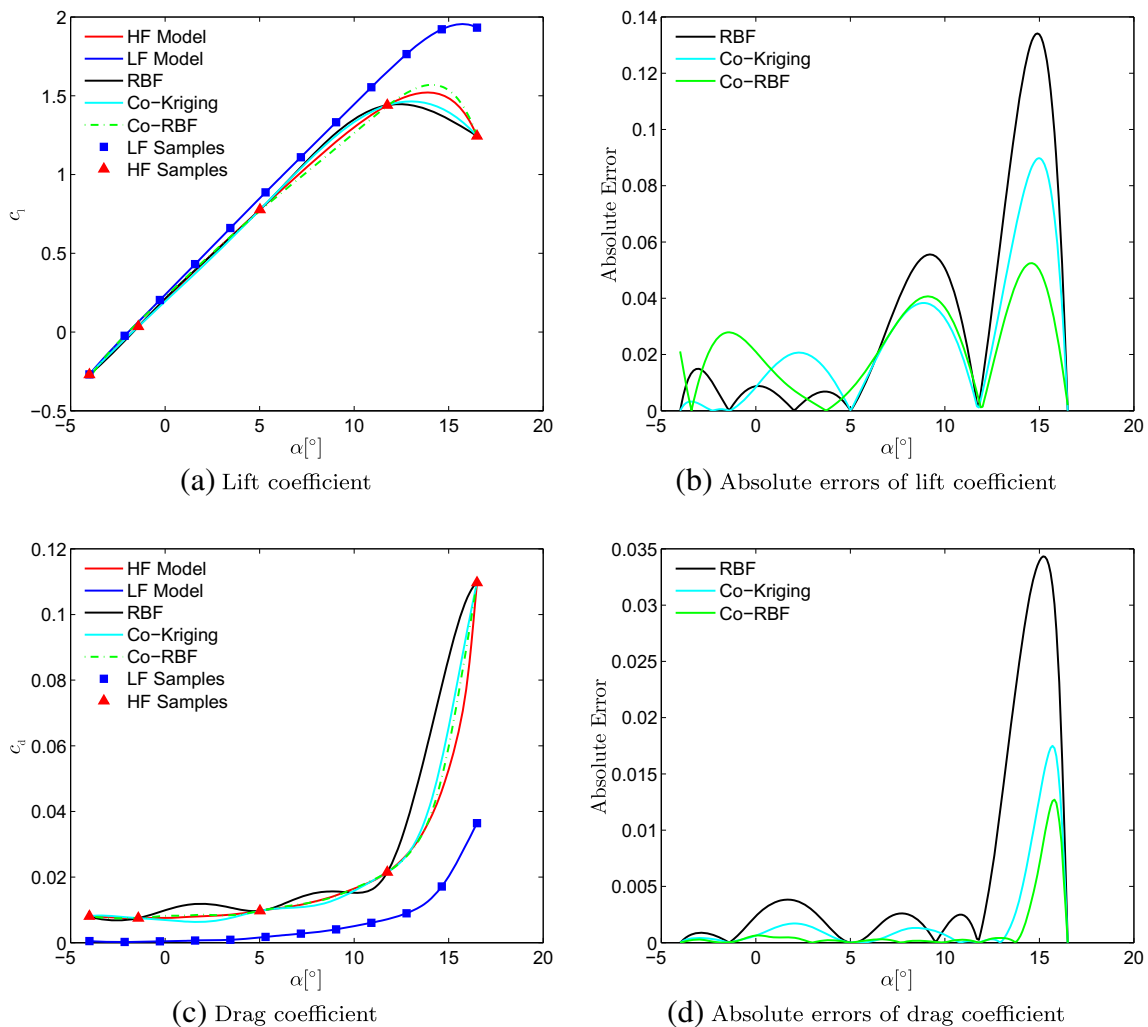


Fig. 9 Comparisons of VFMs and the absolute errors for RAE2822

According to Forrester et al. (2008), 4 HF samples $\mathbf{X}_H = [0, 0.4, 0.6, 1]^T$ and 11 LF samples $\mathbf{X}_L = [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1]^T$ are used to build the surrogate models. And additional 11 samples $\mathbf{X}_V = [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1]^T$ for both HF and LF models, are taken to compute RMSE, MAX and R^2 .

3.1.1 Predicted accuracy analysis

Figure 2a and b show that the RBF surrogate model based on the HF samples performs poor with large error than the other two methods. Co-Kriging performs better than RBF and captures the features of HF model with a prediction curve close to the HF function. However, in the scaling absolute error of Fig. 2c and d, Co-RBF performs even much better than Co-Kriging, since the maximum absolute error is about 0.8 in Co-Kriging while about 5×10^{-6} in Co-RBF. More details about the performance comparison are given in Table 2. The Co-RBF method has a good prediction accuracy with $R^2 = 1.0000$ and $RMSE = 3.1412 \times 10^{-6}$, thus it is a perfect approximation of the HF model for this problem.

3.1.2 Effect of the correction function $\delta(\mathbf{x})$

In the earlier section, Co-RBF performs well, since the correction function $\delta(\mathbf{x})$ is relatively simple. However, what if $\delta(\mathbf{x})$ becomes more complex? Intuitively, all the surrogate methods would perform worse. In what follows, the effect of the correction function is demonstrated and discussed. In the 1D numerical example, the LF model is modified as

$$y_L(x) = 0.5(6x - 2)^2 \sin(12x - 4) + 10(x - 0.5) - 5 + x^3$$

Figure 3 and Table 3 show that the prediction accuracy of RBF keeps the same since it doesn't use the LF model information. However, the performance of Co-Kriging and Co-RBF decreases with the increase of the complexity of $\delta(\mathbf{x})$. In Fig. 3, the accuracy of Co-Kriging decreases much more than that of Co-RBF, while the Co-RBF method still has a good approximation of the HF model.

3.1.3 Effect of linear coefficient β_L

In the previous section, the effect of correction function $\delta(\mathbf{x})$ is discussed. Moreover, the linear coefficient β_L has a significant effect on the prediction accuracy. β_L is assumed to be a constant scaling factor, while there may exist some non-linear component in practice. In what follows, the effect of the nonlinearity is discussed. In the 1D numerical example, the LF model is modified as

$$y_L(x) = (0.5 + 0.1x^2)(6x - 2)^2 \sin(12x - 4) + 10(x - 0.5) - 5$$

Table 7 Comparison of prediction models for C_l

Models	R^2	RMSE	MAX
LF Model	0.8448	0.2255	0.6881
RBF	0.9926	0.0492	0.1338
Co-Kriging	0.9965	0.0339	0.0898
Co-RBF	0.9979	0.0263	0.0524

Figure 4 and Table 4 show that the predictions performances are reduced in different degree. Co-RBF doesn't perform significantly better than Co-Kriging, but still works. A conclusion that the basic linear dependent assumption plays an important role in Co-Kriging and Co-RBF can be drawn. If the assumption is violated, more HF samples should be added to improve the accuracy.

3.1.4 Effect of the variable disturbance

In practice, the relation between HF and LF is unknown. Though the almost linear dependent assumption is common and suitable in most cases, the other possible assumptions also should be taken into consideration to validate the robustness and range of application of the method. Here we assume the LF model comes from the HF model with small variable disturbance. In the 1D numerical example, the LF model is modified as

$$y_L(x) = 0.5(6(x + 0.005) - 2)^2 \sin(12(x + 0.005) - 4)$$

Figure 5 and Table 5 show that a small disturbance in the variable results in a prominent reduction of the prediction performances. Co-Kriging and Co-RBF performs better than RBF, since the VFMs include some information of LF model in some degree. It can be concluded that when the assumption is slightly violated Co-RBF and Co-Kriging still work, but the performance is worse than the case satisfying the assumption.

3.2 Example 2: 2D Branin function

The second numerical example comes from Forrester et al. (2008). It is a modified version of the traditional two-variable Branin function. The LF model

Table 8 Comparison of prediction models for C_d

Models	R^2	RMSE	MAX
LF Model	0.1874	0.0192	0.0733
RBF	0.7270	0.0111	0.0339
Co-Kriging	0.9584	0.0043	0.0169
Co-RBF	0.9829	0.0028	0.0116

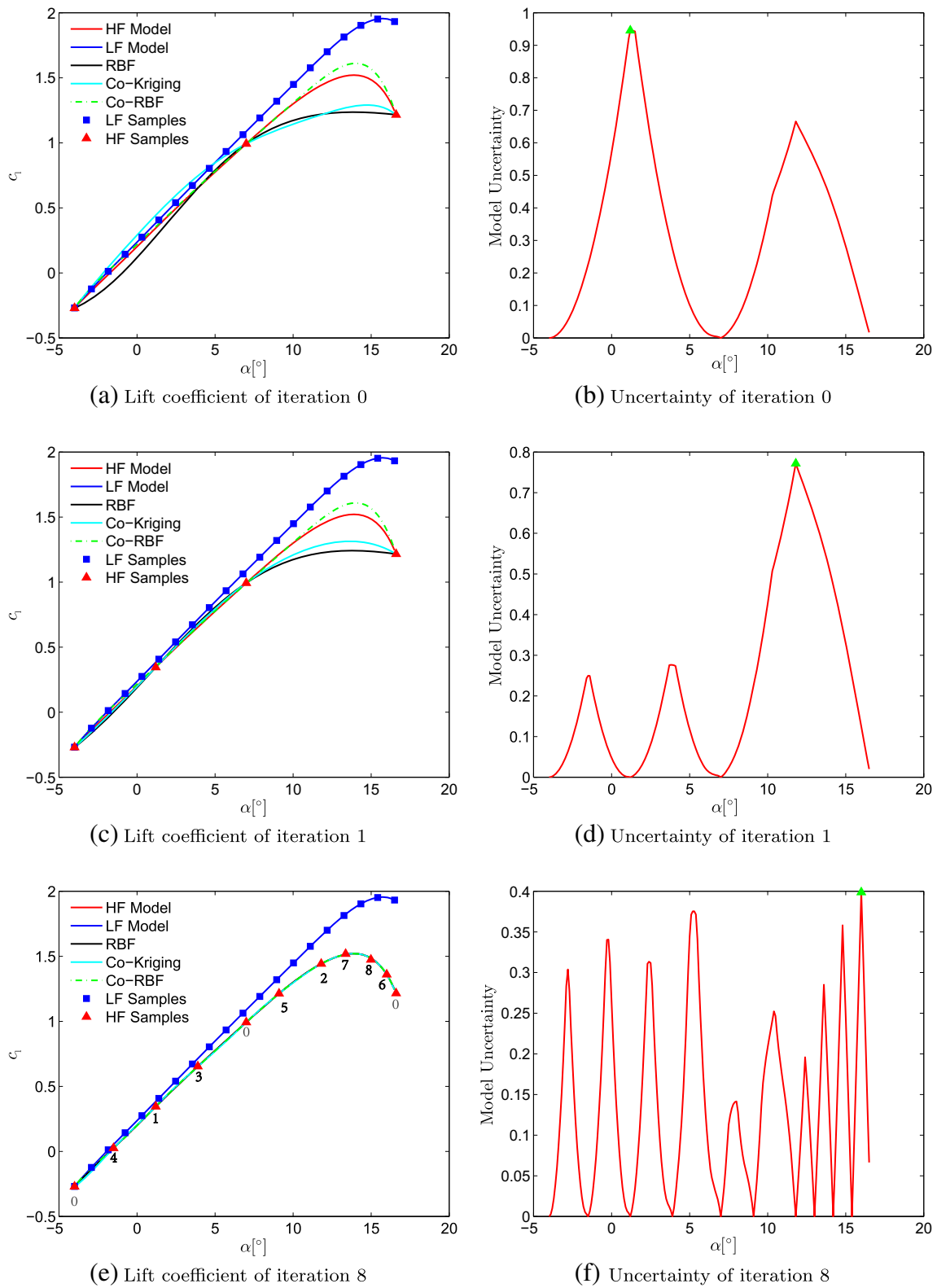


Fig. 10 Iterations of Lift coefficient and the model uncertainty for RAE2822

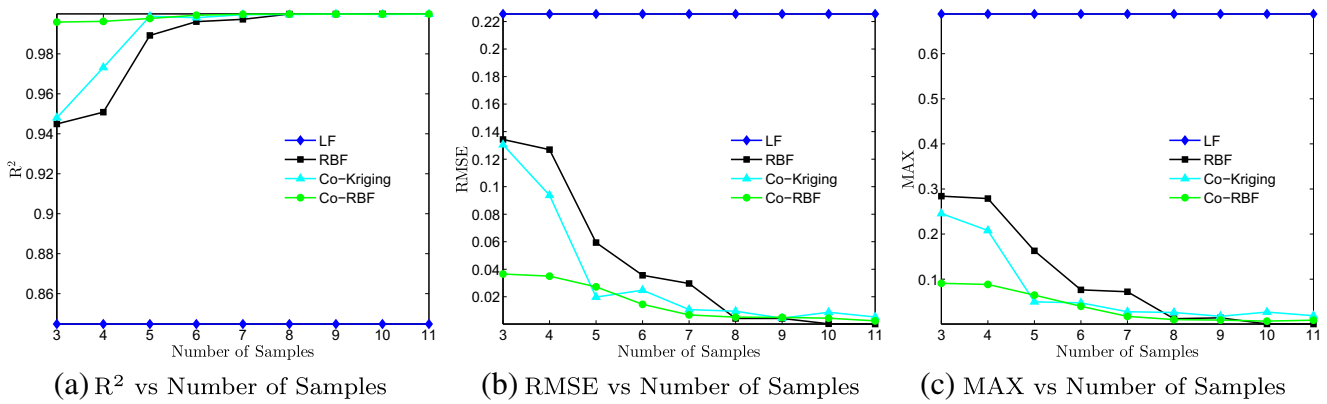


Fig. 11 Prediction performances of different models with increasing samples for RAE2822

is transformed from the HF model and a correction term is added. The HF and LF models are defined as $y_H(\mathbf{x}) = (15x_2 - 5.1/(4\pi^2)(15x_1 - 5)^2 - 6)^2 + 10((1 - 1/8\pi) \cos(15x_1 - 5) + 1) + 5(15x_1 - 5)$, $y_L(\mathbf{x}) = 0.5y_H(\mathbf{x}) + 10(x_1 - 0.5) + 5x_2 - 5$, $x_1 \in [0, 1]$, $x_2 \in [0, 1]$.

16 HF samples (See Fig. 6d) are included by latin hypercube sampling (LHS) (Forrester et al. 2008). 32 LF samples are included, among which 16 are the same with the HF samples and the left 16 are sampled by 4^2 full factorial design (FFD) (See Fig. 6e, f). Moreover, additional 20^2 HF and LF samples by FFD are used for the model validation.

Since the prediction surfaces are similar to the origin HF one, for convenience, we use the contour map to compare the performances of the prediction models. Figures 6 and 7 show the predictions and the absolute errors and Table 6 shows the performances for different models. We can see that the prediction performances of Co-RBF are much better than those of RBF and Co-Kriging, since R^2 is 1, RMSE is 9.5152×10^{-7} and MAX is 5.4481×10^{-6} . Since $R^2 = 1$, we know that the linear dependence assumption holds between LF model and HF model.

3.3 Example 3: aerodynamic coefficients of the RAE2822 airfoil

In the previous section, the 1D and 2D numerical problems are illustrated, and the effects of different factors are discussed. In the literature of surrogate model, the models are under different assumptions. We often validate the performances of models rather than validate the assumptions directly, which means that a good prediction performance illustrates the assumption holds. Next we consider the RAE2822 airfoil as shown in Fig. 8 (Han 2013) and use the Co-RBF to generate VFM models for the aerodynamic lift and drag coefficients (C_l and C_d) as functions of the attack angle α which varies between -4° and 16.5° . Setting the Mach number and the Reynolds number to be 0.2 and 6.5×10^6 , respectively. The HF model is Navier-Stocks code and the

LF model is Euler code. The aerodynamic data comes from Han (2013). This problem includes 5 HF samples, in which the first and last one are fixed and the other 3 are sampled by LHS, and 12 LF samples (See Fig. 9). Meanwhile, 42 validation samples $\mathbf{X}_V = [-4 : 0.5 : 16.5]^T$ are adopted.

Figure 9, Tables 7 and 8 show that Co-RBF performs better than the other two methods, since the absolute error curve is closer to zero and the prediction curve is closer to the lift coefficient curve or the drag coefficient curve, with R^2 to be 0.9979 and 0.9829 respectively.

However, if we want a more accurate prediction model, more HF samples are required to refine the VFM model. The process of adding points is shown in Fig. 10 and the validation errors are shown in Fig. 11. It can be seen from Fig. 10 that the higher the difference between the HF and LF model, the more sample points are added, and the sample distribution is relatively uniform throughout the sample space. Figure 11 shows that the accuracy of RBF, Co-Kriging and Co-RBF increases with the increase of sample points. When the number of samples is less than 8, the accuracy of Co-RBF is higher

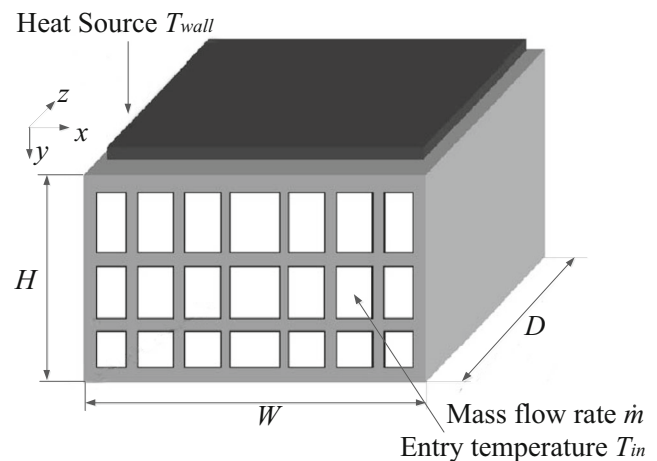


Fig. 12 Physical process for the heat exchanger

Table 9 Comparison of prediction models for the heat exchanger

Models	R ²	RMSE	MAX
LF Model	0.8655	4.431	10.490
Qian et al. (2006)	0.9017	3.795	8.490
Zheng et al. (2013)	0.9183	3.571	7.950
RBF (Zheng et al. 2013)	0.6460	6.990	17.293
RBF (LOOCV)	0.9460	2.807	5.170
Co-Kriging	0.9250	3.308	6.754
Co-RBF	0.9474	2.770	5.646

than that of RBF and Co-Kriging, however, RBF starts to perform better than the VFM method when the number of the sample points is higher than 8. Therefore, the VFMs have high accuracy when the number of the samples is small. With the increase of the sample points, the HF samples provide enough information to capture the features of the HF model. Thus, when enough HF samples are available, it isn't necessary to take into the LF model information.

3.4 Example 4: heat exchanger design problem

This section deals with a heat exchanger (shown in Fig. 12) for an electronic cooling application design problem with Co-RBF. The device dissipates the heat generated by a microprocessor. The influence factors include 4 variables, the entry temperature T_{in} , total mass flow rate \dot{m} , the heat source wall temperature T_{wall} and the solid material thermal conductivity k . For more details about the problem, see Qian et al. (2006).

Two types of simulations, computationally expensive HF FLUENT simulations and fast LF finite difference (FD) simulations are employed to analyze the impact on heat transfer rates. Each FLUENT simulation requires two three orders of magnitude more computing time than the corresponding FD simulation. However, the FLUENT simulations are generally more accurate than the FD simulations. The 22 HF samples and 64 LF samples are used to build

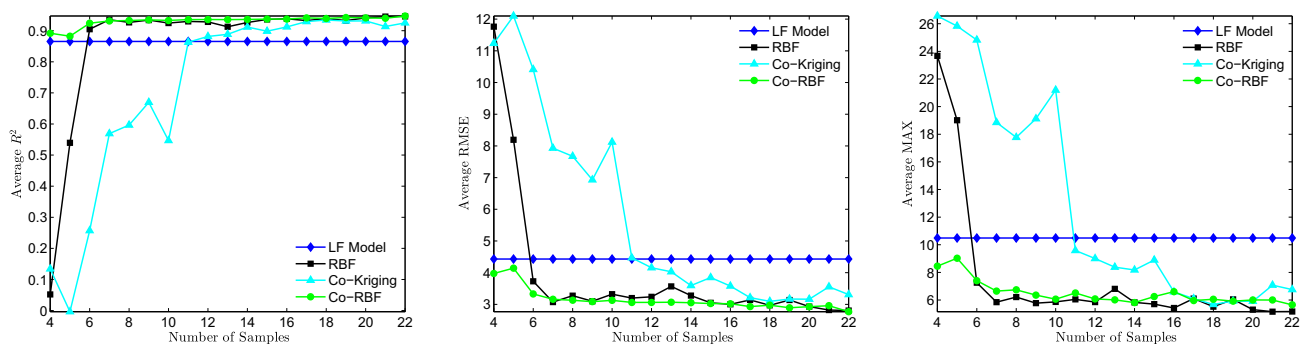
the surrogate models. For validation and comparison, the additional 14 samples are also utilized. For more details about the samples, see Qian et al. (2006).

The prediction performances comparison of the different models for the problem is illustrated in Table 9 and Fig. 13. Table 9 shows the results with 22 HF, 64 LF samples and 14 validation samples. Co-RBF (RMSE = 2.770) improves the RMSE of the LF model about 38.5%, while Co-Kriging, and the methods by Qian et al. (2006) and Zheng et al. (2013) improve that about 26.1%, 14.8% and 20.0%, respectively. To avoid the effect of different sampling locations, we take the average value of 50 times. Figure 13 illustrates the average performances of different samples. We can see that the performance of RBF improves fast with the number of HF samples increasing, due to the weak nonlinearity of the HF model. Co-RBF achieves a lower RMSE with small number (about 6) of samples, while Co-Kriging requires about 18 samples to reach the accuracy. Since the prediction accuracy is small, it illustrates that Co-RBF with few HF samples performs well in this engineering problem.

3.5 Example 5: inverted wing with vortex generators in ground effect

This section deals with an inverted wing with vortex generators in ground effect problem. The influence factors include 2 normalized variables (ranging between 0 to 1) of the wing incidence angle α and ride height h/c . The response is the sectional downforce C_{L_s} . For more details about the problem, see Kuya et al. (2011).

Two types of data, the expensive HF experimental data and relatively cheap LF FLUENT simulations are employed. The HF experimental data are obtained in the 2.1×1.5 m closed-section wind tunnel at the University of Southampton. The LF FLUENT data are obtained by three-dimensional steady Reynolds-averaged Navier-Stokes (RANS) simulations using the Spalart-Allmaras turbulence model. 12 HF samples are used to build the surrogate models. For the LF data two types of sampling designs are



(a) Average R² vs Number of Samples (b) Average RMSE vs Number of Samples (c) Average MAX vs Number of Samples

Fig. 13 Prediction performances of different models with increasing samples for the heat exchanger

Table 10 Model verification at four verification points regarding response of sectional downforce

Models	(0.25, 0.25)	(0.75, 0.25)	(0.25, 0.75)	(0.75, 0.75)	RMSE
HF Model (Kuya et al. 2011)	2.42	1.87	2.6	1.75	0.00
Kriging (Kuya et al. 2011)	1.75	1.74	1.74	1.74	0.55
Co-Kriging(FFD) (Kuya et al. 2011)	2.37	1.69	1.98	1.55	0.34
Co-Kriging(LHS) (Kuya et al. 2011)	2.51	1.56	2.08	1.51	0.33
LF Model	2.59	1.82	2.34	1.98	0.1949
RBF (Optimized with LOOCV)	2.0864	1.6326	1.7999	1.6424	0.4527
Co-RBF(FFD with LF Validation Values)	2.4192	1.7374	2.0363	1.6013	0.2989
Co-RBF(LHS with LF Validation Values)	2.5853	1.6189	2.1101	1.5989	0.2972
Co-RBF(FFD without LF Validation Values)	2.4192	1.7374	2.0363	1.6013	0.2989
Co-RBF(LHS without LF Validation Values)	2.4387	1.6214	2.1160	1.5801	0.2852

employed: full factorial design (FFD) and latin hypercube sampling (LHS) with 25 samples. For validation and comparison, the additional 4 validation samples are also utilized (Kuya et al. 2011).

The comparison of the different prediction performances is illustrated in Table 10. As a reference, the Kriging and RBF methods are constructed only with the HF samples. They are used to highlight how much the LF samples contribute to improving the prediction accuracy. Table 10 shows the model validations with RSME at four validation points. The Kriging and RBF methods show a poor accuracy due to the sparsity and location of the HF samples, leading to the model features not being captured. The Co-Kriging method improves the accuracy by about 40% compared with the Kriging method, while the Co-RBF method improves that about 46% – 48%. For the Co-RBF with the sampling method of FFD, the LF validation samples don't affect the RMSE, because in this problem the prediction locations are sub set the initial LF locations. In Co-RBF without LF validation values, since the LF values are not available, we have to use extra RBF interpolation to calculate the values at the prediction locations. Moreover, the RMSE of LF model with a value of 0.1949 is much better than any of the other prediction models in this problem, which means that the VFMs don't always work better than the LF model. However, when the LF model (FLUENT simulation) is time-consuming, the VFM is necessary.

4 Conclusion

An efficient Co-RBF for variable-fidelity surrogate modelling is proposed in this paper. The HF samples and LF model are combined by a constant scaling factor and additional radial basis functions. According to the analytical example 1 and 2, Co-RBF performs well with a small number of HF samples when the HF model and LF model have an approximate linear relationship. However, in most

cases the link between the HF and LF is unknown, thus some validation samples are required to measure the prediction accuracy of Co-RBF. When the accuracy is not enough, more HF samples are required to refine the surrogate model. In this paper, we add samples where the samples are sparse and where the HF samples and the LF model differ greatly. The engineering example 3 and 4 show that Co-RBF have better accuracy than the alternative methods when the number of the HF samples is small. However, when the HF samples are more and more, the surrogate model (RBF) constructed just with the HF samples have similar accuracy with the VFMs (Co-RBF and Co-Kriging), which means that when the HF samples are enough the LF model doesn't improve the accuracy significantly. Thus, VFMs are well suited for the situations in which the number of the HF samples is small. The engineering example 5 compares Co-RBF with the existing alternative methods, which further shows the effectiveness and superiority of the proposed method. The further research may be the construction of variable-fidelity model in specific application background and taking more physical or empirical information into consideration. In addition, the criteria for adding points in different application backgrounds are worthy of further study.

Acknowledgements The authors gratefully appreciate the support by Fundamental Research Funds for the Central Universities (Grant No. G2016KY0302), National Natural Science Foundation of China (Grant No. 51505385, No. 11572134), and also thank Dr. Hua Su and Dr. Chunna Li for the helpful discussion about the Co-RBF method. Moreover, thanks Dr. Zhao Jing and the anonymous reviewers for their efforts and constructive advice to improve the study.

References

- Bischl B, Wessing S, Bauer N, Friedrichs K, Weihs C (2014) MOI-MBO: multiobjective infill for parallel model-based optimization. Springer International Publishing

- Booker AJ, Dennis JE, Frank PD, Serafini DB, Torczon V, Trosset MW (1998) A rigorous framework for optimization of expensive functions by surrogates. *Struct Multidiscip Optim* 17(1):1–13
- Boopathy K, Rumpfkeil MP (2015) A multivariate interpolation and regression enhanced kriging surrogate model. In: AIAA computational fluid dynamics conference
- Breitkopf P, Coelho RF (2010) Multidisciplinary design optimization in computational mechanics. Iste Ltd, London
- Courrier N, Boucard PA, Soulier B (2016) Variable-fidelity modeling of structural analysis of assemblies. *J Glob Optim* 64(3):577–613
- Eldred M, Giunta A, Collis S (2004) Second-order corrections for surrogate-based optimization with model hierarchies. In: Proceedings of the 11th AIAA/ISSMO multidisciplinary analysis & optimization conference
- Fasshauer GE, Zhang JG (2007) On choosing optimal shape parameters for rbf approximation. *Numer Algor* 45(1–4):345–368
- Forrester AIJ, Keane AJ (2007) Multi-fidelity optimization via surrogate modelling. *Proc R Soc A Math Phys Eng Sci* 463(2088):3251–3269
- Forrester DAJ, Sobester DA, Keane AJ (2008) Engineering design via surrogate modelling: a practical guide. Wiley, West Sussex
- Gano S, Renaud J, Sanders B (2005) Hybrid variable fidelity optimization by using a kriging-based scaling function. *AIAA J* 43(11):2422–2433
- Gano SE, Renaud JE, Martin JD, Simpson TW (2006) Update strategies for kriging models used in variable fidelity optimization. *Struct Multidiscip Optim* 32(4):287–298
- Gorissen D, Couckuyt I, Demeester P, Dhaene T, Crombecq K (2010) A surrogate modeling and adaptive sampling toolbox for computer based design. *J Mach Learn Res* 11(1):2051–2055
- Gutmann HM (2000) A radial basis function method for global optimization. *J Glob Optim* 19(3):201–227
- Haftka RT (1991) Combining global and local approximations. *AIAA J* 29(9):1523–1525
- Haftka RT, Villanueva D, Chaudhuri A (2016) Parallel surrogate-assisted global optimization with expensive functions: a survey. *Struct Multidiscip Optim* 54(1):3–13
- Han ZH, Zimmermann R, Gortz S (2012) Alternative cokriging model for variable-fidelity surrogate modeling. *Aiaa J* 50(5):1205–1210
- Han ZH, Gortz S, Zimmermann R (2013) Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. *Aerosp Sci Technol* 25(1):177–189
- Hastie T, Tibshirani R, Friedman J (2000) The elements of statistical learning
- Kaare AJ, Brandt P, Michael, Syskind P (2012) The matrix cookbook. <http://matrixcookbook.com>
- Keane AJ, Nair PB (2005) Computational approaches for aerospace design: the pursuit of excellence. *Can Med Assoc J* 37(6):351–360
- Kennedy MC, O’Hagan A (2000) Predicting the output from a complex computer code when fast approximations are available. *Biometrika* 87(1):1–13
- Kuya Y, Takeda K, Zhang X, Forrester AIJ (2011) Multifidelity surrogate modeling of experimental and computational aerodynamic data sets. *AIAA J* 49(2):289–298
- Laurenceau J, Sagaut P (2015) Building efficient response surfaces of aerodynamic functions with kriging and cokriging. *Aiaa J* 46(2):498–507
- Leary SJ, Bhaskar A, Keane AJ (2003) A knowledge-based approach to response surface modelling in multifidelity optimization. *J Glob Optim* 26(3):297–319
- Leifsson L, Koziel S (2015) Aerodynamic shape optimization by variable-fidelity computational fluid dynamics models: a review of recent progress. *J Comput Sci* 10:45–54
- Lewis RM, Nash SG (2005) Model problems for the multigrid optimization of systems governed by differential equations. *Siam J Sci Comput* 26(6):1811–1837
- Liem RP, Mader CA, Martins JRRA (2015) Surrogate models and mixtures of experts in aerodynamic performance prediction for aircraft mission analysis. *Aerosp Sci Technol* 43(8):121–151
- Marduel X, Tribes C, Trpanier JY (2006) Variable-fidelity optimization: efficiency and robustness. *Optim Eng* 7(4):479–500
- Qian Z, Seepersad CC, Joseph VR, Allen JK, Wu CFJ (2006) Building surrogate models based on detailed and approximate simulations. *J Mech Des* 128(4):668–677. ASME 2004 Design Engineering Technical Conference, Salt Lake City, UT, 2004
- Rodriguez JF, Prez VM, Padmanabhan D, Renaud JE (2001) Sequential approximate optimization using variable fidelity response surface approximations. *Struct Multidiscip Optim* 22(1):24–34
- Simpson TW, Booker AJ, Ghosh D, Giunta AA, Koch PN, Yang RJ (2004) Approximation methods in multidisciplinary analysis and optimization: a panel discussion. *Struct Multidiscip Optim* 27(5):302–313
- Sobieszcanski-Sobieski J, Haftka RT (1997) Multidisciplinary aerospace design optimization: survey of recent developments. *Struct Optim* 14(1):1–23
- Tetko IV, Livingstone DJ, Luik AI (1995) Neural network studies. 1. comparison of overfitting and overtraining. *J Chem Inf Comput Sci* 35(5):826–833
- Zahir MK, Gao Z (2012) Variable fidelity surrogate assisted optimization using a suite of low fidelity solvers. *Open J Optim* 1(1):8–14
- Zheng J, Qiu H, Zhang X (2012) Variable-fidelity multidisciplinary design optimization based on analytical target cascading framework. *Adv Mater Res* 544:49–54
- Zheng J, Shao X, Gao L, Jiang P (2013) A hybrid variable-fidelity global approximation modelling method combining tuned radial basis function base and kriging correction. *J Eng Des* 24:604–622