Ocean Engineering 243 (2022) 110239

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Multi-fidelity Co-Kriging surrogate model for ship hull form optimization

Xinwang Liu [a,b], Weiwen Zhao [a], Decheng Wan [a,*]

[a] Computational Marine Hydrodynamics Lab (CMHL), School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China

[b] College of Mathematical Sciences, Harbin Engineering University, Harbin, 150001, China

ARTICLE INFO

Keywords:
Multi-fidelity
Hull form optimization
Co-Kriging surrogate model
Potential flow
Viscous flow

ABSTRACT

For the simulation-based hull form optimization design, there are many methods to evaluate the hydrodynamic performance of the hull form. Although the high fidelity of the surrogate model can be guaranteed by evaluating a large number of new sample hulls based on viscous flow theory, the computational cost can be too high. Therefore, in order to release the burden of calculation, based on the traditional single-fidelity Kriging surrogate model, the multi-fidelity Co-Kriging surrogate model gives attention to both high accuracy and efficiency by combining the accuracy advantage of high-fidelity sample evaluation with the efficiency advantage of low-fidelity sample evaluation. This paper first introduces the construction process of the multi-fidelity Co-Kriging surrogate model, and then uses a series of numerical examples to illustrate the advantages of the multi-fidelity Co-Kriging surrogate model compared with the single-fidelity Kriging surrogate model in terms of fidelity and efficiency. Finally, a hull form optimization design for total drag of DTMB-5415 hull at the design speed is given in detail, where the viscous flow theory and potential flow theory are used for the hydrodynamic evaluations of the hull forms to obtain the high- and low-fidelity results respectively. Results show that the multi-fidelity Co-Kriging surrogate model can be established for hull form hydrodynamic performance optimization, which is superior to the single-fidelity Kriging surrogate model in accuracy, and the optimal hull obtained by Co-Kriging surrogate model has a better resistance optimization effect.

1. Introduction

For optimization problems in engineering fields, it is often difficult to find the accurate functional relationship between the objective function and the design variables. In order to seek the optimal solution through the optimization algorithm (especially intelligent optimization algorithm), it is theoretically necessary to evaluate each new sample generated in each iteration by model test or numerical simulation, so that the optimization algorithm, driven by the changing law of the objective function value, can gradually search for the optimal solution. However, for the hull form optimization design problems, numerical simulations or model tests requires high costs. Therefore, the total optimization period will be extremely long, resulting in a significant reduction in optimization efficiency.

Surrogate model construction technology is a very effective way to solve the problem mentioned above. The so-called Surrogate Model (SM), or Approximate Model (AM), refers to the approximate functional relationship between dependent variable (objective function) and independent variables (design variables). The surrogate-based optimization method is to use the surrogate model constructed by sample points generated by the design of experiment with their corresponding objective function values, which replaces the real but difficult-to-find expression of objective function on the design variables, to seek the optimal solution (set).

The construction of the surrogate model is carried out after the sample points are determined in the design space by design of experiment and their performance are evaluated. At present, the mainstream surrogate models include Kriging (Krige, 1951), Response Surface Model (RSM) (Box and Wilson, 1951), Artificial Neural Network (ANN) (Smith, 1993), Radial Basis Function (RBF) (Orr, 1995), and Support Vector Machine (SVM) (Smola and Schölkopf, 2004).

Surrogate models have been widely used in many engineering optimization fields. Peri et al. (2001) introduced the surrogate models into the field of hull form optimization early, and compared the advantages and disadvantages of surrogate models such as RSM, Kriging, ANN, and RBF in detail. Since then, several studies have given examples of hull form optimization based on surrogate model.

There is no doubt that the use of single-fidelity surrogate model

* Corresponding author.
E-mail addresses: liuxinwang@hrbeu.edu.cn (X. Liu), dcwan@sjtu.edu.cn (D. Wan).

https://doi.org/10.1016/j.oceaneng.2021.110239
Received 7 May 2021; Received in revised form 17 October 2021; Accepted 19 November 2021

makes the simulation-based hull form optimization design possible. Campana et al. (2006) used the SBD technique early to optimize the DTMB-5415 hull. Tahara et al. (2011) adopted the Kriging model when optimizing the Delft catamaran. Li et al. (2014) adopted the RSM when optimizing the total drag of a bulk carrier. Chen et al. (2015) used RBF to optimize Delft catamaran 372 model with considering the design-space reduction. Huang et al. (2015) used RBF when optimizing the resistance and seeping performance of Series 60 hull. Yang and Huang (2016) adopted RBF when optimizing Series 60 hull for reduced total resistance at two speeds. Hou (2017) used ANN for EEDI optimization of Wigley hull. Tezdogan et al. (2018) optimised the calm-water total drag of a fishing boat by using SBD technique. Diez et al. (2018) used RBF to optimize the resistance and seakeeping performances of Delft catamaran 372 model considering stochastic conditions. Zong et al. (2018) used second-order RSM to optimize the total drag coefficient of a trimaran. Coppede et al. (2019) used the Kriging model to optimize the total drag of KCS ship. Zhang et al. (2018) used Elman Neural Network to optimize the total drag coefficients of DTMB-5512 and Wigley III hulls in calm water at design speed. Miao et al. (2020) optimized a S60 catamaran for resistance reduction by the change of the demihull shape and separation using Kriging model. Liu and Wan (2020) used Kriging model to optimize the wave-making drag coefficients of a luxury cruise ship under the design and maximum speeds. Serani et al. (2021) optimized DTMB-5415 hull by using RBF to minimize the mean total resistance expected value and maximize the ship operability in a fully stochastic environment.

Compared with several other common surrogate models, Kriging model can give the prediction error (variance) in probability at the same time while giving the prediction value at the predicted point. Therefore, in order to improve its prediction accuracy, the prediction error in probability given by the Kriging model can be used naturally and easily, such as the maximum Expectation Improvement (EI) criterion (Jones et al., 1998), the maximum mean Square Error (MSE) criterion (Sasena et al., 2002), and the maximum Probability Improvement (PI) criterion (Forrester and Keane, 2009).

Even so, for the simulation-based hull form optimization design, if the performance evaluation of each new sample hull is done by the same solver, the surrogate model constructed from the evaluation results obtained can be called a single-fidelity surrogate model. On the contrary, if the evaluation results of solvers with different fidelity are considered together, the final constructed surrogate model is generally called a multi-fidelity surrogate model. Under the premise that the number of samples is large enough, the conventional single-fidelity surrogate model will have relatively high accuracy, which can basically represent the change rule of the objective function in the whole design space, and be used for optimization solution. For the actual hull form optimization problem, especially for the high-dimensional problem, in order to ensure high accuracy of the surrogate model obtained, it is necessary to evaluate a large number of new sample hulls, which in fact will consume a lot of computing time and resources, and the optimization efficiency is limited to an extent.

Multi-fidelity Kriging surrogate model based on Bayesian deviation was first proposed by Kennedy and O 'Hagan (2001), which requires prior information (Qian and Wu, 2008). Co-Kriging surrogate model (Sacks et al., 1989) provides the equivalent form of above Bayesian-based model, which does not require prior information and has good computational characteristics (Forrester et al., 2007; Kuya et al., 2011). Multi-fidelity Co-Kriging surrogate model essentially uses a larger number of low-fidelity sample data to assist a smaller number of high-fidelity sample data for forecasting high-fidelity outputs(results), which can reduce the total calculation cost and make the surrogate model have relatively high accuracy (Kim et al., 2007; Ghoreyshi et al., 2008; Roderick et al., 2014).

For the simulation-based hull form optimization of comprehensive hydrodynamic performances, it is very necessary to establish a multi-fidelity surrogate model, since it can greatly reduce the burden of calculation by obtaining the sample hulls' performance indexes with different fidelity in a faster time period before getting the optimal hull form. In fact, there are different levels of fidelity for the ship's hydrodynamic performance evaluation methods. That is to say, in order to obtain the hydrodynamic performance indexes of the new hull forms, methods with different physical models or numerical discrete scheme can be used. If the two methods are complementary to each other, it is possible to evaluate the performance of new sample hulls with high efficiency and build the surrogate model with high accuracy.

Adaptive sampling technique can also be used in the construction of multi-fidelity surrogate model. Serani et al. (2019b) presented four adaptive sampling methods for multi-fidelity RBF surrogate model, and the stochastic RBF provides the surrogate model prediction and the associated uncertainty, which can be a guide to the new samples. On the one hand, the adaptive sampling can purposefully increase sample points, which may lead to a better performance of the surrogate model with smaller prediction error; on the other hand, the number of the initial points and the new points are hard to determine, and if the adaptive sampling is adding-one-point per cycle, the whole optimization process may take too long, but the one-shot sampling for the high- and low-fidelity samples may lead to their numerical evaluation being carried out almost simultaneously.

Furthermore, derivative-enhanced multi-fidelity surrogate models are also constructed in order to improve the accuracy of the model. Han et al. (2013) developed a derivative-enhanced multi-fidelity Kriging surrogate model combined with a new Generalized Hybrid Bridge Function to improve the efficiency and accuracy. Yamazaki and Mavriplis (2013) developed a derivative-enhanced multi-fidelity Kriging surrogate model based on a direct Kriging formulation, and combined the developed surrogate models with an efficient adjoint CFD gradient evaluation method respectively. Rumpfkeil and Beran (2017) adaptively select training points and use derivative information to enhance the fidelity of the surrogate model. Furthermore, no more than three-dimensional analytic test functions are given to demonstrate their benefits. Zuhal et al. (2021) proposed derivative-enhanced multi-fidelity Kriging surrogate model with a polynomial of undetermined order as the deterministic (trend) function. The airfoil drag minimization problems were respectively given in the end. However, unlike mathematical function or aerodynamic performance optimization, for the hull form hydrodynamic performance optimization, since numerical uncertainty may exist, and the adjoint CFD gradient evaluation method considering free surface is hard to implement (for instance, the VOF method needs to solve the discontinuous phase fraction equation), finite difference method should be used to obtain the gradient information, which needs more additional points that are near the generated sample points. Unfortunately, the stability of the derivative calculation cannot be ensured considering the difference spacing. Therefore, accurately calculate the gradient may occur high computational cost.

More-than-two-fidelity surrogate model can also be established and the numerical noise can be considered in the multi-fidelity surrogate model construction. Serani et al. (2019a) presented the extension of a two-fidelity RBF surrogate model, which has been tested for an analytical test problem, and the optimization of a NACA hydrofoil, and a RoPax ferry. Results showed that the use of three-fidelity RBF surrogate model achieved better results than one- or two-fidelity models. In fact, like multi-fidelity RBF model, multi-fidelity Co-Kriging model can also be extended to three-or-even-more-fidelity one. Wackers et al. (2020) assessed the performance of an adaptive multi-fidelity RBF and Kriging model based on noisy CFD data for design space exploration and design optimization. RANS-based CFD simulations have been performed and three fidelity levels have been used based on an adaptive grid refinement technique. Results showed that using more than one fidelity level reduces the number of high-fidelity samples required for surrogate-based-optimization, while using intermediate fidelity levels has a beneficial effect when dealing with noisy data. Ficini et al. (2021) assessed the performance of an adaptive multi-fidelity surrogate model based on Kriging considering noisy function evaluation. Through five

mathematical test problems, surrogate model with three fidelity level has led to an improvement in comparison with that with one and two fidelities. However, the real noise amplitude is hard to know exactly, even the numerical (calculation) noise has high randomness. That is to say, if we know the noise distribution, we can simply do the compensation for errors, which is unrealistic. Of course, the overall fidelity affected by the noise should be considered and studied.

To sum up, traditional optimization solution based on the single-fidelity Kriging surrogate model usually needs to evaluate a large number of sample points in order to make the surrogate model have high accuracy. In hull form optimization, to be specific, comprehensive hydrodynamic performances of a series of new sample hulls need to be evaluated. Multi-fidelity Co-Kriging surrogate model can give attention to both high efficiency and high precision, where fewer samples are used for high-fidelity simulation and more samples for low-fidelity simulation, but it has not been widely used in the hull form hydrodynamic performance optimization yet. In this paper, two-fidelity Co-Kriging surrogate model is applied in the resistance performance of the DTMB-5415 hull, where the potential flow and viscous flow simulation results are regarded as the low- and high-fidelity data. Before the application, several mathematical functions with different dimensions are tested, showing that the Co-Kriging model has some advantages compared with the single-fidelity Kriging model using static or even adaptive samples. The single-fidelity Kriging and multi-fidelity Co-Kinging model are constructed for the optimization of the DTMB-5415 hull, and the optimization results show that the Co-Kriging model has relatively higher accuracy and costs less. Last but not the least, some recommendation for the decision of the number of the high-fidelity and low-fidelity samples are also given.

2. Basic principles of Co-Kriging surrogate model construction method

Co-Kriging surrogate model can be established by two sets of independent samples with high and low reliability (accuracy). The model can obtain more accurate prediction results by using fewer high-fidelity samples and more low-fidelity samples when the high-fidelity sample points are difficult to obtain or the calculation is expensive (Forrester et al., 2007; Gratiet and Garnier, 2014).

Firstly, the sample data sets under two levels of fidelity are given, namely, the high-fidelity samples $X_e = [x_e^{(1)}, x_e^{(2)}, \cdots, x_e^{(N_e)}]$ with a sample number of N_e, and the low-fidelity samples $X_c = [x_c^{(1)}, x_c^{(2)}, \cdots, x_c^{(N_c)}]$ with a sample number of N_c. Like the Kriging surrogate model, the value of the function at a certain point in the space is regarded as a random process with certain expectation and variance. For Co-Kriging model, we have a set of sample function values

$$Y = \begin{bmatrix} Y_c(X_c) \\ Y_e(X_e) \end{bmatrix} = \left[Y_c\left(x_c^{(1)}\right), Y_c\left(x_c^{(2)}\right), \cdots, Y_c\left(x_c^{(N_c)}\right), Y_e\left(x_e^{(1)}\right), Y_e\left(x_e^{(2)}\right), \cdots, Y_e\left(x_e^{(N_e)}\right)\right]^T \tag{1}$$

Based on the auto-regressive model, it is assumed that $\mathrm{cov}[Y_e(x^{(i)}), Y_c(x^{(i)})|Y_e(x^{(i)})] = 0$, that is, when the high-fidelity function value at $x^{(i)}$ is known, the low-fidelity model at this point cannot provide any more information.

Assume Gaussian processes Z_e and Z_c represent the approximation of high-fidelity and low-fidelity functions respectively. Based on auto-regressive model, the Gaussian processes Z_e can be obtained by the sum of low-fidelity model multiplied by a constant scale factor ρ and the deviation function Z_d, as shown in the following:

$$Z_e(x) = \rho Z_c(x) + Z_d(x) \tag{2}$$

We know that the covariance based on the Kriging surrogate model is

$$\mathrm{cov}\left[Y\left(x^{(i)}\right), Y\left(x^{(j)}\right)\right] = \sigma^2 \psi\left(x^{(i)}, x^{(j)}\right) \tag{3}$$

where σ^2 is the variance, and the correlation matrix is determined by Spatial Correlation Function (SCF), one of whose common forms is the Gaussian SCF:

$$\psi\left(x^{(i)}, x^{(j)}\right) = e^{-\sum_{k=1}^{n_{dv}} \theta_k \left|x_k^{(i)} - x_k^{(j)}\right|^{p_k}} \tag{4}$$

where n_{dv} is the number of design variables, and θ_k, p_k are spatially related parameters. The covariance matrix based on Kriging model is then

$$\mathrm{cov}[\mathbf{Y}(\mathbf{X}), \mathbf{Y}(\mathbf{X})] = \sigma^2 \boldsymbol{\Psi}(\mathbf{X}, \mathbf{X}) \tag{5}$$

Similarly, several sub-blocks of the covariance matrix based on the Co-Kriging model can be given as follows:

$$\begin{aligned} \mathrm{cov}[\mathbf{Y}_c(\mathbf{X}_c), \mathbf{Y}_c(\mathbf{X}_c)] &= \mathrm{cov}[Z_c(\mathbf{X}_c), Z_c(\mathbf{X}_c)] \\ &= \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_c) \end{aligned} \tag{6}$$

$$\begin{aligned} \mathrm{cov}[\mathbf{Y}_e(\mathbf{X}_e), \mathbf{Y}_c(\mathbf{X}_c)] &= \mathrm{cov}[\rho Z_c(\mathbf{X}_e) + Z_d(\mathbf{X}_e), Z_c(\mathbf{X}_c)] \\ &= \rho \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_e) \end{aligned} \tag{7}$$

$$\begin{aligned} \mathrm{cov}[\mathbf{Y}_e(\mathbf{X}_e), \mathbf{Y}_e(\mathbf{X}_e)] &= \mathrm{cov}[\rho Z_c(\mathbf{X}_e) + Z_d(\mathbf{X}_e), \rho Z_c(\mathbf{X}_e) + Z_d(\mathbf{X}_e)] \\ &= \rho^2 \mathrm{cov}[Z_c(\mathbf{X}_e), Z_c(\mathbf{X}_e)] + \mathrm{cov}[Z_d(\mathbf{X}_e), Z_d(\mathbf{X}_e)] \\ &= \rho^2 \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{X}_e, \mathbf{X}_e) + \sigma_d^2 \boldsymbol{\Psi}_d(\mathbf{X}_e, \mathbf{X}_e) \end{aligned} \tag{8}$$

where $\boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_e)$ represents the matrix constituted by the correlation coefficient of $\mathbf{X}_c$ and $\mathbf{X}_e$ in the low-fidelity Kriging model. Then, the overall covariance matrix of the Co-Kriging surrogate model is expressed as follows:

$$\mathbf{C} = \begin{bmatrix} \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_c) & \rho \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_e) \\ \rho \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_e) & \rho^2 \sigma_c^2 \boldsymbol{\Psi}_c(\mathbf{X}_e, \mathbf{X}_e) + \sigma_d^2 \boldsymbol{\Psi}_d(\mathbf{X}_e, \mathbf{X}_e) \end{bmatrix} \tag{9}$$

According to the above formula, the Co-Kriging model has two SCFs, so it has more relevant parameters than the single-fidelity Kriging model, i.e., $\mu_c, \sigma_c^2, \mu_d, \sigma_d^2, \boldsymbol{\theta}_c, \boldsymbol{\theta}_d, \mathbf{p}_c, \mathbf{p}_d, \rho$, which need to be estimated.

Since the low-fidelity sample data is independent from the high-fidelity sample data, the log-maximum likelihood function shown as Eq. (10) (ignoring the constant term) can be used to obtain the parameters in low-fidelity Kriging model:

$$MLE_c = -\frac{N_c}{2} \ln\left(\sigma_c^2\right) - \frac{1}{2} \ln|\boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_c)| - \frac{(\mathbf{y}_c - \mathbf{1}\mu_c)^T \boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_c)^{-1} (\mathbf{y}_c - \mathbf{1}\mu_c)}{2\sigma_c^2} \tag{10}$$

Estimation of μ_c and σ_c^2 can be obtained by taking the derivatives of MLE_c with respect to μ_c and σ_c^2, and set the derivatives to 0, we have:

$$\widehat{\mu}_c = \frac{\mathbf{1}^T \boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_c)^{-1} \mathbf{y}_c}{\mathbf{1}^T \boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_c)^{-1} \mathbf{1}} \tag{11}$$

$$\widehat{\sigma}_c^2 = \frac{(\mathbf{y}_c - \mathbf{1}\widehat{\mu}_c)^T \boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_c)^{-1} (\mathbf{y}_c - \mathbf{1}\widehat{\mu}_c)}{N_c} \tag{12}$$

Substitute the above two expressions into the maximum likelihood function Eq. (10), and the following equation is obtained:

$$MLE_c = -\frac{N_c}{2} \ln\left(\widehat{\sigma}_c^2\right) - \frac{1}{2} \ln|\boldsymbol{\Psi}_c(\mathbf{X}_c, \mathbf{X}_c)| \tag{13}$$

At this point, since $\boldsymbol{\theta}_c$ and $\mathbf{p}_c$ cannot be obtained theoretically, they can only be got by numerical optimization (maximization) of the above likelihood function, single-objective Genetic Algorithm (GA) (Deb and Agrawal, 1995) is applied in this paper.

The above steps are just the construction process of the single-fidelity Kriging surrogate model, which is used to build the approximation of Z_c with the low-fidelity sample data in the Co-Kriging model.

To estimate the rest parameters $\mu_d, \sigma_d^2, \boldsymbol{\theta}_d, \mathbf{p}_d, \rho$, first define

$$\mathbf{d} = \mathbf{y}_e - \rho \mathbf{y}_c(\mathbf{X}_e) \tag{14}$$

Here, the low-fidelity function values $\mathbf{y}_c$ at high-fidelity sample points $\mathbf{X}_e$ are needed. If the high-fidelity sample set $\mathbf{X}_e$ is a subset of the low-fidelity sample set $\mathbf{X}_c$, their $\mathbf{y}_c(\mathbf{X}_e)$ are the real values of the low-fidelity function. Otherwise, if some points in $\mathbf{X}_e$ are unavailable in $\mathbf{X}_c$, their $\mathbf{y}_c(\mathbf{X}_e)$ have to be obtained by the prediction (interpolation) of the low-fidelity Kriging model, and additional interpolation errors may be introduced.

At this point, the log-maximum likelihood function of the approximate function for high-fidelity samples is:

$$MLE_e = -\frac{N_e}{2}\ln\left(\sigma_d^2\right) - \frac{1}{2}\ln|\mathbf{\Psi}_c(\mathbf{X}_c,\mathbf{X}_c)| - \frac{(\mathbf{d}-\mathbf{1}\mu_d)^T\mathbf{\Psi}_d(\mathbf{X}_e,\mathbf{X}_e)^{-1}(\mathbf{d}-\mathbf{1}\mu_d)}{2\sigma_d^2} \tag{15}$$

The estimation of the μ_d and σ_d^2 can be obtained by taking the derivatives of MLE_e with respect to μ_d and σ_d^2, and set the derivatives to 0:

$$\widehat{\mu}_d = \frac{\mathbf{1}^T\mathbf{\Psi}_d(\mathbf{X}_e,\mathbf{X}_e)^{-1}\mathbf{d}}{\mathbf{1}^T\mathbf{\Psi}_d(\mathbf{X}_e,\mathbf{X}_e)^{-1}\mathbf{1}} \tag{16}$$

$$\widehat{\sigma}_c^2 = \frac{(\mathbf{y}_c-\mathbf{1}\widehat{\mu}_c)^T\mathbf{\Psi}_c(\mathbf{X}_c,\mathbf{X}_c)^{-1}(\mathbf{y}_c-\mathbf{1}\widehat{\mu}_c)}{N_c} \tag{17}$$

Substitute the above two expressions into the maximum likelihood function Eq. (15), and the following equation is obtained:

$$MLE_c = -\frac{N_c}{2}\ln\left(\widehat{\sigma}_c^2\right) - \frac{1}{2}\ln|\mathbf{\Psi}_c(\mathbf{X}_c,\mathbf{X}_c)| \tag{18}$$

Likewise, estimation of $\boldsymbol{\theta}_d$, $\mathbf{p}_d$, and ρ can also be obtained by single-objective Genetic Algorithm (GA) in this paper.

In order to get the prediction value of the multi-fidelity Co-Kriging surrogate model, we adopt a method similar to the single-fidelity Kriging model. Therefore, we use the new sample point to be predicted and the existing sample data to seek the most likely value of the function at the to-be-predicted point from the perspective of probability, that is, to maximum the likelihood function $MLE\ \widehat{y}_e(\mathbf{x})$, by keeping the existing estimation surrogate model parameters unchanged. At this time, new sample set is $\mathbf{X} = [\mathbf{X}_c^T, \mathbf{X}_e^T, \mathbf{x}^T]^T$, the function value set is $\mathbf{y} = [\mathbf{y}_c^T, \mathbf{y}_e^T, \widehat{\mathbf{y}}_e^T(\mathbf{x})]^T$, and the spatial correlation matrix of the sample set $\widetilde{\mathbf{C}}$ can be written as

$$\widetilde{\mathbf{C}} = \begin{bmatrix} \widehat{\sigma}_c^2\mathbf{\Psi}_c(\mathbf{X}_c,\mathbf{X}_c) & \rho\widehat{\sigma}_c^2\mathbf{\Psi}_c(\mathbf{X}_c,\mathbf{X}_e) & \rho\widehat{\sigma}_c^2\mathbf{\Psi}_c(\mathbf{X}_c,\mathbf{x}) \\ \rho\widehat{\sigma}_c^2\mathbf{\Psi}_c(\mathbf{X}_c,\mathbf{X}_e) & \rho^2\widehat{\sigma}_c^2\mathbf{\Psi}_c(\mathbf{X}_e,\mathbf{X}_e)+\widehat{\sigma}_d^2\mathbf{\Psi}_d(\mathbf{X}_e,\mathbf{X}_e) & \left(\rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2\right)\mathbf{\Psi}_d(\mathbf{X}_e,\mathbf{x}) \\ \rho\widehat{\sigma}_c^2\mathbf{\Psi}_c(\mathbf{X}_c,\mathbf{x})^T & \left(\rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2\right)\mathbf{\Psi}_d(\mathbf{X}_e,\mathbf{x})^T & \rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2 \end{bmatrix} \tag{19}$$

Define vector $\mathbf{c}$ to be the covariance of $\mathbf{X}$ and $\mathbf{x}$, then the correlation matrix $\widetilde{\mathbf{C}}$ can be rewritten as

$$\widetilde{\mathbf{C}} = \begin{bmatrix} \mathbf{C} & \mathbf{c} \\ \mathbf{c}^T & \rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2 \end{bmatrix} \tag{20}$$

Similar to the single-fidelity Kriging surrogate model, in order to maximize the maximum likelihood function $MLE\ \widehat{y}_e(\mathbf{x})$, the following item needs to be maximized:

$$\max\ f = -\frac{1}{2}(\widetilde{\mathbf{y}}-\mathbf{1}\mu)^T\widetilde{\mathbf{C}}^{-1}(\widetilde{\mathbf{y}}-\mathbf{1}\mu) \tag{21}$$

Substitute Eq. (20) into Eq. (21), we have

$$\max\ f = -\frac{1}{2}\begin{pmatrix} \mathbf{y}-\mathbf{1}\widehat{\mu} \\ \widehat{y}_e(\mathbf{x})-\widehat{\mu} \end{pmatrix}^T \begin{bmatrix} \mathbf{C} & \mathbf{c} \\ \mathbf{c}^T & \rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2 \end{bmatrix}^{-1} \begin{pmatrix} \mathbf{y}-\mathbf{1}\widehat{\mu} \\ \widehat{y}_e(\mathbf{x})-\widehat{\mu} \end{pmatrix} \tag{22}$$

According to the matrix theory, the inverse of the correlation matrix $\widetilde{\mathbf{C}}$ can be obtained, i.e.

$$\widetilde{\mathbf{C}}^{-1} = \begin{bmatrix} \mathbf{C}^{-1}+\mathbf{C}^{-1}\mathbf{c}\left(\rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2-\mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}\right)^{-1}\mathbf{c}^T\mathbf{C}^{-1} & -\mathbf{C}^{-1}\mathbf{c}\left(\rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2-\mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}\right)^{-1} \\ -\left(\rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2-\mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}\right)^{-1}\mathbf{c}^T\mathbf{C}^{-1} & \left(\rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2-\mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}\right)^{-1} \end{bmatrix} \tag{23}$$

In fact, when substituting Eq. (23) into Eq. (22), one term of f has nothing to do with $\widehat{y}_e(\mathbf{x})$: $-\frac{1}{2}(\mathbf{y}-\mathbf{1}\widehat{\mu})^T[\mathbf{C}^{-1}+\mathbf{C}^{-1}\mathbf{c}(\rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2-\mathbf{c}^T\mathbf{C}^{-1}\mathbf{c})^{-1}\mathbf{c}^T\mathbf{C}^{-1}](\mathbf{y}-\mathbf{1}\widehat{\mu})$. Therefore, if we take the derivative of f with respect to $\widehat{y}_e(x)$, this term can be ignored. Set the derivative to 0, we have

$$\frac{-1}{\rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2-\mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}}(\widehat{\mathbf{y}}_e(\mathbf{x})-\widehat{\mu}) + \frac{\mathbf{c}^T\widetilde{\mathbf{C}}^{-1}(\mathbf{y}-\mathbf{1}\widehat{\mu})}{\rho^2\widehat{\sigma}_c^2+\widehat{\sigma}_d^2-\mathbf{c}^T\mathbf{C}^{-1}\mathbf{c}} = 0 \tag{24}$$

The best predicted value $\widehat{y}_e(\mathbf{x})$ of the to-be-predicted sample point, which is just an arbitrary point in the variables space, is given below:

$$\widehat{y}_e(\mathbf{x}) = \widehat{\mu} + \mathbf{c}^T\mathbf{C}^{-1}(\mathbf{y}-\mathbf{1}\widehat{\mu}) \tag{25}$$

It can be proved that, similar to the single-fidelity Kriging surrogate model, the above Co-Kriging surrogate model is an interpolation of the existing high-fidelity sample data. However, it will make regression to the low-fidelity samples without high-fidelity function values in a sense, that is, make regression to the low-fidelity sample data X_c except for X_e.

Similar to the single-fidelity Kriging surrogate model, the estimation error in probability of multi-fidelity Co-Kriging model is approximately

$$\widehat{s}^2(\mathbf{x}) \approx \rho^2\widehat{\sigma}_c^2 + \widehat{\sigma}_d^2 - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{c} + \frac{1-\mathbf{1}^T\mathbf{C}^{-1}\mathbf{c}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}} \tag{26}$$

3. Numerical validations of multi-fidelity Co-Kriging surrogate model

In order to show the high efficiency and high fidelity of the Co-Kriging surrogate model used in this paper in representing complex functions, 4 numerical tests are carried out on typical mathematical functions of different types. Before the surrogate-based-optimization, the surrogate model should be constructed with high accuracy. In fact, if the true function or the high-fidelity surrogate model is given, GA can have good performance in finding global minimum.

1) Case-1

Firstly, a one-variable function with only one global minimum is given (Forrester and Keane, 2009). Its high-fidelity and low-fidelity functions are shown below respectively

$$f_e(x) = (6x-2)^2\sin(6x-2)^2 \tag{27}$$

$$f_c(x) = \frac{1}{2}(6x-2)^2\sin(6x-2)^2 + 10(x-1) \tag{28}$$

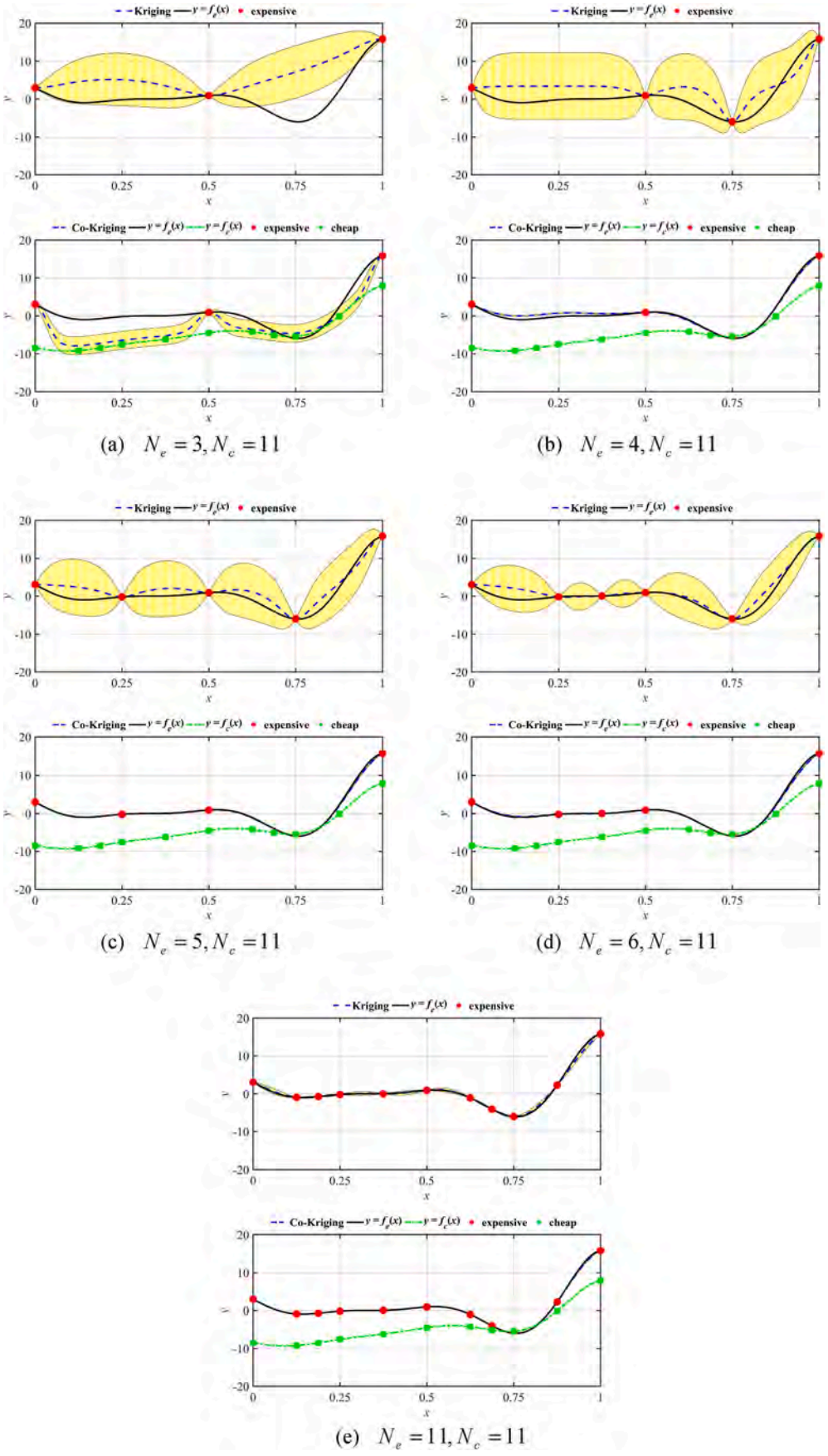

Fig. 1. Influence of different high-fidelity sample numbers on the errors of two surrogate models.

It can be seen that the low-fidelity function is the sum of the compressed high-fidelity function and a linear function, which can actually be analogous to the random error of numerical simulation results.

Through the Sobol (1979) sampling method, 11 points are selected in the range of $[0,1]$, and their low-fidelity function values are calculated. At the same time, 3, 4, 5 and 6 points are selected from the 11 points according to the maximum-minimum distance criterion, and their high-fidelity function values are calculated. Single-fidelity Kriging surrogate model by 3, 4, 5 and 6 high-fidelity samples, and multi-fidelity Co-Kriging surrogate model by 3, 4, 5 and 6 high-fidelity samples together with 11 low-fidelity function samples are constructed respectively. The approximation results of the surrogate models to high-fidelity real function are shown in Fig. 1 below. Among them, the blue dotted line represents the surrogate models, the black solid line represents the high-fidelity function, the green dotted line represents the low-fidelity function, and the half height of the yellow region at each point represents the root mean square error of the surrogate model in probability.

Seen from Fig. 1, in the case of 11 low-fidelity samples, when there are 3 high-fidelity samples, due to the limited high-fidelity function information, the Co-Kriging model basically depends on the low-fidelity samples. Specifically, except that the prediction value of the Co-Kriging surrogate model at the high-fidelity sample points is completely consistent with the real high-fidelity function value, the rest of the Co-Kriging model has a trend towards low-fidelity function. Furthermore, the Kriging model also has a big error compared with high-fidelity function at this time. However, when the sample number of the high-fidelity function is 4, the Co-Kriging model is already very close to the high-fidelity function, but there is a small error. At this time, the overall error of the Kriging model is reduced, but the fidelity of Kriging model is still very poor compared with the Co-Kriging model. When the sample number of the high-fidelity function is 5, the error between the Co-Kriging model and the high-fidelity function is quite small, so it can be considered that the co-Kriging model can forecast the real high-fidelity function very well. When the number of high-fidelity function samples is larger, the overall error of Kriging model becomes smaller and smaller. Finally, when the sample number of the high-fidelity function is 11, both the Kriging and Co-Kriging models are nearly consistent with the high-fidelity function. However, if the Co-Kriging model is constructed by 11 high- and low-fidelity samples, the high efficiency of Co-Kriging model cannot be reflected. Considering the high-fidelity sample actually represents CFD-based numerical simulation result with expensive costs, in order to reduce the computational costs, for the purpose of this example, 5 samples with high fidelity is enough, together with 11 low-fidelity samples, to get high-fidelity Co-Kriging surrogate model.

It can be predicted that for a large number of low-fidelity samples, a certain number of high-fidelity samples are required to make the Co-Kriging surrogate model have a high fidelity. However, too many high-fidelity samples will reduce the computational efficiency, which will weaken the advantages of Co-Kriging model compared with Kriging model.

2) Case-2

Slight difference from the previous case is that, although the high-fidelity function in this case is also a function of one variable with only one global minimum, there are many local minimum values, which can be regarded as a "multimodal" function (De et al., 2015). The high-fidelity and low-fidelity functions are shown below respectively

$$f_e(x) = 3 + \cos\left(5x^2\right) + 0.2\sin(40x) \tag{29}$$

$$f_c(x) = 0.5 + \cos\left(5x^2\right) + 0.2\sin(40x) + 1.5\sin(3x) \tag{30}$$

It can be seen that the low-fidelity function is the sum of shifted high-fidelity function, and a sinusoidal perturbation, which makes the mode of the low-fidelity function more complex.

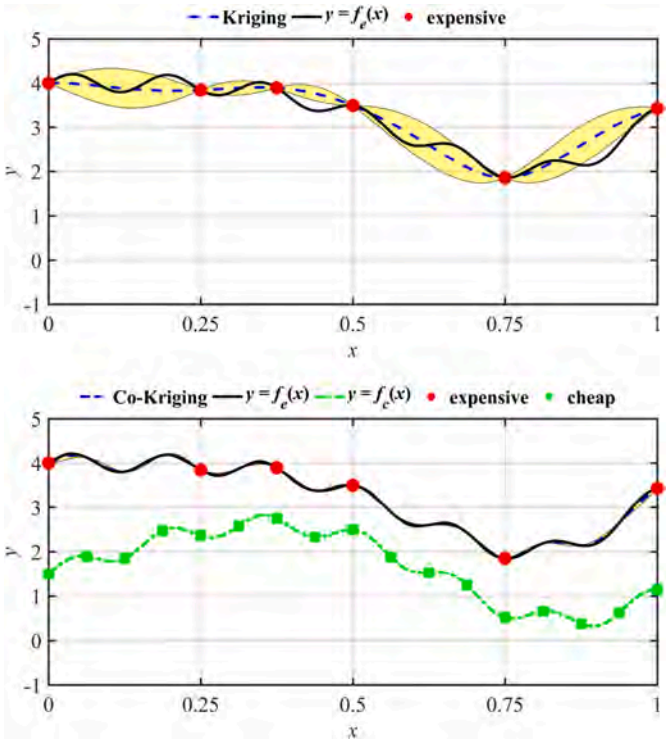

Fig. 2. Comparison of the errors of two surrogate models.

For this case, similar to the above process, after determining the sample number of low-fidelity function, the sample number of high-fidelity function is gradually increased to compare the accuracy of the obtained Co-Kriging surrogate model and the conventional Kriging surrogate model. As the high-fidelity function in this case is more complex, the sample number of the low-fidelity function selected by the uniform design method is increased, that is to say, 17 points are selected in the range of $[0,1]$. Considering accuracy and efficiency, the Co-Kriging model is constructed by using 6 high-fidelity function samples and 17 low-fidelity function samples. The comparison of the approximate function obtained by the two surrogate models and the real high-fidelity function is shown in Fig. 2 below.

It is obviously that, the Kriging model with only 6 high-fidelity function samples has a relatively big error compared with the real high-fidelity function, and there are great differences in the capture of functional modes, making it difficult to globally capture the change trend of this function. Therefore, for similar functions, the constructed Kriging surrogate model with limited sample data may skip some of the local and global minima of the real high-fidelity function. However, the Co-Kriging surrogate model can capture the full information of the high-fidelity function much better without using too many samples.

As mentioned above, it is well known that EI is commonly used in the adaptive sampling. On the one hand, the adaptive sampling can purposefully increase sample points, which may lead to a better performance of the surrogate model with smaller prediction error; on the other hand, the number of the initial points and the new points are hard to determine, and the whole optimization process may take longer, but the one-shot sampling for the high- and low-fidelity samples may lead to their numerical evaluation being carried out almost simultaneously, especially when the high-fidelity data is obtained by CFD evaluations on HPC system, which has been widely used for the CFD-based hull form optimization, compared with the adding-one-point adaptive sampling method like EI. Certainly, if multiple samples are added per cycle (Diez et al., 2019), the whole optimization procedure can be shortened.

Take this case as an example. If we use EI to add additional points to

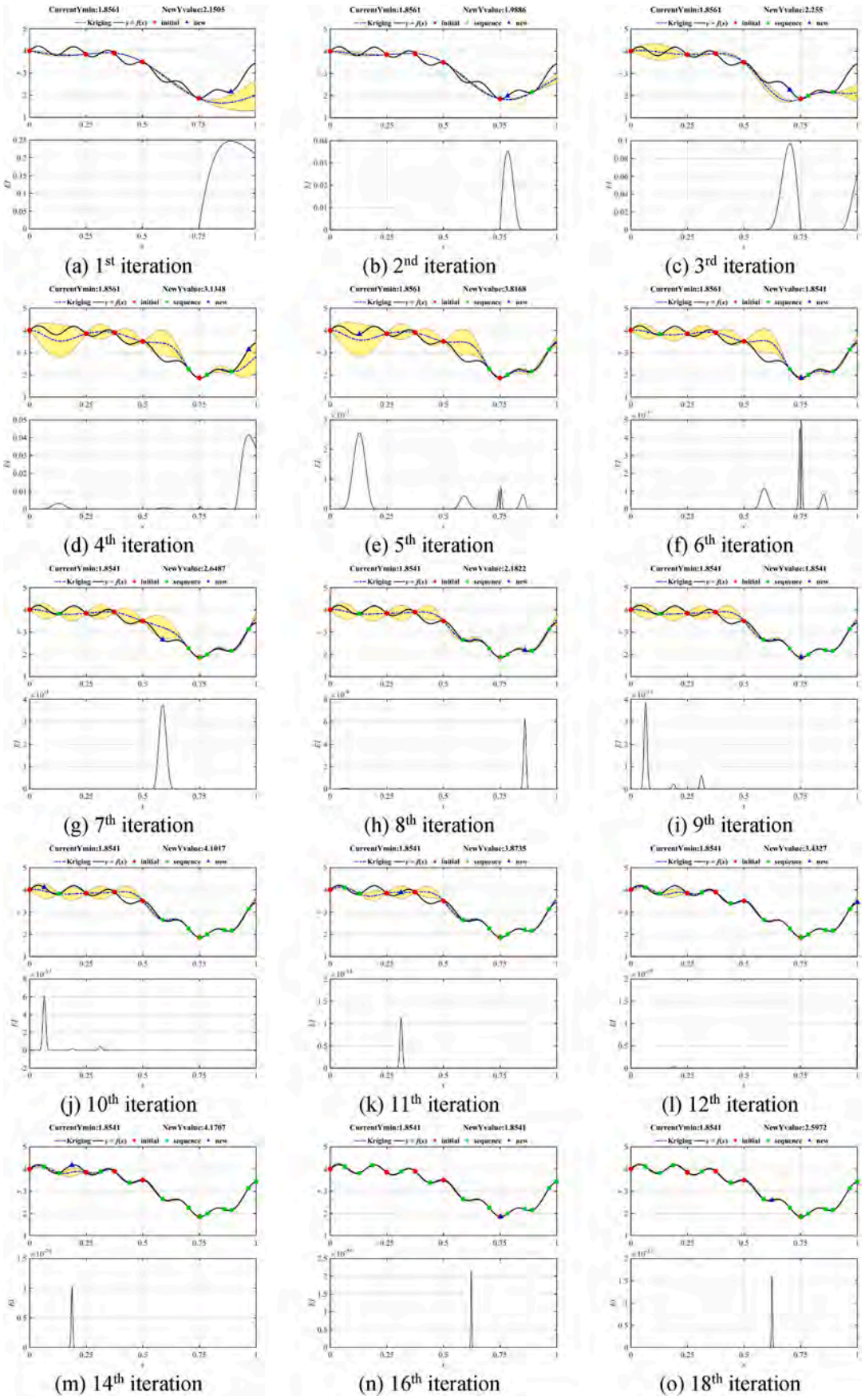

Fig. 3. Updating iterations of the Kriging surrogate model with EI-based adaptive sampling.

the single-high-fidelity Kriging model of the $f_e(x)$, the updating iterations are shown in Fig. 3. It can be seen that if EI is applied to add additional high-fidelity sample points, more high-fidelity data should be obtained, which is less efficient than the Co-Kriging model presented in this paper. It should be noted that, although after 4 iterations the global optimum can be already reached shown from Fig. 3, the overall changing trend of the Kriging model compared with the real function still has a significant difference. Therefore, the EI loop lasts 18 iterations. Therefore, for functions with higher dimension, if we cannot clearly see the changing trend like this case, we may only judge by the EI loop criterion, which may take a relatively long period of time to obtain a high-fidelity surrogate or a reliable optimum.

3) Case-3

The function of this case contains two independent variables (Cai et al., 2017), and the high-fidelity and low-fidelity functions are given respectively

$$f_e(\vec{\mathbf{x}}) = 4(4x_1-2)^2 - \frac{11}{5}(4x_1-2)^4 + \frac{1}{3}(4x_1-2)^6 + (4x_1-2)(4x_2-2) - 4(4x_2-2)^2 + 4(4x_2-2)^4 \tag{31}$$

$$f_c(\vec{\mathbf{x}}) = f_e(0.7\vec{\mathbf{x}}) + x_1x_2 - 65 \tag{32}$$

It can be seen that the low-fidelity function amplifies the high-fidelity function in the two directions of the independent variables, making some information of the high-fidelity function missing to some extent, and adds certain disturbance and overall shifting.

As the high-fidelity function in this case is a two-variable function, the sample number of the low-fidelity function selected by the Sobol sampling method is increased to be 60. Taking accuracy and efficiency into consideration, the Co-Kriging model is constructed by using 30 high-fidelity function samples and 60 low-fidelity function samples. The comparison of the approximate function obtained by the two surrogate models and the real high-fidelity function is shown in Fig. 4 below.

It can be seen that the Kriging surrogate model using only 30 function samples of high fidelity still has some errors compared with real high-fidelity function, especially in the "middle" region of the two-dimensional space with small function values. However, after adding 60 samples of low-fidelity function, the Co-Kriging model can capture almost all the information of real high-fidelity function well. It can be predicted that if only 30 samples of high-fidelity function are used, it is far from enough from the perspective of the surrogate model accuracy, but the continuous addition of high-fidelity sample points will increase a lot of calculation amount.

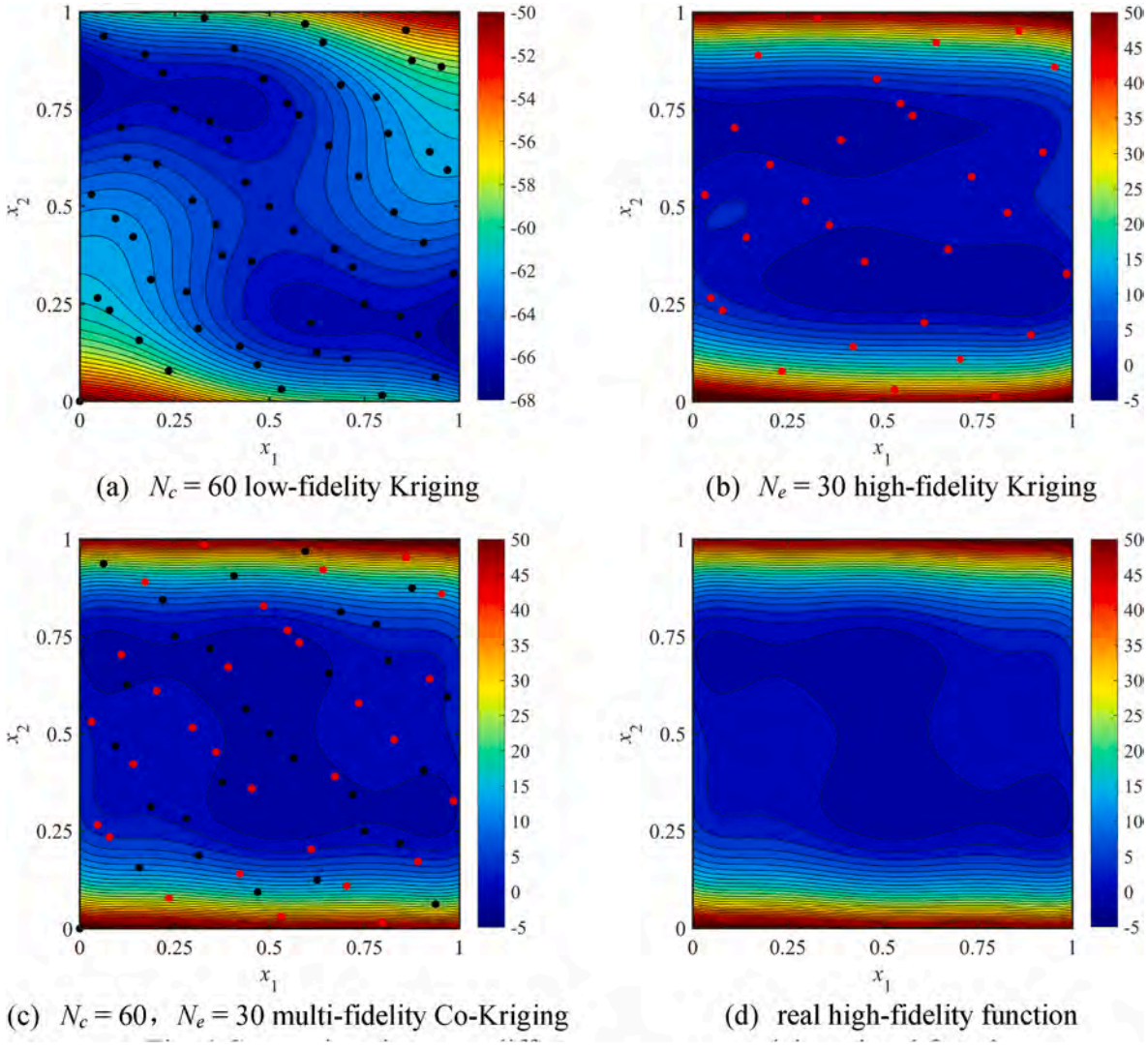

Fig. 4. Comparison between different surrogate models and real function.

4) Case-4

The function of this case contains five independent variables (Cai et al., 2017), and the high-fidelity and low-fidelity functions are given respectively

$$f_e(\vec{\mathbf{x}}) = \sum_{i=1}^{4}\left\{\left[(6x_{i+1}-3)^2-(6x_i-3)\right]^2+(6x_i-4)^2\right\} \tag{33}$$

$$f_c(\vec{\mathbf{x}}) = \sum_{i=1}^{4}\left[0.9(6x_{i+1}-3)^4+2.2(6x_i-3)^2-1.8(6x_i-3)(6x_{i+1}-3)^2 + 0.5\right] \tag{34}$$

As the high-fidelity function in this case is a five-variable function, the sample number of the low-fidelity function selected by the Sobol sampling method is increased to be 100. Taking accuracy and efficiency into consideration, the Co-Kriging model is constructed by using 50 high-fidelity function samples and 100 low-fidelity function samples. After the surrogate models are constructed, for display convenience, let $x_2 = 1$, $x_3 = 1$, $x_5 = 0.5$. The comparison of the approximate function obtained by the two surrogate models and the real high-fidelity function is shown in Fig. 5.

It can be seen that the Kriging surrogate model using only 50 high-

Table 1
Summary of the optimal results for four test cases.

/	Real (High-fidelity) function		Kriging model		Co-Kriging model	
	$\mathbf{x}_{opt}$	f_{min}	$\mathbf{x}_{opt}$	f_{min}	$\mathbf{x}_{opt}$	f_{min}
Case 1	0.757	−6.021	0.750	−5.993	0.757	−6.017
Case 2	0.753	1.854	0.752	1.856	0.753	1.855
Case 3	[0.522, 0.322]	−1.032	[0, 0.705]	−1.276	[0.523, 0.322]	−1.031
Case 4	[0.667, 0.667, 0.667, 0.667, 0.333]	0	[0.447, 0.577, 0.547, 0.458, 0.573]	−35.824	[0.667, 0.667, 0.666, 0.663, 0.335]	0.002

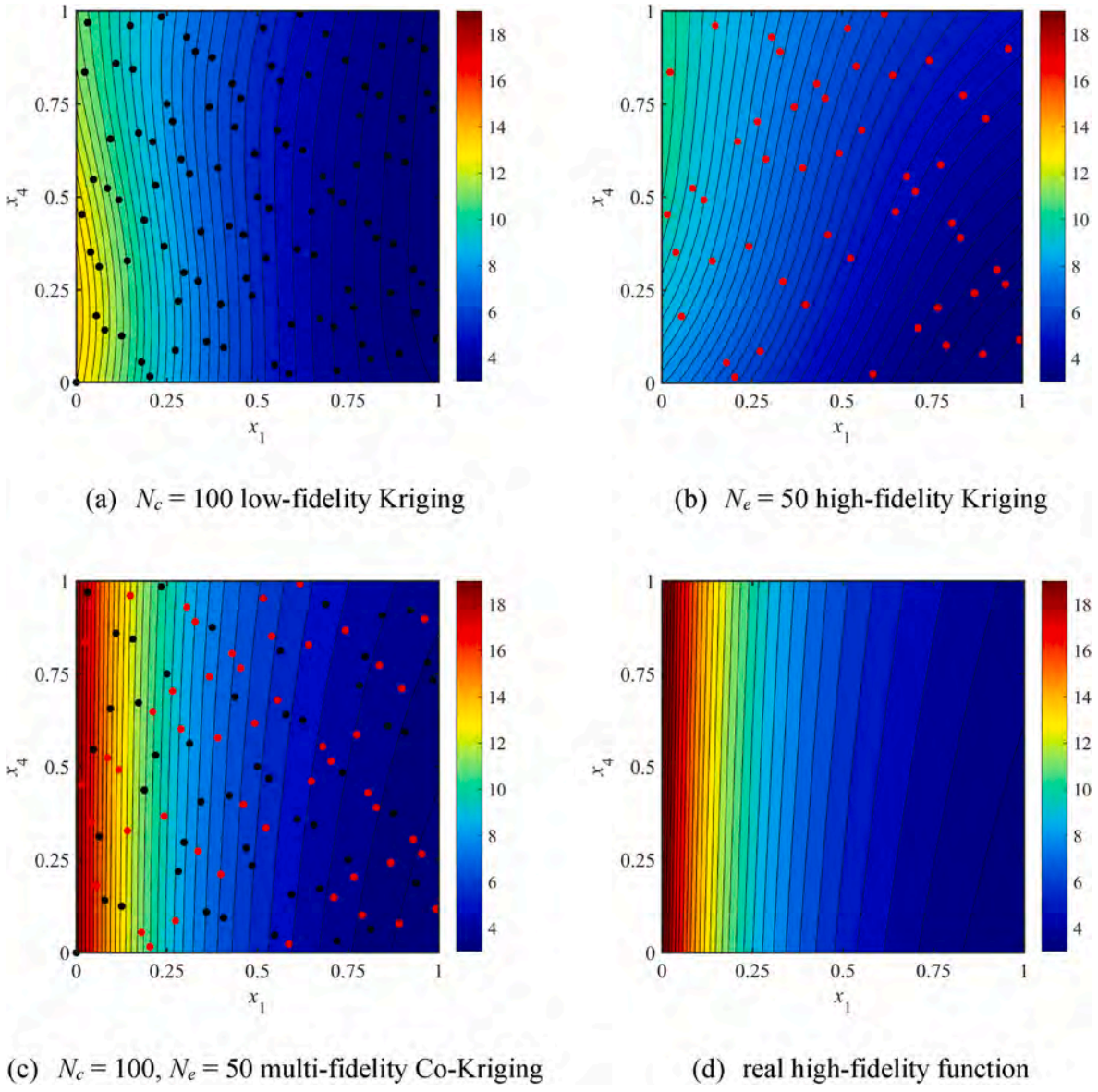

Fig. 5. Comparison between different surrogate models and real function.

fidelity samples has some errors compared with real high-fidelity function. However, after adding 100 low-fidelity samples, the Co-Kriging model can capture almost all the information of real high-fidelity function well. However, the addition of a series of low-fidelity sample points is acceptable if the calculation cost of low-fidelity samples can be relatively small.

From Table 1, we can see that, based on Co-Kriging model, the optimum point and value is much closer to the real optimum in the four cases, especially for the high-dimensional cases such as Case 3 and Case 4. However, since the Kriging model has a relatively large error by using a relatively small number of samples, the real optimum is hard to find.

4. Total drag optimization case of DTMB-5415 hull

4.1. Basic information of the mother ship

DTMB-5415 ship was originally designed for a warship by David Taylor Model Basin in the 1980s. With a sonar dome on the bow and a transom stern, DTMB-5415 has become an internationally recognized standard ship model, and abundant model test results (Olivieri et al., 2001) can be used for the verification of numerical calculation. The ship model being optimized in this paper has the INSEAN model scale of 5.72 m long, and the full scale model of it has 142 m long. This ship belongs to medium- or high-speed ships. Its three-dimensional model is shown in Fig. 6, and main parameters in model scale are shown in Table 2.

Due to the relatively high ship speed and thin overall shape, the free surface wave elevation in calm water is relatively obvious, and the breaking bow wave can even appear at a high speed, in other words, the ratio of wave-making drag to total drag is relatively high. Therefore, it is very important to reduce the wave height of the free surface. On the one hand, it can reduce the wave-making drag; on the other hand, free surface wave can bring about hydrodynamic noise and long wake after the ship stern, so reduce the free surface wave can also enhance its stealth. In order to reduce the bow and stern waves as much as possible, we consider local deformations of the bow part and the global deformation of the entire hull of DTMB-5415 ship to optimize its total calm-water drag at the design speed of Froude number 0.28.

4.2. Definition of optimization problem and hydrodynamic evaluation methods

For the numerical simulation of calm-water drag, in general, there are efficient prediction methods based on potential flow theory, such as in-house solver NMShip-SJTU (Liu et al., 2019) based on the Neumann-Michell (NM) theory (Noblesse et al., 2013), to calculate wave-making drag quickly, and high-fidelity prediction methods based on viscous flow theory, such as in-house solver naoe-FOAM-SJTU (Shen et al., 2015) based on RANS equations.

For the resistance performance evaluation of DTMB-5415 ship, the wave-making drag coefficient can be obtained by NM theory, and 1957 ITTC plate frictional drag coefficient formula can be used to approximate the frictional drag coefficient of the hull, and the total drag based on potential flow theory of the ship model can be obtained according to the ship length, namely

$$R_{t,PF} = \left(C_f + C_w\right)\frac{1}{2}\rho U^2 S \tag{35}$$

If the viscous-flow-based solver naoe-FOAM-SJTU is used, the pressure drag R_p and frictional drag R_f are obtained by integrating the dynamic pressure distribution on the hull surface, and the sum of the two drag components is regarded as the total drag of the ship model based on the viscous flow theory, namely

$$R_{t,VF} = R_p + R_f \tag{36}$$

The above two methods have their own advantages, and Co-Kriging surrogate model can combine the advantages of the two methods. Specifically, more potential flow solution results are regarded as low-fidelity sample data, which can be used to capture the overall relationship between the total drag and the hull form optimization design variables, and fewer viscous flow solution results as high-fidelity sample data. Firstly, two solvers are used to evaluate the calm-water total drag of the initial DTMB-5415 ship hull to analyze and compare the similarities and differences between the calculation results, and the feasibility of using two fidelity results to establish Co-Kriging model. For the potential flow solver, the hull is fixed at the waterline with 0 sinkage and trim, while for the viscous flow solver, the hull has 2 DoF (sinkage and trim).

When NMShip-SJTU is used to calculate the total drag, only the mesh of hull surface and free surface are needed. As shown in Fig. 7 and Fig. 8, the calculation mesh of the hull surface and the free surface around the DTMB-5415 hull are given respectively.

Fig. 9 shows the calculation region size of naoe-FOAM-SJTU when evaluating the calm-water drag, and the summary of the boundary conditions is given in Table 3. For the calm-water total drag evaluation by naoe-FOAM-SJTU solver, the RANS equations are used to address the incompressible viscous flow, and k-ω SST turbulence model is applied. The free surface is captured by Volume of fluid (VOF) method with artificial bounded compression techniques. The RANS and VOF transport equations are discretized by finite volume method (FVM). The PIMPLE algorithm is used to solve the coupled equation of velocity and pressure. The convection terms are approximated by a second-order TVD limited linear scheme, and the diffusion terms are approximated by a second-order central difference scheme. The unstructured meshes are generated by snappyHexMesh, an automatic mesh-generation utility provided by OpenFOAM. This utility generates mesh on an original Cartesian background mesh, splitting hexahedral cells into split-hex

Table 2
Main particulars of DTMB-5415.

Parameter	Symbol and unit	Value
length between perpendiculars	L_{pp} (m)	5.72
Breadth	B (m)	0.768
Draught	T (m)	0.248
molded depth	D (m)	0.772
drainage volume	∇ (m^3)	0.552
wet surface area	S (m^2)	4.861

(a) main view (b) global view

Fig. 6. Geometry model of DTMB-5415.

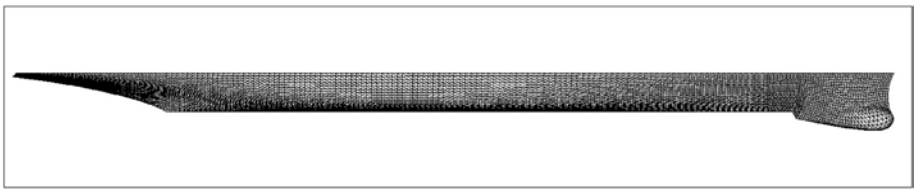

Fig. 7. Mesh of hull surface.

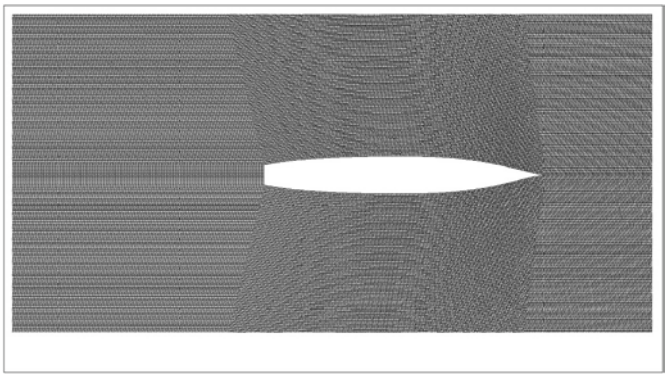

Fig. 8. Free-surface mesh around the hull.

cells. Most of these cells distribute around the free surface and the hull after mesh refinement. Near-wall grid spacing is designed to meet the requirement of turbulence model using wall functions, where the y^+ is around 30. Furthermore, the number of the boundary layers is 5. The global refinement is shown in Fig. 10, and the local refinement near the hull is shown in Fig. 11.

It can be seen from Fig. 12 that locations (phases) of wave systems (including bow wave, shoulder wave and stern wave) obtained by the two solvers are basically the same, but to be honest, specific values of wave peak and trough are different to some extent. For the bow wave and shoulder wave, they are underestimated by potential flow theory compared with the viscous flow theory, which may be due to the fact that the hull has a certain positive trim at the ship bow. Therefore, the bow wave is larger and the wet surface area of the hull is no longer the static wet surface area, resulting in a certain nonlinearity which cannot be considered in the linear potential flow theory. For the stern wave, the damping is very slow in the potential flow theory because the viscosity is not considered, while relatively fast in the viscous flow theory because of the viscosity and numerical dissipation, so it can be seen that the wave elevation behind the hull stern by potential flow solver is larger than that of the viscous flow solver.

According to the NM theory, the non-dimensional wave height along the hull with respect to the non-dimensional ship length comparison results are shown in Fig. 13, showing a similar trend by the two solvers.

It is obvious from Fig. 14 that, for the underwater part of the ship hull, the location and size of high- and low-pressure regions are similar, showing that the pressure distribution and integration has a similar trend. Meanwhile, since only the hull surface under the design waterline is calculated, the dynamic pressure distribution above the design waterline is not presented by potential flow solver, some difference truly occurs near the free surface.

With a mesh configuration similar to the above, the total calm-water drag coefficients for a series of Froude numbers of the ship model can be calculated and compared with the model test results, which are shown in Fig. 15 below. It can be seen that the total drag coefficients evaluated

Table 3
Summary of boundary conditions.

Boundary name	Boundary condition
Inlet	velocity inlet
Atmosphere	no-slip condition
Hull	no-slip wall condition
Symmetry	no-slip condition
Outlet	pressure outlet

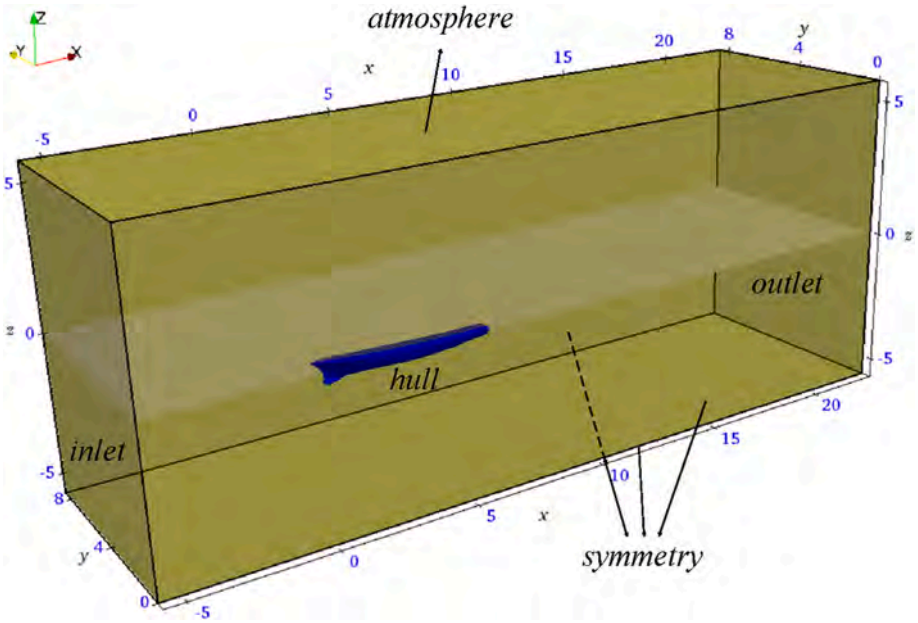

Fig. 9. The size and boundaries of the computational domain.

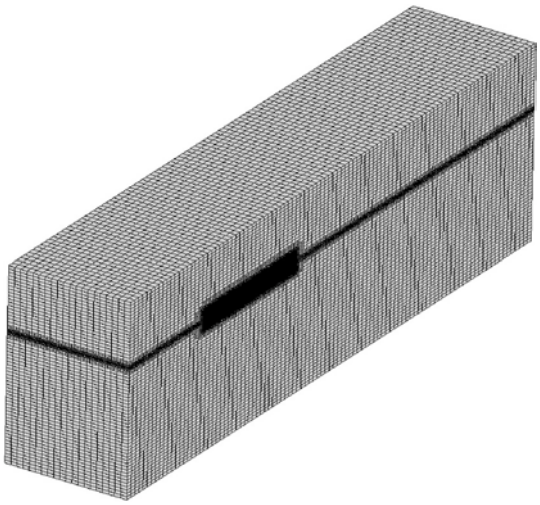

Fig. 10. The refinement settings of the computational domain.

based on the viscous flow theory at several typical speeds are basically consistent with the experimental values within small errors. In addition, the variation rule of total drag coefficient with Fr evaluated based on potential flow theory and the real variation rule by model test of total drag coefficient with Fr are basically consistent on the whole, and there is small deviation between the two especially around Fr = 0.28.

Last but not the least, the V&V of the naoe-FOAM-SJTU are conducted, and the detailed equations in the verification and validation procedure can be seen in Wilson et al. (2001), and the results of the verification and validation parameters are shown in Table 4. Three grid sets distributed around the hull are shown in Fig. 16. The convergence histories of the three grid sets are given in Fig. 17, showing that through resistance evaluation, the mean calm-water total drag can be obtained.

From Table 4, we can obtain that the calculation is monotonic convergent for three grids since the convergence rate R_G is between 0 and 1, and the numerical (grid) uncertainty is 2.50%. Furthermore, since $|E| < U_v$, the combinations of all errors are less than U_v, and the validation is achieved at the level U_v.

4.3. Definition of the design variables

In this paper, the total calm-water total drag of the DTMB-5415 ship at the design speed Fr = 0.28 is optimized. Free Form Deformation (FFD) method (Sederberg and Parry, 1986) is used to carry out the hull form deformation of the ship bow part, and the shifting method (Kim et al., 2011) is used to adjust the Section Area Curve (SAC) of the whole ship, that is, the cross-sections of the front and the aft halves are independently translated along the direction of ship length. The Sobol method is applied to do the design of experiment in the five-dimensional design space. The ranges of design variables are shown in Table 5. It should be noted that all the range values of variables are dimensionless values of the ship model waterline length. Wherein, x_1 corresponds to the movable amount of the red dot along the ship length in Fig. 18 (a), y_1 to the movable amount of the red dot along the ship breadth in Fig. 18 (b), and z_1 to the movable amount of the red dot along the draught in Fig. 18 (c). It should be also noted that the movable control points along the ship breadth are moved symmetrically about the central sheer plane of the hull. These deformation parameters control the length, width and height of the bow sonar dome respectively. In addition, α_{1f} and α_{1a} represent the amplitudes of the modification function of the SAC of the front and aft halves respectively in the shifting method. According to the Sobol method, 50 new sample hulls are generated, and hydrodynamic evaluations are carried out respectively using the two solvers. The computational grid is basically consistent with the grid given above to ensure that the surface grid or body grid in the computational domain corresponding to each new sample hull form is almost the same to reduce model error.

The objective function of this optimization is the calm-water total drag at the sailing Froude number Fr = 0.28, as shown below.

$$\min \ f_{obj} = R_t \tag{37}$$

4.4. Construction of the surrogate models

In order to build the Kriging model or Co-Kriging surrogate model that meets the demand of practical engineering, the maximum high-fidelity sample size is chosen to be 50, and the low-fidelity sample size is also 50, where the high-fidelity sample data is obtained by the viscous flow calculation results, while the low-fidelity sample data is obtained by the potential flow calculation results.

After evaluations of the sample hulls, the single-fidelity Kriging surrogate model and the multi-fidelity Co-Kriging surrogate model of the total drag on the design variables of hull form deformation can be

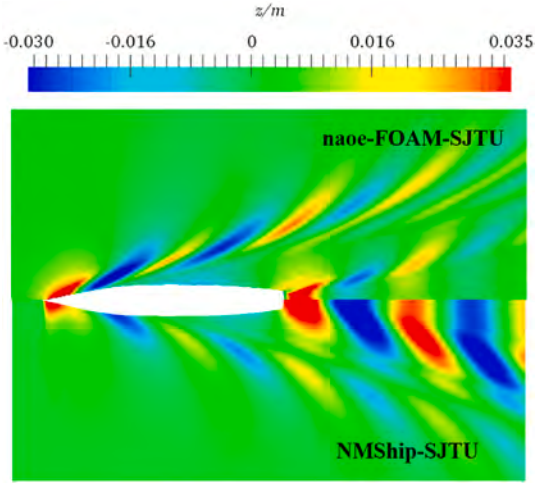

Fig. 12. Comparison of free surface elevation.

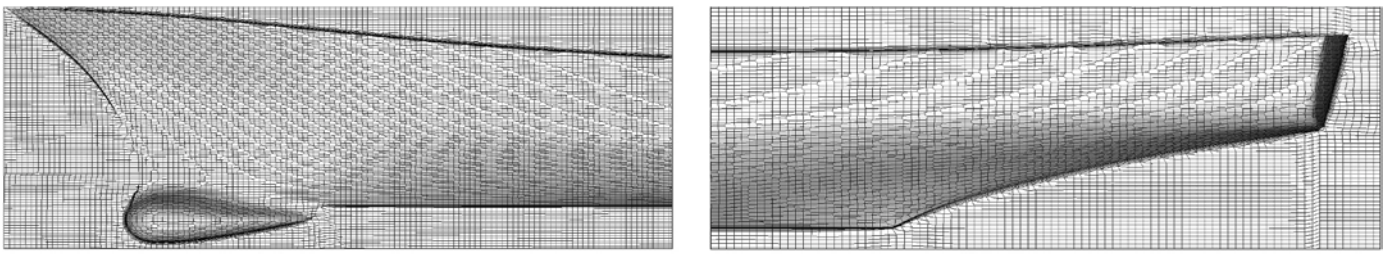

Fig. 11. Local views of meshes around the hull.

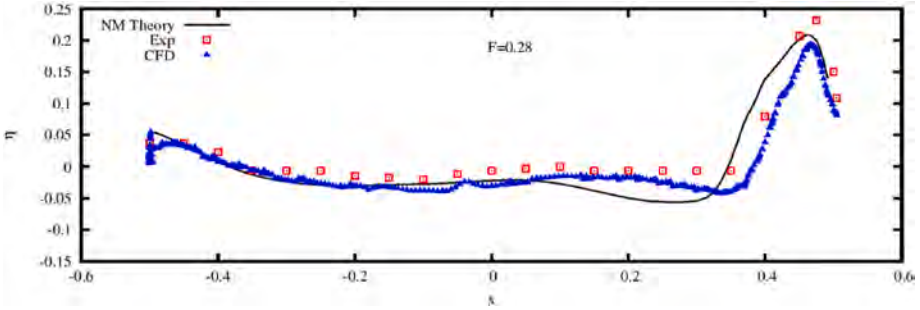

Fig. 13. Comparison of the wave height around hull.

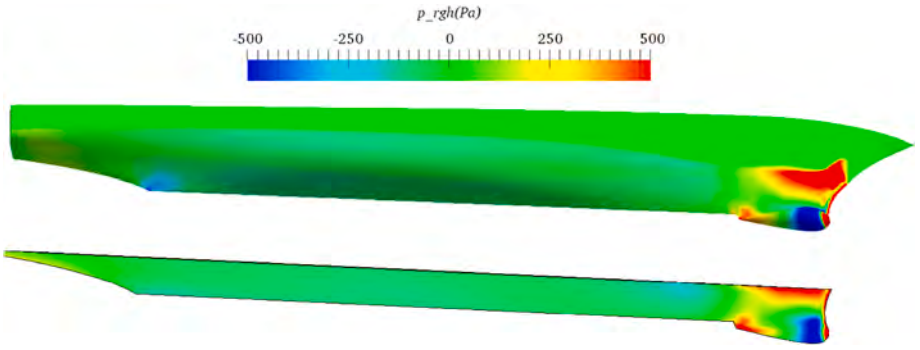

Fig. 14. Comparison of pressure distribution on hull surface (upper: viscous flow; lower: potential flow).

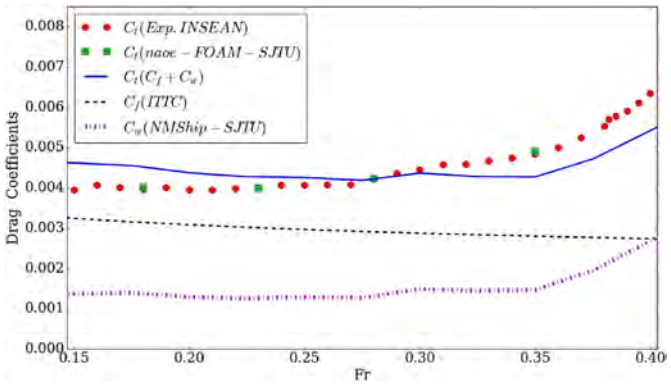

Fig. 15. Comparisons of drag coefficients at different Fr

Table 4
Verification and validation parameters for naoe-FOAM-SJTU calculation.

Grid	3	2	1
Background grid dimensions ($N_x \times N_y \times N_z$)	70 × 14 × 42	100 × 20 × 60	140 × 28 × 84
Total grid number	510,947	1,974,584	4,664,930
S_G (C_t)	3.76×10^{-3}	4.05×10^{-3}	4.18×10^{-3}
ε_{G21}	-1.30×10^{-4}		
ε_{G32}	-2.90×10^{-4}		
R_G	0.448		
P_G	2.315		
δ^*_G	-1.06×10^{-4}		
C_G	1.231		
U_G	1.06×10^{-4}		
D	4.23×10^{-3}		
U_G/D	2.50%		
E	5.22×10^{-5}		
E/D	1.23%		
U_V/D	3.53%		

constructed respectively.

For the single-fidelity Kriging surrogate model using 50 high-fidelity samples, different Spatial Correlation Function (SCF) and deterministic polynomials are used to combine. Gaussian (Laurenceau and Sagaut, 2008) or cubic spline (Yamazaki et al., 2010) functions are the most common forms for SCF, where the hyper-parameter expresses the distance weight for both spatial correlation functions. Furthermore, the deterministic polynomials used in this paper are polynomials of degree 0, 1, and 2. Three main error indexes of each surrogate model can be determined through leave-one-out cross validation, namely the Average Absolute Error (AAE), Maximum Absolute Error (MAE), and Root Mean Square Error (RMSE):

$$AAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{38}$$

$$MAE = \max_{1\le i\le N}|\hat{y}_i - y_i| \tag{39}$$

$$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{40}$$

where N represents the sample number, $\hat{y}_i$ represents the predicted value through leave-one-out surrogate model, and y_i represents the real value of the leave-one-out sample point.

Three main error indexes for the six constructed Kriging models are listed in Table 6 respectively. All things considered, Gaussian spatial correlation function and 2-degree deterministic polynomial are applied to construct the Kriging model with 50 high-fidelity data.

For Co-Kriging surrogate model, in order to facilitate further discussion, the SCF and deterministic regression part (polynomial function) in the Co-Kriging model are Gaussian correlation function and zero polynomial respectively (i.e., corrgauss-regpoly0). Under the premise of determining the sample size of low fidelity 50, the influence of the different high-fidelity sample number for constructing the Co-Kriging

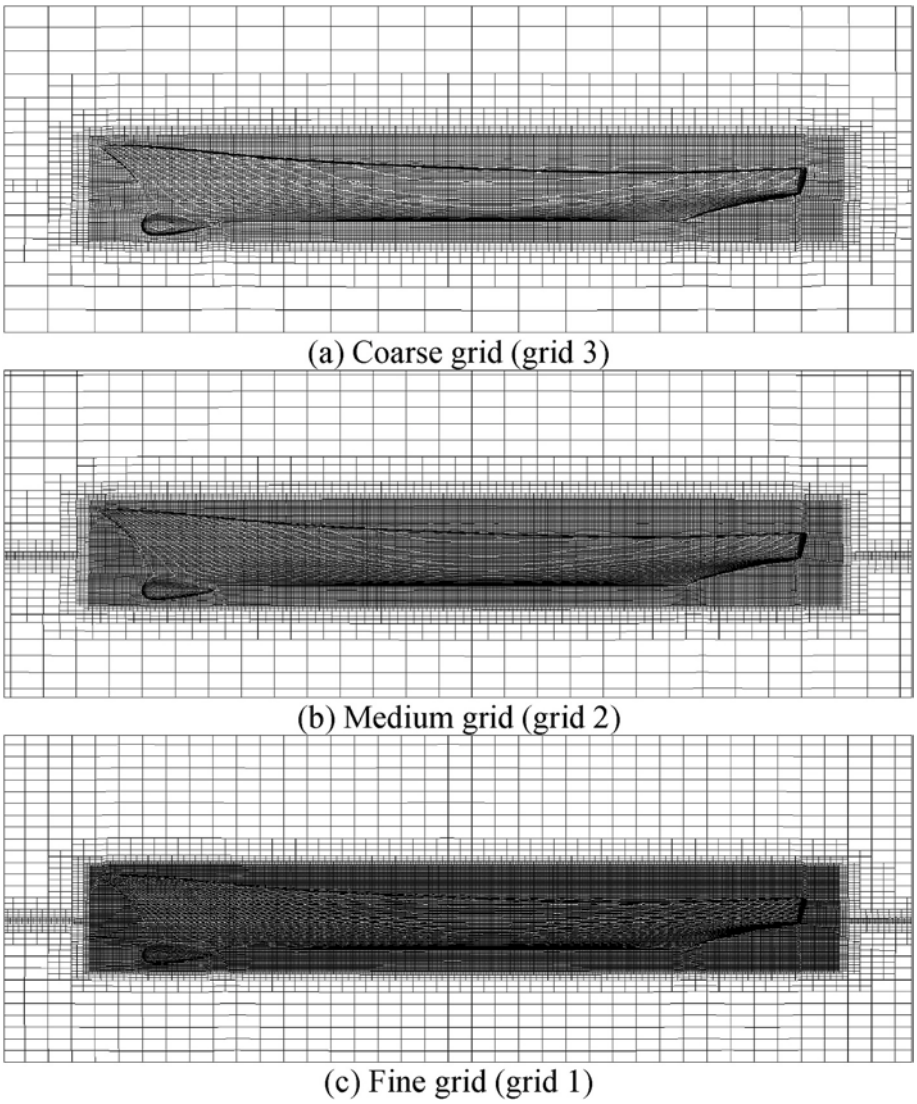

Fig. 16. Three grids for convergence study.

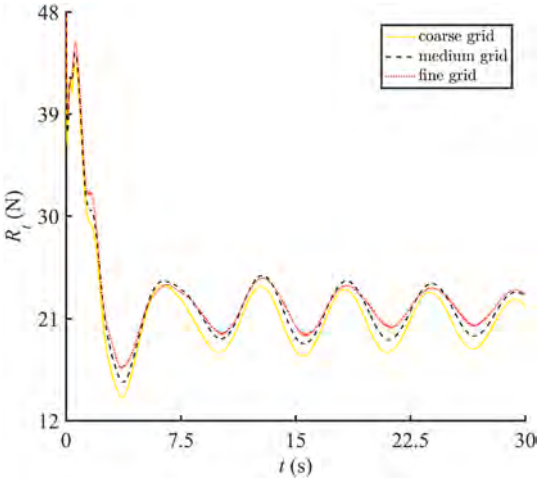

Fig. 17. Convergence histories for total drag of three grid sets.

model on the main error indexes are listed in Table 7. When the high-fidelity data set is a subset of the low-fidelity data set, the generation of some additional prediction errors can be reduced, and the high-fidelity sample data can be fully utilized to correct the low-fidelity sample data. For instance, Fig. 19 shows the 30 high-fidelity samples (red dots) and 50 low-fidelity samples (green squares) distribution in 3-dimensional design space for design variables x_1, y_1, z_1.

Table 5
Design variables and their ranges.

Design variable	Lower bound	Upper bound	Note
x_1	−0.012	0.012	movable amount in x direction
z_1	−0.006	0.006	movable amount in z direction
y_1	−0.015	0.015	movable amount in y direction
α_{1f}	−0.015	0.015	Amplitude of the modification function for the fore half body in shifting method
α_{1a}	−0.015	0.015	Amplitude of the modification function for the aft half body in shifting method

According to the error statistics, with the increase of high-fidelity sample number, main error indicators generally decrease first and then increase. This is mainly because, when high-fidelity sample number is small, the correction of high-fidelity samples to the low-fidelity surrogate model is insufficient, leading to the dominant position of the low-fidelity surrogate model. However, when the high-fidelity sample number increases, high- and low-fidelity samples cooperate with each

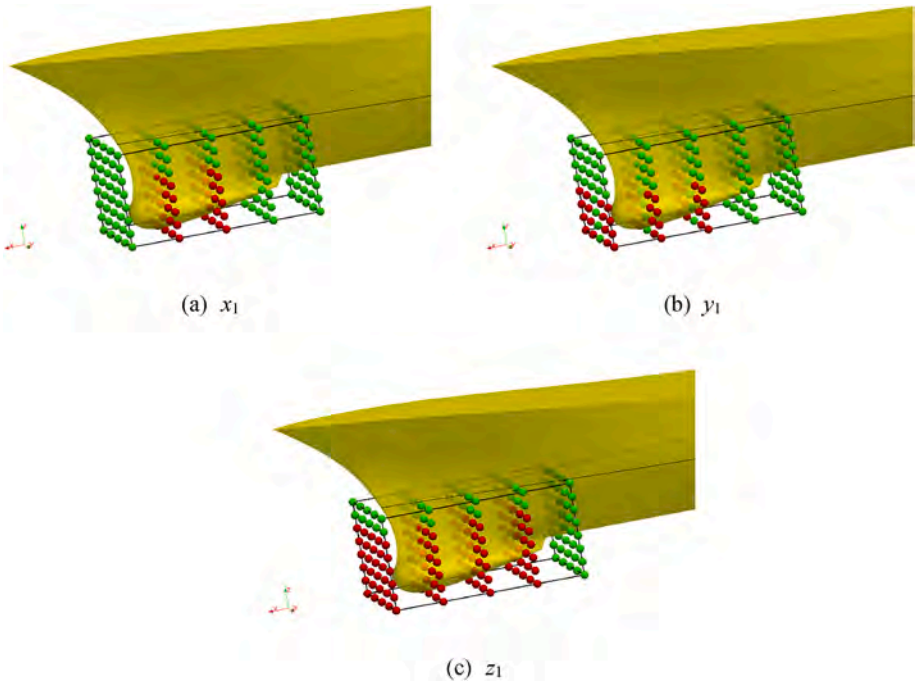

Fig. 18. Bow deformation of DTMB-5415 based on FFD method.

Table 6
Error analysis of Kriging surrogate model.

	regpoly0	regpoly1	regpoly2	
corrgauss	0.3381	0.2866	0.2602	AAE
	2.4490	0.9711	1.0823	MAE
	0.5282	0.3608	0.3358	RMSE
corrspline	0.4021	0.3130	0.2689	AAE
	2.0568	1.1434	1.0896	MAE
	0.5638	0.4173	0.3445	RMSE

Table 7
Error analysis of Co-Kriging surrogate model.

High-fidelity sample number N_e						
20	25	30	35	40	50	
0.6420	0.4299	0.2251	0.2725	0.2628	0.2867	AAE
1.1991	1.3323	0.6524	1.0578	0.8646	1.9322	MAE
0.7401	0.5869	0.3023	0.3684	0.3538	0.4544	RMSE

other, and accuracy of the surrogate model becomes higher. When the high-fidelity sample number increases further, high-fidelity surrogate model leads the dominant position, making the low-fidelity samples difficult to play their role. Therefore, it tends to be the single-fidelity Kriging surrogate model constructed by pure high-fidelity samples. In other words, when the total amount of low-fidelity samples is determined, the number of high-fidelity samples has a small range that makes the multi-fidelity Co-Kriging surrogate model have a relatively high accuracy. Seen from this example, when the low-fidelity sample number is 50, considering the calculation efficiency of hydrodynamic evaluation and the accuracy of the surrogate model constructed, the high-fidelity sample number can be chosen as 30.

4.5. Setting of optimization algorithm

According to the Kriging model and Co-Kriging model determined above, the single-objective Genetic Algorithm (GA) is applied to iteratively solve the optimal hull with minimum total drag. The main

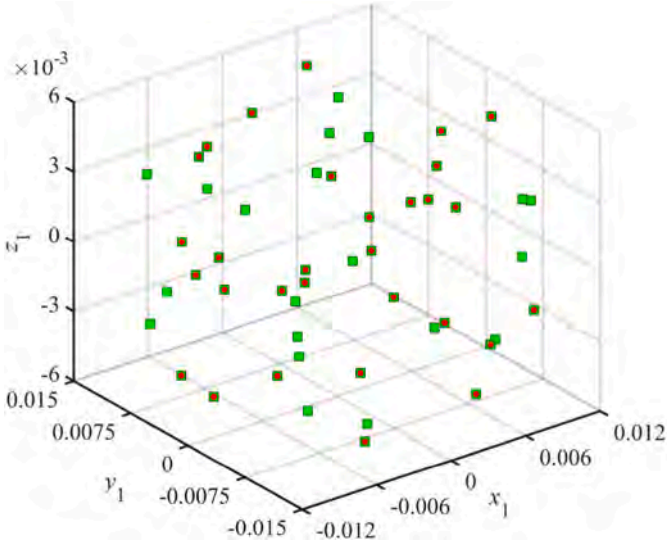

Fig. 19. High-fidelity samples and low-fidelity samples distribution.

Table 8
Single-objective optimization parameters setup of DTMB-5415 ship.

Parameter	Value
Population	50
maximum iteration	300
crossover rate	0.8
mutation rate	0.2

optimization parameters are shown in Table 8. Furthermore, the GA used in this paper has already been validated for a lot of standard optimization problem tests and hull form optimization problems (Wu et al., 2017; Liu et al., 2021a; Liu et al., 2021b), showing that the GA can capture the global minimum of the mathematical functions well.

4.6. Optimization results and analysis

Based on Kriging and Co-Kriging surrogate models, the optimal solutions of this optimization problem can be finally obtained according to GA with the termination criterion that the termination tolerances on fitness function value and constraints are both less than 10^{-6}, respectively. Herein, the optimal hull "Opt-K" is obtained based on Kriging model, while the optimal hull "Opt-CoK" is obtained based on Co-Kriging model. Comparison of transverse hull lines between the two optimal and the initial hulls is shown in Fig. 20.

It can be seen that for "Opt-K″ hull, the volume of sonar dome decreases and the sonar dome moves up as a whole. In addition, the cross section of the front half of the "Opt-K″ hull has a large shifting range, and the cross section of the front quarter of the ship length has a large forward movement, while the rest cross section of 1/4 ship length has a small backward movement. Similarly, the cross section of the aft quarter of the ship length has a large backward movement, while the rest cross section of 1/4 ship length has a small forward movement.

In comparison, the change of "Opt-CoK" hull is small. Specifically, the volume of the sonar dome changes little, but it also moves up as a whole. In addition, the cross section of the front half of "Opt-CoK" hull has a small shifting range, and the cross section of the front quarter of the ship length has a small forward movement, while the rest cross section of 1/4 ship length almost has no movement. The aft part of "Opt-CoK" hull is almost unchanged from the initial hull. The detailed design variables of the two optimal hulls are shown in Table 9.

To be trust, since the GA is stochastic, 100 repetitions of the optimization by GA have been done under the construction of Kriging or Co-Kriging surrogate model, and the minimum of the 100 optimal hulls is chosen as the final optimal hull called "Opt-K″ or "Opt-CoK". The convergence of GA using Kriging model is shown in Fig. 21, indicating that the GA can give steady optimal solution by iterations (generations) less than 300, and the trend is consistent for Co-Kriging model. For the objective function, the expected value of the "Opt-K″ and "Opt-CoK" are 39.500 and 41.020 respectively, and standard deviation of the "Opt-K″ and "Opt-CoK" are 2.595×10^{-6} and 7.430×10^{-8} respectively. It can be seen that the obtained optimal solutions are steady.

For each optimal hull, in order to ensure the optimization effect and further analyze the optimization results, the naoe-FOAM-SJTU solver is used to carry out the viscous flow calculation verification, and the calculation mesh and related configuration parameters are almost consistent with those of the initial hull.

Although Table 9 shows that the single-fidelity Kriging provides a better optimal hull with a drag of 39.500 N than the multi-fidelity Co-Kriging optimal hull with a drag of 41.020 N, after further validation through CFD, the real drags of the two are 42.011 N and 41.001 N respectively, which indicates that the surrogate models have some prediction error. Particularly, at the point (0.015,0.015,0.012,0.006,-0.015), the relatively error of Kriging model is (39.500–42.011)/42.011≈−5.98%, while at the point (0.008,0.001,0.005,0.006,0.004), the relatively error of Co-Kriging model is (41.020–41.001)/41.001 ≈ 0.05%. Therefore, we can say that the Co-Kriging model has relatively high fidelity and can obtain a really better optimum than Kriging model, and the optimization effect for "Opt-CoK" is better, reaching 5%. On the contrary, the Kriging surrogate model prediction for "Opt-K″ hull is not so accurate compared with the CFD result, and the optimization effect is poorer, not reaching 3%.

Fig. 22 and Fig. 23 show the comparison of the free surface wave elevations of the two optimal hulls and the initial hull at the design speed respectively. On the whole, the wave amplitudes of the two optimal hulls are both reduced at the design speed, but there are some differences.

For the "Opt-K″ hull shown in Fig. 22, due to the change of sonar dome's location and size, along with the change of SAC, the phases of the wave systems are changed, that is, the whole bow wave system moves backward, and the amplitudes of the peak and trough near midship is reduced (region 1). In addition, the peak range of stern wave is also reduced (region 2).

For the "Opt-CoK" hull shown in Fig. 23, because of the limited influence of the shifting method, namely the hull form has nearly no change by the shifting method, the bow wave system hardly has any phase shift, but the diverging wave amplitudes of peaks and troughs are significantly lower (region 3). Predictably, the wave-making drag will have a larger decline, so the total drag optimization effect is more obvious.

Fig. 24 shows the comparison of dynamic pressure distribution on the hull surface of two optimal and initial hulls. On the whole, the dynamic pressure distribution on the hull surface of the two optimal hulls are both improved at the design speed, but the effect is also different.

Shown in Fig. 24, for "Opt-K″ hull, the volume of sonar dome decreases, so that the high- and low-pressure regions at the bow part of ship are obviously smaller (region 4). However, due to the effect of the shifting method, the shoulder wave of the ship comes into being, and a low-pressure region of certain amplitude is formed near the midship (region 6). In addition, higher- and lower-pressure regions appear at the ship stern respectively, so that the longitudinal pressure gradient is raised.

For "Opt-CoK" hull, although the high- and low-pressure area at the bow part of the ship do not significantly decrease (region 4), the high-pressure area near the free surface decreases to some extent (region 5), compared with "Opt-K″ hull. Furthermore, the longitudinal pressure gradient around the midship is small enough (region 6) which can be seen form the contour line, and there is no such obvious high- and low-pressure region of the certain amplitude at ship stern like "Opt-K″ hull.

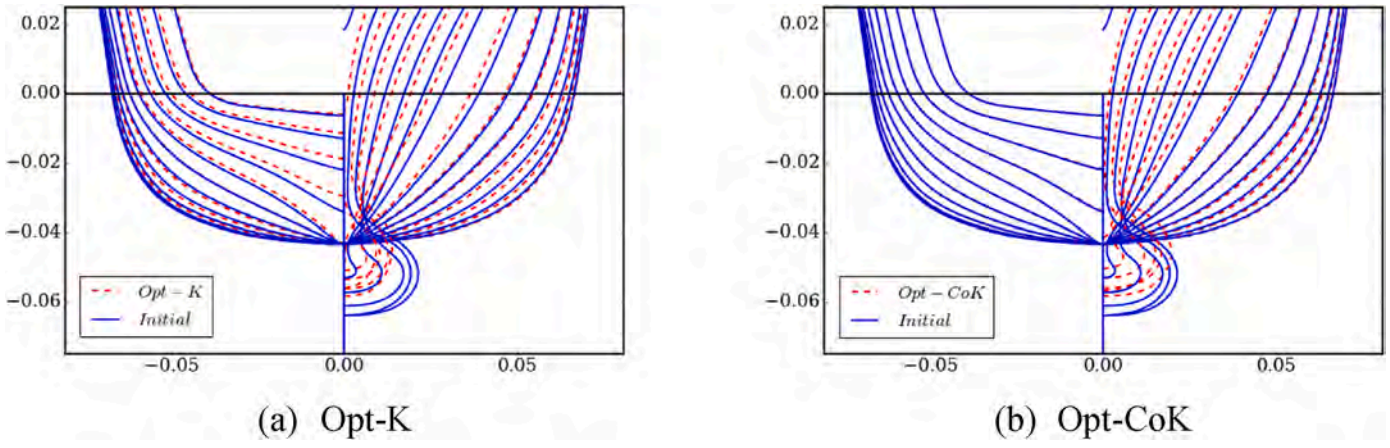

Fig. 20. Comparison of the transverse hull lines of initial and the optimal hulls.

Table 9
Summary of the optimal results, including design variables and objective functions.

/	Design variable value					R_t		
	α_{1f}	α_{1a}	x_1	z_1	y_1	Model prediction (N)	CFD calculation (N)	Reduction ratio
Initial	0	0	0	0	0	/	43.220	/
Opt-K	0.015	0.015	0.012	0.006	−0.015	39.500	42.011	2.80%
Opt-CoK	0.008	0.001	0.005	0.006	0.004	41.020	41.001	5.13%

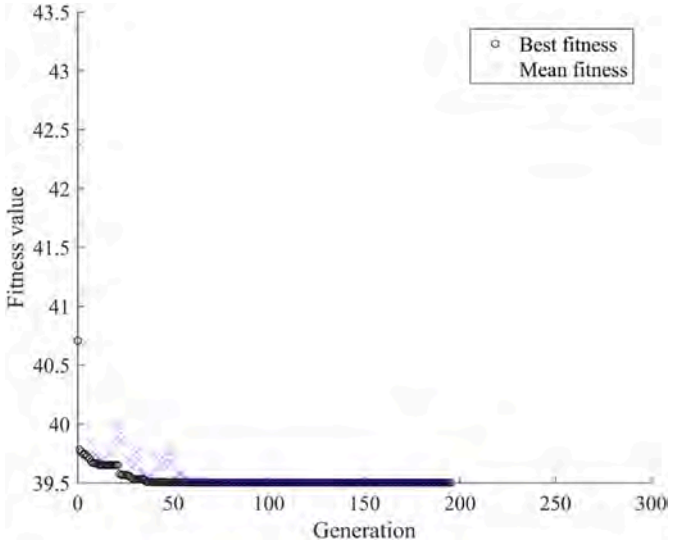

Fig. 21. Convergence of GA using Kriging model.

To sum up, the multi-fidelity Co-Kriging surrogate model can be established through evaluation methods based on the viscous flow and potential flow theory respectively, which is superior to the single-fidelity Kriging surrogate model in accuracy, and the optimal hull obtained by Co-Kriging surrogate model has a better resistance optimization effect.

Last but not least, it should be noted that the Kriging model adopts 50 high-fidelity samples, while the Co-Kriging model adopts 50 low-fidelity samples and 30 high-fidelity samples in this optimization design. In terms of surrogate model accuracy, it is suggested that the number of low-fidelity samples adopted by the multi-fidelity Co-Kriging model should not be less than 10 times the number of design variables, and the number of high-fidelity samples should not be less than 5 times the number of design variables. In addition, within the range of 5–7 times the number of design variables, higher accuracy of the multi-fidelity Co-Kriging model can be achieved, the calculation cost is relatively small at the same time.

In terms of computational efficiency, in the hull form optimization design given above, it is vital to discuss the CPU times needed to obtain the wave-making drag coefficient by NMShip-SJTU solver and total drag coefficient by naoe-FOAM-SJTU.

The NMShip-SJTU solver is run on PC, and the calculation information for a single hull is listed:

(i) CPU information: Intel(R) i7-4790 K @ 2.00 GHz;
(ii) grid number: hull panel about 9700; free-surface panel 296,400;
(iii) iteration: approximately 50 CPU seconds by 1 processor per hull panel velocity potential iteration;
(iv) free-surface calculation: approximately 400 CPU seconds;
(v) total CPU time: approximately 900 CPU seconds;

The naoe-FOAM-SJTU solver is implemented on High Performance Computing (HPC) cluster. For the calm-water total drag calculation of a single hull, the calculation information is listed:

(i) CPU information: Intel(R) Xeon Gold 5120 @ 2.20 GHz;
(ii) grid number: about 1,975,000;
(iii) iteration: approximately 6 CPU seconds by 40 processors per time step $\triangle t = 0.001$ s;
(iv) total CPU time: approximately 195,000 CPU seconds.

For the calm-water drag evaluation of one ship hull, according to the grid quantity and calculation time given above, it is not difficult to see that the time needed to solve the wave drag coefficient by using NM theory (900 s) is only 1/216 of that needed to solve the total drag

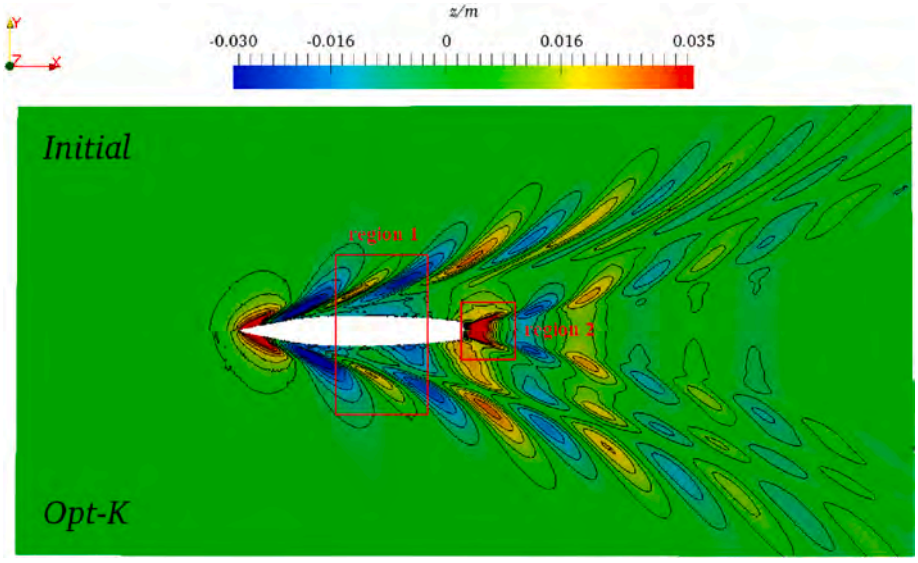

Fig. 22. The wave elevation comparison of the Opt-K and Initial hulls.

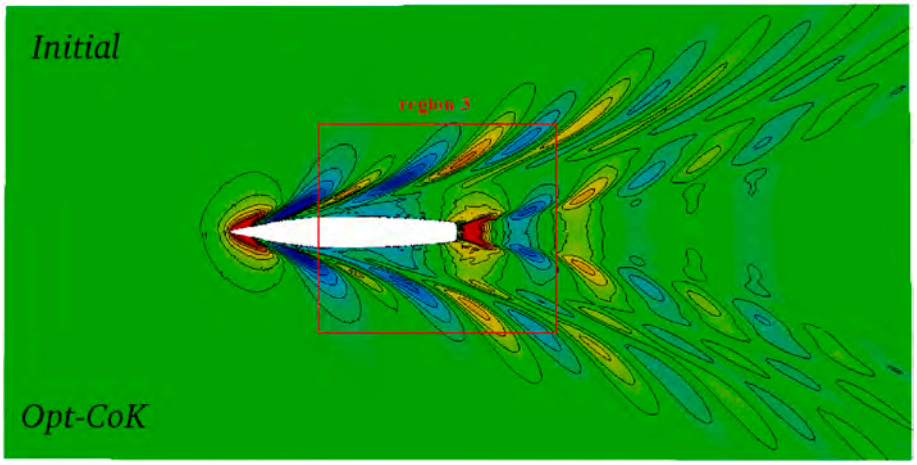

Fig. 23. The wave elevation comparison of the Opt-CoK and Initial hulls.

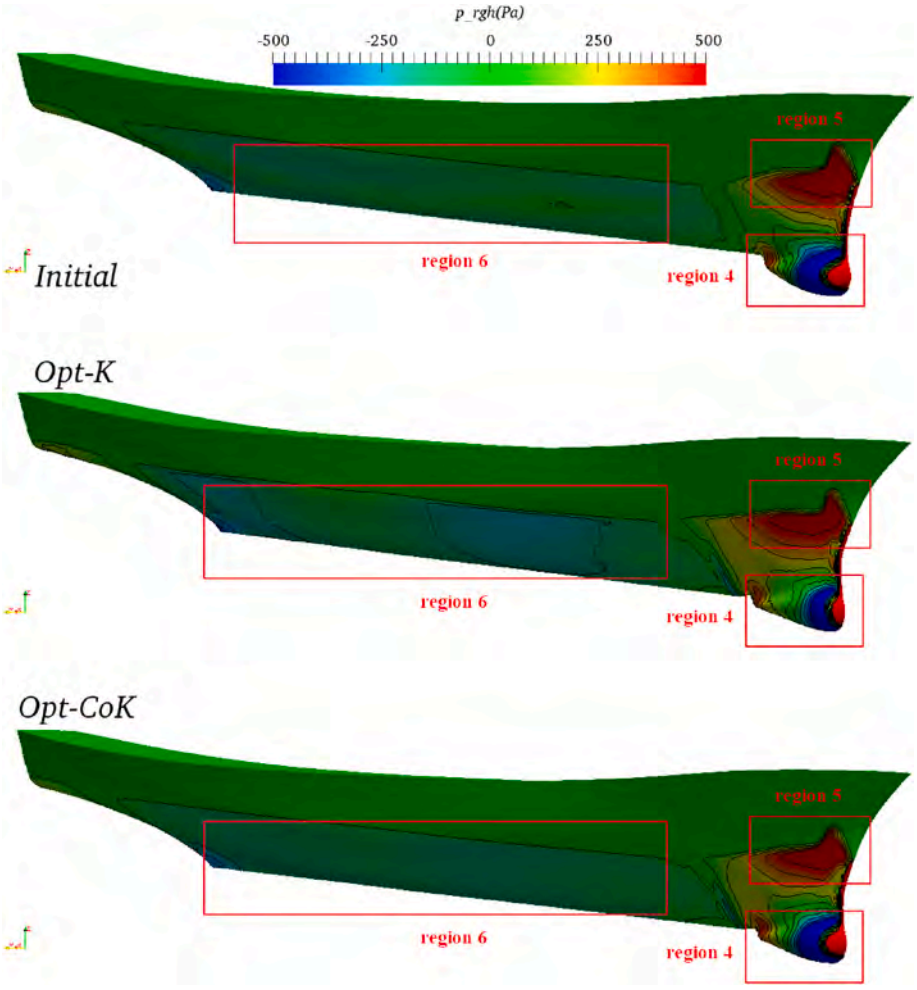

Fig. 24. The hull pressure distribution comparison of the Opt-K, Opt-CoK, and Initial hulls.

coefficient by using the viscous flow solver (195,000 s). It is enough to prove that the use of multi-fidelity Co-Kriging model has superiority for the resistance performance optimization of hull form.

5. Conclusions and future work

For the hull form optimization design based on numerical simulation, the hydrodynamic performance evaluation results of different fidelity can be used to construct surrogate models. For example, the result obtained by potential flow theory can be regarded as a low-fidelity solution, while the result obtained by viscous flow theory can be regarded as a high-fidelity solution. Traditionally, based on the single-fidelity Kriging surrogate model, in order to ensure its high accuracy, a large number of sample points are needed to evaluate, that is,

to do the numerical simulations of a series of new hull forms. Multi-fidelity Co-Kriging surrogate model, on the contrary, can give attention to both high efficiency and high fidelity. To be specific, fewer samples of high fidelity, together with more samples of low fidelity are used to construct the surrogate model, making the total computational cost less or even far less than that of more samples of high fidelity only, and the constructed model can have higher fidelity, which can be used for optimization solution, and the final optimization results can be guaranteed.

In this paper, the basic principle of the multi-fidelity Co-Kriging surrogate model is first introduced, and then a series of numerical examples are given to illustrate the advantages of the multi-fidelity Co-Kriging surrogate model compared with the single-fidelity Kriging surrogate model in terms of efficiency and accuracy. Furthermore, an example of total drag single-objective optimization design of DTMB-5415 hull under the design speed Fr = 0.28 is given, in which the low-fidelity total drag results are obtained by using NMShip-SJTU solver based on the NM potential flow theory, and the high-fidelity total drag results are obtained by using naoe-FOAM-SJTU solver based on the RANS equations. The optimal hull obtained by Co-Kriging surrogate model has a better resistance optimization effect than that of the optimal one by Kriging surrogate model, indicating the superiority of Co-Kriging surrogate model in terms of computational efficiency and fidelity compared with the Kriging surrogate model in the hull form optimization field.

In terms of accuracy, it is recommended that the Co-Kriging surrogate model uses the low fidelity sample with the size of not less than 10 times the number of design variables, and high fidelity sample size with the size of not less than 5–7 times the number of design variables. If the relationship between the objective function and the design variables are more complex, the sample sizes can be increased according to actual demand. It should be noted that Fernández-Godino et al. (2019) reviewed the papers that used multi-fidelity surrogates in the fields of fluid and solid mechanics. As a matter of fact, the authors came to the conclusion that it is hard to get an idea, in terms of cost savings, of when it is useful to invest the additional effort of creating and using multi-fidelity surrogate models, which should be considered in the future work. In addition, due to the limited computational resources, the selection of the numbers of the low- and high-fidelity sample points is not deep discussed in this paper. A parallel research is conducted to study the influence of sample numbers of different levels of fidelity on the accuracy of the Co-Kriging surrogate model constructed by changing the numbers of both the low- and high-fidelity sample points based on numerical cases and actual comprehensive performance optimization designs for the hull form.

Although this paper presents the application of multi-fidelity Co-Kriging surrogate model in hull form optimization, in order to further utilize the high efficiency obtained from lower-fidelity data, multi-fidelity Co-Kriging surrogate model construction technique under more than two accuracy levels of sample data in the field of hull form optimization can be included in the future work. In this paper, since the high-fidelity and low-fidelity data is from viscous and potential flow solver, the intermediate fidelity level is hard to determine. In contrast, three viscous flow solution sets with different grid sizes or time steps can be used as data with three fidelity levels.

Furthermore, the numerical (calculation) noise is not considered as an external influence or even new “design variable”, that is, regarded the low-fidelity and high-fidelity data as a steady result without noise in this paper. In the future, the noise can be considered as an external influence or even new “design variable” to analyze the stability and robustness of the multi-fidelity Co-Kriging surrogate model. Furthermore, the regressive formulation, instead an interpolating one, will be constructed and compared with the interpolating formulations, such as the Co-Kriging model.

CRediT authorship contribution statement

Xinwang Liu: Data curation, Writing – original draft, Visualization, Investigation, Software, Validation. **Weiwen Zhao:** Software, Data curation, Visualization, Investigation, Validation. **Decheng Wan:** (Corresponding author), Supervision, Conceptualization, Methodology, Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Acknowledgments

This work is supported by the National Key Research and Development Program of China (2019YFB1704200), National Natural Science Foundation of China (52131102, 51879159), to which the authors are most grateful.

References

Box, G.E.P., Wilson, K.B., 1951. On the experimental attainment of optimum conditions. J. Roy. Stat. Soc. B 13 (1), 1–45.
Cai, X.W., Qiu, H.B., Gao, L., Shao, X.Y., 2017. Metamodeling for high dimensional design problems by multi-fidelity simulations. Struct. Multidiscip. O 56 (1), 151–166.
Campana, E.F., Peri, D., Tahara, Y., Stern, F., 2006. Shape optimization in ship hydrodynamics using computational fluid dynamics. Comput. Methods Appl. Mech. Eng. 196 (1–3), 634–651.
Chen, X., Diez, M., Kandasamy, M., Zhang, Z., Campana, E.F., Stern, F., 2015. High-fidelity global optimization of shape design by dimensionality reduction, metamodels and deterministic particle swarm. Eng. Optim. 47 (4), 473–494.
Coppede, A., Gaggero, S., Vernengo, G., Villa, D., 2019. Hydrodynamic shape optimization by high fidelity CFD solver and Gaussian process-based response surface method. Appl. Ocean Res. 90, 101841.
De, B.J., Roberts, S., Dwight, R., Mallol, B., 2015. Uncertainty quantification for a sailing yacht hull, using multi-fidelity kriging. Comput. Fluids 123, 185–201.
Deb, K., Agrawal, R.B., 1995. Simulated binary crossover for continuous search space. Complex Syst. 9, 115–148.
Diez, M., Campana, E.F., Stern, F., 2018. Stochastic optimization methods for ship resistance and operational efficiency via CFD. Struct. Multidiscip. Optim. 57 (2), 735–758.
Diez, M., Volpi, S., Serani, A., Stern, F., Campana, E.F., 2019. Simulation-based design optimization by sequential multi-criterion adaptive sampling and dynamic radial basis functions. Advances in Evolutionary and Deterministic Methods for Design, Optimization and Control in Engineering and Sciences. In: Computational Methods in Applied Sciences, 48. Springer, Cham, pp. 213–228.
Fernández-Godino, M.G., Park, C., Kim, N.H., Haftka, R.T., 2019. Issues in deciding whether to use multifidelity surrogates. AIAA J. 57 (5), 2039–2054.
Ficini, S., Iemma, U., Pellegrini, R., Serani, A., Diez, M., 2021. Assessing the performance of an adaptive multi-fidelity Gaussian process with noisy training data: a statistical analysis. Proceedings of AIAA AVIATION Forum.
Forrester, A.I.J., Sóbester, A., Keane, A.J., 2007. Multi-fidelity optimization via surrogate modeling. Proceedings of the royal society A: mathematical. Phy. Eng. Sci. 463, 3251–3269.
Forrester, A.I.J., Keane, A.J., 2009. Recent advances in surrogate-based optimization. Prog. Aero. Sci. 45 (1), 50–79.
Ghoreyshi, M., Badcock, K., Woodgate, M., 2008. Integration of multi-fidelity methods for generating an aerodynamic model for flight simulation. In: Proceedings of 46th AIAA Aerospace Sciences Meeting and Exhibit, 197.
Gratiet, L.L., Garnier, J., 2014. Recursive co-kriging model for design of computer experiments with multiple levels of fidelity. Int. J. Uncertain. Quantification 4 (5), 365–386.
Han, Z.H., Goertz, S., Zimmermann, R., 2013. Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. Aero. Sci. Technol. 25 (1), 177–189.
Hou, Y.H., 2017. Hull form uncertainty optimization design for minimum EEOI with influence of different speed perturbation types. Ocean Eng. 140 (1), 66–72.
Huang, F.X., Wang, L.J., Yang, C., 2015. Hull form optimization for reduced drag and improved seakeeping using a surrogate-based method. Proc. 25th Int. Ocean Polar Eng.
Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient global optimization of expensive black-box functions. J. Global Optim. 13 (4), 455–492.
Kennedy, M.C., O'Hagan, A., 2001. Bayesian calibration of computer models. J. Roy. Stat. Soc. B 63 (3), 425–464.
Kim, H., Yang, C., Jeong, S., Noblesse, F., 2011. Hull form design exploration based on response surface method. Proc. 21st Int. Ocean Polar Eng. 816–827.

Kim, H.S., Koc, M., Ni, J., 2007. A hybrid multi-fidelity approach to the optimal design of warm forming processes using a knowledge-based artificial neural network. Int. J. Mach. Tool Manufact. 47 (2), 211–222.
Krige, D.G., 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. Journal of the Chemical,. Metall. Min. Eng. Soc. South Africa 52 (6), 119–139.
Kuya, Y., Takeda, K., Zhang, X., Forrester, A.I.J., 2011. Multifidelity surrogate modeling of experimental and computational aerodynamic data sets. AIAA J. 49 (2), 289–298.
Laurenceau, J., Sagaut, P., 2008. Building efficient response surfaces of aerodynamic functions with Kriging and Co-kriging. AIAA J. 46 (2), 498–507.
Li, S.Z., Zhao, F., Ni, Q.J., 2014. Bow and stern shape integrated optimization for a full ship by a simulation-based design technique. J. Ship Res. 58 (2), 83–96.
Liu, X.W., Wan, D.C., 2020. Hull form optimization of wave-making resistance in different speeds for a luxury cruise ship. Chin. J. Ship Res. 15 (5), 1–10, 40.
Liu, X.W., Wan, D.C., Chen, G., Hu, C.H., 2019. Wigley hull form optimization with or without bulbous bow. In: Proceedings of the 29th International Ocean and Polar Engineering Conference, pp. 4486–4493.
Liu, X.W., Zhao, W.W., Wan, D.C., 2021a. Linear reduced order method for design-space dimensionality reduction and flow-field leraning in hull form optimization. Ocean Eng. 237, 109680.
Liu, X.W., Zhao, W.W., Wan, D.C., 2021b. Hull form optimization based on calm-water wave drag with or without generating bulbous bow. Appl. Ocean Res. 116, 102861.
Miao, A.Q., Zhao, M., Wan, D.C., 2020. CFD-based multi-objective optimisation of S60 catamaran considering demihull shape and separation. Appl. Ocean Res. 97, 102071.
Noblesse, F., Huang, F.X., Yang, C., 2013. The Neumann-Michell theory of ship waves. J. Eng. Math. 79 (1), 51–71.
Olivieri, A., Pistani, F., Avanzini, A., Stern, F., Penna, R., 2001. Towing Tank Experiments of Resistance, Sinkage and Trim, Boundary Layer, Wake, and Free Surface Flow Around a Naval Combatant INSEAN 2340 Model. The University of Iowa, College of Engineering. Report No. IIHR-TR-421.
Orr, M., 1995. Regularisation in the selection of RBF centers. Neural Comput. 7 (3), 606–623.
Peri, D., Rossetti, M., Campana, E.F., 2001. Design optimization of ship hulls via CFD techniques. J. Ship Res. 45 (2), 140–149.
Qian, P., Wu, C., 2008. Bayesian hierarchical modeling for integrating low-accuracy and high-accuracy experiments. Technometrics 50 (2), 192–204.
Roderick, O., Anitescu, M., Peet, Y., 2014. Proper orthogonal decompositions in multifidelity uncertainty quantification of complex simulation models. Int. J. Comput. Math. 91 (4), 748–769.
Rumpfkeil, M.P., Beran, P., 2017. Construction of dynamic multifidelity locally optimized surrogate models. AIAA J. 55 (6), 1–11.
Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P., 1989. Design and analysis of computer experiments. Stat. Sci. 4 (4), 409–423.
Sasena, M.J., Papalambros, P., Goovaerts, P., 2002. Exploration of metamodeling sampling criteria for constrained global optimization. Eng. Optim. 34 (3), 263–278.
Sederberg, T.W., Parry, S.R., 1986. Free-form deformation of solid geometric primitives. Comput. Graph. 20 (4), 151–160.
Serani, A., Pellegrini, R., Broglia, R., Wackers, J., Visonneau, M., Diez, M., 2019a. An adaptive N-fidelity metamodel for design and operational-uncertainty space exploration of complex industrial problems. In: Proceedings of VIII International Conference on Computational Methods in Marine Engineering MARINE.
Serani, A., Pellegrini, R., Wackers, J., Jeanson, C.E., Queutey, P., Visonneau, M., Diez, M., 2019b. Adaptive multi-fidelity sampling for CFD based optimisation via radial basis function metamodels. Int. J. Comput. Fluid Dynam. 33 (6–7), 237–255.
Serani, A., Stern, F., Campana, E.F., Diez, M., 2021. Hull-form stochastic optimization via computational-cost reduction methods. Eng. Comput. 2, 1–25.
Shen, Z.R., Wan, D.C., Carrica, P.M., 2015. Dynamic overset grids in OpenFOAM with application to KCS self-propulsion and maneuvering. Ocean Eng. 108, 287–306.
Smith, M., 1993. Neural Networks for Statistical Modeling. Von Nostrand Reinhold, New York.
Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14, 199–222.
Sobol, I., 1979. On the systematic search in a hypercube. SIAM J. Numer. Anal. 16 (5), 790–793.
Tahara, Y., Campana, E.F., Peri, D., Stern, F., 2011. Single and multiobjective design optimization of a fast multihull ship: numerical and experimental results. J. Mar. Sci. Technol. 16 (4), 412–433.
Tezdogan, T., Zhang, S.L., Demirel, Y.K., Liu, W.D., Xu, L.P., Lai, Y.Y., Rafet, E.K., Eko, B. D., Incecik, A., 2018. An investigation into fishing boat optimisation using a hybrid algorithm. Ocean Eng. 167, 204–220.
Wackers, J., Visonneau, M., Ficini, S., Pellegrini, R., Serani, A., Diez, M., 2020. Adaptive n-fidelity metamodels for noisy cfd data. Proceedings of AIAA AVIATION Forum.
Wilson, R.V., Stern, F., Coleman, H.W., Paterson, E.G., 2001. Comprehensive approach to verification and validation of CFD Simulations-Part 2: application for rans simulation of a cargo/container ship. J. Fluid Eng. 123 (4), 803–810.
Wu, J.W., Liu, X.Y., Zhao, M., Wan, D.C., 2017. Neumann-Michell theory-based multi-objective optimization of hull form for a naval surface combatant. Appl. Ocean Res. 63, 129–141.
Yamazaki, W., Mavriplis, D.J., 2013. Derivative-enhanced variable fidelity surrogate modeling for aerodynamic functions. AIAA J. 51 (1), 126–137.
Yamazaki, W., Rumpfkeil, M.P., Mavriplis, D.J., 2010. Design optimization utilizing Gradient/Hessian enhanced surrogate model. In: Proceedings of 28th AIAA Applied Aerodynamics Conference, pp. 2010–4363. No.
Yang, C., Huang, F.X., 2016. An overview of simulation-based hydrodynamic design of ship hull forms. J. Hydrodyn. 28 (6), 947–960.
Zhang, S., Zhang, B.J., Tezdogan, T., Xu, L.P., Lai, Y.Y., 2018. Computational fluid dynamics-based hull form optimization using approximation method. Eng Appl Comput Fluid Mech 12 (1), 74–88.
Zong, Z., Hong, Z., Wang, Y., Hefazi, H., 2018. Hull form optimization of trimaran using self-blending method. Appl. Ocean Res. 80, 240–247.
Zuhal, L.R., Zakaria, K., Palar, P.S., Shimoyama, K., Liem, R.P., 2021. Polynomial-chaos–kriging with gradient information for surrogate modeling in aerodynamic design. AIAA J. 59 (8), 2950–2967.