Multi-fidelity model and reduced-order method for comprehensive hydrodynamic performance optimization and prediction of JBC ship

Xinwang Liu [a,b], Decheng Wan [b], Lei Lei [c,d,*]

[a] College of Mathematical Sciences, Harbin Engineering University, Harbin, China
[b] Computational Marine Hydrodynamics Lab (CMHL), School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, China
[c] Wing Robot Limited, Hung Hom, Hong Kong
[d] Department of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, China

ABSTRACT

The multi-fidelity Co-Kriging surrogate model can be applied to combine the accuracy advantage of high-fidelity sample evaluation with the efficiency advantage of low-fidelity sample evaluation. In this paper, the hull form optimization process for resistance and wake performance of a Japan bulk carrier (JBC) hull at design speed is given in detail, where the evaluation results from medium and coarse grids are regarded as the high- and low-fidelity data, respectively. 60 high-fidelity hydrodynamic evaluations have been done to construct the Kriging model, while the Co-Kriging model uses 30 high-fidelity and 60 low-fidelity evaluations with a 25% reduction of the total computation time. The optimization results show that, for the total drag, the Kriging-based optimal hull has a 4.57% reduction, while the Co-Kriging-based optimal hull has 5.67%; for the axial wake fraction reduction at the propeller disk, the Kriging-based optimal hull has a 7.20% reduction, while the Co-Kriging-based optimal hull has 10.37%. Furthermore, in the latter stage of hull form optimization, dimensionality reduction field learning can be performed to fully use the viscous-flow-based calculation results. An accurate and efficient viscous-flow-based wake field learning method is proposed based on the Kriging model and the Proper Orthogonal Decomposition (POD) method with qualitative and quantitative error analysis, which can guide the sensitivity analysis of the design variables, the selection of the design variables and spaces, and new flow field prediction for comprehensive hull form performance optimization.

1. Introduction

When considering comprehensive hydrodynamic performance in hull form design optimization, numerical simulations or model tests incur high costs. Surrogate models such as single-fidelity Kriging and response surface model (RSM) models can significantly reduce computational costs by evaluating a small number of samples and have already been used in many engineering fields for optimization and data-mining (Peri et al., 2001; Granados-Ortiz and Ortega-Casanova, 2021). Among the single-fidelity surrogate models, the Kriging model has unique advantages and has been widely used in hull form design optimization (Liu et al., 2018, 2021b, 2021c, 2021d).

From the perspective of the interpolation method, compared with the traditional interpolation methods, the Kriging model has both global and local spatial statistical properties. In particular, the interpolation coefficients generally do not depend on the spatial location of the predicted point in traditional interpolation methods, but the interpolation coefficients of the Kriging model are closely related to the location of the predicted point, which gives full consideration of the relationship between independent variables in the space. In addition, the robustness of the Kriging model is strong, that is, the existence of small random errors has little impact on the whole Kriging model. From the perspective of model composition, the Kriging model is composed of a parameterized regression model and a non-parameterized random process. Furthermore, since the random process is assumed to be Gaussian, its good properties make the surrogate model overcome the limitations of the general non-parameterized model when processing high-dimensional data.

However, if the evaluation results of numerical simulations or any data source with different fidelity are obtained and considered, the final constructed surrogate model can be regarded as a multi-fidelity surrogate model. For instance, the Co-Kriging surrogate model (Sacks et al.,

Fig. 1. Geometric model of the JBC.

Table 1
Primary parameters of JBC ship.

Parameter	Symbol and unit	Value
Length between perpendiculars	L_{pp} (m)	7
Breadth	B (m)	1.125
Draught	T (m)	0.4125
Molded depth	D (m)	0.625
Drainage volume	∇_0 (m^3)	2.788
Wet surface area	S_0 (m^2)	12.4

1989) provides the equivalent form of a multi-fidelity Bayesian-based model, which does not require prior information and has good computational characteristics (Forrester et al., 2007). The Co-Kriging model is essentially an extension of the single-fidelity Kriging model, which is desirable because it takes advantage of the accuracy of high-fidelity sample data and the efficiency of low-fidelity sample data.

Several multi-fidelity surrogate models have already been proposed and used to optimize engineering designs, such as robust optimizations of airfoils and wings with Mach number uncertainty (Tao and Sun, 2019), buckling optimization of thin-walled variable-stiffness composite cylinders (Guo et al., 2020), aircraft aerodynamic performance optimization using multi-fidelity deep neural networks (Zhang et al., 2021), and thrust optimization of tandem flapping wings with a multi-fidelity Co-Kriging model (Ji et al., 2022).

In hull form optimization field, multi-fidelity surrogate models have been gradually regarded these days. Co-Kriging model was used to improve the seakeeping (Bonfiglio et al., 2018) and resistance (Bonfiglio et al., 2020) performances of Small Waterplane Area Twin Hull (SWATH). Raven and Scholcz (2019) combined potential flow and viscous flow theory to determine the multi-fidelity surrogate model for a fast displacement vessel, which was found successful due to a significant reduction of the number of high-fidelity evaluations, but for a containership, the correlation between low and high-fidelity resistance values was found to be poor. Two multi-fidelity models were assessed for the uncertainty quantification of a roll-on/roll-off passenger ferry sailing in calm water with ship speed and draught uncertainties through viscous flow evaluations by different levels of grids (Piazzola et al., 2020, 2022). Serani et al. (2019a) and Wackers et al. (2020) extended the multi-fidelity surrogate model, which was then used for the shape optimization of a NACA hydrofoil, the DTMB-5415 hull model, and the operational uncertainty quantification of the total drag of a roll-on/roll-off passenger ferry at variable sailing speeds, showing better performance in comparison with its single-fidelity counterpart and dealing with noisy CFD data. Serani et al. (2019b) discussed four different criteria to do the adaptive sampling for multi-fidelity stochastic radial basis functions, which were demonstrated by resistance optimizations of a NACA hydrofoil and DTMB-5415 hull. Pellegrini et al. (2022a) and Pellegrini et al. (2022b) presented and discussed an initialization strategy with a limited training set for adaptive sampling and a required fidelity level to reduce the computational cost of evaluating the initial training set, which was applied to the shape optimization of a NACA hydrofoil.

Machine learning (ML) or deep learning (DL) can provide many techniques to extract information from data that can be translated into knowledge about the underlying fluid mechanics, as long as outcomes

are held to long-standing critical standards (Brunton et al., 2020). Wu et al. (2022) used the unsteady flow field over a foil to train the Conventional Neural Network (CNN) model with different hyperparameter settings based on the mode-decomposing CNN autoencoder to compare the reconstruction effect. Proper Orthogonal Decomposition (POD), also known as Principal Component Analysis (PCA) or Karhunen-Loève Expansion (KLE) (Pearson, 1901; Wold et al., 1987; Jolliffe, 1986) is one of the most classic, but still most widely-used unsupervised reduced order methods. The basic idea of it is to construct a lower-dimensional linear space, while making the projection of the high-dimensional data on it as large as possible. A POD and Multi-Layer Perceptron (MLP)-based neural network technique was employed for fast flow field prediction and applied to two classes of scramjet intakes (Brahmachary et al., 2021). Furthermore, CNN and MLP-based methods were used to predict the physical fields and performance in new flow conditions and geometries given the pixelated shape or limited geometric parameters of a circular cylinder (Lee and You, 2019), supersonic cascade channel (Li et al., 2020), and rotor (Oh et al., 2022).

In the hull form optimization field, some research has focused on applying the reduced-order method to multi-physics field learning in the latter stages of hull form optimization (Serani et al., 2016; Serani and Diez, 2018). However, hydrodynamic performance evaluation is not limited to potential flow theory. For viscous flow calculations, especially three-dimensional calculations, the grid topology relationship of new hull forms may not be entirely consistent, even though there is no unified grid topology relationship for unstructured grids (Liu et al., 2021c).

In this paper, a two-fidelity Co-Kriging surrogate model is first applied to a mathematical test function to verify the superiority of the Co-Kriging model compared with the single-fidelity Kriging model. Then, single-phase and two-phase flow calculation results for the resistance and wake performance of a slow-full Japan bulk carrier (JBC) ship hull are compared and discussed in detail to reduce computational cost. The coarse- and medium-grid-based single-phase flow simulation results are regarded as the low- and high-fidelity data used to construct the Kriging and Co-Kriging surrogate models after comparison to ensure the strong correlation of the results with two fidelity levels. The sample number selection for the Co-Kriging model has been discussed through leave-one-out cross-validation. The optimization results by Kriging and Co-Kriging models are analyzed and compared to show the advantage and feasibility of hull form optimization with the multi-fidelity Co-Kriging model. Furthermore, field learning for viscous flow results can fully use the expensive numerical calculation results in the latter stages of hull form optimization. An accurate and efficient viscous-flow-based wake field learning method is proposed based on the Kriging model and POD method, which is then applied to the sensitivity analysis of the optimization design variables for the JBC ship hull.

2. JBC hull comprehensive performance optimization case

2.1. Basic information about the mother ship

The JBC is a bulk cargo ship that was jointly designed by the National Maritime Research Institute (NMRI), Yokohama National University (YNU), and the Shipbuilding Research Center of Japan (SRC) and it has become an internationally recognized standard ship model with many test results for numerical calculation verification (Bakica et al., 2019). Its three-dimensional model is shown in Fig. 1, and its main parameters at the model scale are shown in Table 1 (Liu et al., 2021a).

The JBC belongs to the low-full ship category because of its low speed and large parallel middle body, resulting in small wave-making and predominant viscous drag. The viscous drag is closely related to the bow and stern shapes. The hull line shrinks considerably when the stern curvature changes rapidly, producing vortex and viscous drag. In addition, the stern shape directly affects the velocity distribution at the propeller disk, which affects the wake performance.
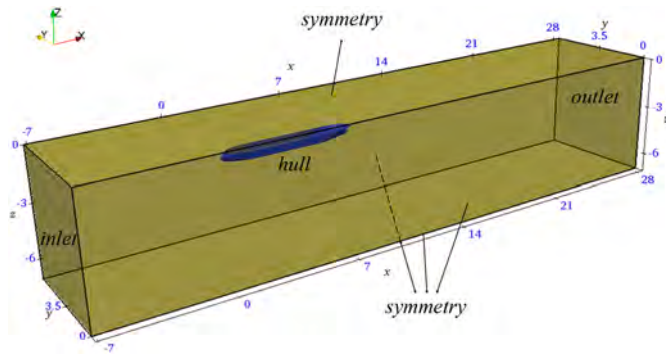
Fig. 2. The size and boundaries of single-phase flow computational domain (Unit: m).

Table 2
Summary of boundary conditions for single-phase flow.

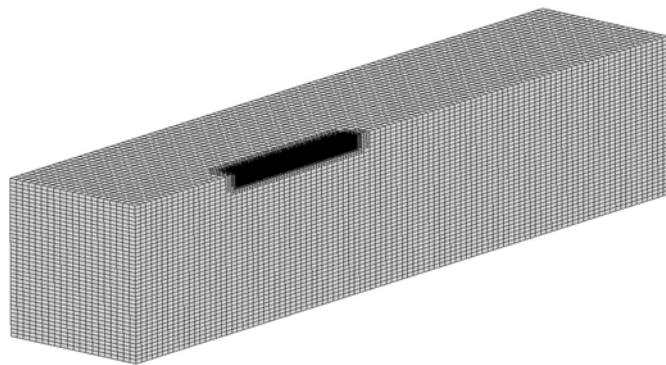| Boundary name	Boundary condition
inlet	velocity inlet
hull	no-slip wall
symmetry	no-slip
outlet	pressure outlet

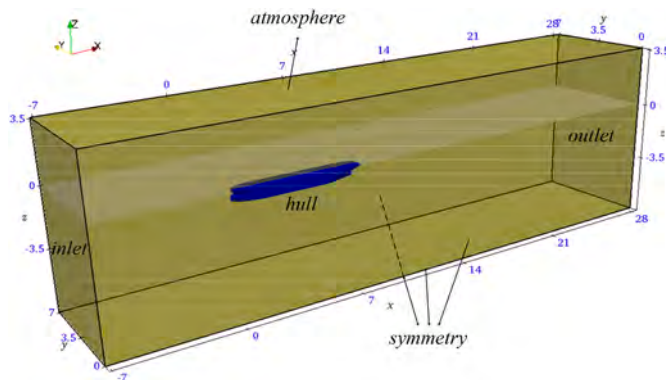**Fig. 3.** The refinement of the computational domain for single-phase flow.

Fig. 4. The size and boundaries of the two-phase flow computational domain (Unit: m).

2.2. Verification of drag and flow fields for numerical calculation

Considering that the wave-making drag is relatively small at the design speed and the free surface wave elevation is not large, it is necessary to consider whether the free surface changes could be neglected in the drag and wake evaluation to reduce the computational

Table 3
Summary of boundary conditions for two-phase flow.

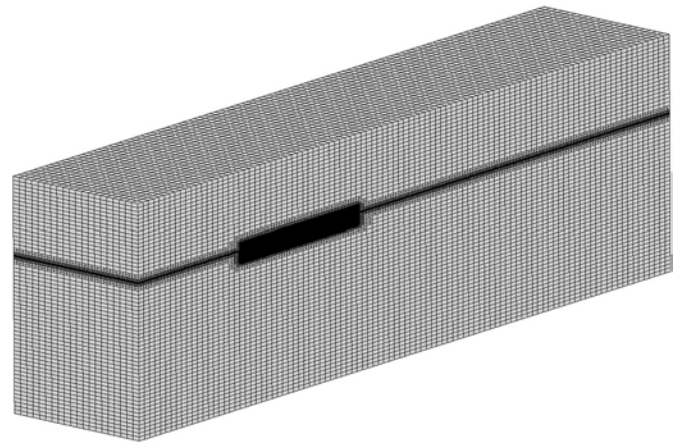| Boundary name	Boundary condition
inlet	velocity inlet
atmosphere	no-slip
hull	no-slip wall
symmetry	no-slip
outlet	pressure outlet

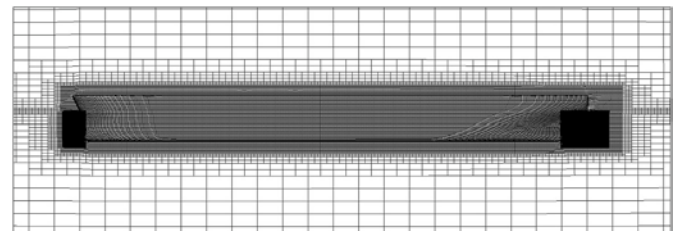**Fig. 5.** The refinement of the computational domain for two-phase flow.

Fig. 6. Grid refinement around the free surface.

costs. This consideration determines whether a single-phase model (sometimes called a double model because the free surface is a symmetry condition), which does not consider the free-surface effect, could be used to evaluate the hydrodynamic performance when doing high-cost optimization.

First, the single-phase and two-phase flows for the initial JBC ship are numerically calculated at model scale and design speed. There are two main differences in the layout of the grids. One difference is that the single-phase flow computational grid domain does not extend above the calm-water free surface (the plane where the ship design waterline is located). The other difference is that multi-level refinement is carried out near the free surface because the two-phase flow computational grid considers the free surface, while the single-phase flow computational grid is not refined near the calm-water free surface.

Fig. 2 shows the single-phase flow grid's computational domain size and boundaries, and Table 2 lists the specific boundary conditions. Grid refinement is mainly concentrated around the hull, especially at the bow and stern. Fig. 3 shows the refinement of the computational domain for single-phase flow. Details of the mesh generation can be found in Liu et al. (2022).

Fig. 4 shows the two-phase flow grid's computational domain size and boundaries, and Table 3 lists the specific boundary conditions. Grid refinement is concentrated mainly on the free surface and around the hull, especially at the bow and stern. Fig. 5 shows the refinement of the computational domain for two-phase flow, and the local refinement near

Table 4
Comparison of calculation and model test results for two-phase flow.

CFD calculation				Model test	Relative error of C_t
				Total drag coefficient C_t	
Pressure drag R_p (N)	Frictional drag R_f (N)	Total drag R_t (N)	Total drag coefficient C_t		
10.24	26.59	36.83	4.267×10^{-3}	4.289×10^{-3}	−0.5%
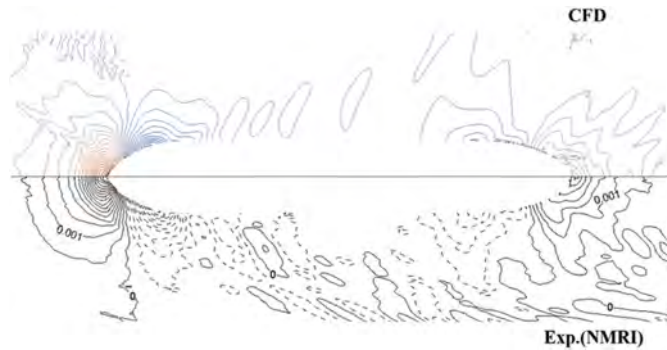
Fig. 7. Comparison of free surface elevations (upper: CFD; lower: experiment).

the free surface is shown in Fig. 6.

The calculation result of calm-water total drag for two-phase flow is obtained by the RANS-based viscous flow calculation platform Open-FOAM with k-ω SST turbulence model, which is compared with model test results in Table 4.

Fig. 7 shows a comparison between the free-surface wave elevation of the JBC in calm water obtained by two-phase flow calculation and the model test. The isoline range of free-surface wave elevation is $-0.008 \leq \frac{z}{L_{pp}} \leq 0.012$, and the interval is $\frac{\Delta z}{L_{pp}} = 0.0005$. Fig. 7 shows that the bow wave, shoulder wave, and stern wave are well captured, and all the main isolines are basically represented. To be honest, there are some differences near the ship's parallel middle body. However, it should be stated that, the wave-making near the parallel middle body is rather small, which is hard to capture, and the total grid number used here is already nearly 6,000,000. Theoretically, through further grid refinement, the calculation result can be better, but the cost can be too high for a series of numerical evaluations.

In addition, Fig. 8 compares the two-phase flow calculation and model test results of the axial velocity distribution at several transverse sections near the stern of the JBC ship. The velocity distributions are generally consistent for the three different sections. In particular, the "curved hook" shape shown in Fig. 8(b) near the propeller shaft is well captured, although the size and position of the calculated "curved hook" shape are slightly lower than the model test result, which is mainly the result of the overestimation of turbulent viscosity because of the near-wall mesh size ($y^+ \approx 30$) and the turbulent wall function used in this calculation. This mesh is adequate in terms of total drag and overall wake performance (such as axial wake fraction) evaluation, considering the computational cost of hull form optimization.

The error between the calculated and the model test total drag coefficient C_t is very small, and the comparisons of the free-surface elevation together with the axial velocity distribution also verify the reliability of the two-phase flow calculation results. However, since the free-surface wave elevation is relatively small, a single-phase flow calculation can be considered for use in optimizing the stern shape and reducing the viscous drag. A comparison between the single-phase and two-phase flow results are given below to illustrate the feasibility of using single-phase flow to evaluate the resistance and wake performance of the JBC ship in this optimization case.

Fig. 9 shows the vorticity iso-surface near the ship model calculated using single- and two-phase flow, where Q equals 20 and is colored by axial velocity. The vorticity distribution in the area below the calm-water free surface is nearly the same, indicating that there is little difference in the viscous pressure drag between the two calculations.

Fig. 10 compares the axial velocity distributions on a section near the propeller calculated by single- and two-phase flow. The velocity distribution is nearly the same, and the main difference is probably because of the hull model's freedom of sinkage and trim during the calculation of two-phase flow, resulting in a certain upward and downward deviation of the velocity distribution contour line. In addition, Fig. 10 also indicates that the small free-surface wave elevation has little influence on the wake near the propeller shaft. Therefore, the influence of the free surface on the wake near the propeller can be ignored.

From the above results, we can conclude that, for this hull form optimization case, the wave-making drag is rather small, and the single-phase flow result can replace the two-phase flow result to construct the total drag and wake fraction surrogate models, and the grid refinement strategy below the free surface we used here is proper and enough, to reduce the calculation cost and accelerate the optimization process.

Validation of the double model grids is then implemented. With the previously used single-phase flow computational grid considered to be the medium grid (S2), the coarse grid (S3) and fine grid (S1) can be generated accordingly. The three grid sets distributed around the hull are shown in Fig. 11, where the background grid size refinement in each coordinate axis has a $\sqrt{2}$-fold relationship, while other conditions keep the same. The grid count for each set is listed in Table 5.

For the RANS-based viscous flow calculation, since the hull is set as a boundary face of the calculation region, and if the grid refinement around the hull is different, the boundary layer of the hull has a difference, so the pressure and velocity distributions cannot be the same, resulting in different total drags and wake fraction reductions. Table 6 shows the calculation results and time for the three models' average total drag and axial wake fraction reduction, when setting calculation time $\Delta t = 0.001$ s and using high performance computing cluster with 40 processors. From Table 6, the calculation is monotonically convergent for three grids because the convergence rate R_G of both the total drag and axial wake fraction are between 0 and 1 (Liu et al., 2021c). Theoretically, a convergent solution can be obtained by performing a series of grid refinements, so the grid layout used is reasonable. Furthermore, the calculation time of the three grids shows that they have a relatively big difference, and if a single-fidelity model is constructed using a medium grid only, the cost can be high enough. Therefore, the Co-Kriging model can be used to comprehensively take advantageages of the finer-grid result as a high-fidelity result, while the coarser-grid result can be regarded as the low-fidelity result to save the total calculation time and resource (Liu et al., 2022).

2.3. New sample hull generation and surrogate model construction

To conduct the surrogate-based optimization of comprehensive performances of the JBC hull, the shape of the initial hull should be changed by the hull form deformation method. In this paper, the free-form deformation (FFD) method (Liu et al., 2018) is used to deform the stern shape, while the shapes of the bow and parallel middle body are not changed. The ranges of the three design variables are shown in Table 7.

By changing each of the three design variables only, the deformation effects are shown in Fig. 12. It should be noted that all the variable ranges are dimensionless values relative to the ship model waterline length. Keeping the green dots fixed in the FFD control lattice, variable x_1 corresponds to the amount of motion of the red (moveable) dots along the ship length in Fig. 12(a)~(c), x_2 to the amount of motion of the red dots along the ship breadth in Fig. 12(d)~(f), and x_3 to the amount of motion of the red dots along the ship breadth in Fig. 12(g)~(i). It should
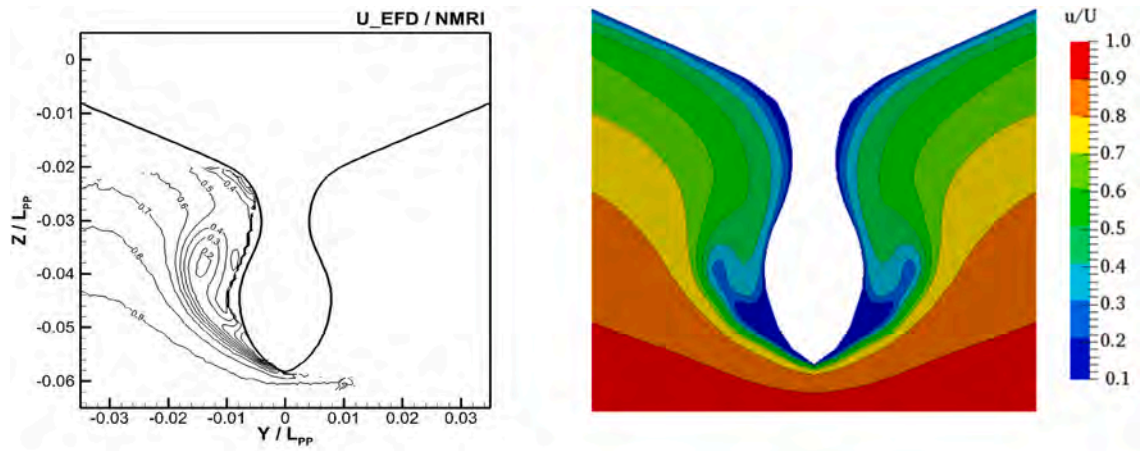
(a) $x/L_{PP} = 0.9625$
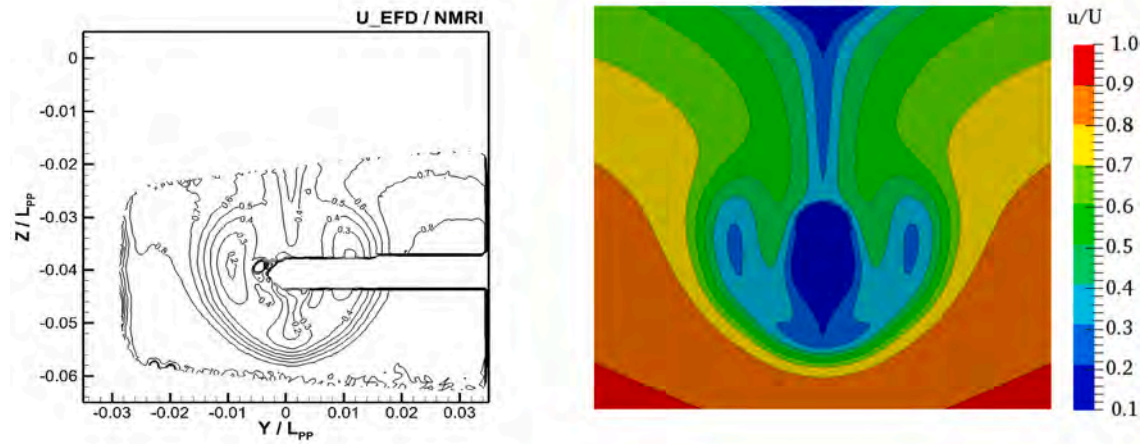
(b) $x/L_{PP} = 0.9843$
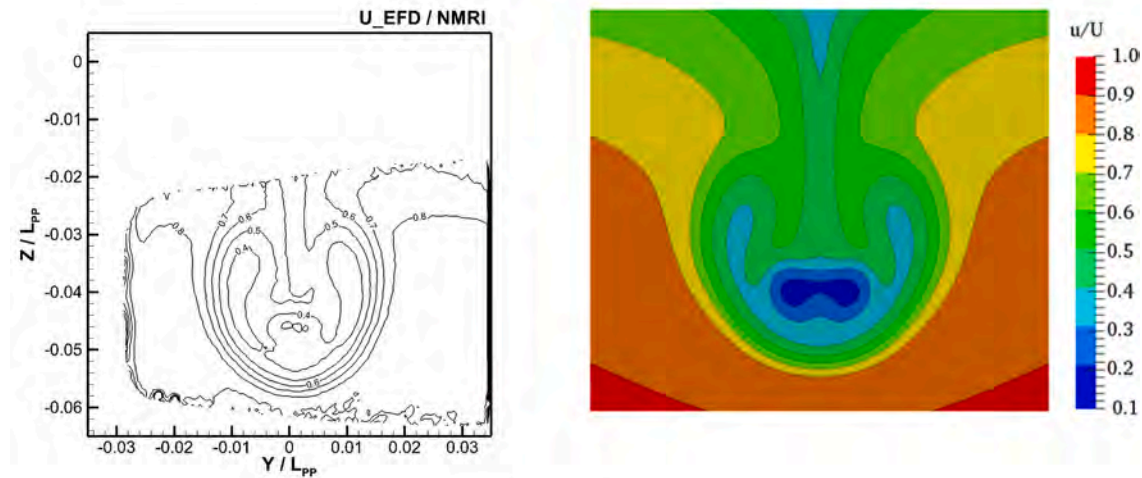
(c) $x/L_{PP} = 1$

Fig. 8. Comparisons of axial velocity distribution (u/U) (left: experiment; right: CFD).
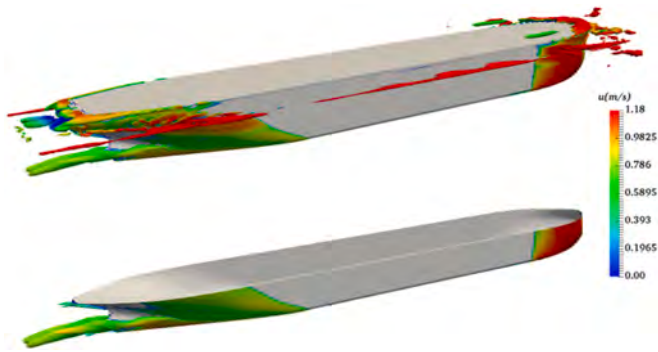
Fig. 9. Comparison of vorticity iso-surface ($Q = 20$) calculated by two methods (upper: two-phase flow; lower: single-phase flow).

also be noted that the movable control points along the ship breadth are moved symmetrically about the central sheer plane of the hull. The three deformation parameters control the degree of contraction of the longitudinal profile of the ship and the degree of contraction of the transverse profile of the two stern sections.

Furthermore, if we locate the red (moveable) points separately, for example, when moving some different single points in *x*, *y*, *z* directions, there will be a lot of design variables, which will definitely cost much more in numerical simulation. Furthermore, the new deformed sample hulls may not have good smoothness or even be in unreal shape, which is not beneficial to the optimization reliability. Therefore, local deformation at the stern of the JBC ship can comprehensively optimize its total drag R_t and reduction of axial wake fraction (1-*w*) at the propeller disk at the design speed $U = 1.179$ m/s (corresponding Froude number Fr = 0.142). In addition to the constraints that are the lower and upper bounds of the three design variables, the relative differences of the wet surface area (*S* and S_0 are for deformed and initial hulls) and drainage volume (∇ and ∇_0 are for deformed and initial hulls) are also constrained to meet the actual design requirements, and the overall optimization problem definition is shown below:
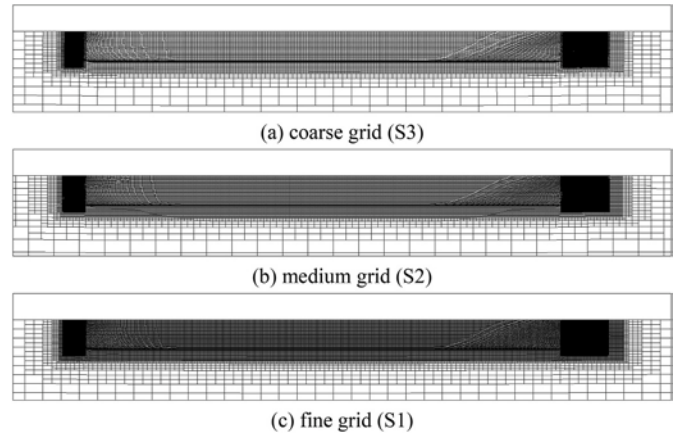

Fig. 11. Grid distributions for uncertainty analysis.

Table 5
Grid counts for the three single-phase sets.

Grid No.	Background grid number	Total grid number
S3	$70 \times 14 \times 28$	878722
S2	$100 \times 20 \times 40$	2519677
S1	$140 \times 28 \times 56$	6747658

Table 6
Grid convergence study for the JBC double-model calculation.

Grid No.	Total drag R_t (N)	Wake fraction reduction 1-*w*	Calculation CPU time (second)
S3	34.152	0.350	200,000
S2	32.702	0.347	750,000
S1	32.156	0.346	2,470,000
R_G	0.377	0.333	–

Table 7
Design variables and their ranges.

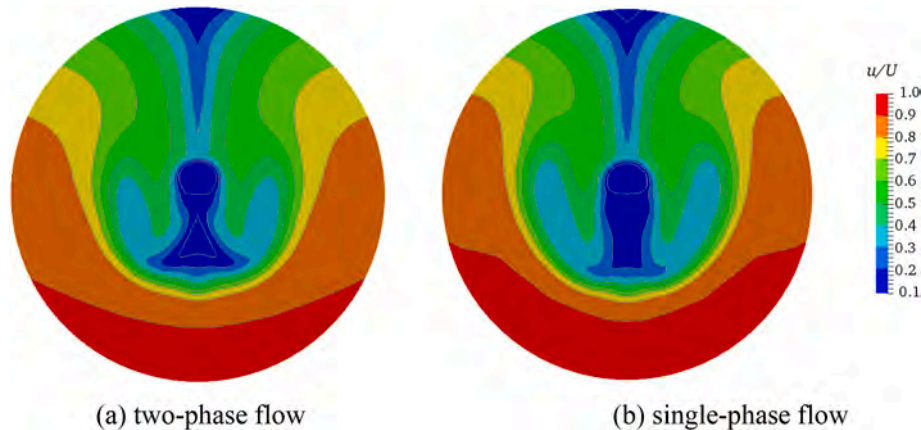| Design variable	Lower bound	Upper bound
x_1	−0.02	0.02
y_1 (x_2)	−0.05	0.05
y_2 (x_3)	−0.03	0.03

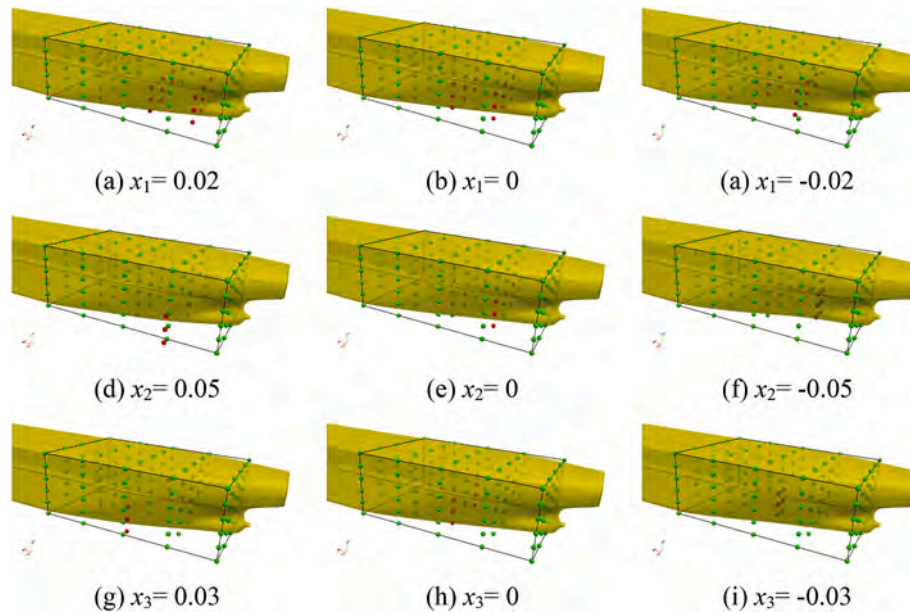**Fig. 10.** Comparisons of axial velocity distribution (*u/U*) near propeller disk.

Fig. 12. Stern deformation diagrams and effects of the JBC based on the FFD method.

design variables x_1, x_2, x_3

$$
\min \begin{cases} f_1 = R_t \\ f_2 = 1 - w = 1 - \dfrac{\overline{U}}{U} \end{cases}
$$

s.t.

$$-0.02 \leq x_1 \leq 0.02,$$

$$-0.05 \leq x_2 \leq 0.05,$$

$$-0.03 \leq x_3 \leq 0.03,$$

$$-1\% \leq \frac{S - S_0}{S_0} \leq 1\%,$$

$$-1.5\% \leq \frac{\nabla - \nabla_0}{\nabla_0} \leq 1.5\%.$$

(1)

where $\overline{U} = \frac{1}{N} \sum_{i=1}^{N} U_i$, and U_i represents the axial velocity of the i-th grid point, and N is the number of the grid points at the propeller disk.

The sample points can be selected using different sampling methods to ensure their uniformity and orthogonality in the design space, such as the Optimal Latin Hypercube Sampling (OLHS) method and Sobol (1979) method, to help ensure the precision of the surrogate model.

For instance, the Sobol method can meet accuracy requirements in constructing relatively high-precision surrogate model for high-dimensional functions, which has been validated in Wigley and DTMB-5415 hull's hydrodynamic performance optimization cases (Liu et al., 2021b, 2021c, 2022). Therefore, Sobol method is adopted as the design of experiments method to determine the design variable values of the new sample hulls. Before using the Co-Kriging model to construct the relationship between the hydrodynamic performance index and the hull design variables, an introduction to the Co-Kriging model (Sacks et al., 1989) is given here.

A two-fidelity Co-Kriging surrogate model can be established by two sets of independent samples with high and low accuracy. First, the sample data sets at two levels of fidelity are given: the high-fidelity samples $\mathbf{X}_e = [\mathbf{x}_e^{(1)}, \mathbf{x}_e^{(2)}, \cdots, \mathbf{x}_e^{(N_e)}]$ consisting of N_e samples and the low-fidelity samples $\mathbf{X}_c = [\mathbf{x}_c^{(1)}, \mathbf{x}_c^{(2)}, \cdots, \mathbf{x}_c^{(N_c)}]$ consisting of N_c samples, so the new sample set is $\mathbf{X} = [\mathbf{X}_c^T, \mathbf{X}_e^T]^T$, and the function value set is $\mathbf{y} = [\mathbf{y}_c^T, \mathbf{y}_e^T]^T$.

Like the Kriging model, the value of the function at a certain point in the space is regarded as a random process with a certain expectation and variance. Assume Gaussian processes Z_e and Z_c represent the approximation of low-fidelity and high-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 below:

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

The covariance based on the Kriging surrogate model is

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

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:

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

where n_{dv} is the number of design variables, and θ_k, p_k are spatially related parameters.

The two-fidelity 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 of the high-fidelity sample data, the log-maximum likelihood can be used to obtain the parameters $\mu_c, \sigma_c^2, \boldsymbol{\theta}_c, \mathbf{p}_c$ in low-fidelity Kriging model, which are just the construction process of the single-fidelity Kriging surrogate model to build the approximation of Z_c with the low-fidelity sample data in the Co-Kriging model.

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

$$\mathbf{d} = \mathbf{y}_e - \rho \mathbf{y}_c(\mathbf{X}_e) \tag{5}$$

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 $\mathbf{y}_c(\mathbf{X}_e)$ values are the real values of the low-fidelity function. At this point, the log-maximum likelihood function of the approximate function for high-fidelity samples can be used to do the estimations similar with the above process.

Finally, the best-predicted value $\hat{y}_e(\mathbf{x})$ of the to-be-predicted sample point, which is just an arbitrary point in the variables space, can be obtained:
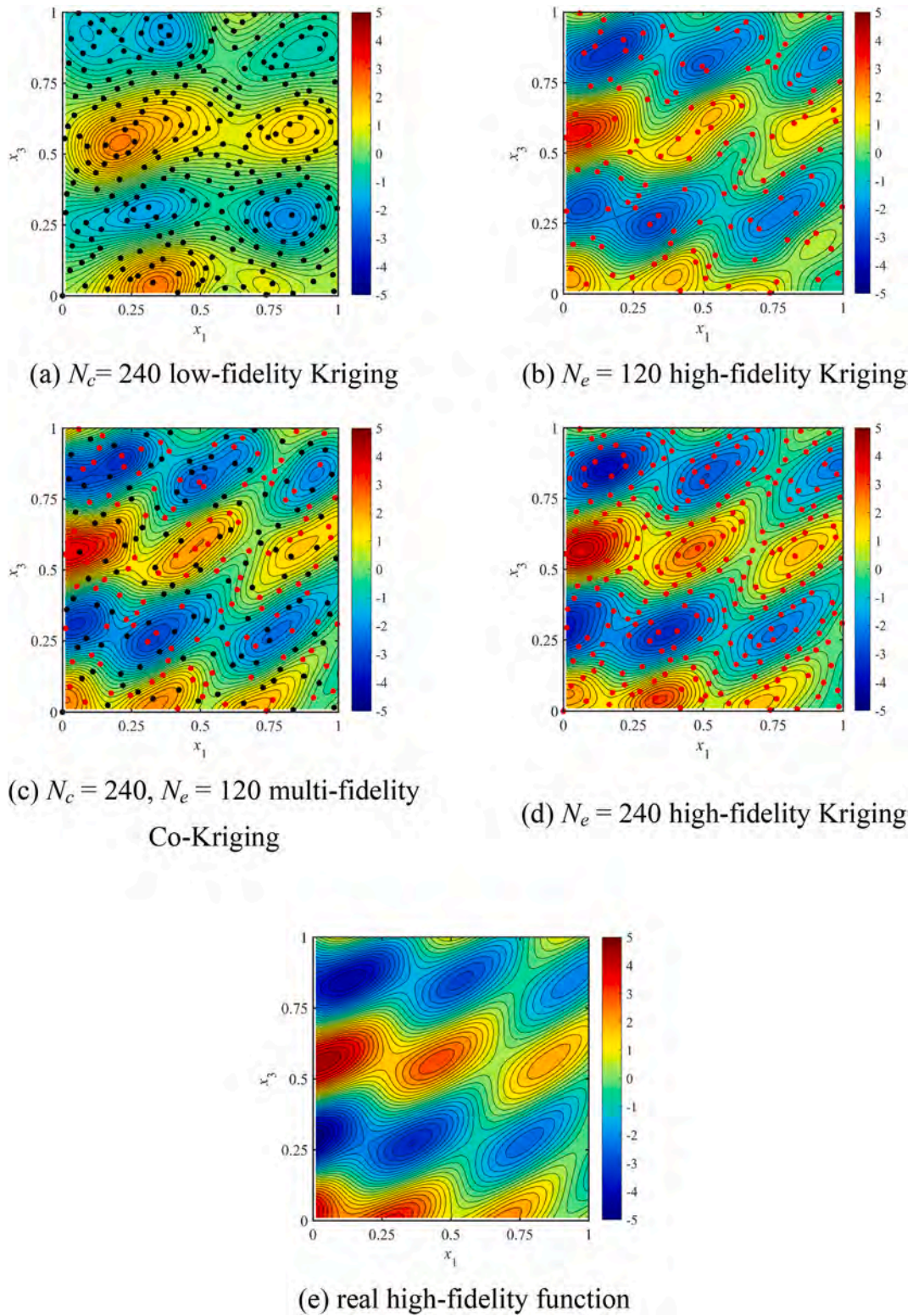
(a) $N_c = 240$ low-fidelity Kriging

(b) $N_e = 120$ high-fidelity Kriging

(c) $N_c = 240$, $N_e = 120$ multi-fidelity Co-Kriging

(d) $N_e = 240$ high-fidelity Kriging

(e) real high-fidelity function

Fig. 13. Comparison between different surrogate models and the real function.

$$\hat{y}_e(\mathbf{x}) = \hat{\mu} + \mathbf{c}^T \mathbf{C}^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu}) \tag{6}$$

where $\hat{\mu}$ is the hyper-parameter obtained by log-maximum likelihood estimation, $\mathbf{c}$ is the covariance of $\mathbf{X}$ and $\mathbf{x}$, and $\mathbf{C}$ is the overall covariance matrix.

Similar to the single-fidelity Kriging surrogate model, it can be proved that the above Co-Kriging surrogate model is an interpolation of the existing high-fidelity sample data. However, it performs a regression to the low-fidelity sample data $\mathbf{X}_c$ except for $\mathbf{X}_e$.

One test function considered here contains three independent variables (Chen et al., 2021), and the high- and low-fidelity functions are given as:

Table 8
Error analysis of surrogate model 1.

	regpoly0	regpoly1	regpoly2	
corrgauss	0.0810	0.0834	0.0488	AAE
	0.4067	1.0634	0.5406	MAE
	0.1096	0.1665	0.0898	RMSE
corrspline	0.0770	0.0588	0.0439	AAE
	0.9236	0.9348	0.5414	MAE
	0.1437	0.1330	0.0856	RMSE

Table 9
Error analysis of surrogate model 2.

	regpoly0	regpoly1	regpoly2	
corrgauss	0.0032	0.0036	0.0023	AAE
	0.0262	0.0476	0.0311	MAE
	0.0058	0.0072	0.0047	RMSE
corrspline	0.0026	0.0026	0.0020	AAE
	0.0493	0.0452	0.0312	MAE
	0.0068	0.0062	0.0046	RMSE

Table 10
Error analysis of surrogate model 3.

Number of high-fidelity samples N_e						
15	20	30	40	50	60	
0.1197	0.1621	0.1343	0.1048	0.0719	0.1070	AAE
0.4028	0.3564	0.3215	0.3749	0.3055	0.3046	MAE
0.1729	0.1539	0.1147	0.1043	0.0829	0.1015	RMSE

Table 11
Error analysis of surrogate model 4.

Number of high-fidelity samples N_e						
15	20	30	40	50	60	
0.0051	0.0037	0.0047	0.0041	0.0041	0.0047	AAE
0.0119	0.0108	0.0153	0.0148	0.0125	0.0123	MAE
0.0064	0.0053	0.0061	0.0052	0.0048	0.0048	RMSE

$$f_e(\overrightarrow{\mathbf{x}}) = e^{1.4(x_2 - x_1)} \cos\left(\frac{7\pi x_3}{2}\right) + \sin[5\pi(x_3 - x_1)] \tag{7}$$

$$f_c(\overrightarrow{\mathbf{x}}) = [1 + 1.4(x_2 - x_1)]\cos\left(\frac{7\pi x_3}{2}\right) + \sin\left[5\pi(x_3 - x_1)^2\right] \tag{8}$$

The low-fidelity function performs a low-order Taylor expansion of some terms and elevates the degree of the other terms in the high-fidelity function. Although the real (high-fidelity) function is ternary, it has strong nonlinearity. Therefore, this example can be considered a true reflection of the difference between results obtained through high- and low-fidelity numerical simulations in the actual hull form optimization process. The low- and high-fidelity function sample sizes selected by the Sobol method are 240 and 120, respectively. Setting $x_2 = 1$, the cloud maps of the function predicted by the Kriging and Co-Kriging surrogate models, together with the real function, are shown in Fig. 13.

Fig. 13 (a) is the surrogate model for the low-fidelity function, which indeed has difference from the true (high-fidelity) function. Furthermore, surrogate model in Fig. 13 (b) has small difference, compared to the real function in Fig. 13 (e), however, the "mode" correspondence between the two is somewhat different. For the Co-Kriging model in Fig. 13 (c), although there are 360 sample points, it should be noted here that it can be regarded as practical evaluation of hydrodynamic performance with different fidelity; therefore, the computation time for the two fidelities is much different. Using less time in total, the Co-Kriging model can give a more realistic reflection of the changing trend of the

objective function. Last but not least, although Kriging model with 240 high-fidelity samples are constructed in Fig. 13 (d), which has smaller difference, the computational cost can be very high for the real hull form optimization case.

Therefore, a multi-fidelity Co-Kriging model with higher accuracy can be established by using fewer high-fidelity samples and more low-fidelity samples. It can be predicted that more reliable optimization results would be obtained using Co-Kriging for hull form optimization problems and that the total calculation costs will be reduced to some degree compared with a Kriging model using more high-fidelity samples. Of course, this depends on the difference in computational cost between the two fidelity samples while comprehensively considering the calculation efficiency and accuracy.

Traditionally, we use 10 times or more new sample hulls to construct single-fidelity surrogate models for hull form optimization problems. Considering that the optimization objective has two in this paper, we decide to use 60 samples to ensure the accuracy of the two single-fidelity Kriging models. Furthermore, the maximum numbers of high-fidelity and low-fidelity samples are also 60 for Co-Kriging models. The high-fidelity sample data are the total drag and wake fraction reduction of each sample hull calculated using the medium grid (S2), while the low-fidelity sample data are the total drag and wake fraction reduction of each sample hull calculated using the coarse grid (S3). Here, it should be noted that the S3 or S2 grid for each new sample hull can be kept similar by using consistent grid settings to minimize systematic errors.

The single-fidelity Kriging model and the multi-fidelity Co-Kriging model can approximate the two objective functions and the design variables using the above-described high- and low-fidelity sample data. The single-fidelity Kriging models for determining the design variables based on total drag and wake fraction reduction are designated as surrogate models 1 and 2, respectively. The multi-fidelity Co-Kriging models based on total drag and wake fraction reduction for determining the design variables are designated as surrogate models 3 and 4, respectively.

Three main error indicators of each surrogate model can be determined through leave-one-out cross-validation to measure the error of the surrogate model with a sample number N_s quantitatively, namely the Average Absolute Error (AAE), Maximum Absolute Error (MAE), and Root Mean Square Error (RMSE):

$$\text{AAE} = \frac{1}{N_s} \sum_{i=1}^{N_s} |\widehat{y}_i - y_i| \tag{9}$$

$$\text{MAE} = \max_{1 \leq i \leq N_s} |\widehat{y}_i - y_i| \tag{10}$$

$$\text{RMSE} = \frac{1}{N_s} \sqrt{\sum_{i=1}^{N_s} (\widehat{y}_i - y_i)^2} \tag{11}$$

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

In total, the three error indicator values for Kriging models 1 and 2, constructed using the 60 high-fidelity samples, are small when the correlation function and deterministic polynomial function are the Gaussian correlation function and zero-order deterministic polynomial ("corrgauss-regpoly0") (see Table 8).

Therefore, to facilitate further discussion, the same correlation function and deterministic polynomial form are adopted for the two Co-Kriging models 3 and 4 (see Table 9). On the premise that the number of low-fidelity samples is 60, the influences of different high-fidelity samples on the main error indicators are discussed, as listed in Tables 10 and 11 for models 3 and 4, respectively.

From the three error indicators of models 3 and 4, it can be seen that, when changing the number of high-fidelity samples from 10 to 60, if the high-fidelity-sample number is small, the error may be relatively big,
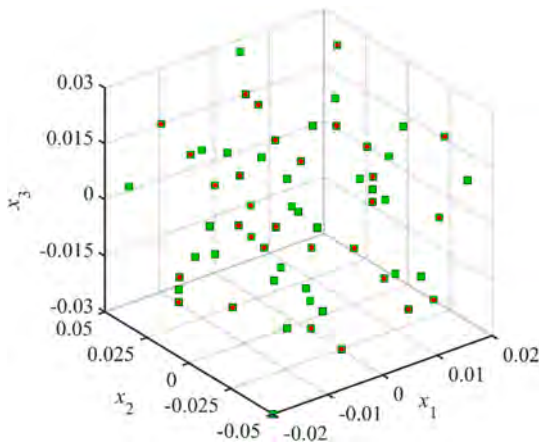
Fig. 14. Sample points for low-fidelity (green) and high-fidelity (red) evaluations. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

while if we choose too many high-fidelity samples, the total cost will be high. Therefore, comprehensively considering the calculation efficiency of hydrodynamic evaluation and the accuracy of the constructed Co-Kriging surrogate model, the number of high-fidelity samples can be chosen as 30. The distribution of high-fidelity (solid red circles) and low-fidelity (solid green squares) sample points is shown in Fig. 14, where the high-fidelity data are obtained from the medium grid (S2) evaluation results, and the low-fidelity data are obtained from the coarse grid (S3) evaluation results. It is not difficult to see that the 30 high-fidelity samples are subsets of the 60 low-fidelity samples and are evenly distributed in the design space to reduce the additional interpolation errors of the Co-Kriging model.

The similarities and differences of the flow field calculation results for two sets of grids are investigated, taking the original hull as an example. Fig. 15 shows the high- and low-fidelity results of axial velocity distribution at the propeller disk. It is evident that the general velocity distribution trend is mostly consistent, with slight differences near the longitudinal plane, such as the hook-shaped low-speed area, and the velocity isolines obtained from the coarse grid are less smooth than from the medium grid.

Fig. 16 presents the high- and low-fidelity results for the vorticity iso-surface distribution ($Q = 20$) around the hull, showing that the general trend is also mostly consistent. The vorticity iso-surface is only slightly different near the ship's stern, and the coarse grid vorticity iso-surface is somewhat less smooth than the medium grid.

To sum up, the coarse and medium grid results' strong correlation is obvious in the flow field comparisons, which is beneficial to ensure the precision of the multi-fidelity surrogate model.

2.4. Optimization results and analysis

According to the Kriging models 1 and 2 and the Co-Kriging models 3 and 4 discussed in the preceding section, the second-generation Non-dominated Sorting Genetic Algorithm (NSGA-II) (Liu et al., 2020), is used to solve the multi-objective hydrodynamic performance optimization problem. Table 12 shows the main optimization parameters.

Based on the constructed Kriging and Co-Kriging surrogate models, optimal Pareto solution sets of the optimization problem can be obtained using the NSGA-II, and each point on the Pareto front (corresponding to each new hull form) has different degrees of performance improvement (objective function value decrease) compared with the initial hull. The Pareto front iteration processes based on Kriging and Co-Kriging surrogate models are given in Fig. 17.

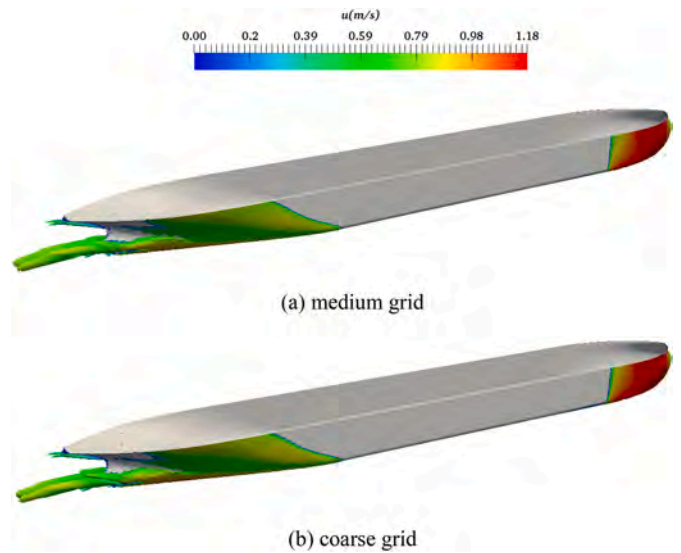Two typical optimal solutions are selected from each Kriging-based

Fig. 16. Comparison of $Q = 20$ iso-surface calculated using two grids.

Table 12
Multi-objective optimization parameters setup for the JBC.

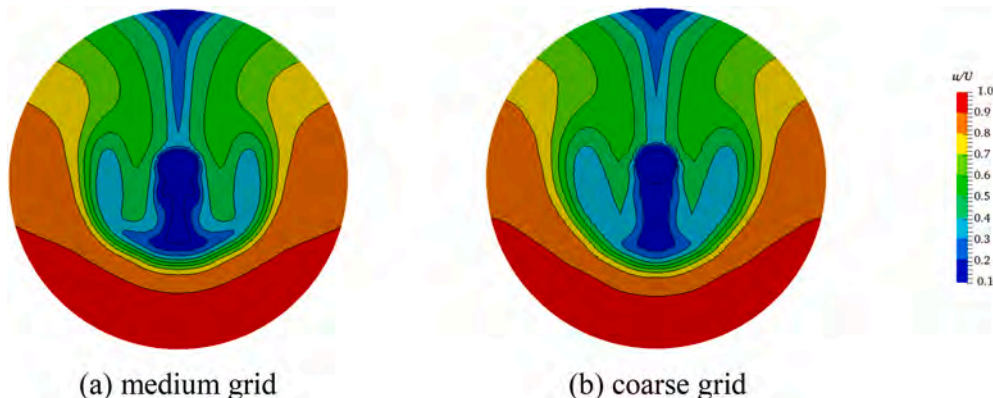| Parameter	Value
Population size	200
Maximum iteration	150
Crossover rate	0.8
Mutation rate	0.2
Optimal population ratio	0.9

Fig. 15. Comparison of velocity distribution at the propeller disk for two grids.
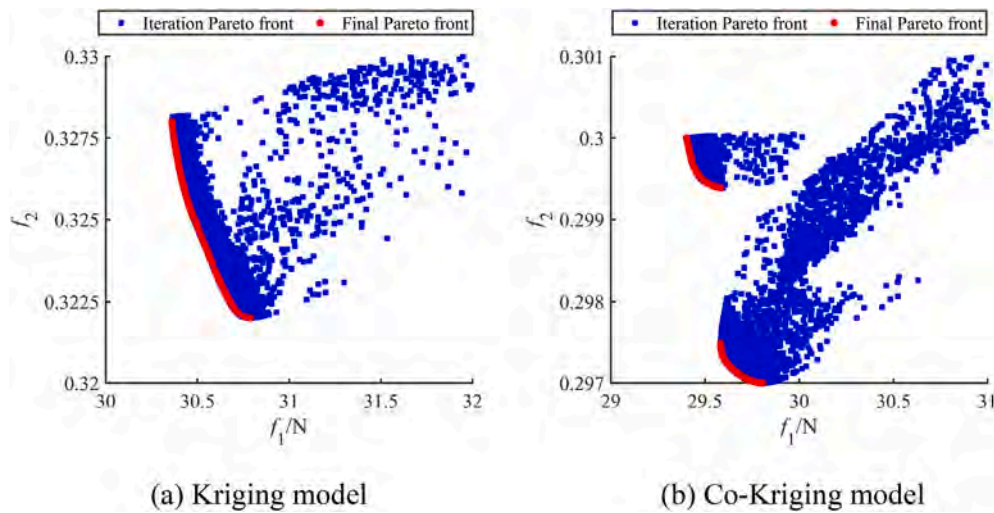
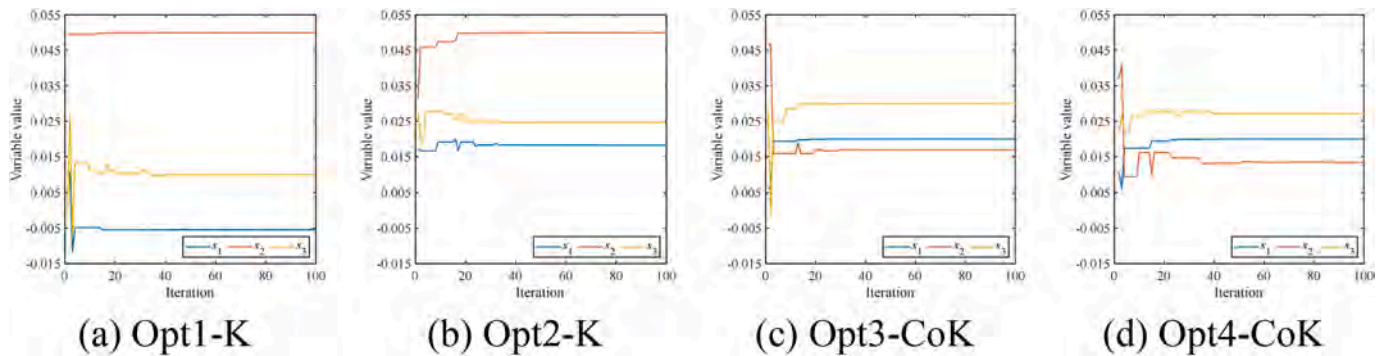**Fig. 17.** Pareto front iteration process.

Fig. 18. Convergence diagram for the three design variables for the four optimal hulls.

and Co-Kriging-based Pareto front to verify the reliability of optimization results, so a total of four optimal hulls (Opt1-K, Opt2-K, Opt3-CoK, Opt4-CoK) can be obtained, and the convergence diagram for the three design variables for the four optimal hulls are given below in Fig. 18.

The initial and optimal hulls' transverse and sheer body plans are compared, as shown in Figs. 19 and 20.

There is a relatively large difference in the hull lines for the two optimal hulls obtained based on the Kriging model (Opt1-K and Opt2-K). However, the difference in hull lines is relatively small for the two optimal hulls based on the Co-Kriging model (Opt3-CoK and Opt4-CoK). However, the profile generally has a contraction, which gradually changes the stern section from a "U" shape to a "V" shape. The detailed design variable values for the four typical optimal hulls are provided in Table 13.

From Table 13, Opt1-K represents the optimal hull based on the Kriging surrogate model 1 for total drag. However, because of the Kriging model error, the CFD prediction result of its total drag is slightly larger than for the Kriging model and even larger than the total drag of Opt2-K. Although a 4.57% drag decrease is achieved, it can be inferred that the Kriging model has a relatively large error in the vicinity of the real optimal solution, so there is still a certain distance between the obtained and the real optimal hulls. Opt2-K represents the optimal hull based on Kriging model 2 for wake fraction reduction. However, because of the Kriging model error, the CFD prediction result of the wake fraction reduction is slightly larger than for the Kriging model, although there is still a reduction of 7.20%, which is conducive to improving the propeller propulsion efficiency.

Opt3-CoK represents the optimal hull based on Co-Kriging surrogate

model 3 for total drag. Although the CFD prediction result of its total drag is slightly larger than for the Co-Kriging model, it has better resistance performance than Opt1-K, and the total drag is optimized by 5.67%. This indicates that the Co-Kriging model can find a better optimal solution than the Kriging model. Opt4-CoK is the optimal hull based on Co-Kriging model 4 for wake fraction reduction. Compared with the initial hull, the wake fraction reduction is optimized by 10.37%, which is better than Opt2-K so that the propeller propulsion efficiency can be significantly improved.

Similar with the mathematical function validation case given above, it can be inferred that the Co-Kriging model can well capture the changing trend of the design space, so the optimal design variables obtained by Co-Kriging model can be regarded to be closer to the real minimum of the objective function and Co-Kriging model better reflects the changing trend around the real minimum. At the same time, the overall computation time has a 25% reduction.

In addition, the wet surface areas of the four optimal hulls Opt1-K, Opt2-K, Opt3-CoK, and Opt4-CoK are 12.37, 12.35, 12.36, and 12.37 m^2, respectively, and the drainage volumes are 2.769, 2.752, 2.763 and 2.765 m^3, respectively. Therefore, the relative changes in wet surface area and drainage volume for the optimal hulls are no more than 0.6% and 1.3%, respectively, which meet the actual optimization requirements.

Fig. 21 compares the vorticity iso-surface distributions around the initial and optimal hulls. The area of the vorticity iso-surface near the ship stern decreases, indicating that the overall vorticity of the ship stern decreases to varying degrees. This leads to a decrease in the ship's viscous pressure drag and total drag, confirming the reliability of
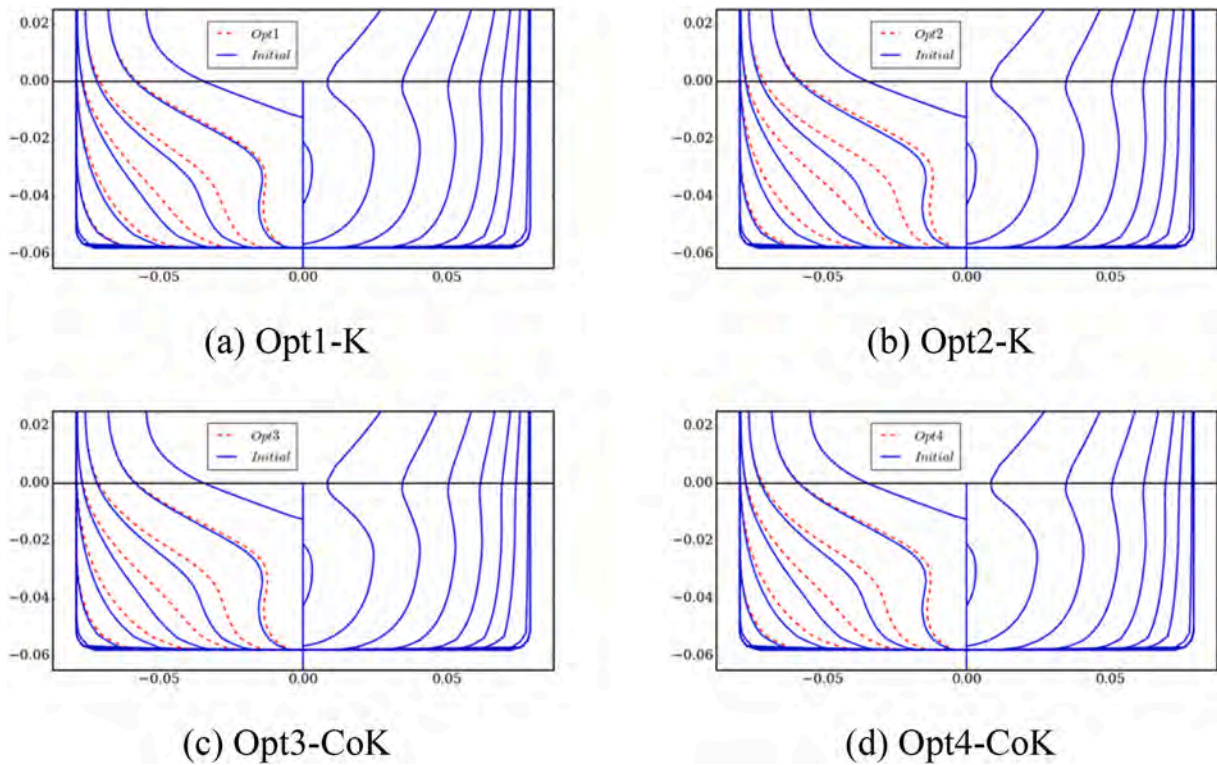
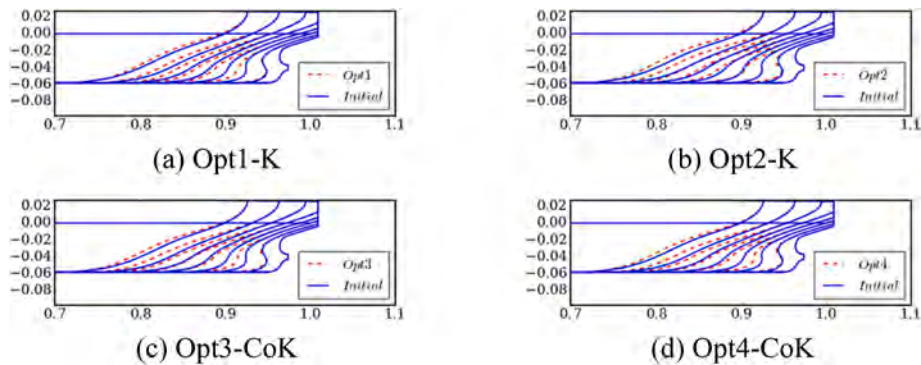**Fig. 19.** Transverse line comparison of the initial and optimal hulls.

Fig. 20. Sheer line comparison of the initial and optimal hulls.

Table 13
Summary of optimal results.

/	Design variable value			Total drag R_t			Wake fraction reduction 1-w		
	x_1	y_1	y_2	Model prediction (N)	CFD calculation (N)	Reduction ratio	Model prediction	CFD calculation	Reduction ratio
Initial	0	0	0	/	32.702	/	/	0.347	/
Opt1-K	−0.005	0.05	0.010	30.361	31.206	4.57%	0.325	0.328	5.48%
Opt2-K	0.018	0.05	0.025	30.787	30.532	6.64%	0.308	0.322	7.20%
Opt3-CoK	0.02	0.017	0.03	29.402	30.847	5.67%	0.300	0.319	8.07%
Opt4-CoK	0.02	0.013	0.027	29.806	31.021	5.14%	0.297	0.311	10.37%

optimization results. Opt3-CoK clearly has a smaller vorticity iso-surface area than Opt1-K.

The dynamic pressure distribution comparison of the initial and optimal hulls is given in Fig. 22. Because the bow shape of the JBC hull has not been changed, the pressure distribution at the ship bow is entirely consistent. However, the low-pressure region's area and amplitude at the stern bottom and contraction section have different degrees of decrease for each optimal hull. This causes the differential

pressure between bow and stern to decrease, resulting in a total drag decrease, also confirming the optimization results' reliability.

Fig. 23 compares the axial velocity distributions at the propeller disk for the initial and optimal hulls. The velocity distributions of the optimal hulls have significant changes compared with the initial hull, and the decreases in area and amplitude of the low-speed region lead to overall average axial velocity increases. Therefore, the wake fraction reduction decreases, conducive to propeller propulsion efficiency, and Opt4-CoK
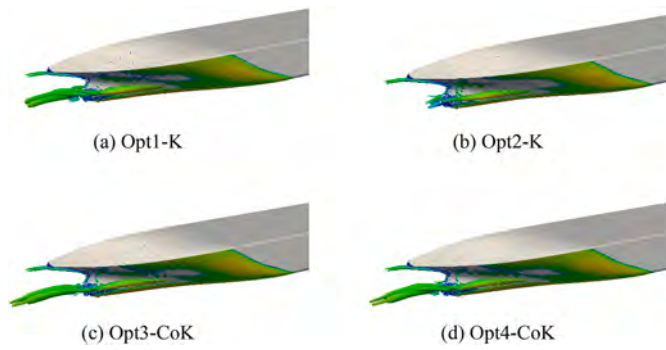
Fig. 21. Comparison of $Q = 20$ iso-surfaces around the optimal hulls.
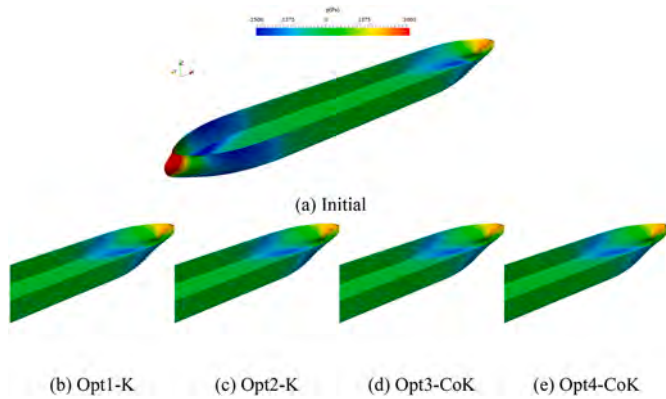

Fig. 22. Comparison of pressure distributions of the initial and optimal hulls.

has the highest average axial velocity.

In summary, the multi-fidelity Co-Kriging surrogate model can be established by using the viscous flow evaluation method with coarse and medium grids. The optimal hulls obtained by the multi-fidelity Co-Kriging model have better comprehensive resistance and wake performance than those from the single-fidelity Kriging model. Therefore, the optimization effects are more significant, and the Co-Kriging model is shown to be a data-driven model by investigating the deviation of the two-fidelity sample data and achieving a relatively high-precision surrogate model.

It should be noted that the Kriging model employs 60 high-fidelity samples, while the Co-Kriging model adopts 60 low-fidelity samples and 30 high-fidelity samples. In terms of surrogate model accuracy, when there are more objective functions, it is suggested that the number of low-fidelity samples should be slightly increased to construct the Co-Kriging model according to actual demand, and the number of high-fidelity samples can be adjusted in the same way.

Last but not the least, in terms of computational efficiency, for this optimization case, a high-fidelity sample takes about 2 days to evaluate, while a low-fidelity sample takes only about half a day. Therefore, the Co-Kriging model constructed in this case has a much lower computational cost while still producing optimal hulls that have much better comprehensive hydrodynamic performance. This again reflects the Co-Kriging model's superiority over the Kriging model in terms of computational efficiency and accuracy.

3. POD-based wake field learning of the JBC ship hull

In the latter stages of hull form optimization, flow field dimensionality reduction learning can be performed to fully use the flow field results of the new sample hulls using linear reduced order methods, such as POD method (Diez et al., 2015; Liu et al., 2021c).

In POD method, for all samples in the database, the average wake field $\bar{x}$ can be got at first, and regard the wake field of any sample hull form as a linear superposition of the average field and an infinite number of unit orthogonal basis modes as follows:

$$x_i = \bar{x} + \sum_{j=1}^{\infty} a_{ij} u_j \tag{12}$$

where

$$a_{ij} = (x_i - \bar{x})^T u_j, \quad (u_i, u_j) = \delta_{ij} \tag{13}$$

If only M (finite) basis modes $u_1, u_2, \cdots, u_M$ are selected, the reconstructed wake field x_i^{rec} can be expressed as

$$x_i^{rec} = \bar{x} + \sum_{j=1}^{M} a_{ij} u_j \tag{14}$$

Because of the truncation of the number of series items, a "cutoff error" $E_M(u_1, u_2, \cdots, u_M)$ can be measured by the square sum of the modulus length of the error vector in Euclidean space, namely

$$E_M(u_1, u_2, \cdots, u_M) = \sum_{i=1}^{N} \left\| x_i - x_i^{rec} \right\|^2 \tag{15}$$

Denote the difference between the wake field of each sample hull form x_i and the average field $\bar{x}$ as $\widetilde{x}_i = x_i - \bar{x}$, and define the matrix $\widetilde{X} = (\widetilde{x}_1, \widetilde{x}_2, \cdots, \widetilde{x}_N)$, then

$$\widetilde{X}\widetilde{X}^T = \sum_{i=1}^{N} (x_i - \bar{x})(x_i - \widetilde{x})^T \tag{16}$$

Assuming that M is certain, the constrained optimization problem on the basis $u_1, u_2, \cdots, u_M$ can be defined and optimized by the Lagrange multiplier method, and we have

$$\widetilde{X}\widetilde{X}^T u_j = \lambda_j u_j \tag{17}$$

That is to say, the basis should be the eigenvectors corresponding to the eigenvalues of matrix $S = \widetilde{X}\widetilde{X}^T$. From singular value decomposition of matrix $\widetilde{X}$:

$$\widetilde{X} = UDV^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T \tag{18}$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, we have

$$\widetilde{X}\widetilde{X}^T u_j = \sigma_j^2 u_j \tag{19}$$

where the basis u_j happens to be the j-th column of the matrix U obtained by singular value decomposition of $\widetilde{X}$, and the reconstructed field can be obtained by Eq. (14).

The dimensionality reduction learning framework for the wake field is shown in Fig. 24.

Constructing a flow field database based on the viscous flow evaluation results for flow field learning by dimensionality reduction methods in the hull form optimization field has great importance and practical significance. Based on the viscous flow performance optimization results for the JBC hull using the medium grid (S2), the viscous wake flow field dimensionality reduction learning process and results are given below.

The flow field database is quite different from potential flow evaluation results because the grid topological relations at the propeller disk for different sample hulls may not be entirely consistent because there are no regular grid topological relations. The volumetric grid is usually unstructured in viscous flow CFD calculations based on the Finite Volume Method (FVM). However, the POD method for free-surface and hull pressure field learning in Liu et al. (2021c) assumes that the grid topological relations in the new sample hull geometric/physical field database are exactly the same. Therefore, when viscous flow results are used to perform flow field learning, it is necessary to preprocess the flow field
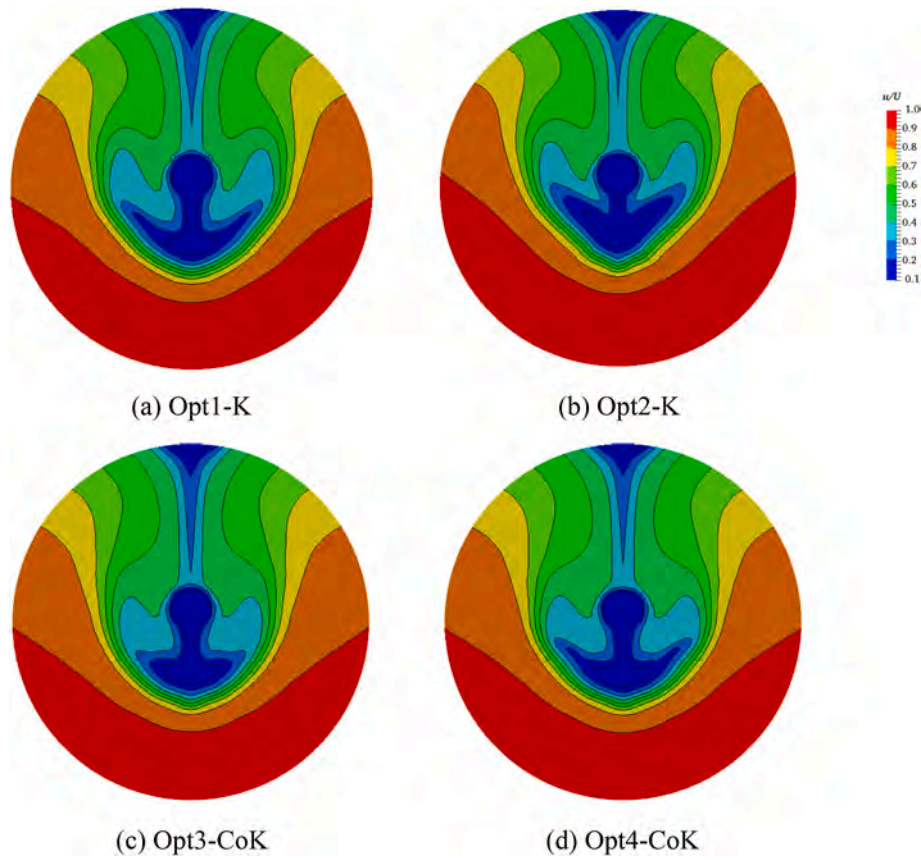
(a) Opt1-K

(b) Opt2-K

(c) Opt3-CoK

(d) Opt4-CoK

Fig. 23. Comparison of velocity distributions at the propeller disk of the optimal hulls.
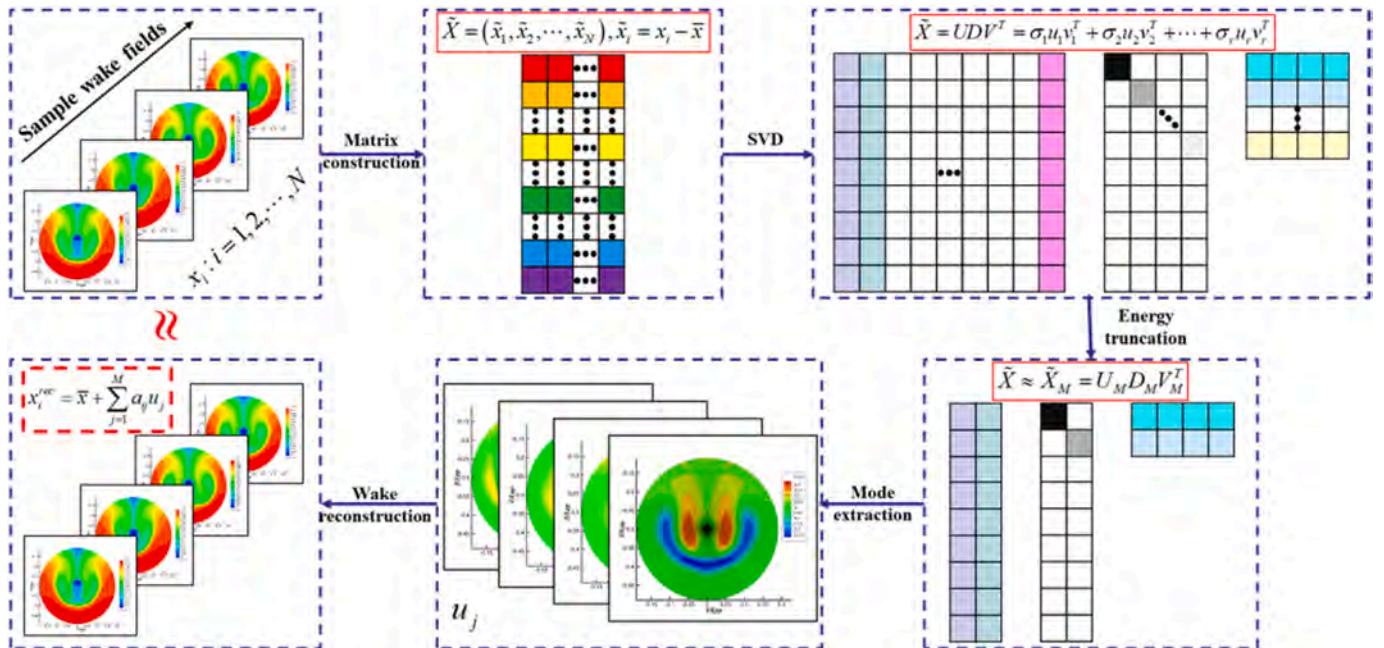

Fig. 24. Framework for POD-based wake field learning.

data of each sample hull to conduct flow field learning and rapid prediction using the POD method.

A new two-dimensional mesh can be established, and a wake field interpolation or other mapping by the Kriging model can be performed from each original (unstructured) grid to a new (structured) grid to ensure that the wake field data for each new sample hull has the same grid topological relationship. Since the axial velocity field region of interest is a circle at the propeller disk, the new grid can be established using polar coordinates.

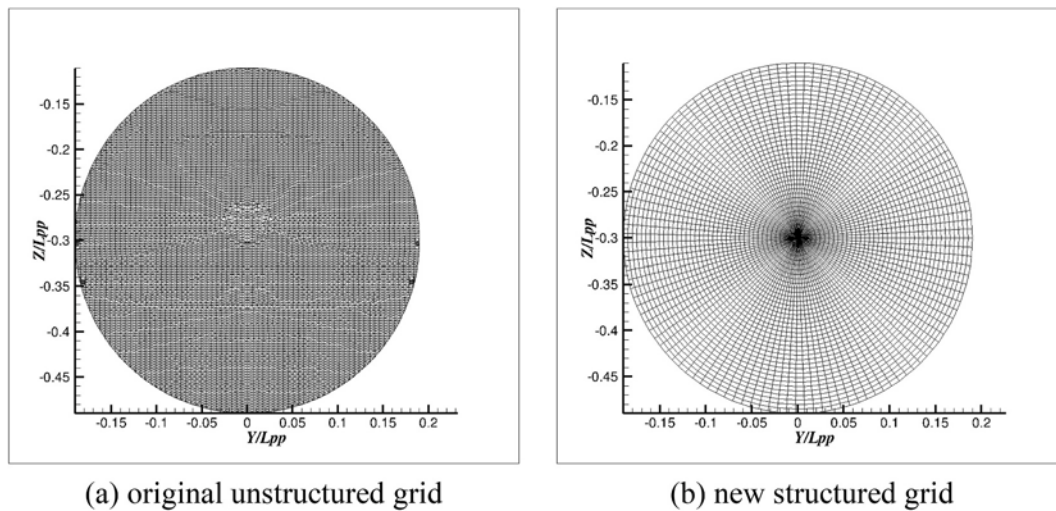Fig. 25 shows sample original unstructured and new structured two-

(a) original unstructured grid

(b) new structured grid

Fig. 25. Original and mapping wake field grids.

(a) original (unstructured) grid
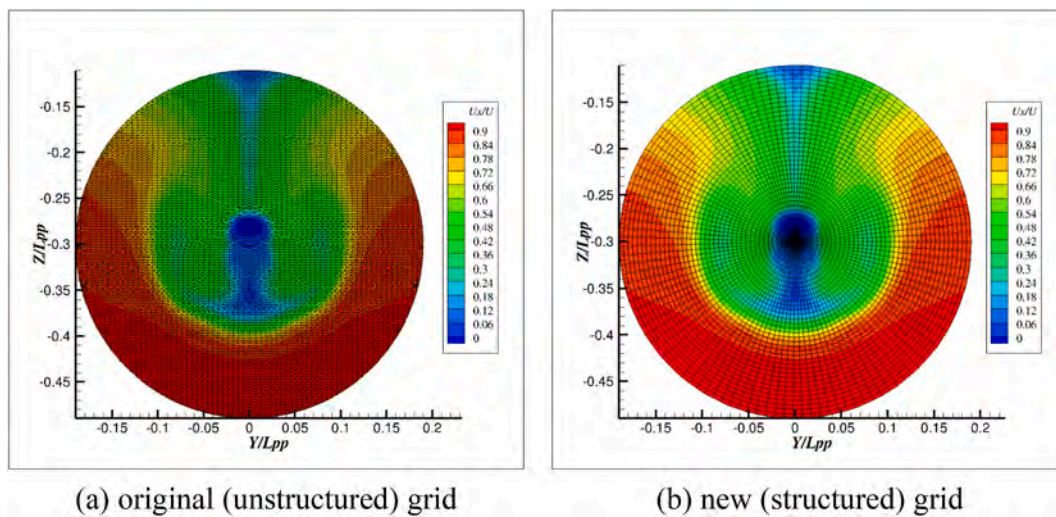
(b) new (structured) grid

Fig. 26. Comparison of wake fields before and after mapping.
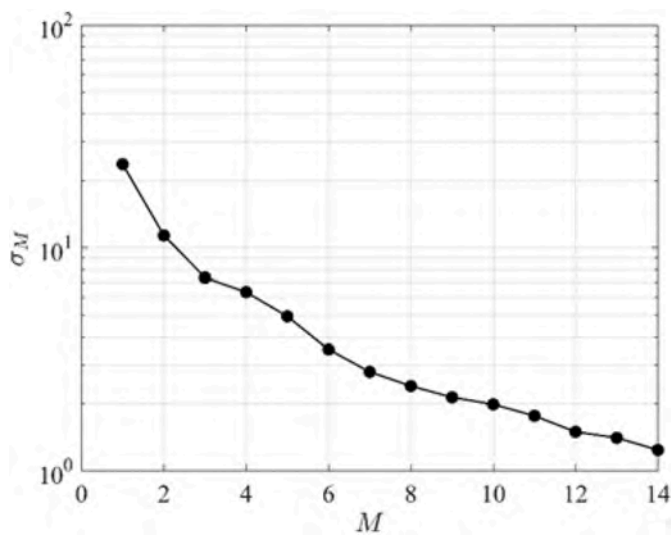

Fig. 27. Changes in the square root of eigenvalues corresponding to each mode.
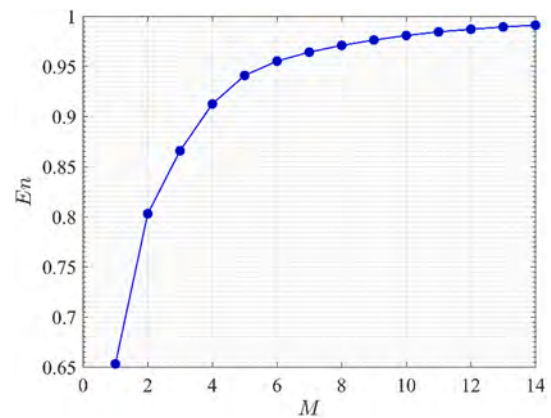

Fig. 28. Cumulative energy of the first several modes.

dimensional grids of the wake field at the propeller disk. Fig. 25(b) shows the "O-type" grid adopting 40 radial × 120 circumferential grid points, totaling 4800 grid points used for dimensionality reduction learning of the wake field.

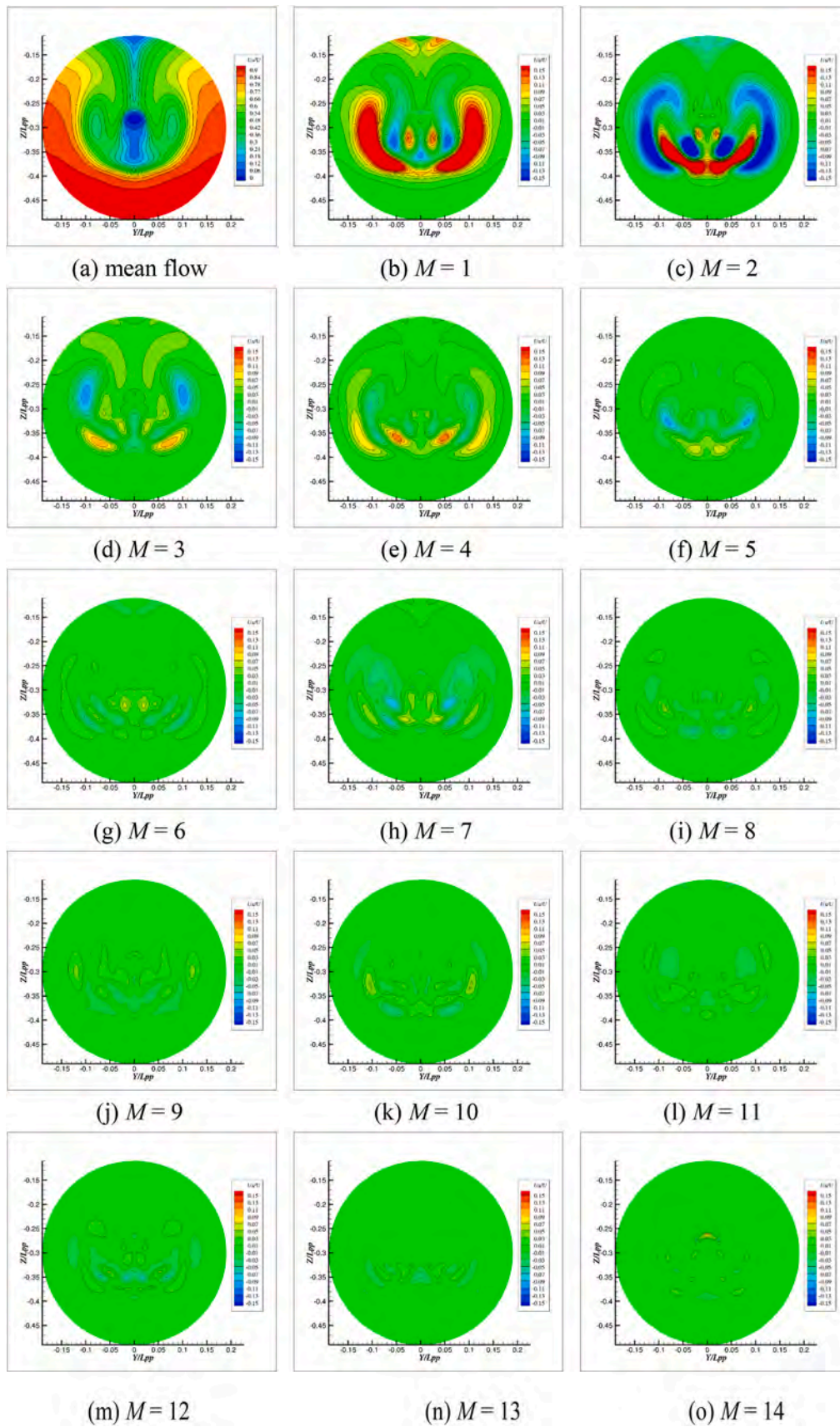**Fig. 29.** Average wake field and fields corresponding to the first 14 modes.
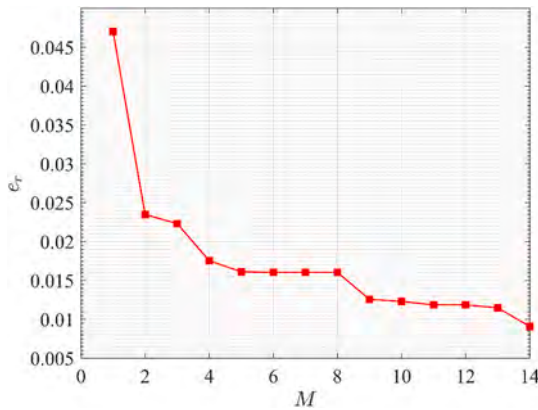
Fig. 30. Relationship between the relative error of the wake field of a new hull and number of modes *M*.

In addition, the physical field is made dimensionless, that is, the axial velocity U_i of each grid point at the propeller disk is normalized by the main flow velocity, which is the sailing speed of the ship $U = 1.179$ m/s.

$$u_i = \frac{U_i}{U} \tag{20}$$

where *N* is the number of the structured grid points.

As mentioned previously, the wake field on the original grid has a strong spatial correlation between grid points, and since the number of the original grid points for a certain velocity field is about 3000, it is quite enough for constructing a high-precision Kriging model that can comprehensively account for the global and local spatial correlation of the data in the region. Therefore, a Kriging model in the propeller disk region can be constructed using the original two-dimensional grid coordinates and their corresponding dimensionless axial velocity values. The dimensionless axial velocity values at the new grid points can then be predicted and mapped. Furthermore, from the comparison figure of the field before and after mapping.

Taking one of the sample hulls as an example, Fig. 26 compares the dimensionless wake field before and after mapping. It can be seen that the wake field is nearly identical before and after mapping, indicating that the error due to the mapping operation proposed in this paper is rather small. As a result, the mapped field can be used to replace the original field to do the POD implementation and field forecasting of a new sample hull.

The wake field evaluation results for 60 sample hulls with the medium grid (S2) are preprocessed as above to establish the wake field database, and the POD-based wake field dimensionality reduction learning is then conducted.

First, the mode order *M* required for flow field learning is determined. Figs. 27 and 28 show the arithmetic square root of the eigenvalues corresponding to each mode contained in the database (arranged from largest to smallest) and the cumulative energy at different modes, respectively.

According to the POD results and the energy contained in each basis mode shown in Fig. 28, the first 14 basis modes capture 99% of the total wake field energy of the new hulls, where each basis mode has the same structured grid topology as the 60 mapped axial velocity fields. The average dimensionless axial velocity field and the first 14 basis modes are shown in Fig. 29.

Except for the average field, the dimensionless axial velocity field corresponding to each basis mode is small in value, and the decrease becomes more evident with increasing mode number. In addition, because the basis modes are orthogonal, it can be seen from Fig. 29 that the positions of high-speed and low-speed regions in each mode are almost "staggered". For example, in Fig. 29(o), a high-speed region appears near the propeller shaft for the 14th mode, while other modes

do not have high-speed regions there. In fact, this phenomenon makes it possible to reconstruct the actual wake field through the linear superposition of each basis mode.

The ratio of the sum of square error of each new sample hull's reconstructed (predicted) dimensionless wake field information by the reduced-order model and its corresponding actual dimensionless wake field information (total error energy) to the sum of squares of each new sample hull's actual dimensionless wake field information (total energy) can be obtained, and the relative error e_r can be defined as the arithmetic square root of that ratio, i.e.,

$$e_r = \sqrt{\frac{\sum_{i=1}^{Np} (u_i - \widetilde{u}_i)^2}{\sum_{i=1}^{Np} u_i^2}} \tag{21}$$

The relationship between the relative error e_r and node number *M* can be obtained for each sample in the database. The relationship is shown in Fig. 30 for one selected sample from the database.

As *M* increases, the relative error of the reconstructed wake field corresponding to the new sample hull decreases monotonically, and when *M* reaches 14, the relative error is less than 1%, indicating that the reconstructed flow field has high reliability.

The absolute errors in the reconstructed wake field of the above new sample hull for different values of *M* are shown in Fig. 31. It is evident that the absolute error of the wake field decreases gradually with increasing *M*. Therefore, the dimensionality reduction learning effect can be considered good when $M = 14$. In other words, based on the wake field database, the axial velocity field at the propeller disk for any new hull form in the design space can be predicted quickly and accurately.

A sensitivity analysis can be carried out for each design variable in the optimization case using the dimensionality reduction learning results. The three design variables are normalized as follows

$$x_i^* = \frac{x_i - x_{i,\min}}{x_{i,\max} - x_{i,\min}}, i = 1, 2, 3 \tag{22}$$

Then, the axial velocity field variation at the propeller disk can be obtained when the three dimensionless design variables x_1^*, x_2^*, x_3^* change from 0 to 1.

The sensitivity analysis is conducted on the normalized design variables x_1^*, x_2^*, x_3^* using 0.1 as the interval. The wake field corresponding to each new hull can be quickly predicted using the POD results instead of the much more expensive numerical simulations, as shown in Figs. 32–34.

Fig. 32 shows that the high-speed regions on the left and right sides of the wake field expand upward, and the low-speed regions are relatively squeezed and become smaller with increasing x_1^*, while the low-speed area near the propeller shaft expands. Similar sensitivity analyses can be performed for other design variables.

Fig. 33 shows sections at the ship stern with increasing x_2^*. In general, the high-speed regions on the left and right sides and at the bottom of the wake field expand upward and toward the propeller shaft, while the contour lines of the hook-shaped low-speed region near the propeller shaft converge with those below the propeller shaft.

It can be seen from Fig. 34 that the high-speed regions on the left and right sides of the wake field expand towards the propeller shaft with increasing x_3^*, while the contour lines of the hook-shaped low-speed region near the propeller shaft converge with those below the propeller shaft. However, the overall wake change with x_3^* is smaller than with x_2^*. Therefore, the sensitivity to x_3^* is somewhat weaker than to x_2^*. In addition, as x_3^* increases, the inflection point of the contour line in the high-speed region tends to draw close from the left and right sides below the propeller shaft to the sides above the propeller shaft.

Again, it should be explained that in the sensitivity analysis, only one design variable value is changed to exclude the effects of the other variables. However, because the design variable values of the new

(a) $M = 1$ (b) $M = 2$ (c) $M = 3$

(d) $M = 4$ (e) $M = 5$ (f) $M = 6$

(g) $M = 7$ (h) $M = 8$ (i) $M = 9$

(j) $M = 10$ (k) $M = 11$ (l) $M = 12$

(m) $M = 13$ (n) $M = 14$
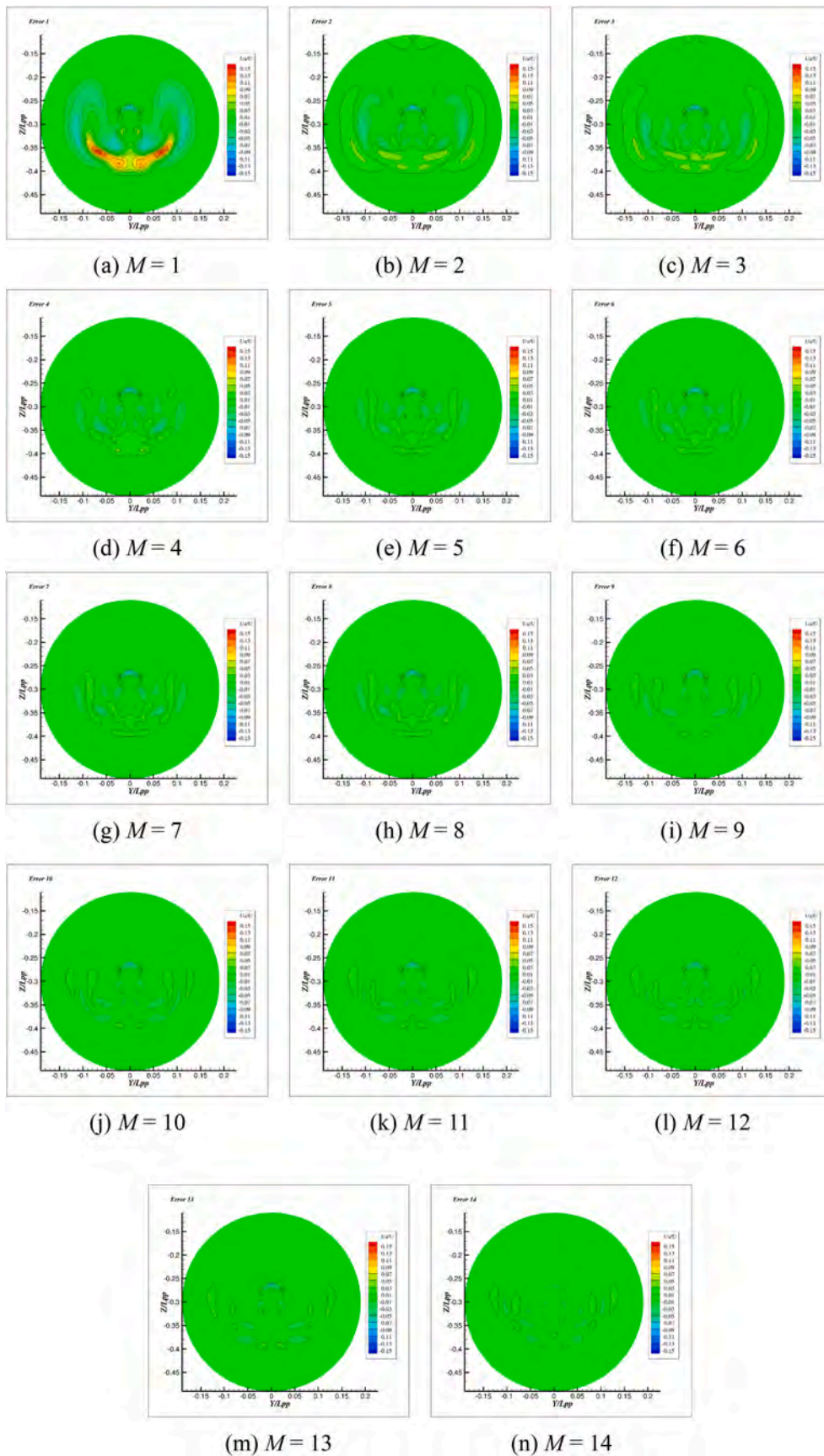
Fig. 31. Absolute error in the reconstructed wake field for different M.

(a) $x_1^* = 0$

(b) $x_1^* = 0.1$

(c) $x_1^* = 0.2$

(d) $x_1^* = 0.3$

(e) $x_1^* = 0.4$

(f) $x_1^* = 0.5$

(g) $x_1^* = 0.6$

(h) $x_1^* = 0.7$

(i) $x_1^* = 0.8$

(j) $x_1^* = 0.9$

(k) $x_1^* = 1$
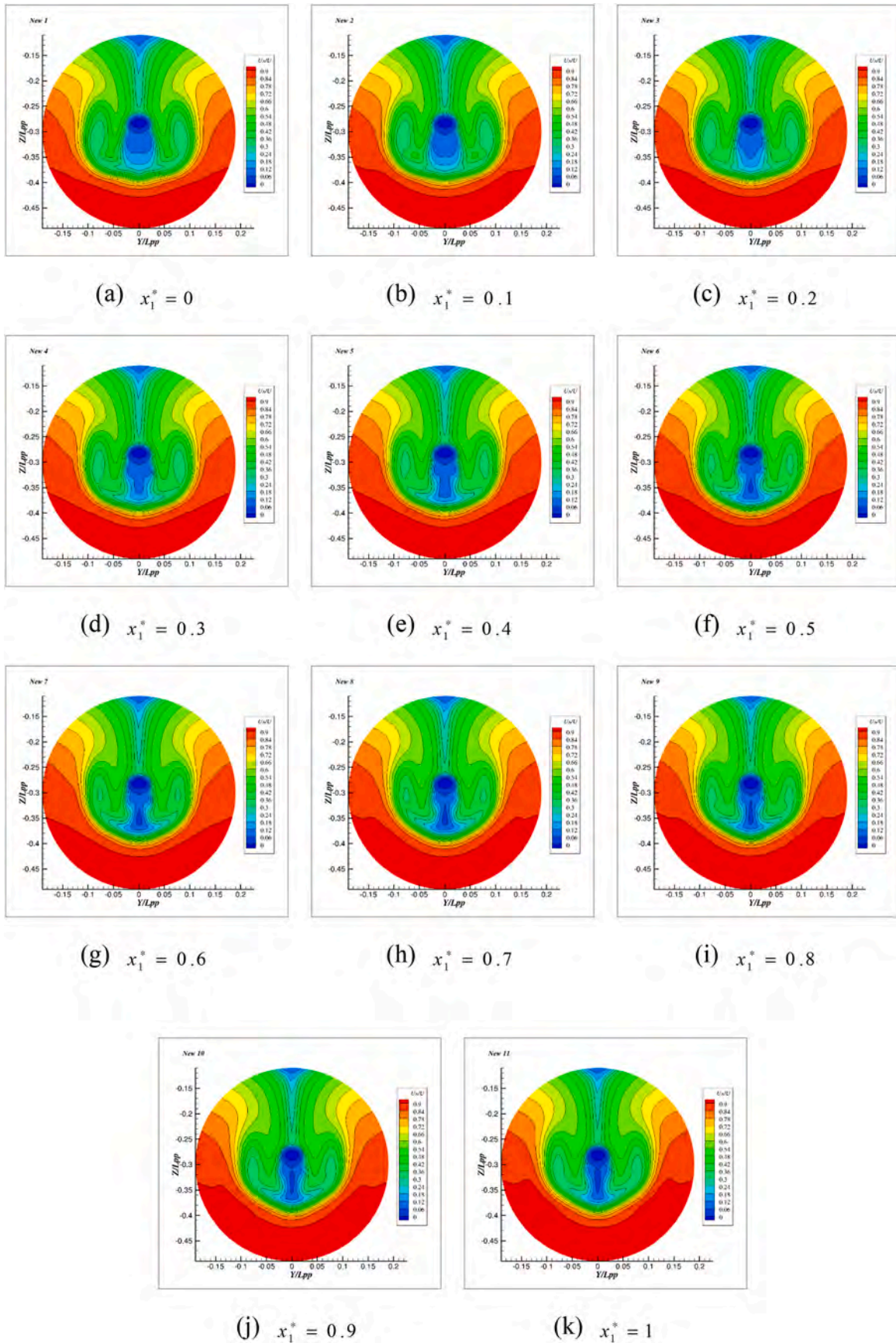
Fig. 32. Variation of wake fields with changes in the first design variable value.

(a) $x_2^* = 0$

(b) $x_2^* = 0.1$

(c) $x_2^* = 0.2$

(d) $x_2^* = 0.3$

(e) $x_2^* = 0.4$

(f) $x_2^* = 0.5$

(g) $x_2^* = 0.6$

(h) $x_2^* = 0.7$

(i) $x_2^* = 0.8$

(j) $x_2^* = 0.9$

(k) $x_2^* = 1$
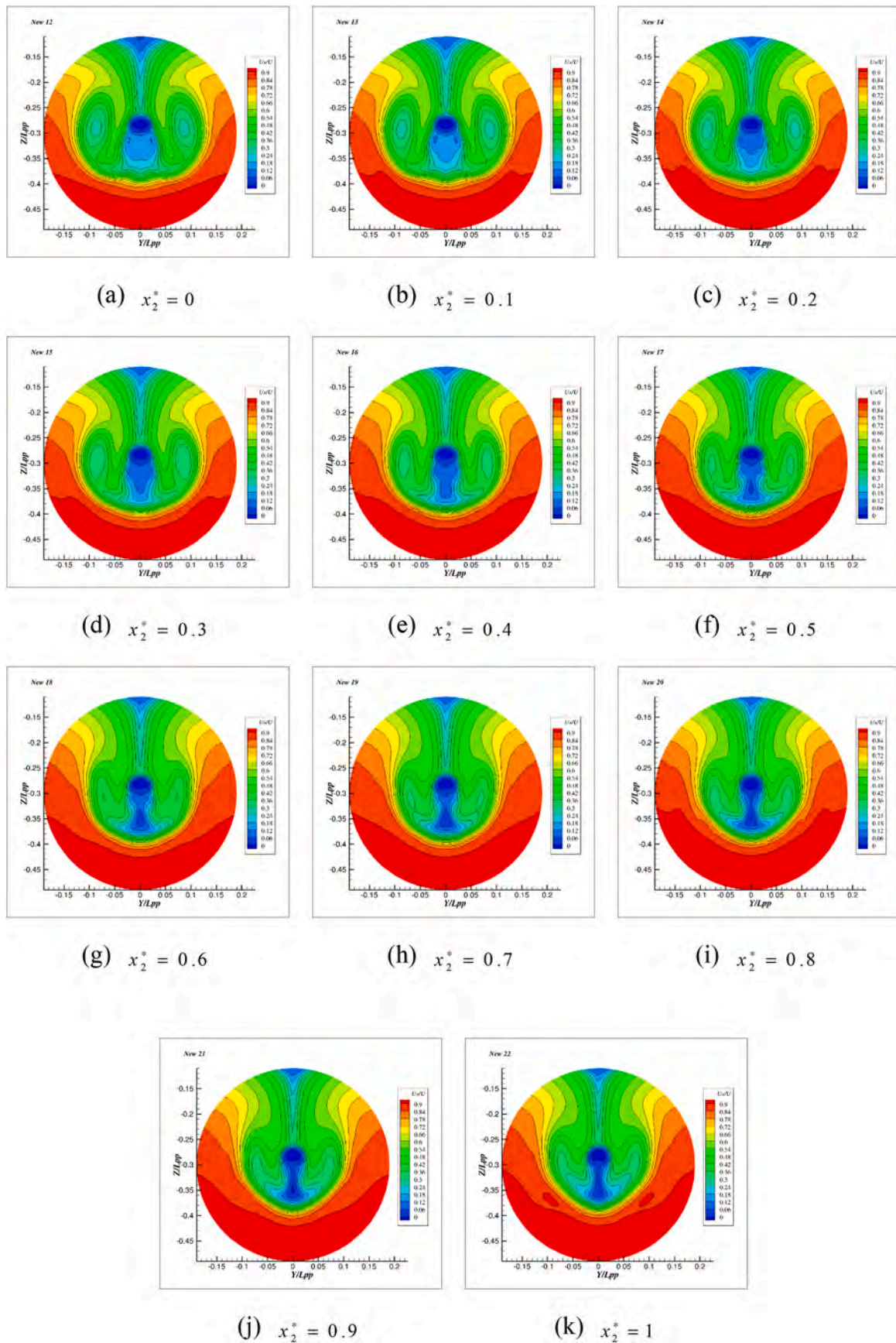
Fig. 33. Variation of wake fields with changes in the second design variable value.

(a) $x_3^* = 0$

(b) $x_3^* = 0.1$

(c) $x_3^* = 0.2$

(d) $x_3^* = 0.3$

(e) $x_3^* = 0.4$

(f) $x_3^* = 0.5$

(g) $x_3^* = 0.6$

(h) $x_3^* = 0.7$

(i) $x_3^* = 0.8$

(j) $x_3^* = 0.9$

(k) $x_3^* = 1$

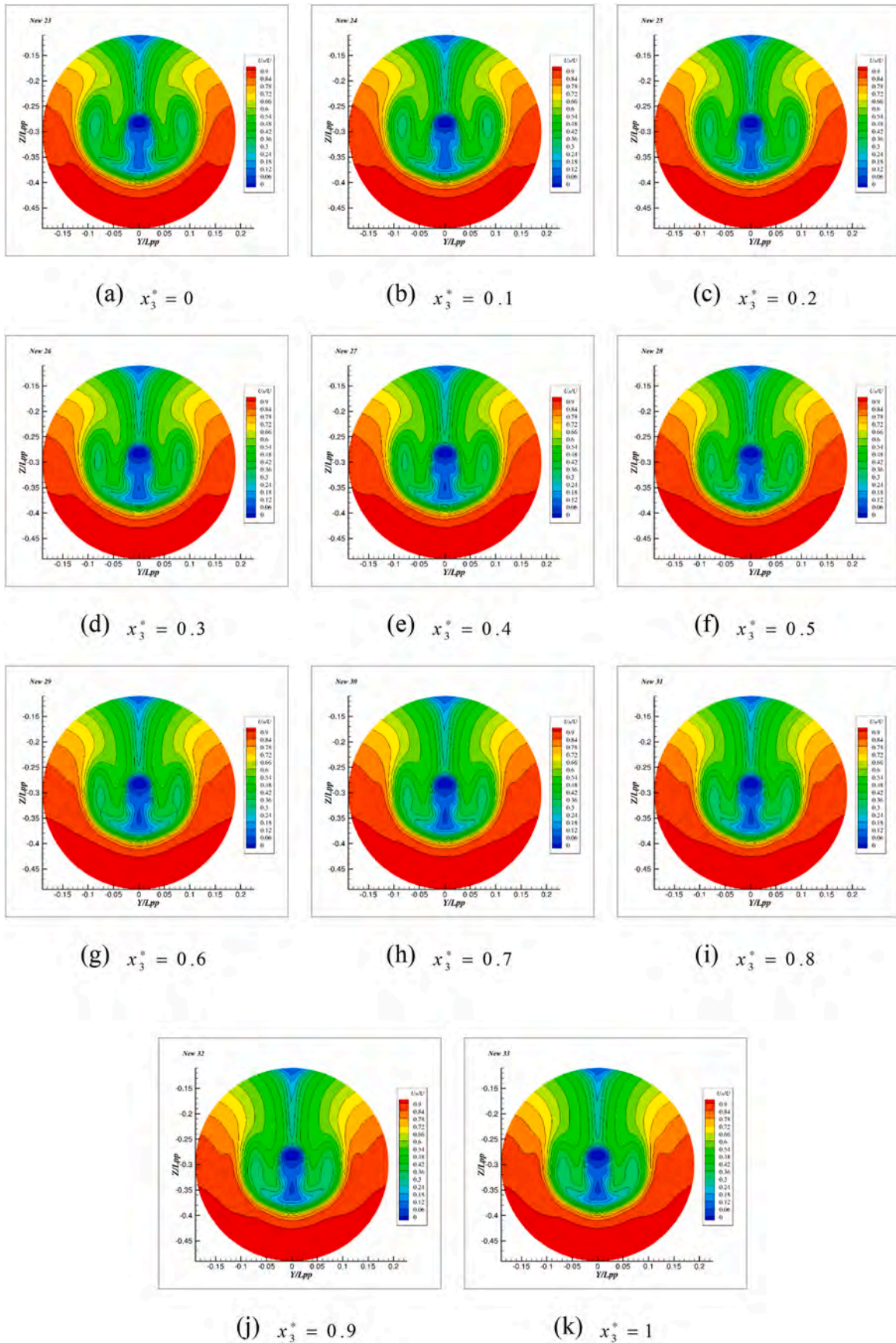**Fig. 34.** Variation of wake fields with changes in the third design variable value.

sample hulls are generally obtained through the space-filling design of experiments method to ensure good uniformity and orthogonality among the samples, the new hulls used for sensitivity analysis are not included in the new $N = 60$ sample hull database of used for batch hydrodynamic performance evaluation. In other words, the flow field information of a series of new hulls used for sensitivity analysis is quickly predicted based on the sample hull database of $N = 60$ and the dimensionality reduction method, which saves significant computational costs in the viscous-flow-based hydrodynamic evaluation of a series of new hulls.

The sensitivity analysis of the optimization design variables above shows how the axial velocity field at the propeller disk changes when each design variable changes within its variation range. This can further guide the selection of the design space and hull form optimization variables and the rapid flow field prediction of any new hull form in the design space.

The wake field learning process is implemented on personal computer, and the calculation information for the wake field database is listed:

 (i) CPU information: Intel(R) i7-4790 K @ 2.00 GHz;
 (ii) grid number: about 452,000;
 (iii) dimensionality reduction calculation: approximately 3600 CPU seconds;
 (iv) wake field prediction for a new sample hull: approximately 20 CPU seconds.

It is not difficult to see that the time needed to construct the Co-Kriging model is much less than that for single-fidelity Kriging model, and the wake field prediction can be very effective rather than doing a new calculation of a new hull.

4. Conclusions and future work

In this paper, data-driven surrogate models, including single-fidelity Kriging and multi-fidelity Kriging models, are constructed and compared in detail. The Co-Kriging model can fully use the relationship between the high- and low-fidelity sample data for a mathematical test function and has a relatively smaller error than the Kriging model. The calm-water resistance and wake performances of the JBC ship at its design speed are then optimized, where the Kriging and Co-Kriging surrogate models are established through the viscous-flow evaluation method using coarse and medium grids. In this case, 60 high-fidelity hydrodynamic evaluations have been done to construct the Kriging model, while the Co-Kriging model uses 30 high-fidelity and 60 low-fidelity evaluations with a 25% reduction of the total computation time, while obtaining better-performed hulls. Therefore, it can be seen that not only the potential and viscous flow results can be regarded as low- and high-fidelity results respectively, but also the coarse- and fine-mesh viscous flow results, to construct Co-Kriging model, which can be successfully applied in the simulation-based hydrodynamic performance optimization of hull forms.

Furthermore, an accurate and efficient viscous-flow-based wake field learning method is proposed based on the Kriging model and the POD method when the grid topological relations for different sample hulls are entirely consistent. As a result, forecasts through dimensionality reduction field learning can be much quicker than evaluating new hull forms using high-fidelity viscous flow calculations, and the sensitivity analysis can be a guidance for the future optimization for the ship wake field. Therefore, this paper's dimensionality reduction learning method for the viscous flow field has significant potential and a wide applicability range.

Meanwhile, we have to say, one bottleneck of the multi-fidelity Co-Kriging model is still the dimensionality of the optimization problem, so only three design variables are selected for JBC hull's drag and wake performance optimization. Moreover, the effectiveness of multi-fidelity

Co-Kriging model may be reduced when the correspondence between the fidelity levels is relatively small, but it's fortunate that the correspondence of the two-fidelity data we use is big enough.

Further study will focus on the viscous flow field learning method for three-dimensional flow field. What's more, sample number choice strategy for the two or more-fidelity data can be studied by comparing different time cost and drag and wake performance gain, to give further elucidation of the multi-fidelity model's advantages.

CRediT authorship contribution statement

Xinwang Liu: Data curation, Software, Visualization, Investigation, Validation, Conceptualization, Funding acquisition, Methodology, Writing – original draft. **Decheng Wan:** Supervision, Conceptualization, Methodology, Writing – review & editing. **Lei Lei:** Data curation, Investigation, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

References

Bakica, A., Gatin, I., Vukcevic, V., Jasak, H., Vladimira, N., 2019. Accurate assessment of ship-propulsion characteristics using CFD. Ocean Eng. 175 (1), 149–162.

Bonfiglio, L., Perdikaris, P., Brizzolara, S., 2020. Multi-fidelity Bayesian optimization of SWATH hull forms. J. Ship Res. 64 (2), 154–170.

Bonfiglio, L., Perdikaris, P., Vernengo, G., de Medeiros, J.S., Karniadakis, G., 2018. Improving swath seakeeping performance using multi-fidelity Gaussian process and Bayesian optimization. J. Ship Res. 62 (4), 223–240.

Brahmachary, S., Bhagyarajan, A., Ogawa, H., 2021. Fast estimation of internal flowfields in scramjet intakes via reduced-order modeling and machine learning. Phys. Fluids 33, 106110.

Brunton, S.L., Noack, B.R., Koumoutsakos, P., 2020. Machine learning for fluid mechanics. Annu. Rev. Fluid Mech. 52, 477–508.

Chen, H., Zhang, Y., Yang, X., 2021. Uniform projection nested Latin hypercube designs. Stat. Pap. 62, 2031–2045.

Diez, M., Campana, E.F., Stern, F., 2015. Design-space dimensionality reduction in shape optimization by Karhunen-Loève expansion. Comput. Methods Appl. Math. 283, 1525–1544.

Forrester, A.I.J., Sóbester, A., Keane, A.J., 2007. Multi-fidelity optimization via surrogate modeling. Proc. Math. Phys. Eng. Sci. 463, 3251–3269.

Granados-Ortiz, F.J., Ortega-Casanova, J., 2021. Machine learning-aided design optimization of a mechanical micromixer. Phys. Fluids 33, 063604.

Guo, Q., Hang, J., Wang, S., Hui, W., Xie, Z., 2020. Buckling optimization of variable stiffness composite cylinders by using multi-fidelity surrogate models. Thin-Walled Struct. 156, 107014.

Ji, T., Jin, F., Xie, F., Zheng, H., Zhang, X., Zheng, Y., 2022. Active learning of tandem flapping wings at optimizing propulsion performance. Phys. Fluids 34, 047117.

Jolliffe, I.T., 1986. Principal component analysis. J. Appl. Stat. 87, 41–64.

Lee, S., You, D., 2019. Data-driven prediction of unsteady flow over a circular cylinder using deep learning. J. Fluid Mech. 879, 217–254.

Li, Y., Chang, J., Kong, C., Wang, Z., 2020. Flow field reconstruction and prediction of the supersonic cascade channel based on a symmetry neural network under complex and variable conditions. AIP Adv. 10, 065116.

Liu, X., Chen, S., Zhao, W., Wan, D., Wang, Y., 2021a. Liutex-based centripetal force field model for improving the resistance and wake performances of JBC ship sailing in calm water. J. Hydrodyn. 33, 494–502.

Liu, X., Wan, D., Chen, G., 2018. Kriging-based surrogate model combined with weighted expected improvement for ship hull form optimization. In: Proceedings of the ASME 37th International Conference on Ocean, Offshore and Arctic Engineering (OMAE), 78388.

Liu, X., Wan, D., Hu, C., 2020. Multi-objective hull form optimization of trimaran's total drag at different speeds. In: Proceedings of the 30th International Ocean and Polar Engineering Conference (ISOPE), pp. 3850–3857.

Liu, X., Zhao, W., Wan, D., 2021b. Hull form optimization based on calm-water wave drag with or without generating bulbous bow. Appl. Ocean Res. 116, 102861.

Liu, X., Zhao, W., Wan, D., 2021c. Linear reduced order method for design-space dimensionality reduction and flow-field learning in hull form optimization. Ocean Eng. 237, 109680.

Liu, X., Zhao, W., Wan, D., 2021d. Optimization of the roll motion for box-shaped hull section with anti-rolling sloshing tanks and fins in beam waves. J. Hydrodyn. 33, 688–697.

Liu, X., Zhao, W., Wan, D., 2022. Multi-fidelity co-Kriging surrogate model for ship hull form optimization. Ocean Eng. 243, 110239.

Oh, S., Lee, S., Son, M., Kim, J., Ki, H., 2022. Accurate prediction of the particle image velocimetry flow field and rotor thrust using deep learning. J. Fluid Mech. 939, A2.

Pearson, K., 1901. On lines and planes of closest fit to systems of points in space. Philos. Mag. A 2 (11), 559–572.

Pellegrini, R., Wackers, J., Broglia, R., Serani, A., Visonneau, M., Diez, M., 2022a. A multi-fidelity active learning method for global design optimization problems with noisy evaluations. Eng. Comput. 1–24.

Pellegrini, R., Wackers, J., Serani, A., Visonneau, M., Diez, M., 2022b. Towards automatic parameter selection for multifidelity surrogate-based optimization. In: Proceedings of the 9th Conference on Computational Methods in Marine Engineering.

Peri, D., Rossetti, M., Campana, E.F., 2001. Design optimization of ship hulls via CFD techniques. J. Ship Res. 45, 140–149.

Piazzola, C., Tamellini, L., Pellegrini, R., Broglia, R., Serani, A., Diez, M., 2022. Comparing multi-index stochastic collocation and multi-fidelity stochastic radial basis functions for forward uncertainty quantification of ship resistance. Eng. Comput. 1–29.

Piazzola, C., Tamellini, L., Pellegrini, R., Broglia, R., Serani, A., Diez, M., 2020. Uncertainty quantification of ship resistance via multi-index stochastic collocation and radial basis function surrogates: a comparison. Proceedings of the AIAA Aviation 2020 Forum 1, 1–23, 3160.

Raven, H.C., Scholcz, T.P., 2019. An assessment of multifidelity procedures for ship hull form optimisation. In: Proceedings of the 8th International Conference on Computational Methods in Marine Engineering.

Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P., 1989. Design and analysis of computer experiments. Stat. Sci. 4 (4), 409–423.

Serani, A., Diez, M., 2018. Shape optimization under stochastic conditions by design-space augmented dimensionality reduction. In: Proceedings of the 19th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference.

Serani, A., Fasano, G., Liuzzi, G., Lucidi, S., Iemma, U., Campana, E.F., Stern, F., Diez, M., 2016. Ship hydrodynamic optimization by local hybridization of deterministic derivative-free global algorithms. Appl. Ocean Res. 59, 115–128.

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 the VIII International Conference on Computational Methods in Marine Engineering.

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, 237–255.

Sobol, I., 1979. On the systematic search in a hypercube. SIAM J. Numer. Anal. 16 (5), 790–793.

Tao, J., Sun, G., 2019. Application of deep learning based multi-fidelity surrogate model to robust aerodynamic design optimization. Aero. Sci. Technol. 92, 722–737.

Wackers, J., Visonneau, M., Serani, A., Pellegrini, R., Broglia, R., Diez, M., 2020. Multi-fidelity machine learning from adaptive- and multi-grid RANS simulations. In: Proceedings of the 33rd Symposium on Naval Hydrodynamics.

Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemometr. Intell. Lab. Sys. 2 (1), 37–52.

Wu, X., Wu, S., Tian, X., Guo, X., Luo, X., 2022. Effects of hyperparameters on flow field reconstruction around a foil by convolutional neural networks. Ocean Eng. 247, 110650.

Zhang, X., Xie, F., Ji, T., Zhu, Z., Zheng, Y., 2021. Multi-fidelity deep neural network surrogate model for aerodynamic shape optimization. Comput. Methods Appl. Math. 373, 113485.