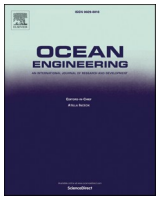
Ocean Engineering 260 (2022) 112077

Contents lists available at ScienceDirect

Ocean Engineering

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



Multi-objective optimization of semi-submersible platforms based on a support vector machine with grid search optimized mixed kernels surrogate model

Yixuan Mao [a,*], Tianqi Wang [a], Menglan Duan [a], Hongyuan Men [b]

[a] College of Safety and Ocean Engineering, China University of Petroleum-Beijing, Beijing, China
[b] Key Laboratory of High Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Keywords:
Multi-objective optimization
SEMI
Surrogate model
SVM
Hydrodynamic response

ABSTRACT

Determination of optimal hull configurations in the semi-submersible platform (SEMI) should account for several objectives. These objectives are pertinent to hydrodynamic performances of SEMI under wave action but also total structure cost. They are often contradictory and cannot achieve the minimum simultaneously. Hence, a group of relative optimal and balanced solutions is introduced as optimization results, called Pareto-optimal solutions. This paper presents a surrogate-assisted technique to seek the optimal configuration of SEMI for minimal heave and roll response and the lightest weight. Design variables samples are generated by means of multidimensional Ladin hypercube design, and then these inputs are employed for hydrodynamic simulation to acquire the response data. To determine the relationship between objectives and hull structure size, Support Vector Machine with Grid Search optimized mixed kernels (SVM-GSM) is constructed as a surrogate model, and triple verification in terms of errors and robustness warrants its reliability. Three categories of Pareto optimal solutions are obtained by Non-dominated Sorting Genetic Algorithm II (NSGA-II), which correspond to three optimal goals. For optimization results, a presented comprehensive verification approach integrates frequency-domain analysis (FD), time-domain analysis (TD), convergence analysis, and main factors screening. This combination renders sufficient and reliable validation to optimization results. Results from FD and TD for SEMI indicate that the optimized effect of Pareto solutions is satisfactory. Besides, the main influence factors in design variables for hydrodynamic response are screened and investigated. Finally, the ranking of the influence degree of each variable is obtained and evaluated. The proposed framework in this paper provides a comprehensive validation idea for the construction of the surrogate model and optimization results for SEMI hull structure optimization.

1. Introduction

As offshore oil and gas exploration gradually transitions to deep water, floating production platforms have been widely employed in various sea areas (API RP 2SK, 2005). Based on different water depths and functional requirements, Floating Production Storage and Off-loading vessels (FPSOs), Tension Leg Platforms (TLPs), Spar, and Semi-submersible platforms (SEMIs) are designed and have achieved a terrific application. As one of the most preferred floating platforms, SEMI not only has larger deck space but also has good adaptability for a range of water depths. Compared to traditional fabrication methods, SEMI is advantageous due to its flexible fabrication location and convenient installation, can be built by various shipyards worldwide and integrated in a yard. Therefore, SEMI plays a more and more significant role in traditional fossil energy and wind energy territory.

Generally, the whole design process of a floating platform can be divided into three phases: conceptual design (or preliminary design), basic design, and detailed design. In conceptual design, the overall size and structure should be determined based on different requirements. Then the subsequent detailed design for different components can be carried out. However, the overall size and hull form of the SEMI platform can have a significant impact on the hydrodynamic performance. The motions induced by waves directly affect subsequent operations, such as the design and installation of a riser system and mooring

* Corresponding author.
E-mail address: 2021310508@student.cup.edu.cn (Y. Mao).

https://doi.org/10.1016/j.oceaneng.2022.112077
Received 25 March 2022; Received in revised form 25 May 2022; Accepted 18 July 2022
Available online 31 July 2022

Table 1
Summary of studies on verification methods in surrogate model and optimization results for platform.

1st Author	Optimization objectives	Approaches	Surrogate model	Verification methods for surrogate model			Optimization model	Verification methods for optimization result		
				Accuracy indicators	Fitting curves	Learning curves		Freq-domain	Time-domain	Main factors screening
Akagi et al. (1984)	Heave amplitude, variable load, displacement	GRG, ISM	No				Yes	●		
Clauss and Birk (1996)	Seakeeping performance (single-objective)	NP, TSM	No				Yes	●		●
Lee et al. (2007)	Fatigue life (single-objective)	GS, SQP	No				Yes	●		
Park et al. (2015)	Structural weight, seakeeping capacity	SA	No				Yes	●	●	
Sugita and Suzuki (2016)	Structural weight, tendon weight	ASA, GA	No				Yes	●		
Zhang et al. (2018)	Tendon tension and structural weight	ANN, NSGA-II	Yes	●			Yes	●		
Qiu et al. (2019)	Structural weight, MPM_{Heave}	ANN, PSO	Yes		●		Yes	●		
Tian et al. (2021)	heave and roll motion, structural weight	ANN, NSGA-II	Yes	●	●		Yes	●		●
Liu et al. (2022)	Total drag (single-objective)	Kriging, GA	Yes	●			Yes			●
This work	Total weight, MPM_{heave} and MPM_{roll}	SVM, NSGA-II	Yes	●	●	●	Yes	●	●	●

arrangement. Therefore, the determination of hull size is based on a well-considered optimization process. In general, an iterative approach is adopted in the practice of engineering to determine the hull structure size, which is called the 'design spiral' (API RP 2T, 1997). However, relying on a traditional experience, this method adopts a continuous trial calculation to update the design parameters until they meet the requirements. Thus, this process has difficulty achieving an optimal hull structure design at a low time-cost.

In order to seek an optimal hull structure size with satisfactory hydrodynamic performance, many endeavors have been made to optimize the hull structure by offshore engineering researchers. Akagi et al. (1984) adopted the generalized reduced gradient algorithm (GRG), when performing the nonlinear multi-objective optimization on SEMI. In the study, interpretive structural modeling (ISM) was adopted to categorize design conditions into some hierarchical levels of heave motion. Variable load and displacement were employed to conduct optimization by GRG. Clauss and Birk (1996) introduced nonlinear programming approach (NP) to maximize the seakeeping performance of TLP and SEMI. The study employed a commercial code WAMIT to evaluate the hydrodynamic performance of different offshore structures. Lee et al. (2007) combined the sequential quadratic programming (SQP) and genetic algorithm (GA) to optimize hull sizes of the TLP. This study simplified the tendon tensions by an equivalent linear spring and considered fatigue life an optimized objective. Park et al. (2015) adopted the simulated annealing algorithm (SA) to optimize the Floating Production Unit (FPU), in which the structural weight and seakeeping capacity were considered as objectives. In the study, the four Pareto-optimal solutions are obtained. Sugita and Suzuki (2016) compared the adaptive simulated annealing (ASA) and GA optimization effects on total hull weight and tendon weight. In the study, the initial design criteria are employed as constraints, and a strong positive correlation between the column parameter and hull weight is found.

With the development of computer science and artificial intelligence, surrogate models provide a flexible and efficient approach to optimization (Pietrenko-Dabrowska and Slawomir Koziel, 2017). Venzon et al. (2013) introduced an approach of the surrogate model with an artificial neural network (ANN) to optimize the structure of SEMI, which can substitute for the direct simulation and reduce time consumption. In the study, the heave response is selected as the objective function. Subsequently, Zhang et al. (2018) presented a TLP optimization strategy based on the surrogate model of the RBF neural network algorithm. In the study, tendon tension and total weight are selected as objectives. Qiu et al. (2019) developed an optimization program based on the particle swarm optimization algorithm using a surrogate model constructed by the RBF neural network algorithm. The study utilized three types of floating platforms to analyze the optimization effect. Emami and Mostafa Gharabaghi (2020) presented a grid search algorithm (GS) to optimize the three objectives, in which minimal total hull structure weight and minimal heave and pitch motions are considered. Tian et al. (2021) combined the evolutionary algorithm combined with RBF neural network surrogate model. Using a weighted function to hydrodynamic motion and total hull structure weight, they converted the problem into single optimization to optimize the hull size. Liu et al. (2022) developed a multi-fidelity co-kriging surrogate model to optimize the hull structure. In the study, recommendations of the selected sample size are given considering the different precision for the surrogate model. For some costly simulation models in some fields, the surrogate models have also been proven to be an effective technique to reduce simulation costs (Koziel and Pietrenko-Dabrowska, 2019; Wu et al., 2019). Koziel and Pietrenko-Dabrowska (2020) proposed a performance-driven surrogate modelling approach for the multi-optimization of the antenna structure. In the study, a surrogate model in a confined domain is constructed, and parameter space dimensionality is reduced based on the principal component analysis (PCA) of Pareto-optimal designs. This methodology achieves considerable cost reduction. In addition, Pietrenko-Dabrowska and Koziel (2019) developed a rapid surrogate-assisted program to yield optimization of narrow- and multi-band antennas. In the study, an accurate surrogate model is established using a small number of training samples, providing an efficient procedure to reduce computational costs. Zhang et al. (2021) proposed the quasi-Newton enhanced differential evolution (DE) to optimize the surrogate model based on the low-fidelity model. In the study, the more accurate minimum values with less computational time demonstrated the validity of the method. These new techniques proposed in recent years regarding the surrogate models undoubtedly provide novel and low-cost solutions for optimization problems in various fields (Koziel and Pietrenko-Dabrowska, 2020).

Due to the particularity and complexity of marine environmental loads, the complete model should be verified by various approaches to reduce contingency. However, the established model in most previous research is verified only by a single method, resulting in difficulty in judging whether the result can provide a representative optimization effect. Table 1 is a summary of methods and verification in previous research. As shown, the verification in most studies is based on a comparison of optimized results with direct simulation in frequency-domain analysis. However, the motion equation in the frequency-domain solution is a linear approximation to nonlinear problems, including the drag

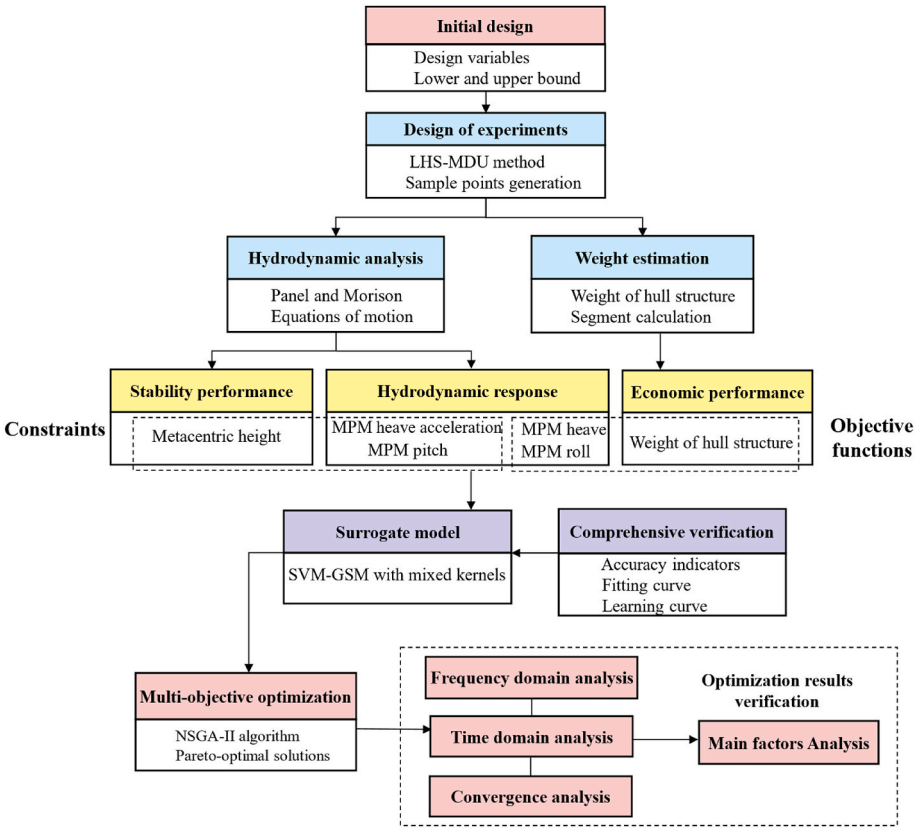


Fig. 1. Flowchart of the overall optimization process.

force, variable surface elevation, and horizontal restoring force (DNVGL–OS–E301). Especially in extreme sea states, as the natural frequency of motion overlaps the wave energy frequency range, second-order frequency forces and higher-order nonlinear effects play a dominant role in the total wave exciting forces. Therefore, single frequency-domain analysis cannot provide a representative optimization effect for verification. In addition, it is insufficient to verify the surrogate model relied on single accuracy indicators in most studies. The surrogate model should also be thoroughly verified or calibrated, which would directly affect the optimization objective.

This work aims to seek optimal hull configurations of SEMI platform hull structure with a comprehensive verification for optimization results. To enhance the computing efficiency, a surrogate model called support vector machine with grid search optimized mixed kernels (SVM-GSM) is constructed to determine the relationship between hydrodynamic performance and design variables. The total hull weight, MPM_{heave}, and MPM_{roll} are considered three optimization objectives that ensure better global motion performances and feasible costs. A comprehensive triple verification for the surrogate model is conducted to improve the insufficiency of verification in the previous study. Subsequently, an evolutionary algorithm, nondominated sorting genetic algorithm-II (NSGA-II), is adopted as the three-objective optimization model to obtain the Pareto optimal solutions, instead of utilizing a weighted function and converting the problem into a single optimization problem. Finally, through frequency and time domain analysis and main factors screening for optimization objectives, the overall verification of the optimization model is ensured. This comprehensive verification approach can provide insights into the establishment and evaluation of the model in future studies.

The rest of the paper is organized as follows: The methodology of obtaining the hydrodynamic motion response and hull weight estimation is introduced in Section 2. In Section 3, the process of surrogate model establishment and verification is introduced, and the multi-objective optimization model is constructed. Section 4 presents the results and discussion, in which the Pareto solutions are shown. In addition, frequency domain and time domain analysis, and main factors screening are employed to verify the optimization results. Finally, Section 5 summarizes the conclusions. The overall optimization process and sequence of steps are shown in Fig. 1.

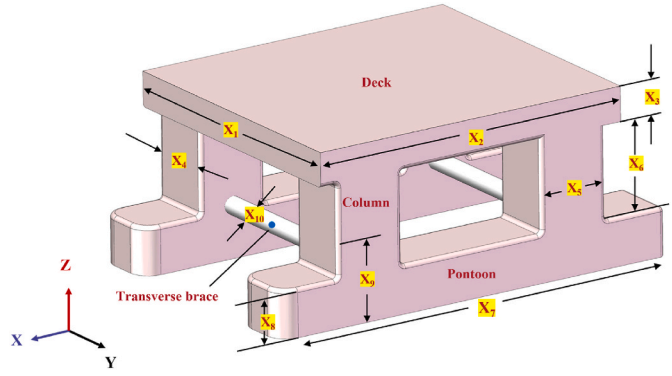


Fig. 2. The sketch of the semi-submersible platform hull structure.

2. Mathematical model

2.1. Design variables of sub-submersible platform

Economic feasibility and environmental loads-resistance are the primary considerations in the early stages of conceptual design for semi-submersible platforms. The shape and size of the platform should be optimized to obtain a safer and more low-cost structure. In the present study, the initial hull structure model of the semi-submersible platform

Table 2
Initial design variables for SEMI.

Design variables	Symbol	Initial value	Lower bound	Upper bound
Deck width	X_1	72.5	70.50	80.50
Deck length	X_2	82.5	78.50	88.50
Deck height	X_3	8.50	7.50	9.50
Column width	X_4	16.50	13.50	16.50
Column length	X_5	17.50	14.00	17.00
Column height	X_6	22.80	19.80	25.80
Pontoon length	X_7	105.8	98.30	113.30
Pontoon height	X_8	11.85	10.85	12.85
Hull draft	X_9	17.00	15.50	18.50
Transverse brace diameter	X_{10}	4.35	4.25	4.45

is determined from the simplified drilling platform in service. As shown in Fig. 2, deck length, deck width, deck height, column length, column width, column height, pontoon length, pontoon height, hull draft, and transverse brace diameter are selected as the design variables. Ten initial design variables are indicated in Table 2, which represent the overall appearance of the deck and hull of the platform. The boundary value is used to control the sample generation in the surrogate model.

In general, anti-rolling and minimum heave motion are the predominant considerations, to ensure the topside operation, crew health, and equipment integrity. Hence, the heave response, roll response, pitch response, heave acceleration, and metacentric height are selected as indicators to describe the stable performance and safety of the platform during the operation period. In addition, the hull structure weight is considered as a cost function and should be minimized. Therefore, this paper selected the heave response, rolling response, and hull structure weight as objective functions. The pitch response, heave acceleration, and metacentric height are selected as constraints to further ensure the rationality of variables in the optimization process of the platform.

2.2. Hydrodynamic performance

In the frame of this research, the wave exciting force is obtained relying on the three-dimensional potential flow theory, which is to compute wave loads acting on large-scale structures. The drag force is based on the Morison equation to consider the effect of viscosity. Subsequently, by solving the 6 degrees of freedom (DoF) motion equation of the frequency domain in numerical computation, the hydrodynamic performance of the SEMI can be obtained.

2.2.1. Panel and Morison model

For large-scale structures, the surrounding flow field can change significantly, and the diffraction effect caused by incident waves should be considered. Due to the relatively little viscosity influence, potential flow theory is adopted, which assumes that the fluid is inviscid and incompressible and flow motion is irrotational. Velocity potential function φ is introduced to better express the fluid motion:

$$\varphi = [\varphi_I(x,y,z) + \varphi_D(x,y,z) + \varphi_R(x,y,z)]e^{-i\omega t} \tag{1}$$

where φ_I, φ_D and φ_R represent the velocity potential by the incident wave, velocity potential by diffracted waves, and velocity potential by radiated waves, respectively.

The first-order diffraction potential should satisfy the Laplace equation, free surface boundary condition, seabed boundary condition, and body surface boundary condition.

Governing by:

$$\nabla^2\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} = 0 \tag{2}$$

On the stable water surface:

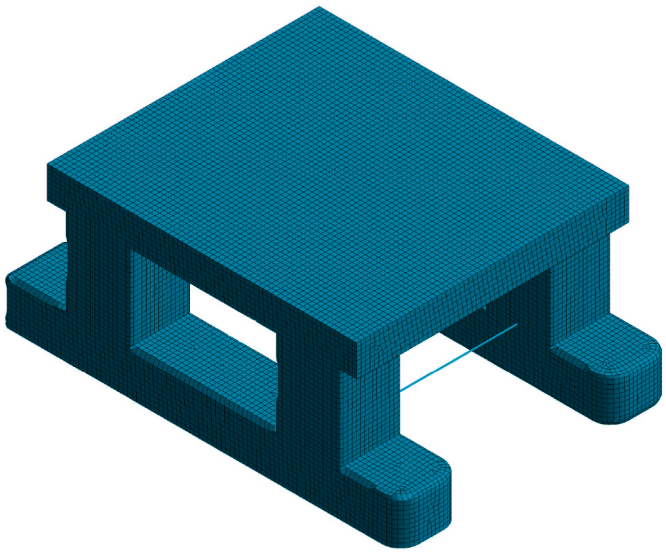


Fig. 3. Panel model with semi-submersible platform for the initial design.

$$\left(-\omega^2 + g\frac{\partial}{\partial z}\right)\varphi = 0, (z=0) \tag{3}$$

On the seabed:

$$\frac{\partial\varphi}{\partial z} = 0, (z=-h) \tag{4}$$

On body surface:

$$\frac{\partial\varphi}{\partial n} = U\cdot n \tag{5}$$

At far-field:

$$\lim_{r\to\infty}\sqrt{r}\left(\frac{\partial}{\partial r} - ik\right)\varphi = 0 \tag{6}$$

where U denotes the hull velocity, n denotes the normal vector of the hull surface, h denotes the water depth, r denotes the distance to the center of the hull, k denotes the wave numbers, g denotes gravity acceleration. Since the source strengths are assumed to be constant, a large number of flat panels divided by the mean wetted surface are utilized to solve the source strengths, and then the hydrodynamic pressure and velocity potential can be obtained. Subsequently, the wave exciting force can be calculated by integrating the hydrodynamic pressure on the mean wetted hull surface. The panel model with semi-submersible platform in this paper is shown in Fig. 3.

For slender structures, the inertia force and drag force wave-induced acting on the slender rod cannot be neglected. The viscous effect and added mass effect dominate the wave loads. In this circumstance, the Morison equation is proposed, and the wave load can be determined by a sum of inertia force in phase with the local flow acceleration and a drag force proportional to the square of the instantaneous flow velocity. The Morison equation can be written as:

$$F = \rho V_0(1+C_A)\dot{v} + \frac{1}{2}\rho\sigma C_D v\left|v\right| \tag{7}$$

where ρ denotes the density of sea water, V_0 denotes the water displacement volume of the slender structure, C_A denotes the added mass, C_D denotes the coefficient of the drag force, and v denotes the flow velocity. By means of the Panel and Morison models, the wave exciting force and drag forces are computed. Subsequently, the motion equation can be solved by these forces.

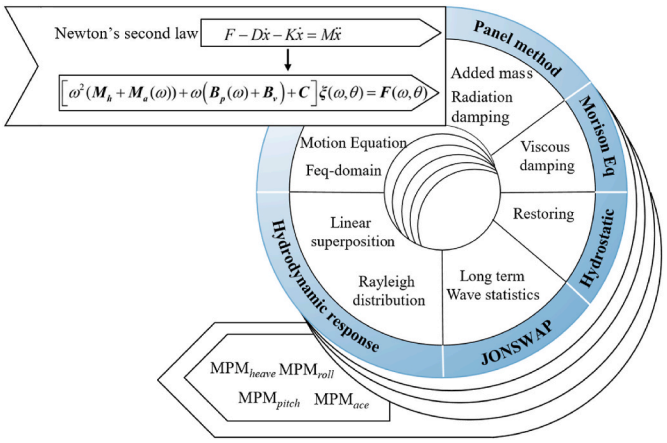


Fig. 4. Flow chart of Hydrodynamic computing in frequency domain equation.

Table 3
Wave parameters with 10-year return sea survival condition.

Return period (year)	Significant height H_s (m)	Peak period T_s (s)	Peak parameter (γ)	Peak parameter (σ_a)	Peak parameter (σ_b)
10	11.1	13.6	2.4	0.07	0.09

2.2.2. Equations of motion

The total wave loads acting on the platform are computed based on the panel model and Morison equation. By solving the motion equation in the frequency domain, the hydrodynamic motion response of the platform can be obtained. The equation can be written as:

$$\left[\omega^2(M_h + M_a(\omega)) + \omega\left(B_p(\omega) + B_v\right) + C\right]\boldsymbol{\xi}(\omega,\theta) = F(\omega,\theta) \tag{8}$$

where $\boldsymbol{\xi}(\omega,\theta)$ denotes the displacement of the rigid body induced by an incident wave with ω frequency and θ wave direction. $\boldsymbol{M_h}$ denotes the mass matrix of the platform, $\boldsymbol{M_a}$ denotes the added mass matrix, $\boldsymbol{B_p}$ denotes the radiation damping matrix, $\boldsymbol{B_v}$ denotes the linearized viscous damping matrix, and $\boldsymbol{C}$ is the hydrostatic restoring matrix. In the equation, $\boldsymbol{M_a}$, $\boldsymbol{B_p}$, $\boldsymbol{B_v}$, and $\boldsymbol{F}$ can be obtained according to the panel method and Morison's equation. Consequently, $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6)$ can be obtained using a commercial program, reflecting the six degrees of freedom motion of the floater. The whole hydrodynamic analysis process is shown in Fig. 4.

To compute the motion responses for irregular waves, the wave spectrum is employed to simulate the wave conditions in a specific sea state. In this study, the survival sea condition of the first 10-year return period is selected from the South China Sea. The wave condition is modeled by the JONSWAP spectrum, for which the parameters are shown in Table 3. Based on the principle of linear superposition, the motion response in irregular waves can be generated by regular waves with different wave amplitudes and wave frequencies:

$$S_r(\omega,\theta) = |H(\omega,\theta)|^2 S_w(\omega) \tag{9}$$

where S_r denotes the motion response for irregular waves, H denotes the RAO in 6 DoF, S_w denotes the wave spectrum. The most probable maximum (MPM) response is utilized to represent the short-term response in a specific sea state, and MPM is determined by:

$$R_{\max} = \sqrt{\left(2\sigma_r \ln\frac{t}{T_2}\right)} \tag{10}$$

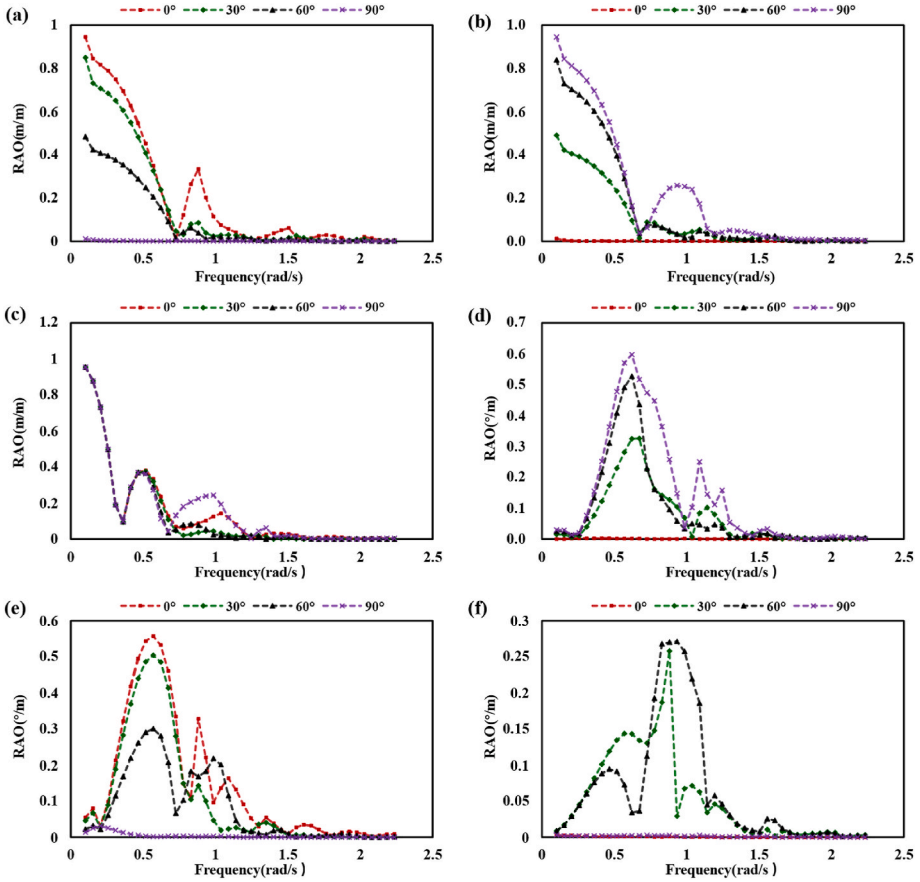


Fig. 5. The RAO curves in different direction waves. (a) Surge RAO; (b)Sway RAO; (c) Heave RAO; (d) Roll RAO; (e) Pitch RAO; (f)Yaw RAO.

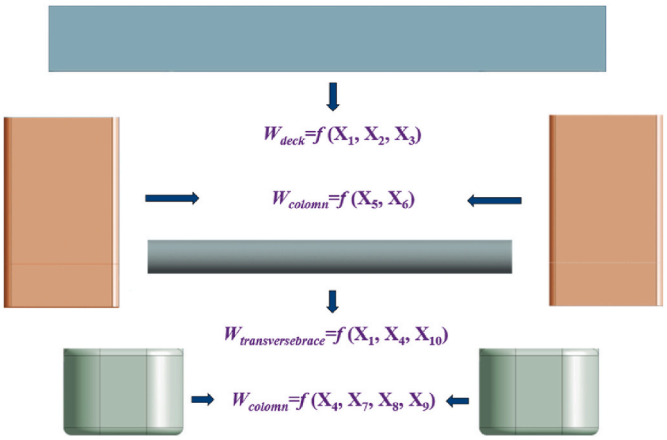


Fig. 6. Segment division for calculation of total hull structure weight.

where T_2 denotes the zero-upcrossing period of the response spectrum, t is the time duration (3 h adopted). And σ_r is the zero moment of response spectrum. In the present study, MPM is used to represent the value of the motion response in a quantitative way.

Fig. 5 shows the RAOs in different direction waves. RAO curves in different wave directions indicate a similar trend. For instance, the heave response in all wave directions seems to be more sensitive to the low-frequency wave, so higher motion amplitudes are shown in the beginning. As the frequency increases, the motion amplitudes decrease and are gradually steady. Particularly, the heave response and roll response have the maximum value when the incident wave direction is 90°. In terms of pitch response, the maximum value is induced by the 0° incident wave. Therefore, in order to ensure the platform safety considering the potential maximum wave loadings, MPM_{heave} and MPM_{roll} with 90° incident wave are calculated as considered objectives, whereas the MPM_{pitch} with 0° incident wave is calculated as constraints.

2.3. wt estimation

The total cost of establishing a semi-submersible drilling platform includes hull structure steel costs, drilling equipment costs, shipyard transportation costs, installation costs, etc. Among these costs, the hull structure steel cost accounts for approximately half, which plays a predominant role in platform build-cost. Therefore, it is significant to reduce steel consumption as much as possible while ensuring safety and stability.

The weight of the hull structure can be divided into four parts, including the pontoon weight, column weight, transverse brace weight, and deck weight. According to the simplified hull shape shown in Fig. 6, each part can be considered the function of the design variables (Park et al., 2015). The single pontoon weight W_p can be calculated by empirical Eq. (11) (Gosain et al., 2017):

$$W_p = 9.4 \times 10^{-3} \times (S_p \times T)^{1.05} \tag{11}$$

where S_p denotes the surface area of the lower floating body, and T denotes the draft, which are determined by:

$$\begin{cases} S_P = 2 \times X_7 \times (X_4 + X_8) \\ T = X_9 \end{cases} \tag{12}$$

where $X_1 \sim X_{10}$ denote the design variables, which are shown in Table 2.

The single column weight W_c is determined by the empirical Eq. (13):

$$W_c = X_6 \times 0.286 \times X_5^{1.612} \tag{13}$$

The transverse brace weight W_t is estimated by the empirical Eq. (14):

$$W_t = L \times 0.405 \times X_{10}^{1.608} \tag{14}$$

where L denotes the transverse brace length, and can be determined by:

$$L = X_1 - 2 \times X_4 \tag{15}$$

The entire deck structure contains main decks, the remaining decks, and the bulkheads. Respectively, the W_{dm}, W_{dr}, and W_{db} are calculated using the following formula (Tian et al., 2021):

$$\begin{cases} W_{dm} = 0.242 \times A_{MDK} - 0.121 \times 10^{-4}(A_{MDK})^2 \\ W_{dr} = 0.054 \times A_{RDK} + 0.162 \times 10^{-4}(A_{RDK})^2 \\ W_{db} = 0.026 \times X_3 \times A_{MDK} - 2.13 \end{cases} \tag{16}$$

where A_{MDK} and A_{RDK} are determined by:

$$\begin{cases} A_{MDK} = X_1 \times X_2 \\ A_{RDK} = 2 \times X_1 \times X_2 \end{cases} \tag{17}$$

Hence, the weight of the entire deck structure W_d can be expressed as:

$$W_d = W_{dm} + W_{dr} + W_{db} \tag{18}$$

In summary, the weight of the entire hull structure can be expressed as:

$$W = 2 \times W_p + 4 \times W_c + 2 \times W_t + W_d \tag{19}$$

The total structure weight W is apparently a multivariate function of the design variable. In the process of optimization, hull structure weight can be considered directly as an explicit objective function.

2.4. Stability performance

The ability of a platform to return to its original position after being tilted by an external force (such as wind and waves) is called stability. As shown in Fig. 7, when the semi-submersible platform is tilted to angle θ by an external force, the center of buoyancy of the platform moves from C_0 to C_1, and metacentric M refers to the intersection of the two

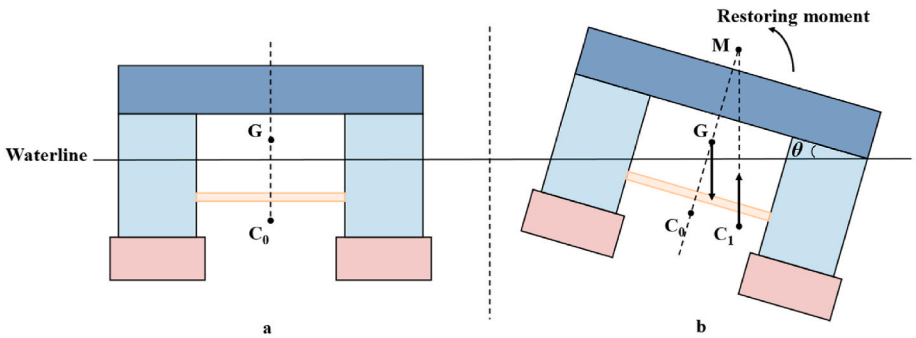


Fig. 7. Initial stability performance analysis (a) stable condition (b) tilted condition.

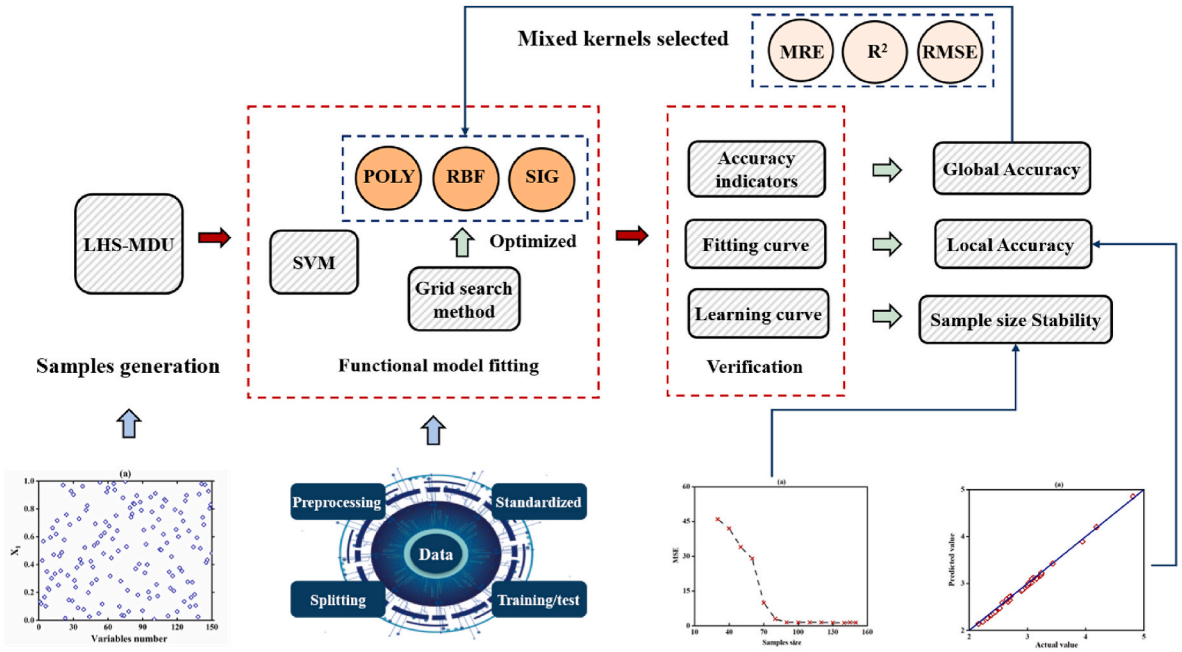


Fig. 8. The establishment and comprehensive verification for whole surrogate model.

buoyancy lines (G is the center of gravity). In this circumstance, gravity and buoyancy are not on the same vertical line, and the restoring moment generated by a couple of forces prompts the platform to return to the original equilibrium position.

The distance between the metacenter M and the center of gravity G is called the metacentric height (GM) of the platform. A great GM provides a good ability for the platform to restore. Therefore, metacentric height can be regarded as an indicator to express the stability of the platform. Since transverse stability has a more significant application in the evaluation of platform safety compared with longitudinal stability, the stability mentioned in the present study refers to transverse stability. GM is considered a constraint condition in the optimization process to ensure that the platform satisfies the stability requirement.

3. Optimization modelling

3.1. Surrogate model

The complex mathematical and physical models in actual engineering often lead to fairly long term calculations, manifested as the nonlinear relationships between multi-dimensional design variables and objective functions. The surrogate model also called the approximate model, has been widely proven to be an efficient technique to solve the issue mentioned above (Anderson and Whitcomb, 1970; Pietrenko-Dabrowska and Koziel, 2019a,b). By generating certain sample points in the design of the experiment, data fitting provides effective ways to shorten the optimization period instead of direct numerical simulation (Matthews et al., 2007; Razavi et al., 2012). The most common form of the surrogate model contains the kriging model (Krige, 1951), the radial basis function model (RBF), the artificial neural network (ANN), and the support vector machine (SVM). In this research, the establishment of a surrogate model is summarized in Fig. 8, referring to the following steps: design of experiments, generation of sample points space, functional model fitting, and accuracy evaluation. A mixed kernel functions SVM model based on grid search optimization and 5-fold cross-validation is selected to construct five specific parameter surrogate models. After determining the relationship between design variables and objective/constraints in surrogate models, triple comprehensive verification is conducted to further ensure the accuracy and robustness of the model.

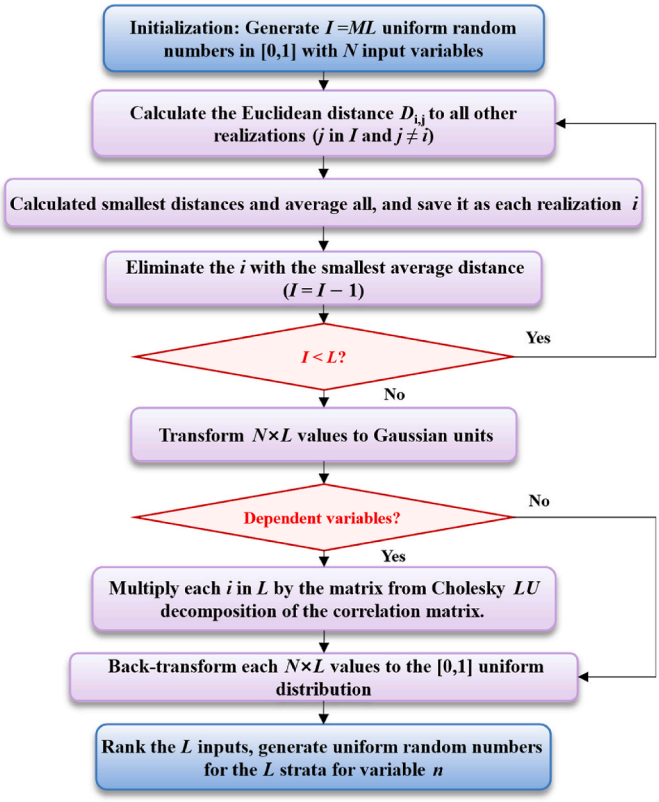


Fig. 9. The process of the sampling points generation by LHD-MDU.

3.1.1. Design of experiment and sampling

DOE (design of experiment) is a mathematical statistics method for experimental arrangement and analysis. DOE has been proven to be an effective technique for making a reasonable test scheme and obtaining ideal results and scientific conclusions on a smaller scale (Wang et al.,

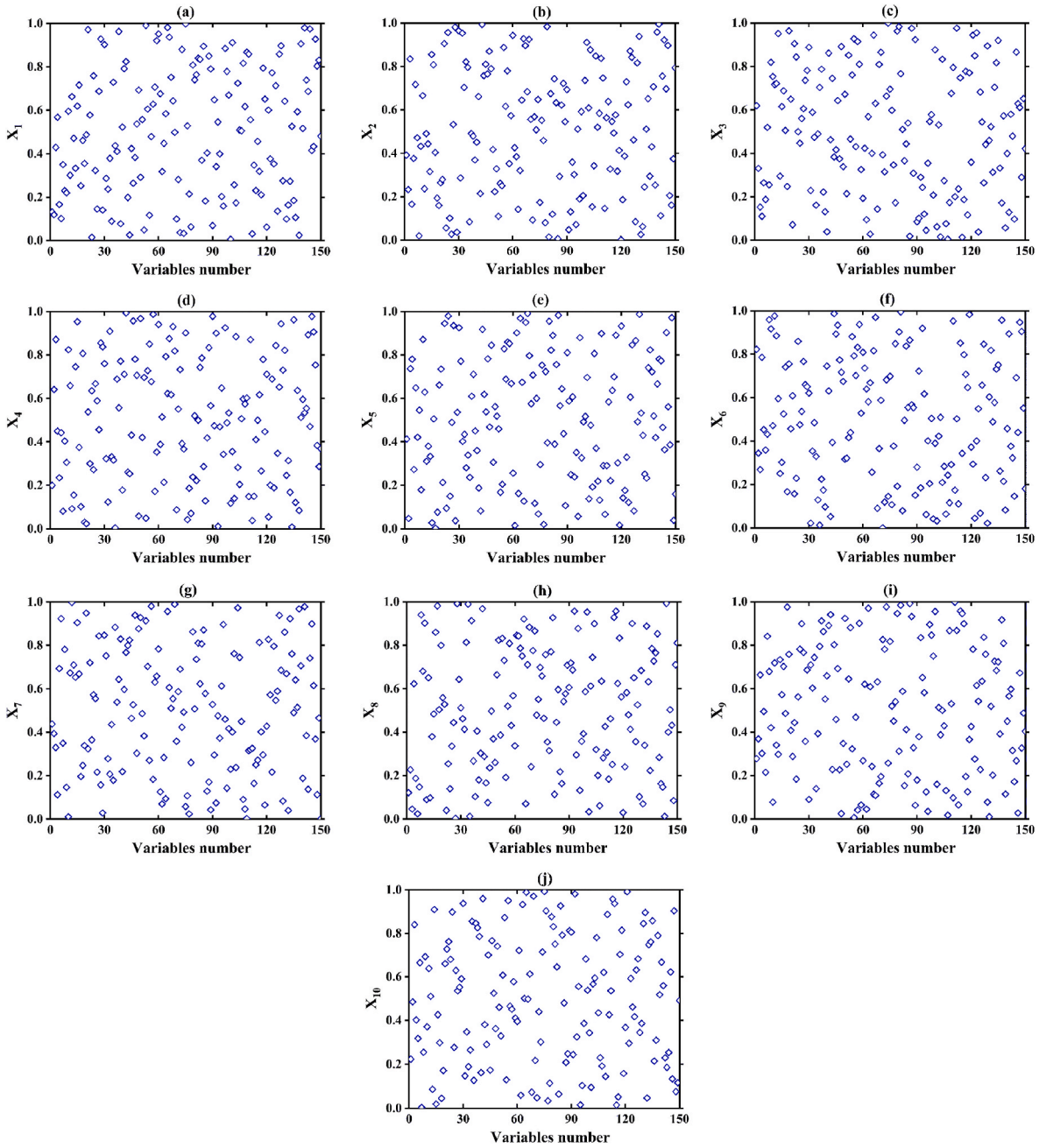


Fig. 10. The generation results in each dimension of samples points. (a) X_1; (b) X_2; (c) X_3; (d) X_4; (e) X_5; (f) X_6; (g)X_7; (h) X_8; (i) X_9; (j) X_{10}.

2020). The common methods for DOE include Full Factorial Designs (FD), orthogonal experimental design (OED), homogeneous design of experiment (HD), Monte Carlo design (MCD), Ladin hypercube design (LHD), and multidimensional uniformity Ladin hypercube design (LHD-MDU).

In the present study, considering that there are ten optimization variables of the SEMI platform, some methods such as FD and OED will lead to a large number of sample generation. Hence, the strategy for choosing the DOE method is to cover the 10-dimensional variable space as much as possible with relatively small sample size. Ladin hypercube method (LH) is a technique of approximately random sampling from a multivariate parametric distribution and has been proven to be an efficient method for multidimensional variable sampling in many studies (Johnson et al., 1990; Leary et al., 2003). In addition, superior prediction accuracy and a simpler constructed procedure make LHD more popular in many studies (Jones and Johnson, 2009; Roshanian and Ebrahimi, 2012; Karolczuk and Kurek, 2022). Hence, LHD can better conform to the core intent of DOE and can be selected as a multivariate sampling method in this study. In recent years, extension from univariate uniformity to multivariate is a typical strategy to optimize the LH algorithm (Joseph and Hung, 2008). The LH-MDU proposed by Deutsch and Deutsch (2012) ensures the improvement and effectiveness with high dimension variables, which has strong adaptability and filling capacity of sample points for multidimensional uniformity.

Fig. 9 shows the flowchart of uniform stratified sampling point generation using LHD-MDU. In this research, LHD-MDU is employed to simultaneously stratify 10 dimensions of the parameter space and generate 150 sample points. Fig. 10 shows the generation results of normalized variables in each dimension. By normalization, the features can be transformed into small scale (0 and 1), and data can also maintain the distribution in the dimension. According to the results, the discrete distribution of the samples in each dimension is satisfactory.

Table 4
The configuration of the laboratory computer.

Configuration	Parameters
CPU	AMD Ryzen 7 5800H with Radeon Graphics
GPU	NVIDIA GeForce RTX 3060 Laptop
System	Windows 10 x64
RAM	16.0 GB
Hard Disk	ZHITAI PC Active 1 TB

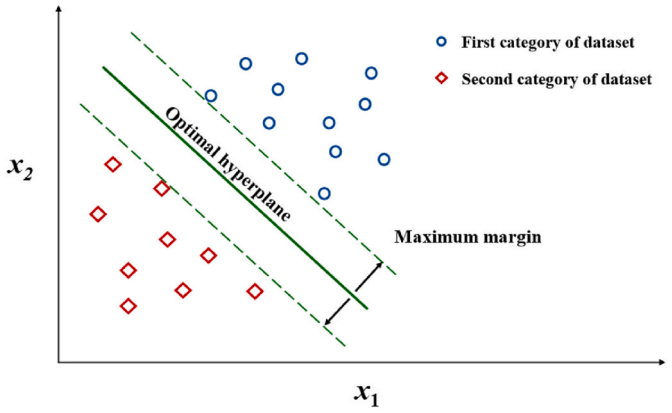


Fig. 11. The schematic diagram of SVM. The red rhombuses: the data classified into the first category; dark blue circle: the data classified into the second category. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Subsequently, hydrodynamic simulation is conducted based on the generated samples to obtain the motion response data and metacentric height, which are used to feed into the surrogate model to obtain the relationship between the optimization objective and design variables. The average time consumption for the hydrodynamic calculation of each group of design variables is 1044 s. The total simulation for 150 samples takes approximately 44 h. All the hydrodynamic calculations are performed on a laboratory computer, for which the configuration is shown in Table 4.

3.1.2. SVM-GSM algorithm

SVM is a generalized linear classifier that performs binary classification of data based on supervised learning (Vapnik et al., 1997). With strict mathematical theory support and strong interpretability, SVM can simplify the usual classification and regression problems. Consequently, SVM has a good application for linear and nonlinear regression with multidimensional variables. Compared with the ANN used in most studies, SVM has a lower risk of overfitting and better suitability in small and medium samples. (Wu et al., 2008; Zhou et al., 2021).

The basic idea of SVM is to determine the hyperplane to classify the data in a training set containing m samples $D = \{(x_1, y_1), (x_2, y_2) \ldots, (x_m, y_m))\}$, as shown in Fig. 11. The optimal hyperplane is the goal of SVM, which divides the data with the maximum margin.

$$\omega^T x + b = 0 \tag{20}$$

where ω determines the direction of the hyperplane. b is the displacement term, which determines the distance between the hyperplane and the origin. Points in the sample that are closest to the hyperplane are called support vectors. The distance r from random point x_i to the hyperplane $\omega^T x + b$ in the sample space is:

$$r = \frac{|\omega^T x_i + b|}{\| \omega \|} \tag{21}$$

The essence of SVM is to maximize the functional margin and find the hyperplane:

$$\begin{aligned} &\min \frac{1}{2} \| \omega \|^2 \\ &y_i\left(\omega^T x_i + b\right) \geq 1, i = 1, 2, \cdots, n \end{aligned} \tag{22}$$

Due to the potential linear inseparable problem, a slack variable ξ_i is introduced to allow partial samples to not be separated. The penalty factor C is introduced to ensure that the samples better satisfy the constraints and decrease the noisy data. Hence, Eq. (22) can be expressed as:

$$\begin{aligned} &\min_{\omega,b,\xi} \frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{m} \xi_i \\ &1 - y_i\left(\omega^T x_i + b\right) - \xi_i \leq 0, \quad \xi_i \geq 0, \quad i = 1, 2, \cdots, m \end{aligned} \tag{23}$$

For the samples in nonlinear problem, kernel functions are proposed to map the data to the higher-dimensional vector space (Talukdar et al., 2020). The function defining this feature map is called the similarity function $\Phi(x)$. In this way, the original low-dimensional calculation can be converted to a high-dimensional space similarity operation. Eq. (24) can be expressed as follows:

$$k(x, x') = \Phi(x)\Phi(x') \tag{24}$$

Then the solution goals of the nonlinear problem with the kernel function can be derived from Eq. (25):

$$\begin{aligned} &\min_{\omega,b,\xi} \frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{m} \xi_i \\ &1 - y_i\left(\omega^T \varphi(x_i) + b\right) - \xi_i \leq 0, \quad \xi_i \geq 0, \quad i = 1, 2, \cdots, m \end{aligned} \tag{25}$$

As shown in Fig. 12, the data samples in 2-dimensional nonlinear space cannot be divided by a specific hyperplane. After the mapping of the kernel function, the original space is mapped to a high dimension,

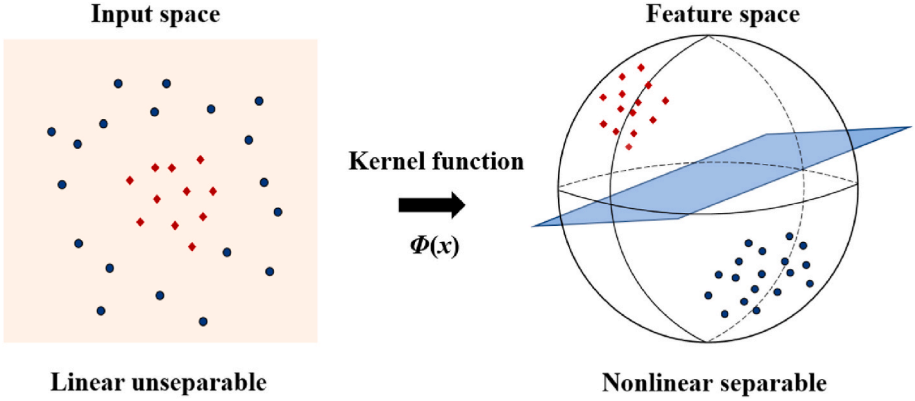


Fig. 12. The characteristic of mapping by kernel functions.

Table 5
The selected kernel functions in mixed kernels SVM.

Kernel function	Expression	Control Parameter
POLY	$K(x_i,x_j) = [\gamma(x_i^T x_j)]^d$	d,γ,C
RBF	$K(x_i,x_j) = \exp(-\gamma \parallel x_i - x_j\parallel^2)$	γ,C
SIG	$K(x_i,x_j) = \tanh(\gamma x_i^T x_j)$	γ,C

and a notable hyperplane can be constructed to divide the data sample in the feature space. The mapping effects depend on the different kernel function types. The most commonly used kernel functions in high-dimensional nonlinear problems are the polynomial kernel (POLY), the Gaussian radial basis function (RBF), and Sigmoid kernel (SIG), which are indicated in Table 5. POLY kernel is very suitable for image processing due to its good global property, while RBF has the widest application due to its good prediction accuracy and adaptability. SIG is widely used in complicated nonlinear problems as a nonlinear function of neurons. This paper selects the three kernels to build the mixed kernel SVM model based on their predicted effect.

The grid search (GS) is an exhaustive search method for specific parameters. As control parameters number in kernels is no more than three, this research is based on possible combination results to generate a grid and update the parameter until the best value in a specific parameter range is found. This process relied on continuous iteration to obtain the fitness score, which is tested by 5-fold cross-validation. In the present study, the control parameters in three kernels are optimized by GS. Fig. 13 shows the process of GS method and 5-fold cross-validation.

3.1.3. Verification of surrogate model

In most previous studies, the verification of the surrogate models was based on a single approach, resulting in an incomplete evaluation of the robustness and stability of model. To improve these issues and enhance the reliability of the surrogate model, this study conducted a triple comprehensive verification of the surrogate model, manifested as accuracy indicators, fitting curves, and learning curves.

Accuracy indicators, including mean relative error (MRE), root mean square error (RMSE), and complex correlation coefficient (R^2), are used to quantitatively evaluate the predicted accuracy based on the original response value. They are expressed as Eqs. (26)–(28), respectively.

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - y_{pi}|}{|y_i|} \tag{26}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_{pi})^2} \tag{27}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - y_{pi})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{28}$$

where y_i, y_{pi} and $\bar{y}$ represent the original sample response values, predictive values, and average original sample values, respectively. MRE and MSE reflect the error level between the actual value and predicted value, which is expected to be as low as possible so that a higher predicted accuracy can be ensured for the surrogate model. R^2 is the index that represents the quality of the model; the closer the value is to 1, the higher the accuracy of the model will be.

Generally, MRE and RMSE are supposed to be lower than 0.2, and R^2 is supposed to be greater than 0.9 for adequate prediction accuracy. According to 5-fold cross-validation, 120 samples and 30 samples are selected as the training set and test set, respectively. Table 6 shows the results of three accuracy indicators in different kernel SVM-GSs. Discrepancy induced by the three kernels are found to result in different predicted accuracy. For MPM_{heave}, MPM_{roll}, and MPM_{pitch}, the RBF kernel has the best prediction effect, whereas the POLY kernel and SIG kernel show better prediction effects on MPM_{acce} and GM, respectively. Consequently, based on the corresponding optimal predicted effect, the surrogate models are established by SVM-GS with three mixed kernels, MPM_{heave}, MPM_{roll}, and MPM_{pitch} adopt the RBF kernel, MPM_{ace} and GM adopt the POLY and SIG kernel, respectively.

However, accuracy indicators are only a reflection of global predicted accuracy, and the local predicted effect cannot be obtained by this method. Fitting curves are adopted to address this issue. The fitting curves can intuitively evaluate the predicted local effect according to the direct comparison of the predicted and actual values for test samples. Fig. 14 shows the fitting curves for five specific parameter surrogate models. All data points are densely distributed on both sides of the 45-degree line, and no individual points significantly deviate from the 45-degree line. Hence, an excellent local prediction effect is guaranteed in the surrogate model.

Subsequently, learning curves are selected to demonstrate the accuracy stability according to changing the sample size. As the sample size increases, if the mean square error (MSE) of the model decreases rapidly and eventually reaches a steady value, then selected sample size can be determined have reach a suitable value. Fig. 15 shows the learning curves of five specific parameter surrogate models. Basically,

Table 6
Accuracy indicators in the different kernels.

Kernels	Accuracy indicators	MPM_{heave}	MPM_{roll}	MPM_{pitch}	GM	MPM_{acce}
POLY	RMSE	0.1156	0.1070	0.0865	0.0712	0.0652
	MRE	0.1024	0.0951	0.0817	0.0854	0.0521
	R_2	0.9104	0.9252	0.9371	0.9317	0.9668
RBF	RMSE	0.0681	0.0654	0.0891	0.0967	0.1243
	MRE	0.0595	0.0573	0.0742	0.0801	0.0974
	R_2	0.9694	0.9509	0.9661	0.9327	0.9275
SIG	RMSE	0.9861	0.1847	0.1277	0.0687	0.1497
	MRE	0.0864	0.1689	0.1050	0.0633	0.1274
	R_2	0.9158	0.8847	0.9296	0.9702	0.9099

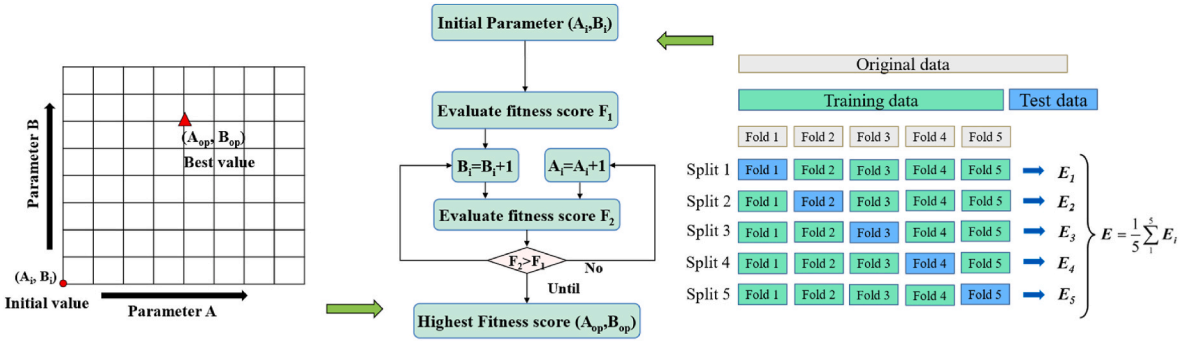


Fig. 13. The schematic diagram of Grid search method and 5-fold cross-validation.

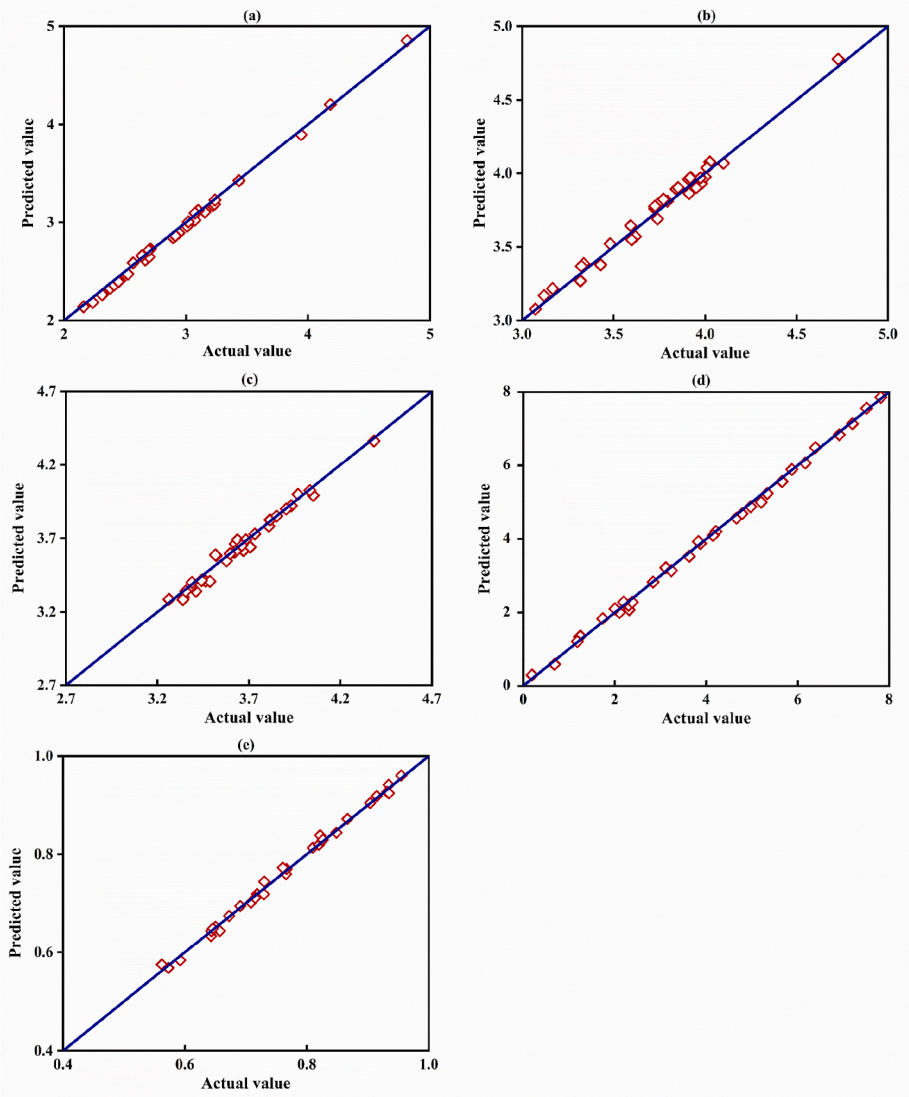


Fig. 14. The fitting curves of five parameter specific surrogate model. (a) MPM_{heave}; (b) MPM_{roll}; (c) MPM_{pitch}; (d) GM; (e) MPM_{acce}.

when the sample size exceeds the 100 groups, the MSE of the five specific parameter surrogate models is very low and remains stable. Consequently, the 150 samples in the present study can be determined to provide sufficient stable accuracy and robustness for five surrogate models.

3.2. Multi-objective optimization model

Multi-objective problems commonly exist in engineering fields. In most scenarios, due to the contradiction among the objectives, each objective cannot reach the optimal value simultaneously. For these problems, a single global optimal solution cannot be obtained. Instead, the goal pursued is to find a compromise situation between the objectives, which is a group of relative optimal and balanced solutions called Pareto-optimal solutions (or non-dominated solutions). In this research, the weight of hull structure (W), MPM_{heave} and MPM_{roll} are selected as three objective functions.

3.2.1. Non dominated sorting genetic algorithm-II (NSGA-II)

The genetic algorithm is a kind of evolution algorithm inspired by species evolution in genetics. The essential idea of the genetic algorithm is to imitate the crossover and mutation in species evolution to finish the selection process based on the survival of the fittest. NSGA-II is the most popular multi-object genetic algorithm due to its predominant fast-running and high efficiency (Deb, 2011). Different from traditional genetic algorithms, non-dominant sorting algorithms and elitist strategies are utilized to improve computing efficiency and accuracy. In addition, the rank and crowding distance are used to reflect the fitness value of chromosomes in a population, and the crowding degree and crowding comparison operators are introduced to ensure that the individuals can uniformly expand to the entire Pareto domain (Deb et al., 2005; Sivaraj and Ravichandran, 2011).

The process of NSGA-II is as follows: First, initialize parent population P_0, and use crossover and mutation operations to generate child population Q_0. Second, conduct non-dominated sorting to R_0, which is formed by P_0 and Q_0. Then, rank different levels of non-dominated solution sets Z_1, Z_2, Z_3 … Third, sort the ranked non-dominated solution set by the crowding distance, and the top N solutions are obtained according to the fitness level to constitute the parent population of the next iteration. Fourth, repeat the above three steps until the solution results converge. The description of the main principle in optimization is shown in Fig. 16. Solution members in Z_1, are better than the Z_2 and Z_3 due to the non-dominated ranking. Especially, if the members number in Z_1 cannot reach the population size, all members in Z_1 will be chosen into P_1, and the rest of the members in P_1 will come from Z_2 and Z_3 in the same ways as Z_1, until the P_1 members equal the population size. The

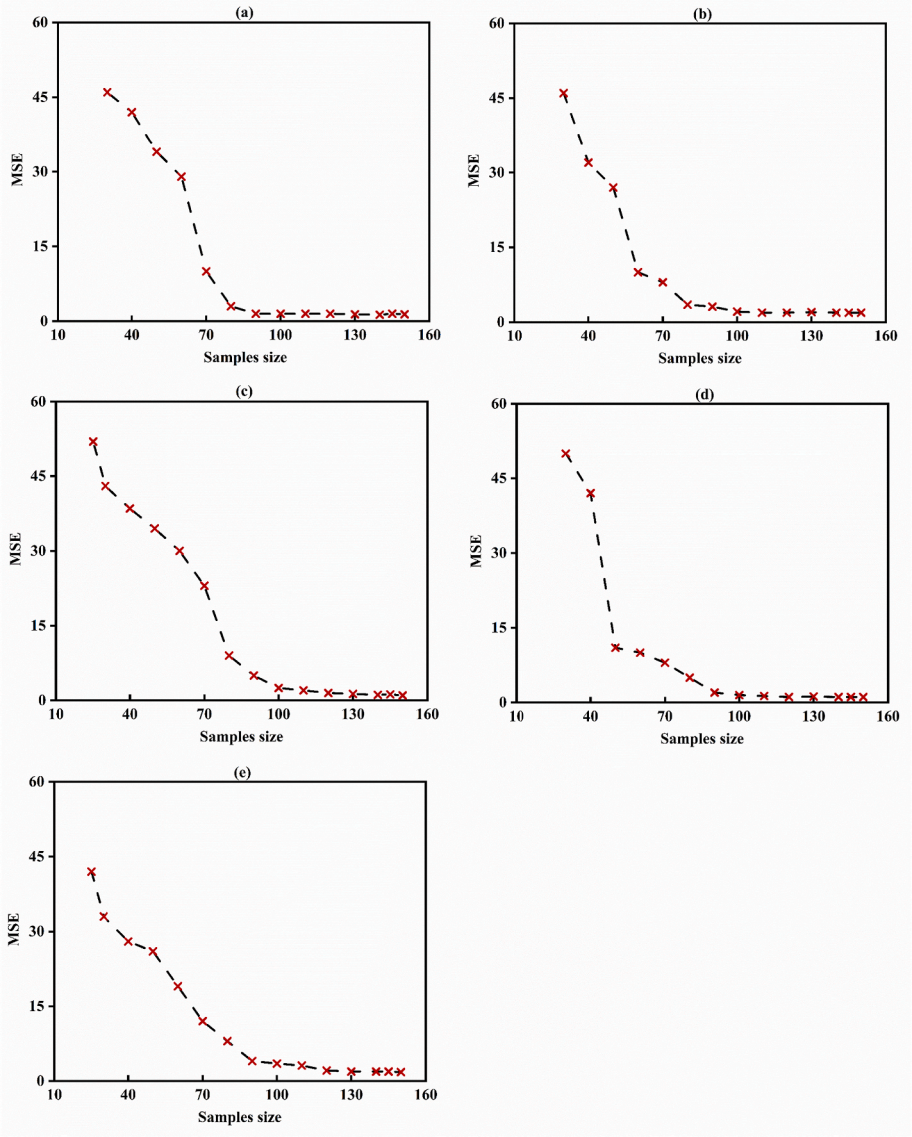


Fig. 15. The learning curves of five specific parameters surrogate model (a) MPM_{heave}; (b) MPM_{roll}; (c) MPM_{pitch}; (d) GM (e) MPM_{acce}.

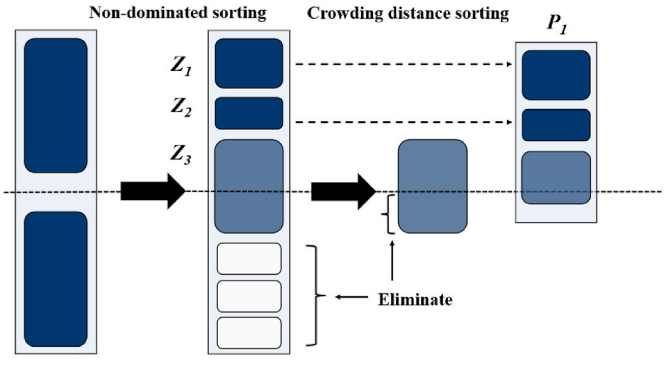


Fig. 16. The description of main principle in NSGA-II optimization.

prior members ranked in Z_3 will be accepted, while the remainder will be rejected and eliminated. This process is based on Python to achieve the iteration until finding the final non-dominate solutions.

3.2.2. Objective function and constraints

The three objectives function F(x) with the weight of hull structure (W), MPM_{heave}, and MPM_{roll}, can be written as:

$$\min : F(x) = f[W, MPM_{heave}, MPM_{roll}] \tag{29}$$

$$\text{s.t.} \begin{cases} GM > 6.25m \\ MPM_{acce} < 0.85m/s^2 \\ MPM_{pitch} < 6.0^{\circ} \end{cases} \tag{30}$$

MPM_{heave} and MPM_{roll} reflect the hydrodynamic performance and are supposed to be minimized to ensure the stability and safety of the platform. W reflects the economy of hull form, and it should be minimized to decrease the cost of the hull. In addition, considering the particular operating conditions and personnel safety, motion response should be limited by setting the constraints to Eq. (30). In this paper, the constraints are the transverse metacentric height, maximum pitch response, and maximum heave acceleration, which are expressed in Eq (30). The constraint condition is added by penalty function methods (Angermann and Wang, 2007; Nielsen et al., 2008), which can be expressed as:

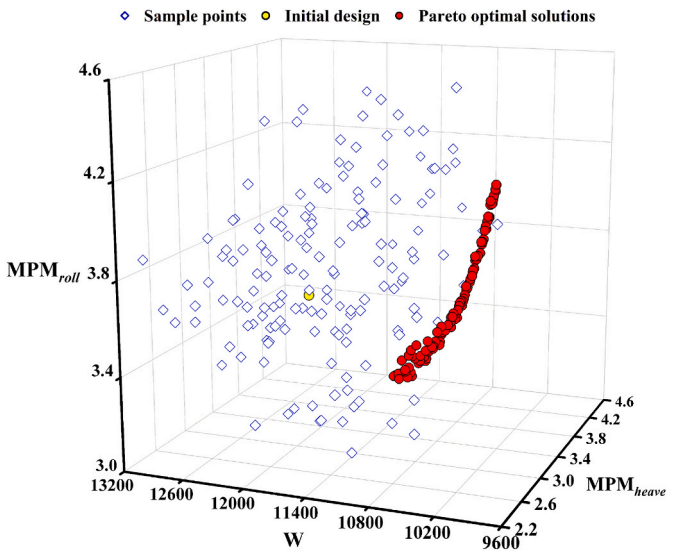


Fig. 17. The Pareto solutions set of optimization results. Red dots: Pareto optimal solutions; blue rhombuses: generated sample points in surrogate model; yellow dot: initial design of the platform. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

$$P(x) = \sum_{i=1}^{k} \max[0, -g_i(x)]^2 + \sum_{j=1}^{l} \left|h_j(x)\right|^2 \tag{31}$$

where g_i and h_j are the ith inequality constraint and jth equation constraint, respectively. Then, the Eq. (29) can be written as:

$$\min : F(x, \sigma) = f([W, MPM_{heave}, MPM_{roll}]) + \sigma P(x) \tag{32}$$

where σ is a very large positive number, called penalty factor. In this way, the optimization problem with constraints can be transformed into an unconstrained problem.

4. Results and discussion

4.1. The results of pareto-optimal solution

NSGA-II conducts the optimization based on the determined relationship between the design variable and the objectives (or constraints) in five specific parameter surrogate models, and the population size and maximum generation are both set to 200. Finally, 84 non-dominated Pareto-optimal solutions are obtained, and the optimization results are shown in Fig. 17. As presented, the yellow dot represents the initial design of the platform, and the blue rhombuses represent the generated samples in surrogate model, which constitute the under optimization space. The red dots are non-dominated solutions, which indicate the results after optimization. As the goal in this study, the red dots reflect the Pareto optimal state, which means there is no more space to further optimize. The two-dimensional graph, which is shown in Fig. 18, is utilized to better describe the non-dominated solutions. The non-dominated solutions distributed in two-dimensional space indicate some correlation between each objective. As the total hull steel weight decreases, MPM for heave motion and roll motion get increasing, which reflects that the lightweight design in the platform is contradictory to the minimum heave and roll motion responses. However, a positive relationship between the MPM of heave motion and MPM of roll motion is apparent in Fig. 18(c), which means that a good performance in heave motion also positively influences the roll motion of the semi-submersible platform.

As shown in Table 7, nine groups of typical optimal solutions are selected and divided into three optimized feature categories based on the minimum objective criterion. Category one emphasizes the cost of the platform, considering the smallest total hull steel weight. And

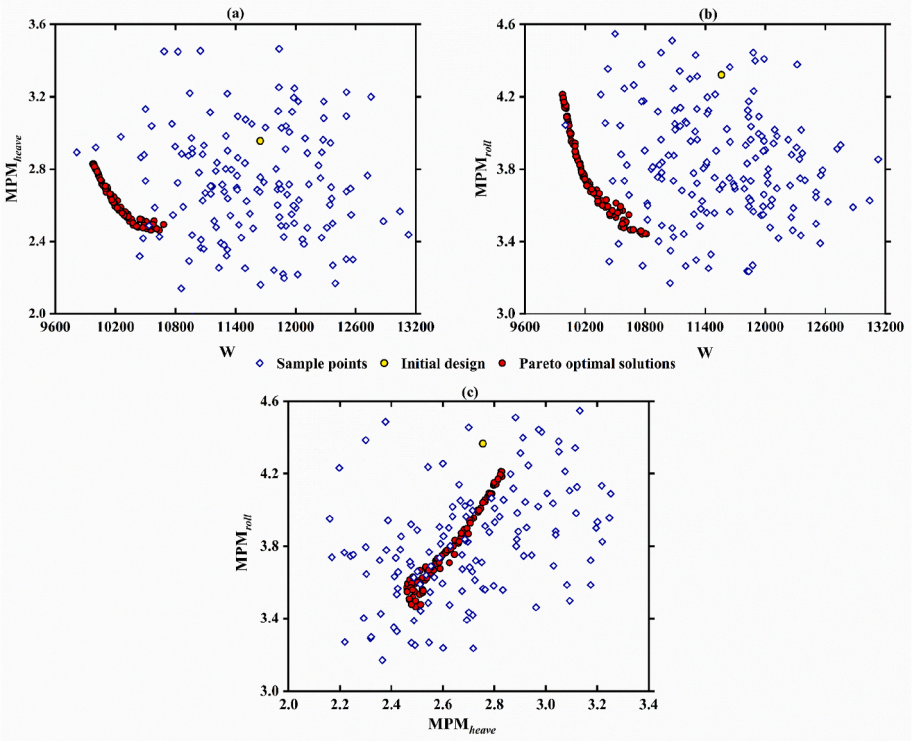


Fig. 18. Pareto solutions set in two dimensions (a) Hull steel weight and MPM$_{heave}$; (b) Hull steel weight and MPM$_{roll}$; (c) MPM$_{heave}$ and MPM$_{roll}$. Red dots: Pareto optimal solutions; blue rhombuses: generated sample points in surrogate model; yellow dot: initial design of the platform. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 7
The results of Pareto-optimal solutions in three typical categories.

		Initial	Opt-Category 1			Change Ratio 1	Opt-Category 2			Change Ratio 2	Opt-Category 3			Change Ratio 3
			No.1	No.2	No.3		No.4	No.5	No.6		No.7	No.8	No.9	
Objectives	W(t)	11646.698	9974.974	9981.332	9981.341	**−14.317%**	10573.457	10637.650	10554.267	−9.086%	10771.098	10811.765	10797.327	−7.327%
	MPM_{heave}(m)	2.956	2.829	2.823	2.829	−4.364%	2.460	2.462	2.463	**−16.723%**	2.651	2.636	2.644	−10.566%
	MPM_{roll}(°)	4.245	4.211	4.194	4.184	−1.146%	3.570	3.547	3.592	−15.909%	3.440	3.441	3.445	**−18.916%**
Design Variables	X_1	72.500	73.362	73.375	73.369	1.198%	74.889	74.884	74.893	3.295%	75.026	75.020	75.020	3.497%
	X_2	82.500	78.500	78.502	78.570	−4.819%	79.142	79.894	78.917	−3.857%	81.782	81.790	81.790	−0.864%
	X_3	8.500	7.516	7.516	7.509	−11.604%	8.103	8.103	8.103	−4.671%	8.074	8.090	8.090	−5.867%
	X_4	16.500	13.507	13.508	13.507	−18.137%	13.815	13.815	13.815	−16.273%	13.624	13.625	13.624	−17.443%
	X_5	17.500	17.001	16.996	17.000	−2.863%	16.999	16.996	17.000	−2.867%	16.985	16.998	17.000	−2.890%
	X_6	22.800	19.806	19.811	19.806	−13.124%	21.391	21.295	21.391	−6.302%	20.364	20.349	20.349	−10.731%
	X_7	105.800	98.308	98.378	98.307	−7.060%	98.346	98.658	98.346	−6.947%	98.473	98.345	98.317	−7.015%
	X_8	11.850	10.854	10.853	10.854	−8.408%	10.850	10.853	10.853	−8.422%	10.850	10.868	10.868	−8.338%
	X_9	17.000	16.158	16.173	16.172	−4.896%	17.400	17.400	17.400	2.353%	17.282	17.293	17.226	1.767%
	X_{10}	4.350	4.268	4.268	4.268	−1.962%	4.254	4.254	4.254	−2.207%	4.393	4.391	4.391	0.964%

category two and three emphasize the safety and stability of platform based on minimum MPM_{heave} and MPM_{roll}, which also correspond to their optimized objective.

Table 7 shows the average change ratio for nine non-dominated solutions compared with the initial design. The discrepancy in the degree of reduction among the three categories can be recognized. Category one focuses on the weight, which is reduced by 14.317%, and the MPMs for heave motion and roll motion are reduced by 4.364% and 1.146%, respectively. And category two focuses on heave, which is reduced by 16.723%, and the weight and MPM_{roll} are reduced by 9.086% and 15.909%, respectively. For category three, the focused MPM_{roll} is reduced by 18.916%, and the MPM_{heave} and weight are reduced by 10.566% and 7.327%, respectively.

Fig. 19(a) shows the constraint value of nine typical optimal solutions in three categories. Three constraints can be seen to satisfy the limit after optimization, resulting in satisfactory performance in platform stability.

$[X]_{10}$ denotes the variables with ten dimensions; $[Y]_5$ denotes the five outputs value by surrogate model.

Fig. 20 shows the running time during four stages, including design variables acquisition, numerical simulation of samples, data training in the surrogate model, and optimization process. The numerical simulation of the samples stage is the most time-consuming. Each group design variable lasts 1044 s, and the total simulation process for all samples requires 2610 min. The remaining three stages for LHS-MDU samples generation, surrogate model construction, and optimization procedure take less than half an hour to complete. Compared to the design spiral that relies on massive iteration and thousands of numerical simulations or even longer (Pawling et al., 2017), this approach based on the surrogate model can quickly find some groups of optimal design solutions under the same sea conditions.

4.2. Frequency and time domain analysis

Based on the optimization results, the optimized design variables in three categories can be employed to establish the simulation model. The direct numerical simulation is conducted to verify the optimization results in three typical categories. Among these objectives, the total hull structure weight is an explicit objective function of the design variable, hence the simulation results are consistent with the results after optimization by the surrogate model. Fig. 21 shows the heave RAOs of the initial design and No.1, No.4, and No.7 Pareto optimal solutions in frequency domain analysis. As the wave energy is concentrated mainly in the 10–20 s, the second peak significantly affect the heave motion. Consequently, the decrease of the area under the second peak can predominantly reflect the optimized degree. As shown, heave motion and roll motion all decrease in three selected optimal groups, in which No.4 has the best heave motion optimized effect, whereas No.7 shows the most notable optimized effect for roll motion.

Fig. 19(b) shows the constraint value in optimization results. The constraints conditions are also satisfied in the numerical simulation. Table 8 shows the comparison between the optimized values by surrogate model and the values by direct numerical simulations. There are slight differences between the values obtained from the direct numerical simulation and those obtained by the surrogate model. The cause of errors may be due to the collective effect of errors in the surrogate model and finite element analysis. On the whole, the maximum relative errors in the three groups of non-dominated solutions are less than 5%, and the optimization model can provide satisfactory accuracy for the conceptual design phase.

In extreme sea state, the motion response should be simulated in time domain analysis, which can reflect actual motion for floaters to ensure maximum safety. As shown in Fig. 22, the simulations for No.4 and No.7 optimal designs are conducted in the time domain for comparison with the initial design. The heave motion and roll motion decrease remarkably in No.4 and No.7, respectively, which demonstrates that a sufficient

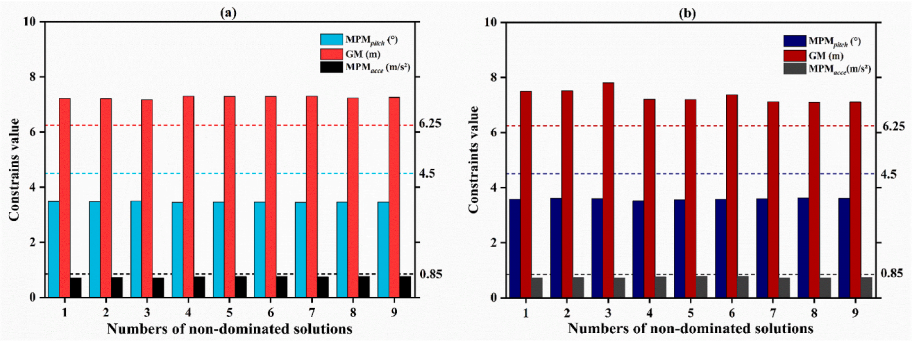


Fig. 19. The constraints values for nine non-dominated solutions. (a) the optimization results (b) the direct simulation results (Red dotted line: the limited value of GM; Blue dotted line: the limited value of pitch; Black dotted line: the limited value of MPM_{acce}.). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

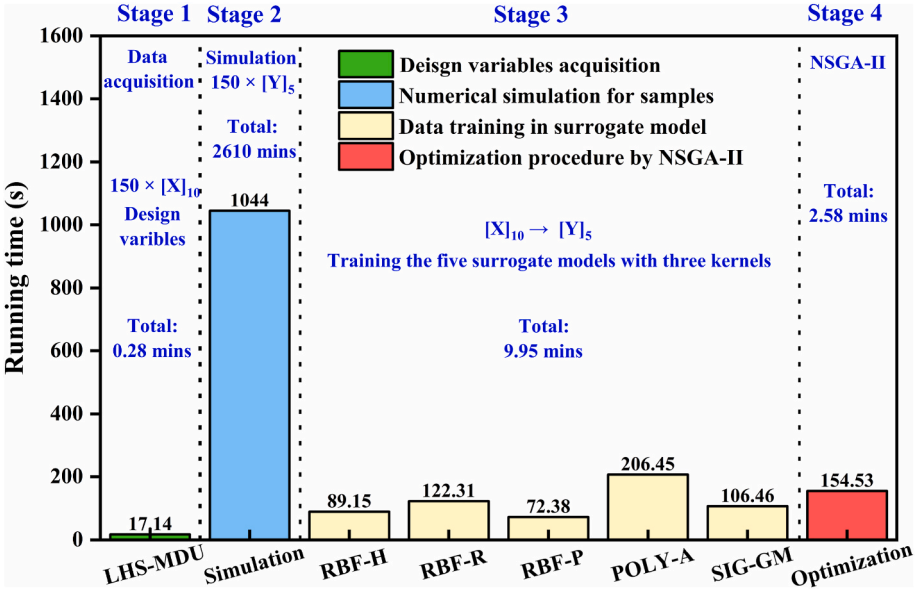


Fig. 20. The time consumption for four stages, including data acquisition, numerical simulation, surrogates model construction, and optimization procedure. RBF-H: MPM_{heave} by SVM with RBF kernel; RBF-R: MPM_{roll} by SVM with RBF kernel; RBF-P: MPM_{pitich} by SVM with RBF kernel; POLY-A: MPM_{acce} by SVM with POLY kernel; SIG-GM: metacentric height by SVM with SIG kernel.

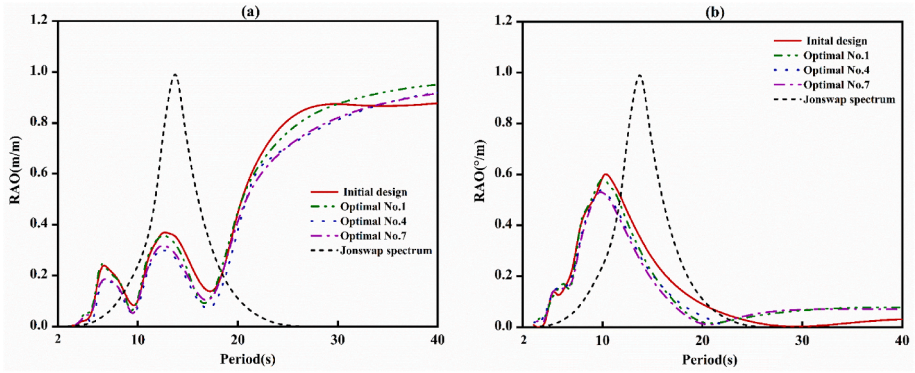


Fig. 21. The RAOs comparison between initial design and optimized design in frequency domain analysis. (a) heave RAOs; (b) roll RAOs.

optimized effect can also be recognized in actual motion.

In addition, the iteration process of ten variables in NSGA-II optimization is indicated in Fig. 23. With the iteration increasing, the value of all design variables tends to converge. It can be found that the final converged value corresponds to the value of variables in three typical optimized categories, which means that the selected three typical categories have a sufficient representative meaning for the optimization results.

Table 8
Comparison of optimized value and numerical value in frequency domain analysis. SM-value: the value obtained by surrogate model; NS-value: the value obtained by direct numerical simulation. δ_r: relative errors.

	Output	Optimal No.1			Optimal No.4			Optimal No.7		
		SM-value	NS-value	δ_r	SM-value	NS-value	δ_r	SM-value	NS-value	δ_r
Objectives	W	9974.974	9974.974	0	10573.457	10573.457	0	10771.098	10771.098	0
	MPM_{heave}	2.829	2.910	2.784%	2.462	2.404	2.413%	2.651	2.753	3.705%
	MPM_{roll}	4.211	4.353	3.262%	3.573	3.441	3.836%	3.440	3.289	4.591%
Constraints	MPM_{pitch}	3.486	3.674	4.073%	3.441	3.512	2.022%	3.442	3.554	3.151%
	GM	7.222	7.398	2.379%	7.294	7.215	1.095%	7.296	7.122	2.443%
	MPM_{acce}	0.703	0.731	3.830%	0.741	0.754	1.724%	0.739	0.763	3.145%

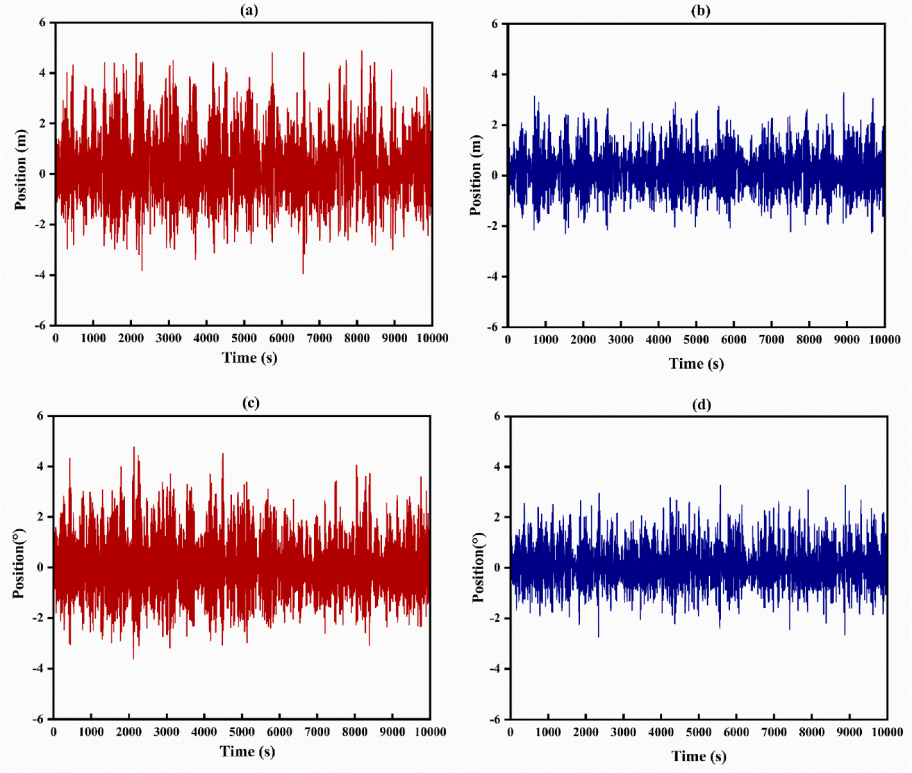


Fig. 22. Comparison of motion response between initial design and optimized design in time domain analysis. (a) heave motion for initial design; (b) heave motion for No.4 optimal design; (c) roll motion for initial design; (d) roll motion for No.7 optimal design.

4.3. Screening and evaluation of the main factors

The Pearson correlation coefficient, Spearman correlation coefficient and Kendall correlation coefficient methods were employed to analyze and evaluate the main factors of structural weight, MPM_{heave}, and MPM_{roll}. As three important coefficients in statistics, they are used to measure the degree of correlation.

The Pearson coefficient is mainly utilized to reflect the degree of linear correlation (Andri et al., 2014), and can be defined as:

$$R_P(X_i, Y_i) = \frac{\sum_{i=1}^{k}(X_i - X_m)(Y_i - Y_m)}{\sqrt{\sqrt{\sum_{i=1}^{k}(X_i - X_m)^2} \times \sqrt{\sum_{i=1}^{k}(Y_i - Y_m)^2}}} \tag{33}$$

where X_i and Y_i represent the datasets of two variables, X_m and Y_m denote the mean values of X_i and Y_i, respectively, k is the total number of test samples.

The Spearman coefficient is usually utilized to evaluate the nonlinear correlation by solving the rank of the data matrix (Musarat et al., 2020), and can be expressed as:

$$R_S(X_i, Y_i) = 1 - \frac{6\sum_{i=1}^{k}(X_{ir} - Y_{ir})^2}{k(k^2 - 1)} \tag{34}$$

where X_i, Y_i, and k are the same as Eq. (33), X_{ir} and Y_{ir} denote the ranks of X_i and Y_i.

Kendall correlation coefficient is another alternative method to evaluate the nonlinear correlation and can be more efficient for dealing with ranking problems. The value can be calculated by:

$$R_K(X_i, Y_i) = \frac{2\sum_{i=1}^{k}\sum_{j=1}^{k}\varepsilon_i}{k(k-1)} \tag{35}$$

where X_i, Y_i, and k are the same as Eq. (33) and Eq. (34), ε_i denotes the criterion function, and the value is determined by the rank of sample points.

The values of the three correlation coefficients are all between −1 and 1. And a negative value means a negative correlation, while a positive value indicates a positive correlation. The closer the absolute value is to 1, the higher the correlation between variables. Many studies

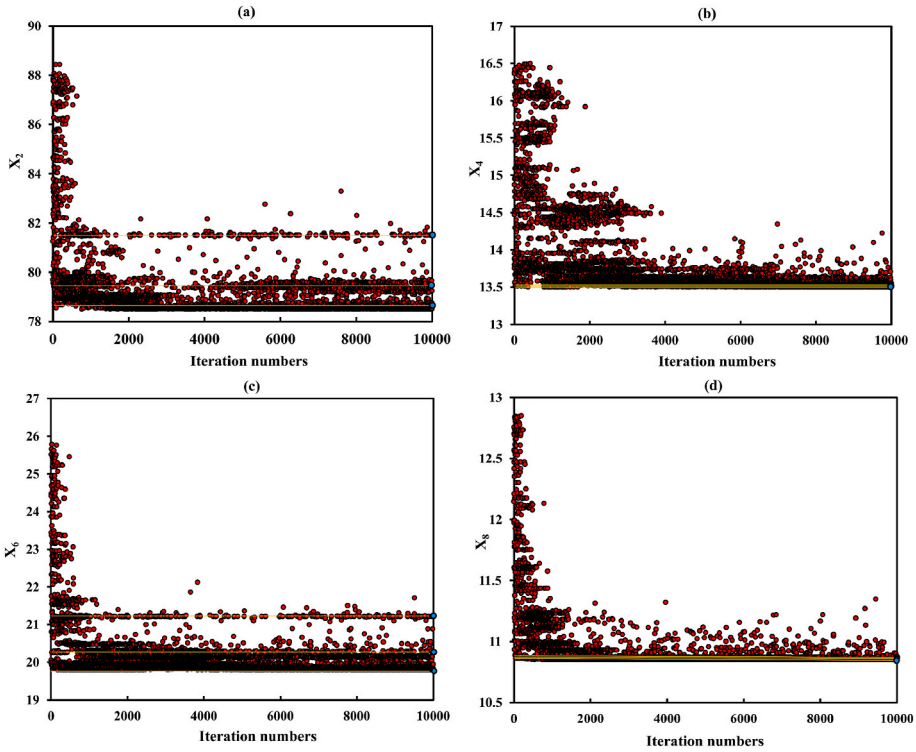


Fig. 23. The iteration process of four representative design variables. (a) X_2; (b) X_4; (c) X_6; (d) X_8.

Table 9
The correlation coefficients of output objective.

Output	Correlation coefficients	Design Variables									
		X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
W	Pierson	0.4487	0.4516	0.2394	0.4060	0.5874	0.3200	0.3180	0.0076	0.1343	0.2021
	Spearman	0.4532	0.4423	0.2475	0.3850	0.5587	0.3116	0.3129	−0.0107	0.1105	0.1915
	Kendall	0.3068	0.2995	0.1645	0.2607	0.3905	0.2115	0.2083	−0.0058	0.0803	0.1291
	Ave-value	0.4029	0.3978	0.2171	0.3505	0.5122	0.2810	0.2797	−0.0030	0.1084	0.1742
MPM_{heave}	Pierson	−0.1267	0.1459	−0.0511	−0.1066	−0.1733	−0.1144	−0.1011	0.0972	−0.2503	0.0749
	Spearman	−0.2751	0.0679	−0.0718	−0.0803	−0.1728	−0.1171	−0.1778	0.1709	−0.3219	0.0144
	Kendall	−0.1936	0.0412	−0.0403	−0.0538	−0.1144	−0.0793	−0.1258	0.1156	−0.2194	0.0052
	Ave-value	−0.1985	0.0850	−0.0544	−0.0802	−0.1535	−0.1036	−0.1349	0.1279	−0.2639	0.0315
MPM_{roll}	Pierson	−0.0457	−0.0423	−0.1929	−0.0070	−0.2544	−0.3740	0.0051	0.3187	−0.5365	−0.1144
	Spearman	−0.0170	−0.0140	−0.2265	−0.0297	−0.1845	−0.3273	0.0037	0.3213	−0.5677	−0.1207
	Kendall	−0.0112	−0.0162	−0.1504	−0.0144	−0.1233	−0.2273	0.0024	0.2217	−0.3949	−0.0835
	Ave-value	−0.0246	−0.0241	−0.1899	−0.0170	−0.1874	−0.3096	0.0037	0.2872	−0.4997	−0.1062
GM	Pierson	0.0174	0.0177	−0.0940	−0.0448	−0.0335	0.0681	0.0836	0.1264	0.1133	−0.1249
	Spearman	0.1636	0.0280	0.0772	0.1188	0.5421	0.1853	0.1659	0.5199	0.1181	0.0259
	Kendall	0.1190	0.0152	−0.0551	0.0762	0.3799	−0.1234	−0.1093	0.3665	0.0813	−0.0185
	Ave-value	0.1000	0.0203	0.0754	0.0799	0.3185	0.1256	0.1196	0.3376	0.1043	0.0564
MPM_{pitch}	Pierson	−0.1227	−0.1049	−0.0741	0.1487	−0.1532	−0.1783	0.3463	0.4356	−0.4611	−0.0406
	Spearman	−0.1600	−0.1365	−0.0521	0.1928	−0.1542	−0.1820	0.3886	0.4011	−0.4063	−0.0362
	Kendall	−0.1058	−0.1024	−0.0459	0.1206	−0.1038	−0.1262	0.2651	0.2783	−0.2815	−0.0259
	Ave-value	−0.1295	−0.1146	−0.0574	0.1540	−0.1371	−0.1622	0.3334	0.3717	−0.3830	−0.0342
MPM_{acce}	Pierson	0.1146	0.1598	−0.0164	−0.2130	−0.3084	−0.2340	−0.0683	0.2309	−0.4556	−0.0810
	Spearman	0.1190	0.1420	−0.0416	−0.2056	−0.3837	−0.2467	−0.1007	0.2701	−0.4267	−0.0237
	Kendall	0.0724	0.0988	−0.0265	−0.1426	−0.2418	−0.1660	−0.0649	0.1837	−0.2961	−0.0115
	Ave-value	0.1020	0.1335	−0.0282	−0.1871	−0.3113	−0.2156	−0.0780	0.2282	−0.3928	−0.0387

have pointed out that the three coefficients have different advantages and disadvantages in some situations (Bonett and Wright, 2020; Fan et al., 2014); hence, the three coefficients are often used together to evaluate the correlation between variables (Li et al., 2021).

As listed in Table 9, three types of the correlation coefficients between the output objective and input design variables are calculated, and the averages of the three coefficients are regarded as comprehensive evaluation results. Three constraints in the study are also analyzed by three correlation coefficients, which are shown in Table 9. As shown in Fig. 24, the main influence factors can be recognized intuitively based on the rank of each variable correlation coefficient. For total hull structure weight, X_5 and X_1 are the top two ranks, which means that column length and deck width provide the most contribution to weight and are the main influence factors in all variables. In terms of the heave and roll response, X_9 and X_1, X_9 and X_6 have the highest correlation, respectively. The correlation coefficient of X_9 is negative, indicating that the heave and roll are inversely related to the hull draft. In addition, although X_{10} has a low ranking as an influence factor in most cases, X_{10}

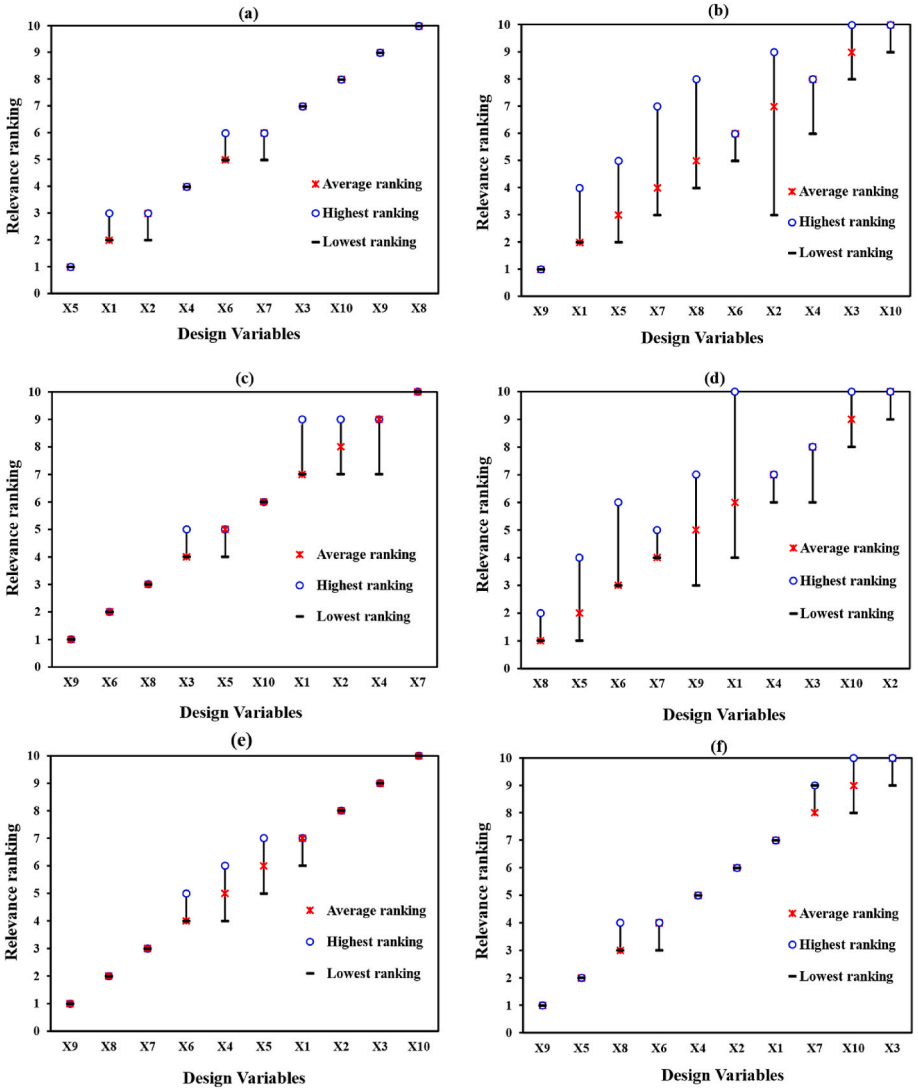


Fig. 24. Comprehensive ranking of the correlation in influence factors. (a) hull steel weight; (b) heave; (c) roll; (d) metacentric height; (e) pitch; (f) heave acceleration.

still plays a nonnegligible role in metacentric height, roll response, and total weight, which means that some studies directly exclude transverse brace diameter from design variables, lacking a certain scientific rigor.

Based on the above analysis, the two design variables that can significantly affect each objective are screened. Fig. 25 shows the relationship between the output objectives and the two design variables. Through the mapping plane constituted by the variables, the color trend from red to purple can be observed as the value of the output objective decreases. A larger deck width, column height, and hull draft will lead to worse heave and roll performance, and hence should be minimized while considering the hydrodynamic response in the conceptual design phase of the platform. In addition, the total weight indicates strong sensitivity to deck width and column length, which should be curtailed to ensure the economy.

5. Conclusions

In this work, a multi-objective optimization is conducted for the semi-submersible platform. The economy and motion stability are considered optimized goals, including the lightest hull structure weight, maximum heave motion, and maximum roll response. Ten design variables of the hull constitute the variables space to construct the surrogate model for optimization. The transverse metacentric height, maximum pitch response, and maximum heave acceleration are employed as the constraints.

(1) A surrogate model based on the SVM-GSM was established to shorten the time-consuming and improve the efficiency instead of the direct numerical simulation. A triple verification approach, including the accuracy indicators, learning curves, and fitting curves, ensure the accuracy and sample size stability of the model.

(2) Based on the different emphasis on optimal goals, three categories of optimal solutions were obtained by NSGA-II multi-objective optimization. The results show that each goal achieves some degree of optimization in three categories. The optimization results are verified by frequency and time domain analysis, and main factors screening, indicating a satisfactory optimized effect.

(3) The main influence factors of the optimized goals were screened and evaluated. The column length and deck width, significantly affect the total weight, and are positively correlated with the weight. Both the heave and roll are negatively correlated with hull draft, column height, and deck width. In addition, although the transverse brace diameter has a low influence in most cases, it

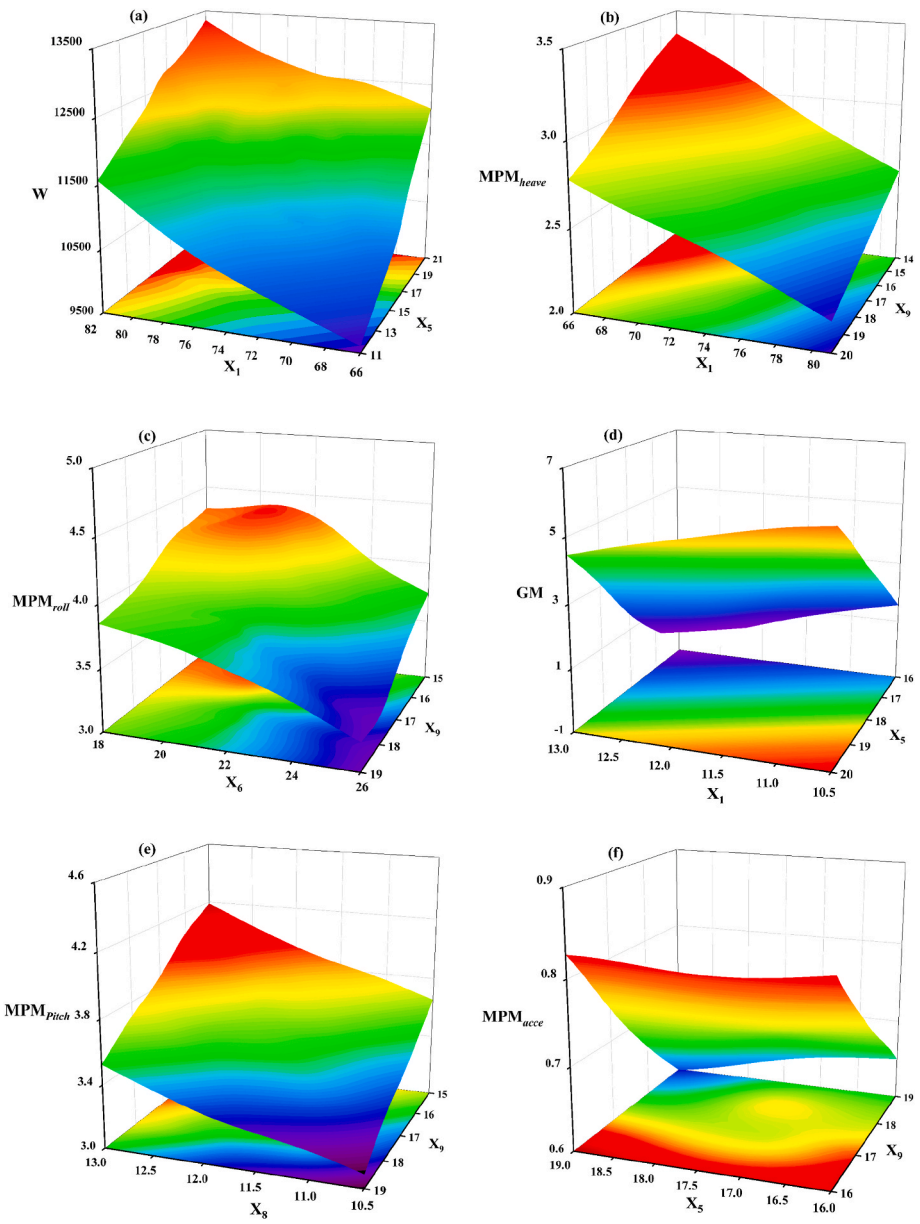


Fig. 25. The relationship of the objectives and constraints with the two main influence factors. (a) hull steel weight; (b) heave; (c) roll; (d) metacentric height; (e) pitch; (f) heave acceleration.

should not be excluded from design variables in the optimization process due to a nonnegligible impact.

CRediT authorship contribution statement

Yixuan Mao: Conceptualization, Methodology, Software, Validation, Writing – original draft. **Tianqi Wang:** Software, Validation, Investigation. **Menglan Duan:** Supervision, Writing – review & editing. **Hongyuan Men:** Software, Formal analysis.

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.

Acknowledgments

This work was supported by National Key Research and Development Program of China (grant number 2016YFC0303701). The authors express sincere thanks to the anonymous reviewers, who are dedicated to providing valuable comments and suggestions, and enhancing the quality of this article. The support is gratefully acknowledged, as are the assistances of Prof. Chen An, Prof. Yu Zhang, Mr. Min Jun Lee from China University of Petroleum, Beijing.

References

Akagi, S., Yokoyama, R., Ito, K., 1984. Optimal design of semisubmersible's form based on systems analysis. J. Mechanisms, Transmissions, and Automation in Design 106 (4), 524–530.

Anderson, M.J., Whitcomb, P.J., 1970. Design of experiments. Concise Encycl. Stat. 2 (Sept), 33–53.

Andri, K., Rahman, N., Shahedan, N.F., 2014. Interrelationship analysis of geopolymer components using pearson correlation technique. Appl. Mech. Mater. 567, 417–421.

Angermann, L., Wang, S., 2007. Convergence of a fitted finite volume method for the penalized Black Scholes equation governing European and Americanoption pricing. Numer. Math. 106, 1–40.
API, R.P.2T., 1997. Recommended Practice for Planning, Design and Constructing Tension Leg Platforms, second ed.
API, R.P.2S.K., 2005. Recommended Practice for Design and Analysis of Stationkeeping Systems for Floating Structures, third ed.
Bonett, D., Wright, T., 2020. Sample size requirements for estimating pearson, kendall and spearman correlations. Psychometrika 65. https://doi.org/10.1007/BF02294183, 2000.
Clauss, G.F., Birk, L., 1996. Hydrodynamic shape optimization of large offshore structures. Appl. Ocean Res. 18 (4), 157–171.
Deb, K., 2011. Multi-objective Optimization Using Evolutionary Algorithms. Wiley Press.
Deb, K., Mohan, M., Mishra, S., 2005. Evaluating the E-domination based multi-objective evolutionary algorithm for a quick computation of pareto-optimal solutions. Evol. Comput. 13 (4), 501–525.
Deutsch, J.L., Deutsch, C.V., 2012. Latin hypercube sampling with multidimensional uniformity. J. Stat. Plann. Inference 142 (3), 763–772, 2012.
Emami, A., Mostafa Gharabaghi, A.R., 2020. Introducing a simple and reliable multi-objective optimization method to estimate hull dimensions of a semi-submersible rig. J. Mar. Eng. 16, 28–40.
Fan, D., Meng, D., Xu, D., 2014. Survey of research process on statistical correlation analysis. Mathematical Model. Appl. 3 https://doi.org/10.3969/j.issn.2095-3070.2014.01.002, 2014.
Gosain, G.D., Sharma, R., Kim, T.W., 2017. An optimization model for preliminary stability and configuration analyses of semi-submersibles. Trans RINA Part A: Int. J. Marit. Eng. 159 (3), 249–270.
Johnson, M.E., Moore, L.M., Ylvisaker, D., 1990. Minimax and maximin distance designs. J. Stat. Plann. Inference 26, 131–148.
Joseph, V.R., Hung, Y., 2008. Orthogonal-maximin Latin hypercube designs. Stat. Sin. 18, 171–186.
Karolczuk, A., Kurek, M., 2022. Fatigue life uncertainty prediction using the Monte Carlo and Latin hypercube sampling techniques under uniaxial and multiaxial cyclic loading. Int. J. Fatig. 160, 106867 https://doi.org/10.1016/j.ijfatigue.2022.106867.
Koziel, S., Pietrenko-Dabrowska, A., 2019. Reliable data-driven modeling of high-frequency structures by means of nested kriging with enhanced design of experiments. Eng. Comput. 36 (7), 2293–2308. https://10.1108/EC-02-2019-0054, 2019.
Koziel, S., Pietrenko-Dabrowska, A., 2020. Design-oriented computationally-efficient feature-based surrogate modelling of multi-band antennas with nested kriging. Int. J. Electr. Commun. AEU 120, 15302. https://doi.org/10.1016/j.aeue.2020.153202.
Krige, D.G., 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. J. Chem. Metall. Min. Eng. Soc. South Africa 52 (6), 119–139.
Leary, S., Bhaskar, A., Keane, A., 2003. Optimal orthogonal-array-based Latin hypercubes. J. Appl. Stat. 30 (5), 585–598.
Lee, J.Y., Koo, B.J., Clauss, G., 2007. Automated design of a tension leg platform with minimized tendon fatigue damage and its verification by a fully coupled analysis. Ship Technol. Res. 54, 11–27.
Li, Z., Gao, X., Lu, D., 2021. Correlation analysis and statistical assessment of early hydration characteristics and compressive strength for multi-composite cement paste. Build. Mater. 310, 125260 https://doi.org/10.1016/j.conbuildmat.2021.125260.
Liu, X., Zhao, W., Wan, D., 2022. Multi-fidelity Co-Kriging surrogate model for ship hull form optimization. Ocean. Eng. 243, 110239.
Musarat, M.A., Alaloul, W.S., Liew, M.S., Maqsoom, A., Qureshi, A.H., 2020. Investigating the impact of inflation on building materials prices in construction industry. J. Build. Eng. JOBE, 101485. https://doi.org/10.1016/j.jobe.2020.101485. 30.
Nielsen, B.F., Skavhaug, O., Tveito, A., 2008. Penalty methods for the numerical solution of American multi-asset option problem. J. Comput. Appl. Math. 222 (1), 3–16.
Park, Y., Jang, B.S., Kim, J.D., 2015. Hull-form optimization of semi-submersible fpu considering seakeeping capability and structural weight. Ocean. Eng. 104, 714–724.
Pawling, R., Andrews, A., Percival, V., 2017. A study into the validity of the ship design spiral in early stage ship design. J. Ship Prod. Des. 33, 81–100. https://doi.org/10.5957/jspd.2017.33.2.81, 02.
Pietrenko-Dabrowska, A., Koziel, S., 2017. Expedited yield optimization of narrow and multi-band Antennas using performance driven surrogates. IEEE Access. https://doi.org/10.1109/ACCESS.2020.3013985.
Pietrenko-Dabrowska, A., Koziel, S., 2019a. Computationally-efficient design optimization of antennas by accelerated gradient search with sensitivity and design change monitoring. IET Microw., Antennas Propag. 14 (2), 165–170. https://doi.org/10.1049/iet-map.2019.0358, 2020.
Pietrenko-Dabrowska, A., Koziel, S., 2019b. Accelerated multiobjective design of miniaturized microwave components by means of nested kriging surrogates. Int. J. RF Microw. Computer-Aided Eng., e22124 https://doi.org/10.1002/mmce.22124.
Qiu, W., Song, X., Shi, K., Zhang, X., Yuan, Z., You, Y., 2019. Multi-objective optimization of semi-submersible platforms using particle swam optimization algorithm based on surrogate model. Ocean. Eng. 178, 388–409.
Razavi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. WaterResour. Res. 48, 32.
Roshanian, J., Ebrahimi, M., 2012. Latin hypercube sampling applied to reliability-based multidisciplinary design optimization of a launch vehicle. Aero. Sci. Technol. 28, 297–304. https://doi.org/10.1016/j.ast.2012.11.010.
Sivaraj, R., Ravichandran, D.T., 2011. A review of selection methods in genetic algorithm. Int. J. Eng. Sci. Technol. 3 (5).
Sugita, T., Suzuki, H., 2016. A study on tlp hull sizing by utilizing optimization algorithm. J. Mar. Sci. Technol. 21 (4), 611–623.
Talukdar, S., Singha, P., Shahfahad, Mahato S., Praveen, B., Rahman, A., 2020. Dynamics of ecosystem services (ESs) in response to land use land cover (LU/LC) changes in the lower Gangetic plain of India. Ecol. Indicat. 112, 14.
Tian, X., Sun, X., Liu, G., Xie, Y., Chen, Y., Wang, H., 2021. Multi-objective optimization of the hull form for the semi-submersible medical platform. Ocean. Eng. 230, 109038.
Vapnik, V.N., Golowich, S.E., Smola, A., 1997. Support vector method for function approximation, regression estimation, and signal processing. Adv. Neural Inf. Process. Syst. 281–287, 1997.
Venzon, R.Z., Tancredi, T.P., deAndrade, B.L.R., 2013. Hull optimization of semisubmersible with seakeeping criteria evaluated with neural network response surface. In: Proceedings of the 12th International Symposiumon Practical Design of Ships and Other Floating Structures. Changwon, Korea, pp. 944–952. PRADS2013.
Wang, Q., Nakashima, T., Lai, C., Mutsuda, H., Kanehira, T., Konishi, Y., Okuizumi, H., 2020. Modified algorithms for fast construction of optimal Latin-hypercube design. IEEE Access 8, 191644–191658.
Wu, X.D., Kumar, V., Quinlan, J.R., Ghosh, J., Yang, Q., Motoda, H., et al., 2008. Top 10 algorithms in data mining. Knowl. Inf. Syst. 14, 1–37.
Wu, X., Peng, X., Chen, W., Zhang, W., 2019. A developed surrogate-based optimization framework combining HDMR-based modeling technique and TLBO algorithm for high-dimensional engineering problems. Struct. Multidiscip. Optim. 60, 663–680. https://doi.org/10.1007/s00158-019-02228-4.
Zhang, X., Song, X., Qiu, W., Yuan, Z., You, Y., Deng, N., 2018. Multi-objective optimization of tension leg platform using evolutionary algorithm based on surrogate model. Ocean. Eng. 148, 612–631.
Zhang, Z., Chen, H.C., Cheng, Q.S., 2021. Surrogate-assisted quasi-Newton enhanced global optimization of antennas based on a heuristic hypersphere sampling. IEEE Trans. Antenn. Propag. 69 (5), 2993–2998. https://doi.org/10.1109/tap.2020.3031474.
Zhou, C., Chase, J.G., Rodgers, G.W., 2021. Support vector machines for automated modelling of nonlinear structures using health monitoring results. Mech. Syst. Signal Process. 149, 15.