Improved Control of Propeller Ventilation Based on POA-XGBoost and Ship Dynamics/Control Model

Shengping Ma [a*], Yu Ding [a], Congbiao Sui [a]

[a] Harbin Engineering University, Harbin, China
*mashengping@hrbeu.edu.cn

Abstract

Under adverse sea conditions, propeller ventilation caused by in-and-out water can decrease the reliability of the ship power grid and the lifespan of the propulsion shaft system. Predicting the development of propeller ventilation severity while identifying it can contribute to improving propeller ventilation control. In this study, the eXtreme Gradient Boosting (XGBoost) algorithm combined with a ship dynamics/control model is proposed as a propeller ventilation identification and prediction method. Meanwhile, the Pelican optimization algorithm (POA), particle swarm optimization (PSO), and genetic algorithm (GA) are applied to determine the optimal hyperparameters of the XGBoost algorithm. The results indicate that the method can effectively identify the current propeller ventilation state and predict whether a full ventilation state will occur after experiencing a partial propeller ventilation state. The comparison results indicate that the POA has a better optimization effect on the XGBoost algorithm for propeller ventilation identification and prediction. The method proposed in this study provides crucial technical support for the effective switching of propulsion control strategies for ship electric propulsion systems under adverse sea conditions.

Keywords: Adverse sea condition, Propeller ventilation, Propulsion control switching strategy, POA-XGBoost model, Propeller ventilation identification and prediction method.

1. INTRODUCTION

Under adverse sea conditions, the dynamic behavior of the propeller in waves causes fluctuations in the ship-integrated power system. In particular, propeller ventilation caused by frequent in-and-out-of-water movements leads to fluctuations as high as 80–100% of the rated load [1]. Propeller ventilation can not only accelerate the wear and tear of the ship's propulsion equipment but also lead to significant fluctuations in the voltage and frequency of the shipboard power network caused by an imbalance between the "source" and "load" power [2]. Severe fluctuations in the shipboard power network can adversely affect the operational performance and efficiency of electrical equipment and even lead to accidents where the entire ship loses power. To solve this issue, Smogeli et al. [3] proposed an anti-spin control strategy that can switch between adverse sea conditions and normal conditions for propulsion control strategies. This switch was based on the propeller ventilation state identified by the ventilation identification module. Without considering the predicted development of ventilation severity, ineffective propulsion control strategy switching may occur. Therefore, predicting the severity of ventilation while identifying it is important for ensuring the safety and stability of ship operations.

Machine Learning (ML) algorithms have been widely applied in the identification of propeller ventilation in recent years. Califano et al. [4] and Luca Savio et al. [5] adopted a support vector machine and a single Kalman filter to identify whether propeller ventilation occurred, both of which have achieved high identification accuracy. Gao et al. [6] and Zhang et al. [7] proposed adopting an evidence reasoning rule based on the Adaboost. ML and maximum likelihood evidential reasoning rule to identify the partial ventilation state. This is a transitional state between the non-ventilated and fully ventilated states. However, there are some limitations to the existing methods. They can only identify the current propeller ventilation state as the moment for switching propulsion control strategies without considering whether a full ventilation state will occur after experiencing the partial propeller ventilation state. If the propeller does not fully transition from a partial ventilation state to a full ventilation state, it quickly returns to the non-ventilation state. This can result in the frequent and ineffective switching of propulsion control strategies under adverse sea conditions. The propeller may return to the non-

DOI: https://doi.org/10.59490/moses.2023.662

ventilation state before completing the switch to the propulsion control strategy, or it immediately needs to switch back to the normal condition control strategy after completing the switch to the adverse sea condition control strategy. The method proposed in this study can identify the current propeller ventilation state and predict whether a full ventilation state will occur after experiencing the partial propeller ventilation state. The propulsion control strategy is switched only when the full ventilation state is predicted, resulting in a reduction in ineffective switching of propulsion control strategies.

To address the aforementioned issues, this study proposes predicting the severity of ventilation development and identifying it to enhance the control of propeller ventilation. If the propeller reached the full ventilation state, the propulsion control strategy was switched. If it does not, it does not switch. Propeller ventilation identification and prediction are classification problems. Currently, the main algorithms used for classification include KNN, neural networks, SVM, and decision trees. The XGBoost algorithm proposed by Chen [8] is based on the CART regression tree and belongs to the boosting ensemble learning method. By incorporating L2 regularization terms and second-order derivatives into the objective function of the gradient augmented decision tree, the ability to generalize and avoid overfitting of XGBoost is improved. The XGBoost algorithm has an excellent ability to handle large-scale data. The comparison results of the KNN, BP neural network, SVM, and XGBoost on the problem of propeller ventilation identification and prediction by the author indicate that XGBoost achieves better performance. However, the XGBoost algorithm without parameter optimization has a low degree of fit with the existing dataset, which leads to its poor generalization performance and adaptability [9]. Therefore, this study adopts POA, PSO, and GA to optimize the hyperparameters of XGBoost.

This paper is organized as follows. In Section 2, the ship dynamics/control model is developed to simulate the dynamic response of a ship propulsion system under adverse sea conditions of different wind and wave conditions. In Section 3, the current and speed signals of the propulsion motor are collected, and their characteristics are extracted for training the POA-XGBoost, PSO-XGBoost, and GA-XGBoost algorithms. These algorithms were designed to accurately identify and predict the propeller ventilation. Section 4 focuses on the identification and prediction performance of the model. Finally, the conclusions and recommendations for future research are presented in Section 5.

2. SHIP DYNAMICS/CONTROL MODEL

To simulate the dynamic characteristics of ship electric propulsion systems under adverse sea conditions, an integrated simulation model including a mechanical module, electrical module, and hydrodynamic module is introduced in this section, as shown in Fig. 1. The integrated simulation model is mainly categorized into three main subsystems: electrical propulsion system, the 3 DOF (degree of freedom) motion system, and the sea state interference system. The electrical propulsion system includes a propulsion motor, propulsion controller, and propeller models. The 3 DOF motion system can calculate the ship heading angle and speed according to the interaction between the thrust generated by the propeller, pod angle, and external interference of wind and waves. The sea state interference system, which includes the first-order wave force model, second-order wave force model, wind force model, and in-and-out-of-water model, was used to simulate the external environmental interference of ships under adverse sea conditions.

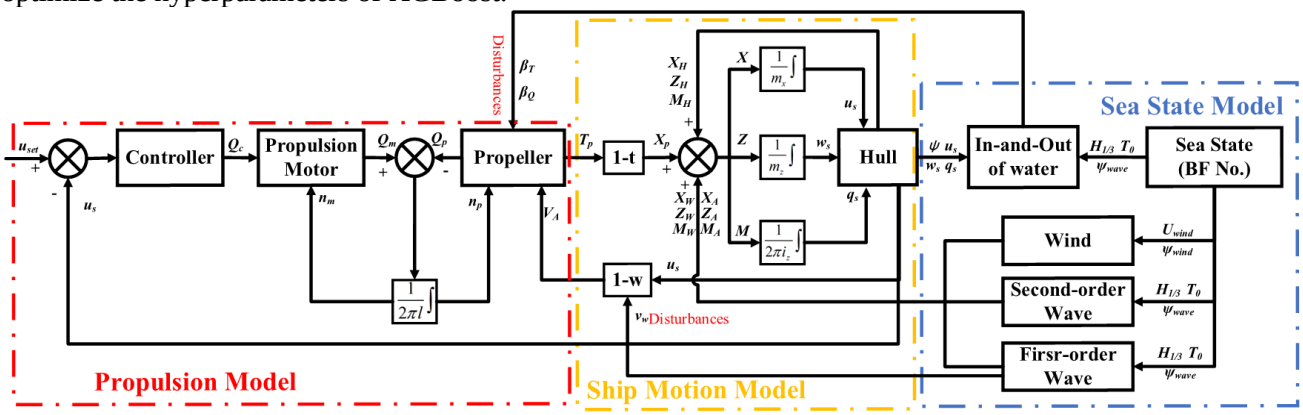


Figure 1: block diagram of the integrated simulation model structure

2.1 Electric propulsion system

2.1.1 The propulsion motor model

Because the electrical time constant is much smaller than the mechanical time constant in an electric propulsion system, the first-order equation in. Equation (1) is used to simplify the motor dynamic equation [10].

$$\frac{dQ_m}{dt} = \frac{1}{T_m} \cdot (Q_c - Q_m) \tag{1}$$

where Q_c is the target torque provided by the controller, Q_m is the output torque of the propulsion motor, T_m is the time constant of the motor.

The torque balance equation between the propeller and the propulsion motor can be simplified using (2).

$$I_s \cdot \frac{d\omega}{dt} = Q_m - Q_p - k_\omega \cdot \omega \tag{2}$$

where I_s is the moment of inertia, k_ω is the friction coefficient of the transmission shaft, ω is the propeller angular velocity, Q_p is the propeller torque.

2.1.2 Propeller model

When determining the geometric parameters of a fixed-pitch propeller, the torque and thrust coefficients of the propeller are dependent only on the advance coefficient. The torque and thrust of the propeller are expressed by (3).

$$\begin{cases} Q_p = \mathrm{sgn}(n) \cdot \beta_Q \cdot K_Q \cdot \rho_w \cdot n^2 \cdot D^5 \\ T_p = \mathrm{sgn}(n) \cdot \beta_T \cdot K_T \cdot \rho_w \cdot n^2 \cdot D^4 \end{cases} \tag{3}$$

where T_p is the propeller thrust; K_Q and K_T represent the propeller torque coefficient and the propeller thrust coefficient, respectively; β_Q and β_T represent the torque loss coefficient and the loss thrust coefficient, respectively, which are used to characterize the variation of propeller torque and thrust with propeller submergence; ρ_w is the water density; n and D represent the revolution speed and the diameter of the propeller, respectively.

The goal of ship propulsion control is to maintain the stability of a ship's velocity. A motor speed control strategy was adopted in this study.

2.2 3 DOF ship motion system

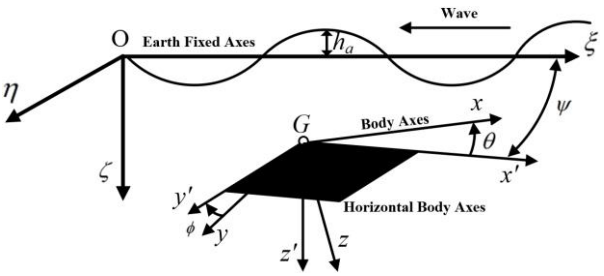


Figure 2: ship motion coordinate system [11]

As shown in Fig. 2, a ship motion coordinate system [11], which includes the earth-fixed coordinate system *O-ξηζ*, the body-fixed coordinate system *G-xyz* and the motion coordinate system *G-x'y'z'*,' is developed to explore the effects of wind and waves on ship forces in adverse sea conditions.

In this study, it was assumed that the forces generated by the wind and waves on the ship were in the same direction. The ship sails against wind and waves under adverse sea conditions. Therefore, only surge, heave, and pitch motions were considered in this study, whereas sway, roll, and yaw motions were ignored.

According to the separation concept of MMG school [12], surge motion can be expressed by (4).

$$(m + m_{11}) \cdot \left(\frac{du}{dt} + q \cdot w \right) = X_H + X_P + X_A + X_W \tag{4}$$

where m and m_{11} are the mass of the ship and the added mass of the entrained water in the longitudinal direction, respectively, u is the ship longitudinal velocity, w is the ship vertical velocity, q is the angular speed of the pitch motion. The subscripts *H*, *P*, *A* and *W* represent the external forces and moments contributed by the hull, pod thruster, wind, and waves, respectively.

According to [13], a simplified response equation derived from a semi-analytical approach was proposed by Jørgen Juncher Jense et al. to calculate the pitch and heave motion of monohull ships induced by waves, which can be expressed by (5).

$$\begin{cases} 2\dfrac{k \cdot d}{\omega^2} \cdot \dfrac{d^2 w}{dt^2} + \dfrac{A^2}{k \cdot B \cdot \alpha^3 \cdot \omega} \cdot \dfrac{dw}{dt} + w = Z_p + Z_A + Z_W \\ 2\dfrac{k \cdot d}{\omega^2} \cdot \dfrac{d^2 \theta}{dt^2} + \dfrac{A^2}{k \cdot B \cdot \alpha^3 \cdot \omega} \cdot \dfrac{d\theta}{dt} + \theta = M_p + M_A + M_W \end{cases} \tag{5}$$

where θ is ship trim angle, B is breadth, d is draught, L is length, $k = 2\pi/\lambda$ is the wave number, λ is the wave length, V_s is the ship velocity, χ_w is the relative wave direction that is taken as π when moving against the waves, Z is the vertical force, M is the longitudinal moment, A is the sectional hydrodynamic damping, which can be modelled by the dimensionless ratio between the incoming and diffracted wave amplitudes [14].

2.3 Sea state interference system

2.3.1 Wind force model

As recorded in [15], the force and moments generated by the wind usually have a significant impact on the surge, sway, and yaw motions of ships on the horizontal plane, with little impact on the heave, pitch, and roll motions. Therefore, only the longitudinal force generated by the wind was

considered in this study, which can be calculated using the steady velocity of the ship.

$$X_A = 1/2 \cdot \rho_A \cdot U_A^2 \cdot C_{XA}(\chi_A) \cdot A_X \quad (6)$$

where ρ_A is the air density, U_A is the relative wind speed, C_{XA} is the longitudinal wind coefficient, χ_A is the relative wind direction, A_X is the projected frontal area of the ship above water.

2.3.2 First-order waves force model

The Froude-Krenov hypothesis states that the pressure distribution in regular waves is not influenced by the presence of ships [13]. The first-order wave force generated by regular waves can be expressed by (7) by considering the ship to be a hexahedron.

$$\begin{cases} X_W^1 = \dfrac{-4\rho_A \cdot g \cdot h_a}{k^2 \cdot \sin^2 \chi_W} \cdot \left(1 - e^{-k \cdot d}\right) \cdot \sin\left(\dfrac{k \cdot L \cdot \cos \chi_W}{2}\right) \\ \cdot \sin\left(\dfrac{k \cdot B \cdot \sin \chi_W}{2}\right) \cdot \sin(\omega_e \cdot t) \\ Z_W^1 = h_a \cdot \kappa \cdot \sqrt{(1 - k \cdot d)^2 + \left(\dfrac{A^2}{k \cdot B \cdot \alpha^2}\right)^2} \\ \cdot \dfrac{2}{k_e \cdot \cos \chi_W} \cdot \sin\left(\dfrac{k_e \cdot L}{2}\right) \cdot \cos(\omega_e \cdot t) \\ M_W^1 = h_a \cdot \kappa \cdot \sqrt{(1 - k \cdot d)^2 + \left(\dfrac{A^2}{k \cdot B \cdot \alpha^2}\right)^2} \\ \cdot \left[\sin\left(\dfrac{k_e \cdot L}{2}\right) - \dfrac{k_e \cdot L}{2} \cdot \cos\left(\dfrac{k_e \cdot L}{2}\right)\right] \cdot \sin(\omega_e \cdot t) \end{cases} \quad (7)$$

where χ_w is the relative wave direction, h_a is the amplitude of the wave, k is the wave number, λ is the wavelength, ω_e is the frequency of encounter, k_e is the effective wave number, κ is the Smith correction coefficient.

2.3.3 Second-order waves force model

According to [15], the second-order wave force has a significant impact on ship surge motion, with little impact on heave and pitch motions. The longitudinal force generated by second-order waves can be expressed as (8).

$$X_W^2 = \rho_w \cdot g \cdot H_{1/3}^2 \cdot \left(B^2/L\right) \cdot C_{XW}(U, T_v, \chi_W) \quad (8)$$

where g is the gravity acceleration, $H_{1/3}$ is the significant wave, approximately 1.6 times the average wave height, T_v is the averaged wave period, C_{XW} is the longitudinal second-order force coefficient.

2.3.4 In-and-out-of-water model

Under adverse sea conditions, the loss of thrust and torque in the propeller caused by the in- and out-of-water effects can be expressed by the thrust loss factor and torque loss factor, respectively. Equation (9) was used to calculate the thrust loss factor.

$$\beta_T = \begin{cases} 0, & \dfrac{h}{R} \le -0.48 \\ 1 - 0.675\left(1 - 0.769\dfrac{h}{R}\right)^{1.258}, & -0.48 < \dfrac{h}{R} < 1.3 \\ 1, & 1.3 < \dfrac{h}{R} \end{cases} \quad (9)$$

Where, h is the propeller submergence; R represents the radius of propeller.

The torque loss factor can be expressed as the thrust loss factor, as shown in (10).

$$\beta_Q = (\beta_T)^m \quad (10)$$

Where, m is a coefficient with the value of 0.85 in this paper.

2.3.5 Propeller wake fluctuation model

The inflow velocity of the propeller continuously fluctuates under the combined influence of ship speed, maneuvering, and waves [16]. This study comprehensively considers the influence of the ship maneuvering motion on the average inflow velocity of the propeller and the influence of regular waves on the oscillation inflow velocity of the propeller. The inflow velocity of the propeller is expressed as (11).

$$\begin{aligned} V_a = {} & u \cdot (1 - w_p) + \alpha \cdot \omega \cdot h_a \cdot \exp(-k \cdot z_p) \\ & \cdot \cos \chi_W \cdot \cos(\omega_e \cdot t - k \cdot x_p \cdot \cos \chi_W) \end{aligned} \quad (11)$$

where x_p and z_p represent the longitudinal and vertical positions of the propeller, respectively, in the ship-fixed coordinate system.

2.4 Model validation

Because the ship dynamic/control model is used to simulate the dynamic response of the ship propulsion system, the simulation accuracy must be validated.

By comparing the experimental results obtained from the propeller model test in the deep-water towing tank of the benchmark cruise ship, the accuracy of the ship surge motion model was validated, as shown in Fig. 3.

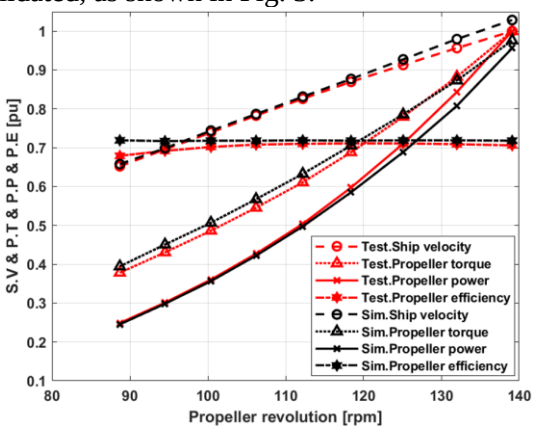
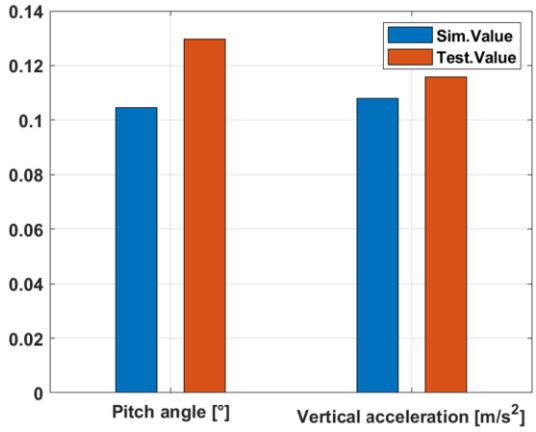
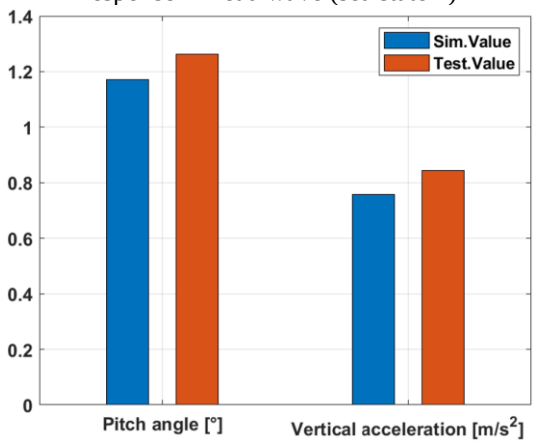


Figure 3: Validation of ship surge motion model

According to the data from the ship seakeeping test recorded in [17], the ship pitch motion, heave motion, and first-order wave models were validated, as shown in Fig. 4.



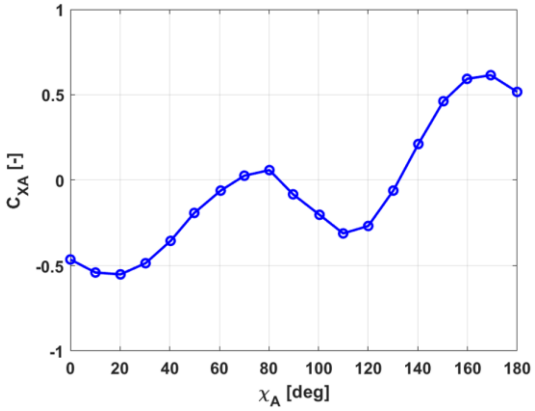
(a) Double significant amplitude of cruise motion response in head wave (sea state 4)



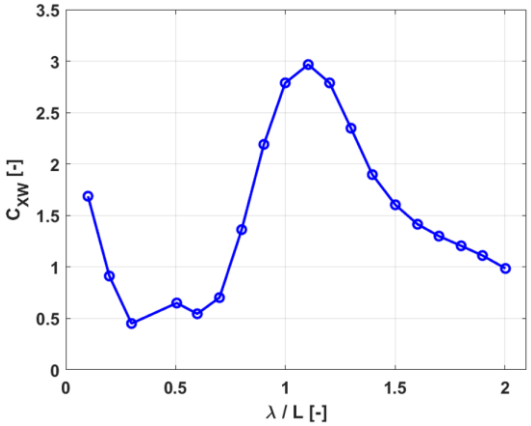
(b) Double significant amplitude of cruise motion response in the head wave (sea state 6)

Figure 4: Validation of pitch motion, heave motion and first-order waves models

To simulate the longitudinal forces caused by the wind and second-order wave forces, the longitudinal and second-order wave force coefficients are shown in Figure 5. Note that the coefficients recorded in [15] were directly used in this study. The uncertainties caused by directly using these coefficients for the cruise ship are assumed to be acceptable for the following reasons: first, both ships are cruise ships with similar hull shapes; second, the longitudinal wind and second-order wave force coefficients are nondimensional.



(a) Longitudinal wind force coefficient



(b) Longitudinal second-order waves force coefficient

Figure 5: Wind and second-order waves force coefficients

3. THE PROPELLER VENTILATION IDENTIFICATION AND PREDICTION METHOD BASED ON XGBOOST

3.1 Design of the ventilation identification and prediction model

There are challenges in directly measuring the changes in propeller performance parameters under adverse sea conditions. Because the propulsion motor is directly connected to the propeller, its state parameters of the propulsion motor can be used for propeller ventilation identification and prediction.

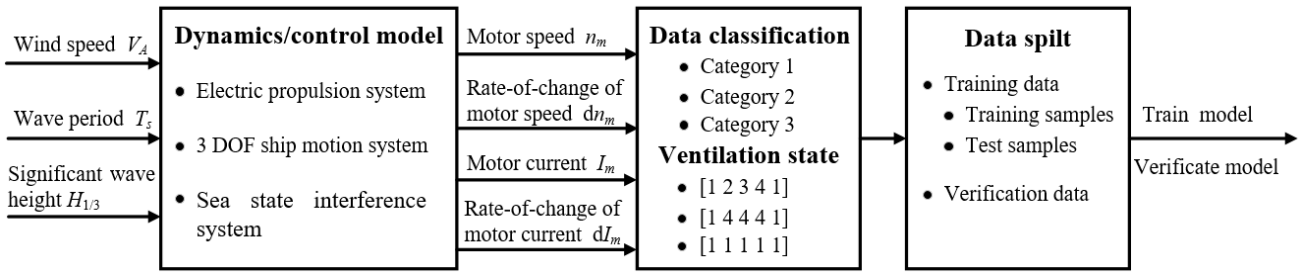


Figure 6: Block diagram of the identification and prediction method

Wang et al. [18] used propulsion torque data to train a model for identifying propeller ventilation, while Zhang [7] incorporated both propulsion torque and speed data. The parameters of the propulsion motor can be divided into electrical and mechanical parameters. Comprehensively considering both the electrical and mechanical parameters may result in better performance. Gao [6] utilized the propulsion torque data and the root mean square value of the electric current. However, to a certain extent, the electric current of the motor varies with torque. In other words, they exhibited the same change trend. Therefore, the electrical and mechanical parameters of the propulsion motor were selected in this study, specifically, the root mean square of the current and speed signals.

The design process of the propeller ventilation identification and prediction model is illustrated in Fig. 6. The dynamics/control model is introduced in Section 2. First, in the simulation model, the ship velocity was set to 6 kn. The operational parameters of the ship propulsion motor under 32 different wind and wave conditions were obtained by adjusting the wind and wave parameters, such as the wind speed, wave period, and significant wave height, as presented in Table 1. The wind speed can be set to 19.0 m/s and 21.0 m/s. The wave period can be set to 9.0 s, 11.0 s, 13.0 s and 15.0 s. The significant wave height can be set to 4.5 m, 5.0 s, 5.5 m and 6.0 m. Then, 32 different conditions can be obtained by cross combining the values of above parameters. The propulsion motor parameters include the motor speed, rate of change of the motor speed, root mean square of the motor current, and rate of change of the root mean square of the motor current. Second, according to the maximum ventilation level that the propeller can develop, 32 sets of simulation results were divided into three categories. Different propeller ventilation states are set for the three categories, which are introduced in Section 3.2. Then, the 32 sets of simulation results were divided into training data and verification data, among which the training data were further divided into training and test samples. Training data were used to train and test the learning and prediction performance of the intelligent algorithm. Verification data were used to demonstrate the generalization ability of the trained intelligent algorithm. Finally, POA-XGBoost, PSO-XGBoost, and GA-XGBoost are developed to achieve a function that can identify the current ventilation state and predict the development of ventilation severity. POA, PSO, and GA were used to select the optimal hyperparameters of the XGBoost algorithm.

Table 1. Input parameters and applications of the 32 sets of data.

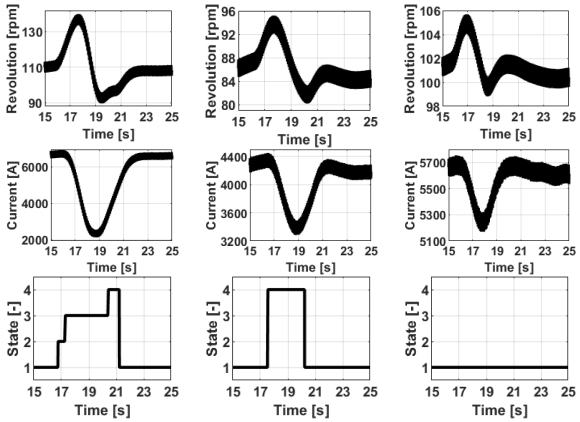
NO.	Wind speed (m/s)	Wave period (s)	Significant wave height (m)	Application
1	19.0	9.0	4.5	Training
2			5.0	Training
3			5.5	Training
4			6.0	Training
5		11.0	4.5	Test
6			5.0	Training
7			5.5	Training
8			6.0	Training
9		13.0	4.5	Training
10			5.0	Training
11			5.5	Training
12			6.0	Training
13		15.0	4.5	Training
14			5.0	Training
15			5.5	Training
16			6.0	Training
17	21.0	9.0	4.5	Training
18			5.0	Training
19			5.5	Training
20			6.0	Training
21		11.0	4.5	Training
22			5.0	Training
23			5.5	Training
24			6.0	Training
25		13.0	4.5	Training
26			5.0	Test
27			5.5	Training
28			6.0	Training
29		15.0	4.5	Training
30			5.0	Training
31			5.5	Test
32			6.0	Training

3.2 Data classification and ventilation state

The method of setting different propeller ventilation states for different categories proposed by the author in [19] will be used again in this study.

According to the maximum ventilation level that the propeller can develop, 32 sets of simulation results were divided into three categories. Category 1 indicates the occurrence of full ventilation. Category 2 indicates that the maximum ventilation level within a wave period corresponds to a partial ventilation. Category 3 indicated that ventilation did not occur.

Different propeller ventilation states were set for different categories. State 1 indicates that the propeller is in a nonventilated state. State 2 indicates that the propeller is currently in a partial ventilation state, and the full ventilation state is predicted to occur immediately. State 3 indicates that the propeller is in a fully ventilated state. State 4 indicates that the propeller is currently in the partial ventilation state; however, the full ventilation state is predicted not to occur immediately.



(a) Category 1 (b) Category 2 (c) Category 3

Figure 7: Different propeller ventilation states setting based on different categories [19]

3.3 Hyperparameters optimization of XGBoost

3.3.1 Pelican optimization algorithm

The POA developed by Pavel Trojovský [20] in 2020 is a random heuristic algorithm. The POA simulates the behavior and strategies of pelicans during hunting and is divided into two stages: moving towards prey and winging on the water surface. In the stage of moving towards prey, the pelicans determine the position of the prey and then move towards the determined area. It is worth mentioning that the position of the prey is randomly generated in the search space, leading to an increase in the exploration ability of the POA in the exact search of the problem-solving space. In the stage of winging on the water surface, after reaching the surface of the water, the pelicans spread their wings to move the fish upward and then collected prey in their throat pouch. This behavior can lead to more prey being caught in the attacked space. A POA can increase its local search capability by simulating its behavior. POA has better global search ability and best local area recognition ability in engineering problems, which has been proven in several reports [21], [22].

3.3.2 Particle swarm optimization

The PSO, developed by Eberhart and Kennedy [23] in 1995, is an evolutionary computation. PSO is a simplified model established using swarm intelligence, which was initially inspired by the regularity of bird clustering activities. The PSO simulates the sharing of information among individuals in an animal population. This process leads to an evolutionary progression from disorder to order in the problem-solving space, ultimately resulting in identification of the optimal solution.

3.3.3 Genetic algorithm

The GA, proposed by Holland in 1973 [24], is a modern intelligent algorithm that draws inspiration from Darwin's theory of evolution and Mendel's theory of genetics. The aim is to simulate the survival of the fittest and natural genetic mechanisms observed in the biological world. The GA transforms practical problems into evolutionary problems by starting with an initial population and generating new populations through repeated genetic operations of selection, crossover, and mutation until the termination conditions are met. Because of its advantages of good robustness, strong global optimization ability, and no limitations on derivatives and function continuity, GA is often adopted in the field of objective optimization.

3.3.4 XGBoost algorithm

The prediction model of XGBoost can be expressed by (12).

$$\hat{y}_i = \sum_{t=1}^{K} f_t(x_i),\ f_t \in F \tag{12}$$

where x_i represents the i-th input sample, $\hat{y}_i$ is the predicted value of the i-th sample, f_t represents the t-th decision tree, F is the set of all regression trees.

The objective function of the XGBoost algorithm consists of a loss function and regularization, which can be expressed as (13).

$$obj = \sum_{i=1}^{m} l(y_i, \hat{y}_i) + \sum_{j=1}^{K} \Omega(f_i),\ f_i \in F \tag{13}$$

where y_i represents the true value of the i-th sample, $\sum_{i=1}^{m} l(y_i, \hat{y}_i)$ is the loss function, and $\sum_{j=1}^{K} \Omega(f_i)$ is the regularization function, which can be specifically represented by (14).

$$\Omega(f_i) = \gamma \cdot T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \tag{14}$$

where T is the number of leaf nodes in the decision tree, ω represents the leaf node output score of each decision tree, λ is the coefficient of the leaf node, γ is the punish regularization term for the leaf weights.

In the XGBoost algorithm, the error generated by the model combined with the previous (t-1) trees is used as a reference to build the t-th tree, leading to a decrease in the value of the loss function. Therefore, the objective function can be rewritten as (15):

$$obj^t \approx \sum_{i=1}^{m} \left(g_i \cdot f_t(x_i) + \frac{1}{2} h_i \cdot f_t^2(x_i) \right) + \gamma \cdot T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \tag{15}$$

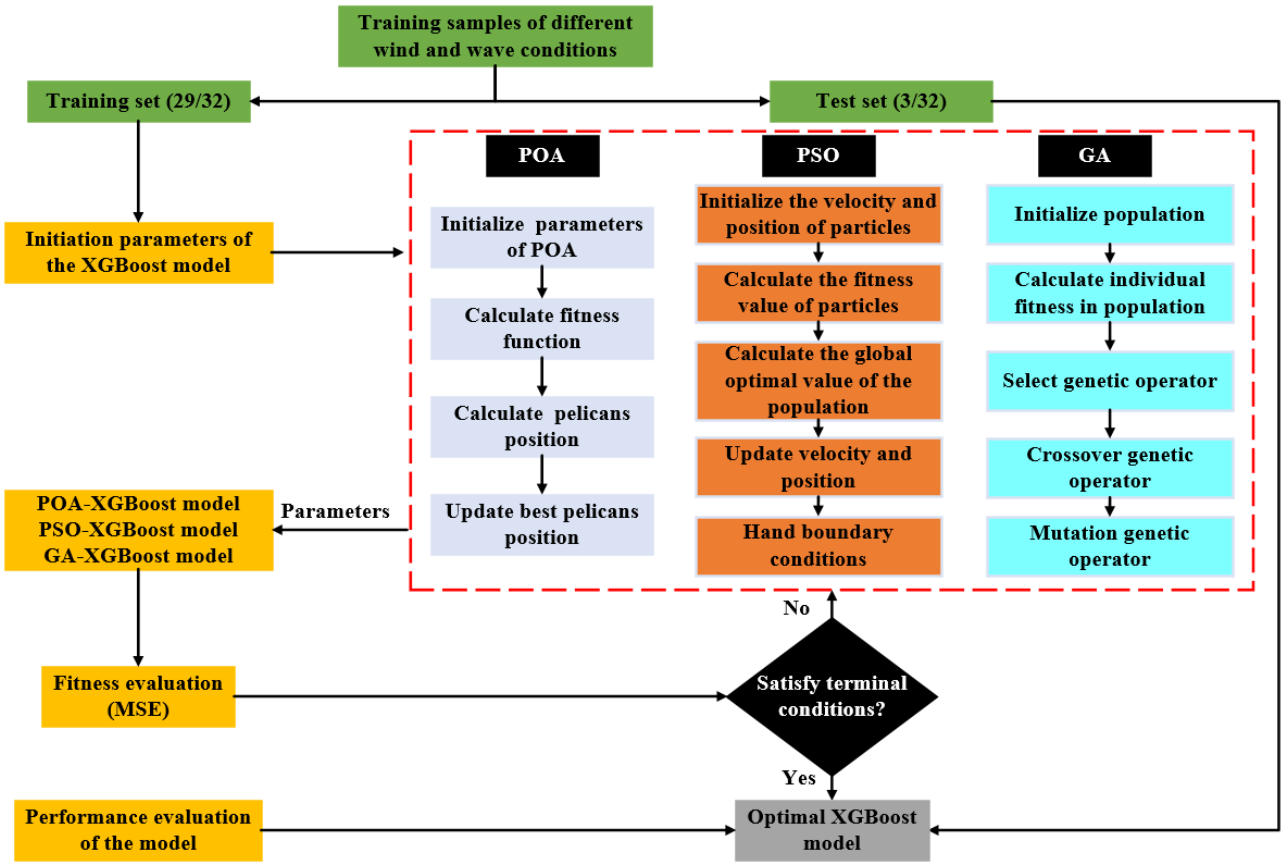


Figure 8: Propeller ventilation identification and prediction process

where $g_i = \frac{\partial l\left(y_i, \hat{y}_i^{(t-1)}\right)}{\partial \hat{y}_i^{(t-1)}}$ and $h_i = \frac{\partial^2 l\left(y_i, \hat{y}_i^{(t-1)}\right)}{\partial^2 \hat{y}_i^{(t-1)}}$ represent the first- and second-order derivatives of the objective function, respectively.

When the tree structure is determined, the optimal weight is obtained by setting its first-order derivative to 0.

$$\omega_j^* = -\frac{G_j}{H_j + \lambda} \tag{16}$$

The optimized objective function can be obtained by substituting the optimal weight expressed in Equation (16) into the objective function.

$$obj^t = -\frac{1}{2}\sum_{j=1}^{T}\left(\frac{G_j^2}{H_j + \lambda}\right) + \lambda \cdot T \tag{17}$$

Where, $G_j = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$ and $I_i = \left\{i \middle| q(x_i) = j\right\}$ are the sample sets that falls into leaf node j.

3.3.5 POA- XGBoost, PSO- XGBoost, and GA- XGBoost

The 32 sets of data were divided into training and test datasets. Meanwhile, the 32 sets of data were divided into three categories. Therefore, three sets of data, including one set from each category, were selected as test data, whereas the remaining 29 sets of data were selected as training data.

Table 2. Hyperparameters to be optimized.

Hyperparameter	Meaning
max_depth	Maximum depth of trees
min_child_weight	Minimum sum of the instance weights contained in child nodes
n_estimators	Number of boosted trees
alpha	Regular term of weight L1
lambda	Regular term of weight L2
gamma	Minimum loss reduction required to make a further partition
subsample	Sampling rate of training samples
colsample_bytree	Column sampling rate of features when building each tree

In this study, POA, PSO, and GA were selected to adjust the hyperparameters of the XGBoost algorithm. The baseline algorithm for propeller ventilation identification and prediction was XGBoost, whereas the POA, PSO, and GA were used to search for the optimal hyperparameters of the Xgboost algorithm. The hyperparameters that need to be optimized are listed in Table 2, which include max_depth, min_child_weight, n_estimators, alpha, lambda, gamma, subsample, and colsample_bytree.

The modeling process of the propeller ventilation identification and prediction model

based on POA-Xgboost, PSO-Xgboost, and GA-Xgboost is shown in Fig. 8. POA, PSO, and GA are used to optimize the XGBoost algorithm to find the hyperparameters that can maximize the classification performance of the XGBoost algorithm. The classification error of the XGBoost algorithm was used as the fitness function of the POA, PSO, and GA. The optimal hyperparameters of the Xgboost algorithm can be obtained by comparing the fitness values to those obtained by continuous iteration and update.

4. RESULT AND ANALYSIS

The training data consisted of 29 sets of simulation results comprising 20,967 sample points. Among these, 14,667 sample points were randomly selected to form the training samples, and the remaining 6,300 were selected to form the test samples.

The MAPE, MSE, and MAE values of the XGBoost, POA-XGBoost, PSO-XGBoost, and GA-XGBoost models are listed in Table 3. It can be seen that the POA, PSO, and GA can improve the classification accuracy of Xgboost, whereas the POA performs better for propeller ventilation identification and prediction problems.

Table 3. MAPE, MSE and MAE values of the different Xgboost model.

Category	Algorithm	MAPE	MSE	MAE
Category 1	Xgboost	0.045	0.332	0.156
	POA-Xgboost	0.028	0.122	0.093
	PSO-Xgboost	0.037	0.234	0.117
	GA-Xgboost	0.036	0.298	0.122
Category 2	Xgboost	0.134	0.684	0.228
	POA-Xgboost	0.144	0.594	0.198
	PSO-Xgboost	0.119	0.612	0.204
	GA-Xgboost	0.177	0.720	0.240

To further demonstrate the generalization ability of the POA-Xgboost model for propeller ventilation identification and prediction, the remaining three sets of verification data were used as the inputs for the trained model. Meanwhile, the GA and PSO are used to optimize the hyperparameters of XGBoost through the same process as POA-XGBoost. The identification accuracy of the POA-XGBoost model was compared with the results of the XGBoost, GA-XGBoost, and PSO-XGBoost models, as shown in Fig.7 and Fig. 8. Note that owing to the extremely obvious features of category 3, all models in this study can be easily identified and have a high identification accuracy. The results for Category 3 are not displayed here.

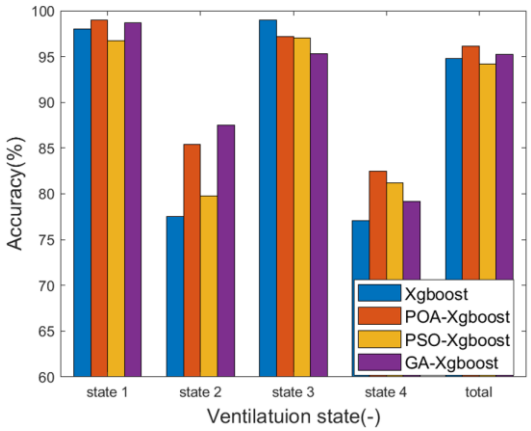


Figure 10: Comparison results identification accuracy for category 1

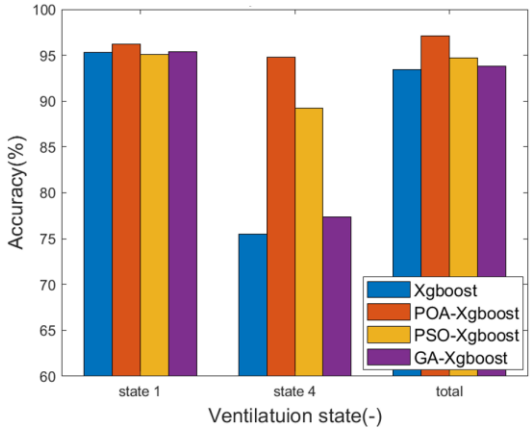


Figure 11: Comparison results identification accuracy for category 2

There is a sample imbalance problem for the remaining verification data of both Category 1 and Category 2. For category 1, the proportion of sample points of state 1 to all samples was 57.6%, 5.8% for state 2, 26.6% for state 3, and 10.0% for state 4. For Category 2, the proportion of the sample points of State 1 to all samples was 90.2%, while that of State 4 was 9.8 %. Accurately identifying states 2 and 4 is most important, as it relates to the switching of propulsion control strategies. From Fig. 10 and Fig. 11, it can be seen that the POA-XGBoost model has the highest total identification accuracy among all the models for both category 1 and category 2. For category 1, although the GA-XGBoost model achieves better identification accuracy for state 2, POA-XGBoost performs better when considering the identification accuracy of states 2 and 4 comprehensively. For category 2, the identification accuracy of state 4 of the POA-XGBoost model is far higher than that of the other models.

The test samples were used to test the accuracy of the trained POA-XGBoost model, as shown in Fig. 6. A confusion matrix is presented in Table 3. From Table 4, it can be seen that the total identification accuracy of POA-XGBoost for the

test samples was 99.38%. This indicates that the POA-XGBoost model has excellent learning ability.

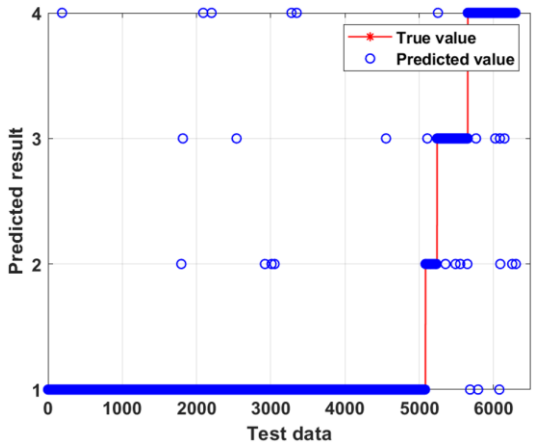


Figure 12: Comparison results between test samples and predicted values

Table 4. Error metrics of POA-XGBoost model for test samples.

		Prediction results				Total	Acc (%)
		D_1	D_2	D_3	D_4		
True values	D_1	5074	2	3	4	5083	99.8
	D_2	3	150	3	0	156	96.2
	D_3	0	4	406	3	413	98.3
	D_4	7	3	7	631	648	97.4

As state 1 has a large sample size and high prediction accuracy, it was excluded from the calculation of the confidence interval for the prediction results, as shown in Table 5. The confidence interval results further demonstrate the effectiveness of the POA-XGBoost model.

Table 5. Confidence interval of prediction accuracy of POA-XGBoost model.

	Category 1	Category 2
Mean error (-)	0.964	0.934
Standard deviation (-)	0.022	0.018
90% CI	[0.950,0.977]	[0.916,0.952]
95% CI	[0.947,0.980]	[0.912,0.956]
99% CI	[0.943,0.985]	[0.905,0.963]

5. CONCLUSION

Considering the important role of propeller state identification and prediction in the safety of ships sailing in adverse sea conditions, as well as the powerful classification performance of the XGBoost algorithm, a method of combining the simulation model with the XGBoost algorithm to design a propeller ventilation identification and prediction model is proposed in this paper. Then, POA, PSO, and GA are adopted to optimize the hyperparameters of the XGBoost algorithm. The results indicate that the proposed method can achieve the function of predict whether a full ventilation state will occur after experiencing a partial propeller ventilation state, which can improve the control of ship propeller ventilation in adverse sea conditions. Meanwhile, the POA has a better optimization effect on the XGBoost algorithm for propeller ventilation identification and prediction.

The proposed method provides a basis for switching propulsion control strategies to improve the stability of ship-integrated power systems. Therefore, in the next step, a model predictive control strategy will be developed based on the propeller ventilation identification and prediction model. We will further develop a propeller ventilation effect prediction model to improve its real-time performance.

REFERENCES

[1] J. Hou, J. Sun, and H. Hofmann, "Adaptive model predictive control with propulsion load estimation and prediction for all-electric ship energy management," *Energy*, vol. 150, pp. 877-889, Jan. 2018.

[2] H. Alafnan *et al*, "Stability Improvement of DC Power Systems in an All-Electric Ship Using Hybrid SMES/Battery," *IEEE Transactions on Applied Superconductivity*, vol. 28, no. 3, pp. 1-6, Nov. 2018.

[3] Ø. N. Smogeli *et al*, "The concept of anti-spin thruster control," *Control Engineering Practice*, vol. 16, no. 4, pp. 465-481, Apr. 2008.

[4] A. Califano, S Steen, "Identification of ventilation regimes of a marine propeller by means of dynamic-loads analysis," *Ocean Engineering*, vol. 38, pp. 1600-1610, Oct. 2011.

[5] L. Savio, S. Steen, "Identification and analysis of full-scale ventilation events," *International Journal of Rotating Machinery*, vol. 2012, pp. 1-19, Jul. 2012.

[6] H. Gao, L. Liao, and Y. He, "Improved control of propeller ventilation using an evidence reasoning rule based Adaboost. M1 approach," *Ocean Engineering*, vol. 209, pp. 107329, Aug. 2020.

[7] X. Zhang, X. Xu, and X. Xu, "Intelligent Sea states identification based on maximum likelihood evidential reasoning rule," *Entropy*, vol. 22, no. 7, pp. 770, Jul. 2020.

[8] T. Chen, and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785-794.

[9] Y. Zhang, B. Feng, and Y. Chen, "Fault diagnosis method for oil-immersed transformer based on XGBoost optimized by genetic algorithm," *Electric Power Automation Equipment*, vol. 41, no. 2, pp. 200-206, 2021.

[10] M. Blanke, K. Lindegaard, and T. I. Fossen, "Dynamic model for thrust generation of marine propellers," *Ifac Proceedings Volumes*, vol. 33 no. 21, pp. 353-358, Aug. 2000.

[11] H. Masami, and K. Yoon-Soo, "A New Coordinate System and the Equations Describing Manoeuvring Motion of a Ship in Waves," *Journal of the Japan Society of Naval Architects & Ocean Engineers*, vol. 1993, no. 173, pp. 209-220, Jan. 2010.

[12] Z. Ayaz, O. Turan, and D. Vassalos, "A 6 DOF manoeuvring model for controlled ship motions of pod-driven ships in astern seas," *Ifac Proceedings Volumes*, vol. 36, no. 21, pp. 193-198, Sep. 2003.

[13] J. J. Jensen, A. E. Mansour, and A. S. Olsen, "Estimation of ship motions using closed-form expressions," *Ocean Engineering*, vol. 31, no. 1, pp. 61-85, Jan. 2004.

[14] Y. Yamamoto, K. Sugai, and H. Inoue, "Wave loads and response of ships and offshore structures from the viewpoint of hydroelasticity," in *Advances in Marine Structures Conference*, 1986.

[15] T. Fujiwara, "Cruising Performance of a Large Passenger Ship in Heavy Sea," in *The Sixteenth International Offshore and Polar Engineering Conference*, 2006.

[16] M. Ueno, Y. Tsukada, and K. Tanizawa, "Estimation and prediction of effective inflow velocity to propeller in waves," *Journal of Marine Science and Technology*, vol. 18, pp. 339-348, Feb. 2013.

[17] X. Zhou *et al*. "Research on cruise ship navigation simulation under the action of wave field (in Chinese)," *Ship science and technology*, vol. 43, no. 9, pp. 80-84, 2021.

[18] W. Hao, S. Fossen, and H. Fang, "Towards Data-driven Identification and Analysis of Propeller Ventilation," in *Mts/ieee Oceans Conference 2016*, 2016.

[19] S. Ma, Y. Ding *et al*. "Identification and Prediction of Propeller Ventilation Based on Stacked LGBM-LSTM-RF-XGB Model,". *Submitted to a journal (Ocean Engineering)*, 2023.

[20] P. Trojovský, and M. Dehghani. "Pelican optimization algorithm: A novel nature-inspired algorithm for engineering applications," *Sensors*, vol. 22, no. 3, pp. 855, Jan. 2022.

[21] N. Alamir, S. Kamel, and T. F. Megahed *et al*. "Developing hybrid demand response technique for energy management in microgrid based on pelican optimization algorithm," *Electric Power Systems Research*, vol. 214, pp. 108905, Jan. 2023.

[22] Y. R. Naidu. "Multi-objective Pelican Optimization Algorithm for Engineering Design Problems," in *International Conference on Distributed Computing and Intelligent Technology: Springer*, 2023, pp. 362-368.

[23] J. Kennedy, and Russell Eberhart, "Particle swarm optimization," *Proceedings of Icnn 95-th international Conference on Neural Networks: Ieee*, 1995, pp. 1942-1948.

[24] S. Mirjalili, "Genetic algorithm," *Evolutionary Algorithms and Neural Networks: Theory and Applications*, vol. 780, pp. 43-55, 2019.