Optimized SVM-Driven Multi-Class Approach by Improved ABC to Estimating Ship Systems State

HUI CAO [ID][1], JUNDONG ZHANG [ID][1], XU CAO[2], RAN LI[2], AND YIRU WANG[2]
[1]Marine Engineering College, Dalian Maritime University, Dalian 116026, China
[2]Dalian Shipping Vocational and Technical College, Dalian 116052, China

Corresponding author: Hui Cao (lustersoft@gmail.com)

ABSTRACT In the intelligent ship field, with the upgrading of ship maintenance mode, the human-centered system maintenance will be gradually replaced by the artificial intelligence decision methods. To improve the training speed and testing accuracy of the state estimation model, an optimized Support Vector Machine (SVM) driven approach by Improved Artificial Bee Colony (IABC) was proposed to solve the global parameters optimization problem. First, the IABC method was achieved from three aspects: nectar source initializing, employed bee global neighborhood searching, and scouts mutation neighborhood searching. Second, the multi-class SVM with one-against-one classifiers was selected, and the best global parameters were achieved by the IABC. Third, the optimized SVM model was adopted in the testing to verify the effectiveness of state estimation. Finally, the elaborated methodology was applied to two actual ship systems to get the analysis results. The effectiveness was verified by using two examples. The results show the following: the IABC optimized SVM can obtain the global optimal parameters at a faster speed than the traditional ABC optimized method; the IABC optimized method can help the training start with better initial parameters, and get a higher classification accuracy rate than the traditional ABC optimized method. Based on the comparative analysis results, the IABC optimized SVM shows an obvious advantage of parameter optimization in the training process, and it can also significantly improve the model training efficiency and achieve a higher state estimation accuracy. The optimized SVM by IABC is an effective state estimation method in ship systems.

INDEX TERMS Marine engineering, support vector machine, artificial bee colony, state estimation.

I. INTRODUCTION

With the development of the intelligent ship field, more sophisticated sensors in marine engine room systems were installed to collect data. To fully exploit the meaning of information and realize an intelligent upgrade of ship monitoring and maintenance mode, it is necessary to study some specific methods to improve the reliability and accuracy of state identification in the engine room systems.

In the traditional view, ship maintenance has been considered an unnecessary high expenditure area, and advanced monitoring methods have not yet been widely applied [1]. Nevertheless, the intelligent maintenance in ships has been made in the past years and is rapidly progressing [2]. For this, Hountalas [3] developed a diesel engine performance

The associate editor coordinating the review of this manuscript and approving it for publication was Zhe Xiao [ID].

model that can account for both normal and faulty conditions. Chandroth [4] proposed a method combined vibration data with cylinder pressures data for the condition monitoring of the main engine. Accordingly, the thermodynamic model of marine engine has been developed and used to achieve condition monitoring [5], [6]. Accordingly, Li *et al.* [7] developed a self-learning algorithm to perform fault diagnosis in the combustion system of a marine diesel engine. In addition, Dikis and Lazakis [8] developed a machinery risk analysis tool that can perform condition monitoring and maintenance support by combining real-time machinery information. Furthermore, Gkerekos *et al.* [9], [10] developed a self-learning model for the condition monitoring of ship machinery based on vibration measurements and further performed a more thorough optimization step for deriving a more robust model. Kowalski *et al.* [11] used various signals that are emitted by the engine as a data input for a pattern classifier, and presented

a one-vs-one extreme learning ensemble-based fault diagnosis method in marine 4-stroke diesel engines. Wang *et al.* 12] adopted the LASSO (Least Absolute Shrinkage and Selection Operator) regression algorithm to implement feature variables selection, and proposed a novel ship fuel consumption prediction model. Coraddu *et al.* [13] and Cipollini *et al.* [14] introduced a regression method to achieve a component degradation estimate in a marine combined diesel-electric and gas propulsion plant. Cipollini *et al.* [15] then further used the same dataset to compare supervised and unsupervised algorithms for fault detection. Raptodimos and Lazakis [16] investigated an unsupervised neural network to achieve running state monitoring of a two stroke marine diesel engine by identifying clusters containing data representing abnormal states. Zhang *et al.* [17] used a single-parameter threshold method and radar chart analysis method to evaluate the health status of ships' power equipment, and converted the evaluation results into health score values. Cheliotis *et al.* [18] introduced a data condition and performance hybrid imputation method for energy efficient operations of marine systems, which combined data mining with first-principle knowledge and was implemented in a dataset acquired from the main engine system of an oceangoing vessel. Zhang *et al.* 19] established a real-time state evaluation system for the power quality of an electric propulsion ship using the fuzzy analytical hierarchy process method, which strengthened the supervision for the ship electric power system. Liu *et al.* [20] applied the fault diagnosis method combining rough set and optimized Support Vector Machine (SVM) to the state identification of the marine main engine, and achieved good results on the multi-category classification problem.

Effective state estimation methods in the ship systems can provide decision support for proactive maintenance and improve the automation and intelligence of ships. Existing methods have disadvantages such as poor parameter optimization ability and low model training efficiency. Especially when faced with complex large-scale engine room electromechanical systems, the generalization and versatility of these methods have been challenged.

This article aims to present an SVM-driven methodology to achieve a multi-state estimation for the ship systems. To improve the performance of state estimation, the method will present an optimized SVM model, which can be trained via the data obtained online and labeled in different system states. To enhance the training speed and testing accuracy of the state estimation model, an improved Artificial Bee Colony (ABC) method will be studied to find the global optimal parameters in the training process. To verify the effectiveness of state estimation, two ship systems will be selected to obtain comparative analysis results compared with the traditional optimization method, and some objective conclusions will be drawn finally.

II. PROPOSED METHODOLOGY

The methodology elaborated in this article includes a preprocessing method for the acquired dataset (part one), an IABC
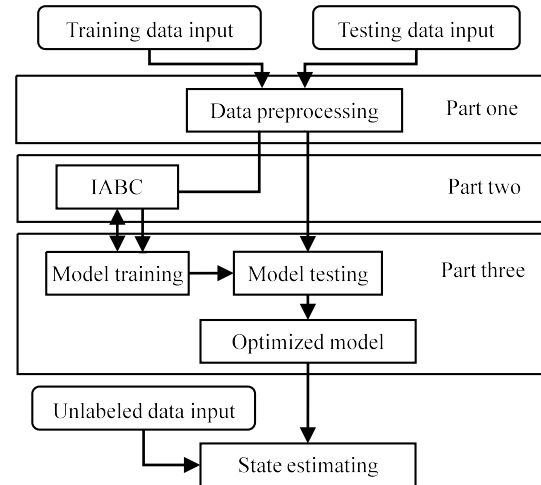

FIGURE 1. Visual representation of the proposed methodology.

method for global parameter optimization (part two), and an SVM driven multiple classification approach using the optimized global parameters (part three) [21]. The flowchart of the proposed methodology is presented in Fig. 1.

A. RAW DATA PREPROCESSING
The data required for SVM model training can be processed by the following steps:

Step 1: Abnormal data filtering.

Due to recording inconsistencies, human error, or sensor faults, the input training data x_i should be filtered out, if its elements with values beyond $\mu \pm 3\sigma$, where μ is the mean value of each element and σ is the standard deviation. According to the 3σ rule, most abnormal data points of the training dataset will be filtered out without affecting the vast majority of normal points.

Step 2: Dataset splitting.

The remaining dataset is split into training (70%) and testing (30%) sub-datasets. The former is used to train the model and optimize parameters, and the latter is used to evaluate the generalization ability of the model. To verify the performance of the SVM model, the K-folding cross-validation method [22] is performed on the training sub-datasets to generate the model training dataset and validation dataset.

Step 3: Dataset normalizing.

Due to the different dimensions of the raw data, it is necessary to normalize the training and testing sub-datasets for the continuity of model training and learning. Formula (1) can be used to normalize the original data [23].

$$y = 2 \times \frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1$$
$$\text{if} \quad x_{\max} = x_{\min}, \quad \text{then} \quad y = x_{\min} \qquad (1)$$

where y is the value after sample normalization, x is the raw value of the sample sequence, and x_{max}, x_{min} are the maximum and minimum values in the sample set.

B. IMPROVED ABC ALGORITHM

The ABC method is a global optimization algorithm based on the intelligent foraging behavior of honey bee swarm, proposed by the Karaboga team in 2005 [24]. The ABC can take all kinds of possible solutions existing in the solution space as the nectar sources, that is, the position of each nectar source represents a possible solution of the problem, and its quality is measured via its fitness value [25]–[27]. To enhance the global optimization capability, we have made some improvements over this algorithm, and call it an Improved ABC (IABC) in this article. The processing of the IABC algorithm is shown in Fig. 2 and the relevant description is as follows.

Step 1: Population parameters initialization.

Set the maximum number of iterations and searches as N_I and N_R. Generate employed bees and onlookers corresponding to the number of nectar sources. If the number of nectar sources representing solutions of the optimization problem is N, and the dimension of each solution is D, the N nectar sources can be generated randomly as

$$\boldsymbol{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,D}), \quad i = 1, 2, \ldots, N \quad (2)$$

where i is the nectar source ID.

If the initial solutions are not reasonable, they will have a great influence on the optimization performance of the whole algorithm. Therefore, to increase the diversity of the initial populations, make the initial solutions distribute in the solution space evenly, an unlearning improvement strategy [28], [29] will be adopted as follows:

First, generate randomly N nectar sources via (3).

$$\widehat{x}_{i,d} = x_{\min,d} + s_d \left(x_{\max,d} - x_{\min,d} \right), \quad d = 1, 2, \ldots, D \quad (3)$$

where d is the dimension of nectar source I, $x_{\min,d}$ and $x_{\max,d}$ mean the upper and lower bounds of d dimension value, and s_d is the random number among [0,1].

Second, generate N inverse nectar sources via (4).

$$\check{x}_{i,d} = x_{\min,d} + x_{\max,d} - \widehat{x}_{i,d} \quad (4)$$

Finally, by merging the two types of nectar sources, calculating and sorting their fitness, the best N nectar sources as the initial populations can be selected via (5).

$$\begin{cases} j = \arg \max_{j=0,\ldots,2N} fit \left((\widehat{\boldsymbol{x}} \cup \check{\boldsymbol{x}})_j \right) \\ \boldsymbol{x}_i = \left(\widehat{\boldsymbol{x}} \cup \check{\boldsymbol{x}} \right)_j \quad i = 1, 2, \ldots, N \end{cases} \quad (5)$$

where: *fit* is the fitness computing function.

Step 2: Employed bees neighborhood search phase.

In the traditional ABC method, the employed bees will search new nectar source via (6), and update the position of the nectar source according to the search results to form a new nectar source $\boldsymbol{x}_i$ as a new solution.

$$x_{i,d}^{(k)} = x_{i,d}^{(k-1)} + r \left(x_{i,d}^{(k-1)} - x_{j,d}^{(k-1)} \right), \quad j \neq i \quad (6)$$
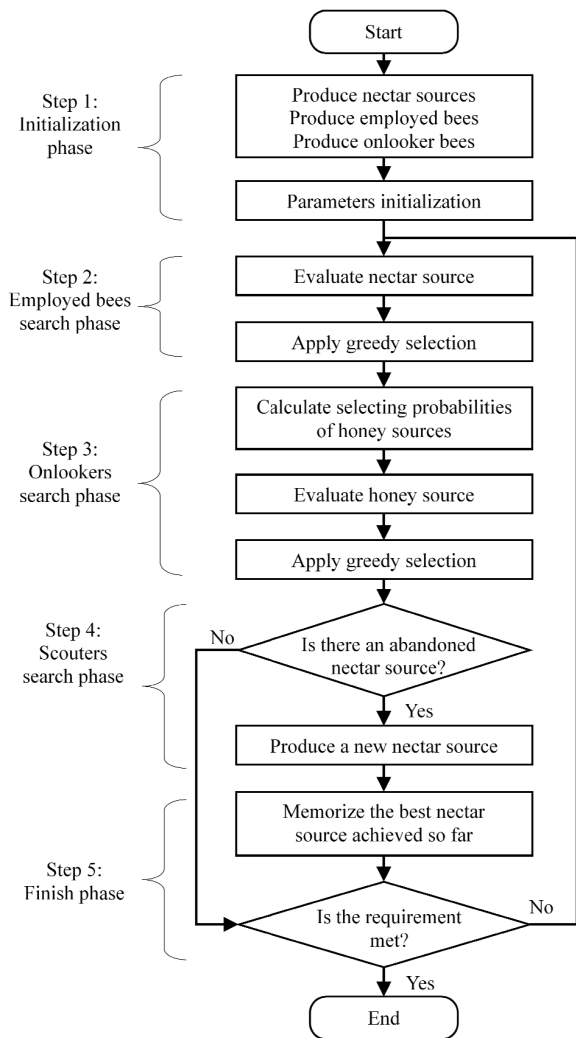

FIGURE 2. Visual representation of the process of IABC algorithm.

where k is the number of searches, i or j is the nectar source ID, d is the dimension of the nectar source, and r is the random number among $[-1,1]$.

It can be seen that (6) does not include the global optimal value, so the algorithm has insufficient searching ability near the global optimal value, which can lead to a slow update speed and easily fall into a local optimal solution. Therefore, the neighborhood search method can be improved by introducing a global search factor [30] as

$$x_{i,d}^{(k)} = x_{i,d}^{(k-1)} + r \left(x_{i,d}^{(k-1)} - x_{j,d}^{(k-1)} \right) + y, \quad j \neq i$$

$$y = s_1 \left(x_{\text{gbest},d} - x_{i,d}^{(k-1)} \right) + s_2 \left(x_{\text{cbest},d} - x_{i,d}^{(k-1)} \right) \quad (7)$$

where r is the random number among $[-1,1]$, s_1 and s_2 are the random numbers among [0,1], $x_{gbest,d}$ is the d-th element of the global best solution found so far, and $x_{cbest,d}$ is the d-th element of the best solution in the current iteration.

After the new nectar source $\boldsymbol{x}_i$ is generated, the fitness f_i is calculated via (8) accordingly. Then, by comparing with

the original nectar source, the greedy selection strategy is adopted to retain the information of the nectar source i with greater fitness.

$$f_i = fit(\boldsymbol{x}_i) = \begin{cases} 1/(1 + t_i) & \text{if } t_i \geq 0 \\ 1 + |t_i| & \text{if } t_i < 0 \end{cases} \quad (8)$$

where t_i is the objective function value of the i-th nectar source $\boldsymbol{x}_i$ corresponding to the optimization problem.

Step 3: Onlookers neighborhood search phase.

After the employed bees complete their search task, they will return to share the nectar source information (solution and fitness) with the onlookers. Each onlooker will calculate the preference probability via the fitness according to the roulette wheel selection method (9) to select the nectar source.

$$p_i = f_i \left/ \sum_{n=1}^{N} f_n \right. \quad (9)$$

where p_i is the preference probability of the nectar source i, and f_i is the fitness of it.

After selecting the nectar source, the onlookers adopts the same improved searching and updating method as the employed bee, and uses the greedy selection strategy to generate a new nectar source i and a fitness f_i.

Step 4: Scouts neighborhood search phase.

If the fitness of a nectar source $\boldsymbol{x}_i$ is not improved within N_R search times, the nectar source will be discarded, and the corresponding employed bee will transform into a scout, then, a new alternative nectar source will be searched via (10).

$$\boldsymbol{x}_i^{(0)} = \boldsymbol{x}_{\min} + \boldsymbol{r} \left(\boldsymbol{x}_{\max} - \boldsymbol{x}_{\min} \right) \quad (10)$$

where the superscript 0 means the nectar source i is a new original nectar source, $\boldsymbol{x}_{min}$ and $\boldsymbol{x}_{max}$ mean the upper and lower bounds of D-dimension value of nectar source, and $\boldsymbol{r}$ is a D-dimension random number among $[-1,1]$ for each dimension.

Similar to (6), the search method (10) does not include the global optimal value. To improve the searching ability near the global optimal value, the search method can be updated as follows:

$$\hat{\boldsymbol{x}}_1 = \boldsymbol{x}_i^{\text{rand}} = \boldsymbol{x}_{\min} + \boldsymbol{s}_1 \left(\boldsymbol{x}_{\max} - \boldsymbol{x}_{\min} \right) \quad (11)$$

$$\text{if } \sum_{d=1}^{D} \boldsymbol{s}_{1,d} \leq \frac{D}{2},$$

$$\hat{\boldsymbol{x}}_2 = \boldsymbol{x}_i^{\text{mutation1}} = \boldsymbol{x}_i + \boldsymbol{s}_2 \left(1 - \frac{n_{\mathrm{I}}}{N_{\mathrm{I}}} \right)^D \left(\boldsymbol{x}_{\text{gbest}} - \boldsymbol{x}_i \right) \quad (12)$$

$$\text{if } \sum_{d=1}^{D} \boldsymbol{s}_{1,d} > \frac{D}{2}, \quad (13)$$

$$\hat{\boldsymbol{x}}_3 = \boldsymbol{x}_i^{\text{mutation2}}$$

$$= \boldsymbol{x}_i + \boldsymbol{s}_3 \left(1 - \frac{n_{\mathrm{I}}}{N_{\mathrm{I}}} \right)^D \left(\boldsymbol{x}_{\text{cbest}} - \boldsymbol{x}_i \right)$$

$$\boldsymbol{x}_i^{(0)} = \hat{\boldsymbol{x}}_j = \arg \max_{j=1,2,3} fit \left(\hat{\boldsymbol{x}}_j \right) \quad (14)$$

where $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, $\boldsymbol{s}_3$ are all D-dimension random number among [0,1], and n_I is the number of the current iteration.

After finding a new nectar source $\boldsymbol{x}_i$ via (14), the scout will change back to an employed bee, and go to step 2 of the next iteration to update the nectar source.

Step 5: End of optimization.

If the number of iterations reaches the maximum number N_I, or the fitness of a certain nectar source reaches the threshold preset, the optimization process will end, and the best nectar source found so far will be memorized and output to the SVM model for testing.

C. SVM CLASSIFICATION METHOD

The SVM method was first proposed by Cortes and Vapnik in 1995 [31]. It is based on the Structural Risk Minimization (SRM) principle rooted in the statistical learning theory, and can give a better generalization ability. In the SVM, the SRM is achieved through a minimization of the upper bound of the generalization error 32].

1) LINEAR BINARY SVM CLASSIFIER

Given training sample set $\boldsymbol{D} = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2),\ldots, (\boldsymbol{x}_M, y_M)\}$, $y_i \in \{-1,1\}$, M is the number of samples. The samples are assumed to have two classes namely positive class with the labels $y_i = 1$ and negative class with the labels $y_i = -1$. For the samples, it is possible to determine a hyperplane $f(\boldsymbol{x}) = 0$ to separate them.

$$f(\boldsymbol{x}) = \boldsymbol{W}^T \boldsymbol{x} + b = 0 \quad (15)$$

where $\boldsymbol{W}$ is the normal vector, which determines the direction of the hyperplane, and b is a scalar.

Via the sign equation $f(\boldsymbol{x})$, a distinct separating hyperplane should satisfy the constraints (16) to classify input data in either positive class or negative class.

$$\begin{cases} f(\boldsymbol{x}_i) = 1 & \text{if } y_i = 1 \\ f(\boldsymbol{x}_i) = -1 & \text{if } y_i = -1 \end{cases} \quad (16)$$

Further, the constraints (16) can also be presented as

$$y_i f(\boldsymbol{x}_i) = y_i (\boldsymbol{W}^T \boldsymbol{x}_i + b) \geq 1, \quad i = 1, 2, \ldots, M \quad (17)$$

The optimal hyperplane of two data sets is required to have a maximum distance between itself and the nearest data. The nearest data used to define the margin are called support vectors. The maximum distance is called the maximum margin, and can be obtained as a solution to the following optimization problem:

$$\min_{\boldsymbol{W},C,\xi} \quad \frac{1}{2} \|\boldsymbol{W}\|^2$$

$$\text{s.t.} \quad y_i(\boldsymbol{W}^T \boldsymbol{x}_i + b) \geq 1, \quad i = 1, 2, \ldots, M \quad (18)$$

By introducing the Lagrange multiplier $\alpha_i \geq 0$, the optimization can be simplified to an equivalent Lagrange dual

problem:

$$\min_{W,b,\alpha} L(W, b, \alpha) = \frac{1}{2} \|W\|^2 + \sum_{i=1}^{M} \alpha_i \left(1 - y_i(W^T x_i + b)\right)$$

$$= \frac{1}{2} \|W\|^2 - \sum_{i=1}^{M} \alpha_i y_i(W^T x_i + b) + \sum_{i=1}^{M} \alpha_i \tag{19}$$

For the optimal point, via the saddle-point equations (20), the results can be obtained as shown in (21).

$$\frac{\partial L(W, b, \alpha)}{\partial W} = 0, \quad \frac{\partial L(W, b, \alpha)}{\partial b} = 0 \tag{20}$$

$$W = \sum_{i=1}^{M} \alpha_i y_i x_i, \quad \sum_{i=1}^{M} \alpha_i y_i = 0 \tag{21}$$

Then, the dual optimization formula (22) can be derived by inserting (21) into (19).

$$\max_{\alpha} L(\alpha) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

$$\text{s.t.} \quad \sum_{i=1}^{M} \alpha_i y_j = 0, \quad \alpha_i \geq 0, \quad i = 1, 2, \dots, M \tag{22}$$

By solving the dual optimization problem, α_i can be solved, which is corresponding to the training sample (x_i, y_i). Because (22) has inequality constraints, the above process satisfies the Karush-Kuhn-Tucker (KKT) condition, for any training sample (x_i, y_i) and α_i, the conclusion can be derived as follows:

$$\begin{cases} \alpha_i \geq 0, \\ y_i f(x_i) \geq 0, \\ \alpha_i (y_i f(x_i) - 1) = 0. \end{cases} \Rightarrow \begin{cases} \alpha_i = 0, \\ \text{or} \\ \alpha_i > 0 \quad \text{and} \quad y_i f(x_i) = 1. \end{cases} \tag{23}$$

where the x_i, which corresponding to $\alpha_i > 0$, is called a support vector, and most of the other training samples will be abandoned after training due to the fact that they do not contribute to the classification model, so by inserting α_i into (21), W can be solved, and by inserting W into (15), b can further be calculated via

$$b = \frac{1}{M_{SV}} \sum_{i=1}^{M_{SV}} (W^T x_i - y_i) \tag{24}$$

where M_{SV} is the number of support vectors, x_i is the support vector, and via the calculated W, b, the SVM classification decision model can be expressed as

$$f(x) = W^T x + b = \sum_{i=1}^{M_{SV}} \alpha_i y_i x_i^T x + b \tag{25}$$

In most situation, the training sample set D is linearly inseparable, that is, not all samples can meet the constraints, and there are always some samples cannot be divided correctly. To solve this problem, the soft margin and kernel function can be introduced in the process of SVM classification.

2) SOFT MARGIN SVM

The soft margin allows the SVM classifier to make errors in some samples. By introducing the soft interval mechanism, the optimization problem (18) can be expressed as

$$\min_{W,C,\xi} \frac{1}{2} \|W\|^2 + C \sum_{i=1}^{M} \xi_i$$

$$\text{s.t.} \begin{cases} y_i(W^T x_i + b) \geq 1 - \xi_i & i = 1, 2, \dots, M \\ \xi_i \geq 0, & i = 1, 2, \dots, M, \end{cases} \tag{26}$$

where ξ_i is called slack variable, which presents the distance between the margin and the examples x_i that lying on the wrong side of the margin, and C is called penalty factor, the larger C, the more intolerable the existence of errors, the easier to over-fit, and vice versa.

Similar to (18), the equivalent Lagrange dual function of (27) can be obtained via the Lagrange multiplier method.

$$\min_{W,b,\alpha,\xi,\mu} L(W, b, \alpha, \xi, \mu)$$

$$= \frac{1}{2} \|W\|^2 + C \sum_{i=1}^{M} \xi_i$$

$$+ \sum_{i=1}^{M} \alpha_i \left(1 - \xi_i - y_i(W^T x_i + b)\right) - \sum_{i=1}^{M} \mu_i \xi_i \tag{27}$$

where $\alpha_i \geq 0$, $\mu_i \geq 0$ are Lagrange multipliers. For the optimal point, let the partial derivative of L as follows:

$$\frac{\partial L(W, b, \alpha)}{\partial W} = 0, \quad \frac{\partial L(W, b, \alpha)}{\partial b} = 0, \quad \frac{\partial L(W, b, \alpha)}{\partial \xi} = 0 \tag{28}$$

Then, the following results can be obtained

$$W = \sum_{i=1}^{M} \alpha_i y_i x_i, \quad \sum_{i=1}^{M} \alpha_i y_i = 0, \quad C = \alpha_i + \mu_i \tag{29}$$

So, the dual optimization problem can be derived by inserting (29) into (27) as

$$\max_{\alpha} L(\alpha) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

$$\text{s.t.} \quad \sum_{i=1}^{M} \alpha_i y_j = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, M \tag{30}$$

Similar to (22), the process of the soft margin SVM satisfies the KKT condition too, and can be shown as

$$\begin{cases} \alpha_i \geq 0, \ \mu_i \geq 0, \\ y_i f(x_i) - 1 + \xi_i \geq 0, \\ \alpha_i (y_i f(x_i) - 1 + \xi_i) = 0, \\ \xi_i \geq 0, \ \mu_i \xi_i = 0 \end{cases} \Rightarrow \begin{cases} \alpha_i = 0, \\ \text{or} \\ \alpha_i > 0 \quad \text{and} \\ y_i f(x_i) = 1 - \xi_i. \end{cases} \tag{31}$$

where x_i represents a support vector when the Lagrange multiplier $\alpha_i > 0$.

According to (29), (31): if $\alpha_i < C$, then $\mu_i > 0$, and $\xi_i = 0$, that is, the sample is just on the boundary of the maximum margin; if $\alpha_i = C$, then $\mu_i = 0$, so further if $\xi_i \leq 1$, the sample will fall within the maximum margin, and if $\xi_i > 1$, the sample will be misclassified. It can be seen that the model of soft margin SVM is only related to the support vectors, so the decision model can be expressed as same as (25).

3) NON-LINEAR SVM

The SVM classification method mentioned above is very efficient, but many data sets are not linearly separable in most cases. For the non-linear classification tasks, a non-linear vector function $\boldsymbol{\Phi}(\boldsymbol{x}) = (\Phi_1(\boldsymbol{x}), \Phi_2(\boldsymbol{x}), \ldots, \Phi_l(\boldsymbol{x}))$ can be used to map the n-dimensional input vector $\boldsymbol{x}$ to a l-dimensional feature space ($l > n$). The mapping can make the input data become linearly separable in l-dimensional space. This leads to the non-linear decision formula (32) as

$$f(\boldsymbol{x}) = \text{sign}\left(\sum_{i=1}^{M_{\text{SV}}} \alpha_i y_i \left(\boldsymbol{\Phi}^{\text{T}}(\boldsymbol{x}_i), \Phi(\boldsymbol{x})\right) + b\right) \quad (32)$$

Because the dimension of feature space may be very high or even infinite, it is usually difficult to directly calculate the dot product of the feature space mappings of the original data points. The problem can be solved by using the kernel function κ, stated as

$$\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \left\langle \boldsymbol{\Phi}(\boldsymbol{x}_i), \boldsymbol{\Phi}(\boldsymbol{x}_j) \right\rangle = \boldsymbol{\Phi}^{\text{T}}(\boldsymbol{x}_i), \Phi(\boldsymbol{x}_j) \quad (33)$$

By introducing the kernel function κ, the dual formula and the non-linear decision formula can be rewritten as shown as follows:

$$\max_{\alpha} L(\alpha) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_i \alpha_j y_i y_j \kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) \quad (34)$$

$$\text{s.t.} \quad \sum_{i=1}^{M} \alpha_i y_j = 0, \qquad 0 \leq \alpha_i \leq C, \qquad i = 1, 2, \ldots, M$$

$$f(\boldsymbol{x}) = \text{sign}\left(\sum_{i=1}^{M_{\text{SV}}} \alpha_i y_i \kappa(\boldsymbol{x}_i, \boldsymbol{x}) + b\right) \quad (35)$$

The commonly used kernel functions include polynomial kernel, Gaussian kernel and sigmoid kernel functions. Among them, the Gaussian kernel function (36) has the advantages of fast convergence, nonlinear mapping and less parameters [33], and will be adopted in this article.

$$K\left(x_i, x_j\right) = \exp\left(-\left\|x_i - x_j\right\|^2 / 2\sigma^2\right) \quad (36)$$

where σ is the kernel parameter, the larger σ is, the less support vectors are identified, and vice versa.

4) MULTI-CLASS SVM

The discussion above is binary classification and the label values are 1 and -1. In ship systems, however, there are always more than two classes of system states that need to
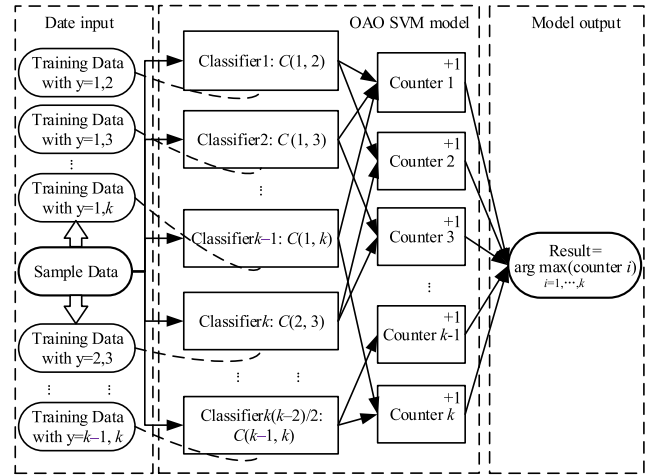

FIGURE 3. Structure of the OAO multi-class SVM model.

be estimated. Therefore, the multi-class classification method will be further discussed based on the binary classification method. There are many ways to solve multi-class SVM classification, such as directed acyclic graph (DAG), binary tree (BT), one-against-one (OAO), and one-against-all (OAA) methods [34]. Among them, the OAO method is a widely used and verified multi-class method that has a shorter training duration and higher training accuracy, when the number of sample labels is not large [35], [36].

If k is the number of classes, there will be $k(k-2)/2$ binary classifiers in OAO model, and each one trains data from two classes. For the training data $\boldsymbol{D}_{ij} = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \ldots, (\boldsymbol{x}_T, y_T)\}$, $y_t \in \{-1, 1\}$, the optimization problem of the binary classifier $C(i, j)$ in OAO can be expressed as

$$\min_{\boldsymbol{W}^{ij}, b^{ij}, \xi_t^{ij}} \quad \frac{1}{2} \left\|\boldsymbol{W}^{ij}\right\|^2 + C \sum_{t=1}^{T} \xi_t^{ij}$$

$$\text{s.t.} \begin{cases} (\boldsymbol{W}^{ij})^{\text{T}} \boldsymbol{\Phi}(\boldsymbol{x}_t) + b^{ij} \geq 1 - \xi_t^{ij}, & \text{if } y_t = i \\ (\boldsymbol{W}^{ij})^{\text{T}} \boldsymbol{\Phi}(\boldsymbol{x}_t) + b^{ij} \leq -1 + \xi_t^{ij}, & \text{if } y_t = j \\ \xi_t^{ij} \geq 0 & i = 1, 2, \ldots, M \end{cases} \quad (37)$$

where $\boldsymbol{x}_t$ is the training data, y_t is the label of $\boldsymbol{x}_t$, and T is the number of the training data $\boldsymbol{D}_{ij}$. The equivalent Lagrange dual function of (37) can be obtained via the Lagrange multiplier method. Similar to (35), the decision formula of binary SVM between class i and class j is

$$f(\boldsymbol{x}) = \text{sign}\left(\sum_{t=1}^{M_{\text{SV}}} \alpha_t^{ij} y_t \kappa(\boldsymbol{x}_t, \boldsymbol{x}) + b^{ij}\right) \quad (38)$$

After the training of each classifier, the process of multi-classification is shown in Fig. 3. In the process, the voting strategy will be adopted: for each unknown sample, all classifiers constructed will be used, and each classifier votes on its classification results via decision formula (38); if the result is identified as the i-th class, the number of votes
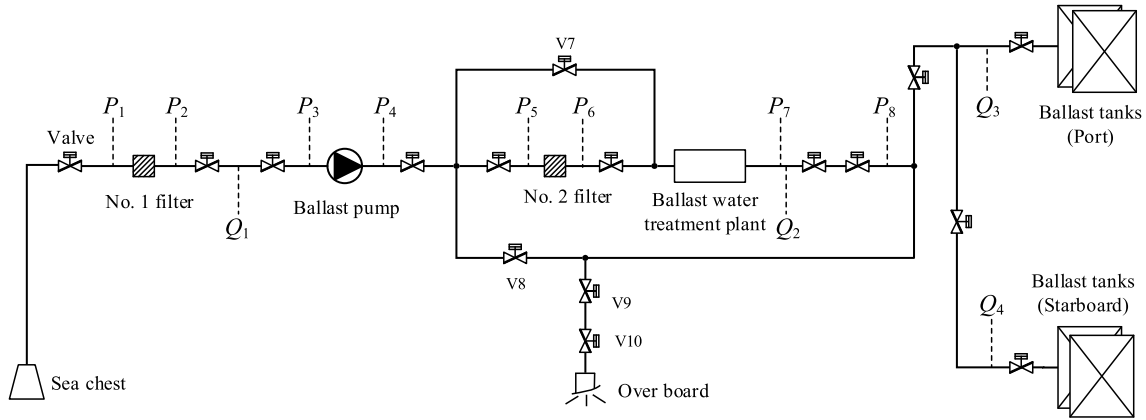
FIGURE 4. The layout of the BW data acquisition and monitoring points.

TABLE 1. Parameter description and unit.

ID	Parameter description	Unit
P_1	No. 1 filter suction pressure	bar
P_2	No. 1 filter discharge pressure	bar
P_3	Ballast pump suction pressure	bar
P_4	Ballast pump suction pressure	bar
P_5	No. 2 filter suction pressure	bar
P_6	No. 2 filter discharge pressure	bar
P_7	Ballast water treatment plant outlet pressure	bar
P_8	No. 12 valve outlet pressure	bar
Q_1	Ballast pump flow	m³/h
Q_2	Ballast water treatment plant flow	m³/h
Q_3	Port ballast tank main pipe flow	m³/h
Q_4	Starboard ballast tank main pipe flow	m³/h
P_m	Ballast pump motor output power	kW

in the *i*-th class will be increased by 1, and the output is the highest cumulative number of votes obtained via the ''MaxWins'' strategy, which is the prediction classification of this unknown sample; in case that two classes have identical votes, the one with the smaller index will be selected.

D. IABC OPTIMIZED SVM METHOD

It can be seen from (36), (37) and (38) that the key parameters that affect the performance of the SVM classifier are the penalty factor C and the kernel parameter σ. For that, the IABC method will be adopted to optimize the two critical parameters as follows [37]:

Step 1: Raw data preprocessing.

According to the steps of the data preprocessing mentioned above to achieve the data process.

Step 2:IABC algorithm initializing.

Set the following parameters: the size of the bee colony, the number of nectar sources, the maximum number of searches for each nectar source N_R, the maximum number of iterations N_I, and the value ranges of the penalty factor C and the kernel function parameter σ.

Step 3: Objective function setting.

To evaluate the quality of the SVM classifiers, the training (70%) sub-dataset preprocessed will be adopted as input data
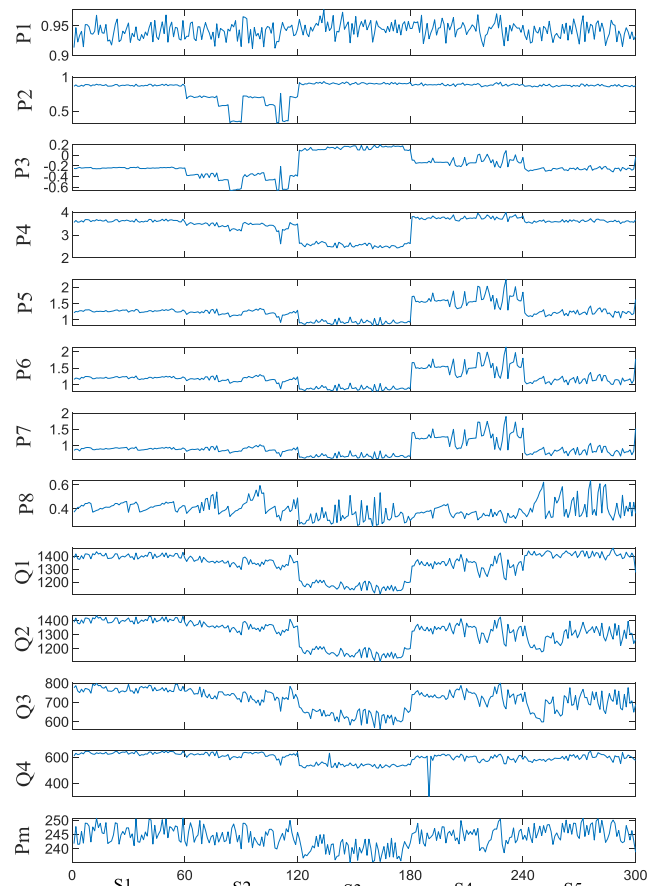

FIGURE 5. Sample data trends with 5 states.

in the SVM model training, and the objective function of SVM will be set according to classification accuracy as

$$t = f_t(v_{acc}) = 1 - v_{acc} \qquad (39)$$

where v_{acc} is the classification accuracy in the range [0, +1], and t is the target value of the optimization process, the smaller the better, 0 is the best.

TABLE 2. Descriptive statistics of the data samples.

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	Q_1	Q_2	Q_3	Q_4	Pm
count	300	300	300	300	300	300	300	300	300	300	300	300	300
mean	0.942	0.831	−0.183	3.386	1.255	1.191	0.913	0.394	1334.3	1310	713.32	595.69	244.12
std	0.013	0.131	0.203	0.438	0.25	0.246	0.236	0.072	92.34	87.092	55.842	40.14	3.552
min	0.912	0.326	−0.662	2.397	0.826	0.788	0.583	0.23	1107.2	1107.5	562.29	297.31	235.08
25%	0.9339	0.8639	−0.275	3.35	1.134	1.049	0.767	0.344	1304.7	1236.9	665.19	571.83	241.84
50%	0.9433	0.8795	−0.232	3.572	1.248	1.178	0.874	0.383	1363.1	1335.2	727.65	607.03	244.16
75%	0.9516	0.8955	−0.086	3.662	1.314	1.247	0.948	0.434	1402.5	1380.2	757.6	624.92	246.59
max	0.9773	0.9352	0.2025	3.986	2.24	2.14	1.897	0.635	1462.3	1434.4	800.19	651.64	250.78

A. Initial nectar sources produced.

B. Initial inverse nectar sources.

C. Nectar sources merged and sorted.

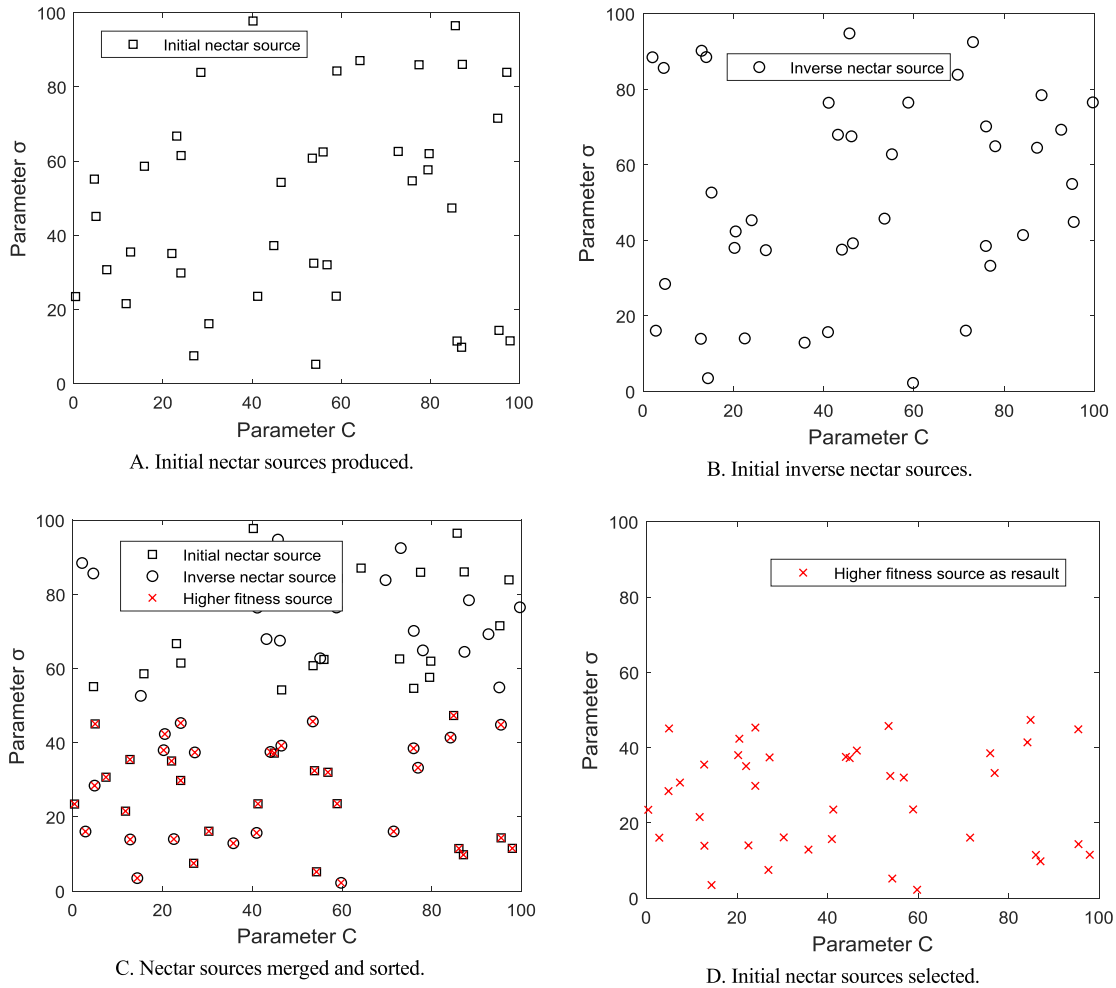D. Initial nectar sources selected.

FIGURE 6. Initial nectar sources producing and selecting.

Step 4: Model training and parameters optimizing.

Adopt the IABC method mentioned above to optimize the SVM model parameters C, σ. According to the nectar source fitness (8), all the possible solutions can be optimized through the respective searching activities of the three honeybees.

Step 5: Model classification performance checking.

The obtained optimal parameters C, σ in the previous step, namely the global optimal nectar source, are used to construct the SVM model. Then, the testing (30%) sub-dataset will be used to evaluate the generalization ability of the model.

III. METHODOLOGY APPLICATION

Two example systems of a real ship were selected to verify the effectiveness of the proposed method. The first is a Ballast Water system (BW), which as an important auxiliary system for adjusting the floating state of ships, and the second is a central Cooling Water system group (CW), which combines a Sea Water cooling system (SW), a Low Temperature Fresh Water cooling system (LTFW), and a High Temperature Fresh Water cooling system (HTFW). Both example systems have a scattered equipment layout, complex pipe network,
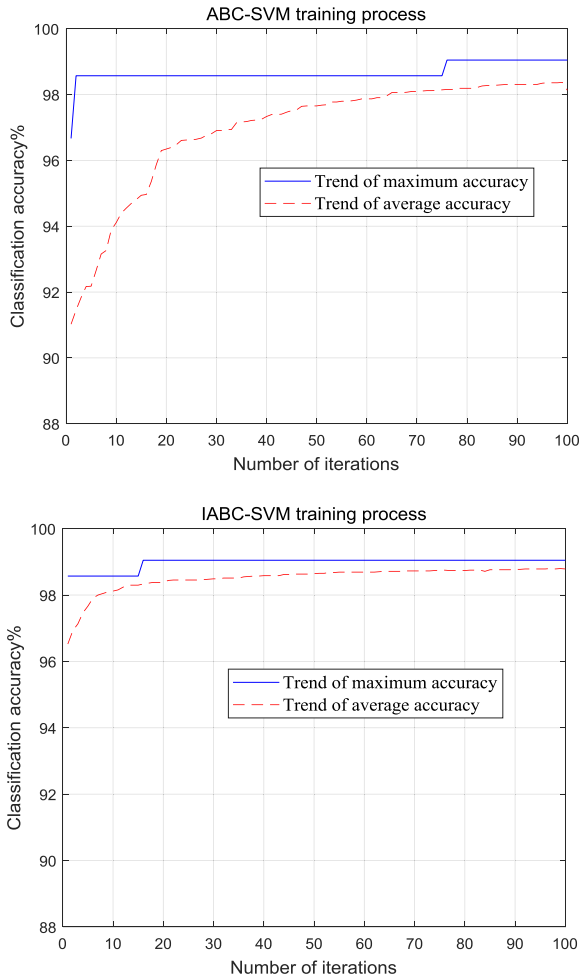
FIGURE 7. Parameter optimization process.

FIGURE 8. Multi-class SVM state estimation results in testing.

and numerous valves, so it is difficult to perform manual maintenance during system operation. But the purpose of method verification here is different. The former focuses on the recognition of fault states, while the latter focuses on the recognition of normal states. The data required for model training and testing was acquired through the online Data Acquisition System (DAS) of the ship in different system states.

A. WORKING STATE IDENTIFICATION OF BW

1) STATE DEFINITION AND PARAMETER SELECTION

The working states of the BW were be summarized as S1 to S5: normal working state (S1), sea chest filter blocked (S2), ballast pump wear (S3), valve stuck (S4), and sea water pipeline leakage (S5). Based on the system working principle and the abstract system layout diagram, it can be derived that the system states have a strong relationship with the following system parameters: the fluid pressure monitoring points P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8; the flow monitoring points Q_1, Q_2, Q_3, Q_4; the ballast pump motor output power P_m. The acquisition points of all parameters are shown in Fig. 4, and the description of each parameter is shown in Table 1.
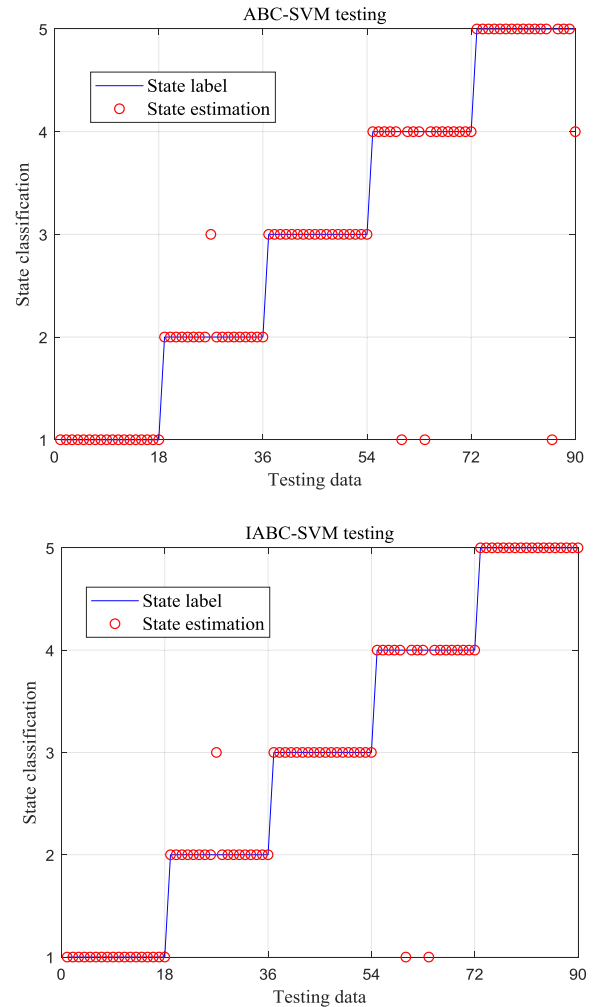
2) DATA ACQUISITION AND PREPROCESSING

According to the 5 system states, after abnormal data filtering via 3σ rule, there are 300 samples acquired through DAS left, where the data trends according to 5 states are shown in Fig. 5. The static feature statistics are shown in Table 2, including the mean, minimum and maximum values of each feature, the standard deviation, and the values at different (25%, 50%, and 75%) quantiles. By data splitting, the number of samples for each state is 60, including 42 (70%) training samples and 18 (30%) testing samples, and due to the different dimensions of the state data, used (1) to further achieve data normalization.

3) IABC INITIALIZATION

Set the maximum number of iterations $N_I = 100$, the maximum number of searches for each nectar source $N_R = 50$, the number of nectar sources $N = 40$, the optimization range of parameters C, σ is 0–100, and generate employed bees and onlookers corresponding to the number of nectar sources.

Adopt the unlearning improvement strategy via (3), (4) and (5) to generate 40 initial nectar sources as shown in Fig. 6.
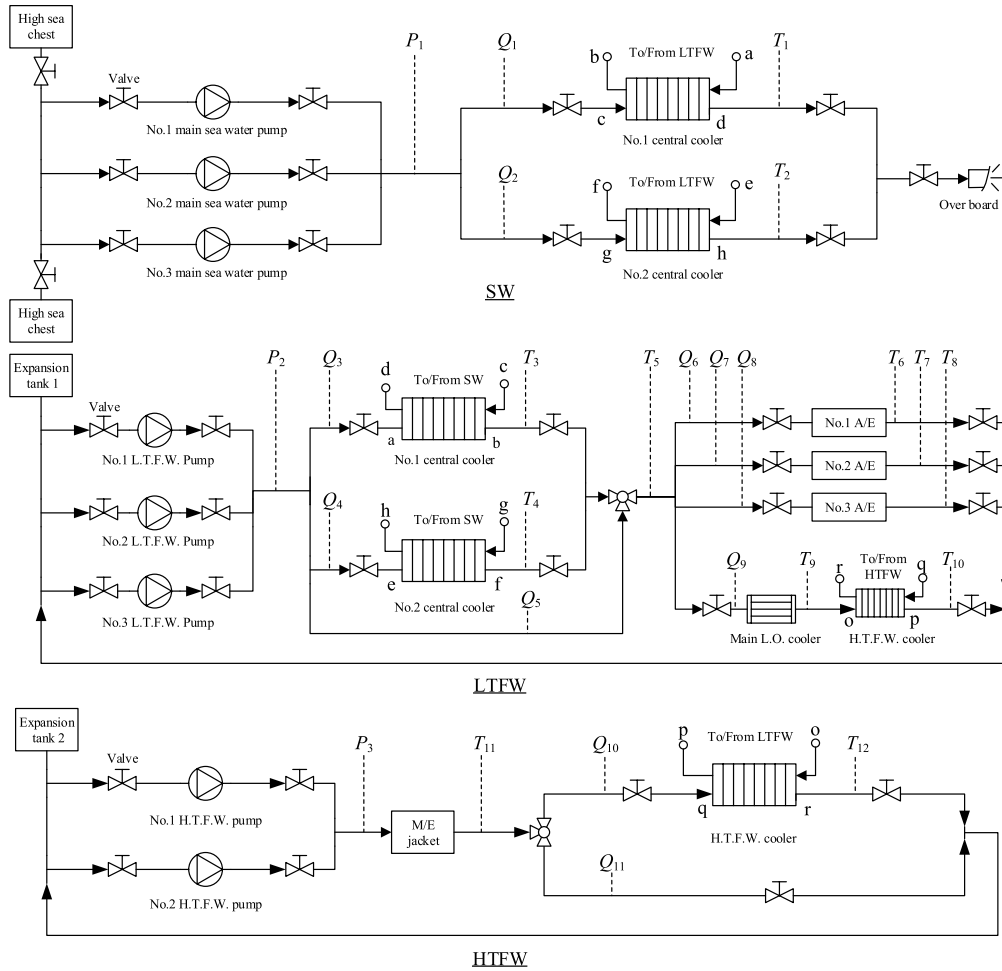
FIGURE 9. The layout of the CW data acquisition and monitoring points.

4) SVM INITIALIZATION

Set the K-folding parameter $k = 6$, adopt OAO classifier (38) as the basic system state estimation model, and select the Gaussian function (36) as the kernel function, so the penalty factor C and the kernel parameter σ will be the optimized parameters via IABC.

5) MODEL TRAINING AND PARAMETERS OPTIMIZING

According to the nectar source fitness (8), the solution (C, σ) can be optimized through the respective searching activities of the three kinds of honeybee from the initial nectar sources selected (as shown in Fig.6 D). The trends of the classification accuracy between IABC optimized SVM (IABC-SVM) and traditional ABC optimized SVM (ABC-SVM) are shown in Fig. 7, and more details are shown in Table 3.

6) CLASSIFICATION PERFORMANCE TESTING

The obtained optimal parameters C, σ in the previous step were used to construct the SVM model. Then, the testing (30%) sub-dataset was used to evaluate the generalization ability of the model in the process of testing. The results of

TABLE 3. Comparison of training performance.

	ABC-SVM	IABC-SVM
Training time	38.77s	23.35s
Initial mean accuracy	91.02%	96.52%
The best mean accuracy	99.05%	99.05%
Cross validation accuracy	89.05%	98.67%
Best C	67.53	75.92
Best σ	0.71	0.74

TABLE 4. Comparison of testing performance.

	ABC-SVM	IABC-SVM
Optimal parameters	(67.53, 0.71)	(75.92, 0.74)
S1 estimation/label	18/18	18/18
S2 estimation/label	17/18	17/18
S3 estimation/label	18/18	18/18
S4 estimation/label	16/18	16/18
S5 estimation/label	16/18	18/18
Accuracy	94.44% (85/90)	96.67% (87/90)

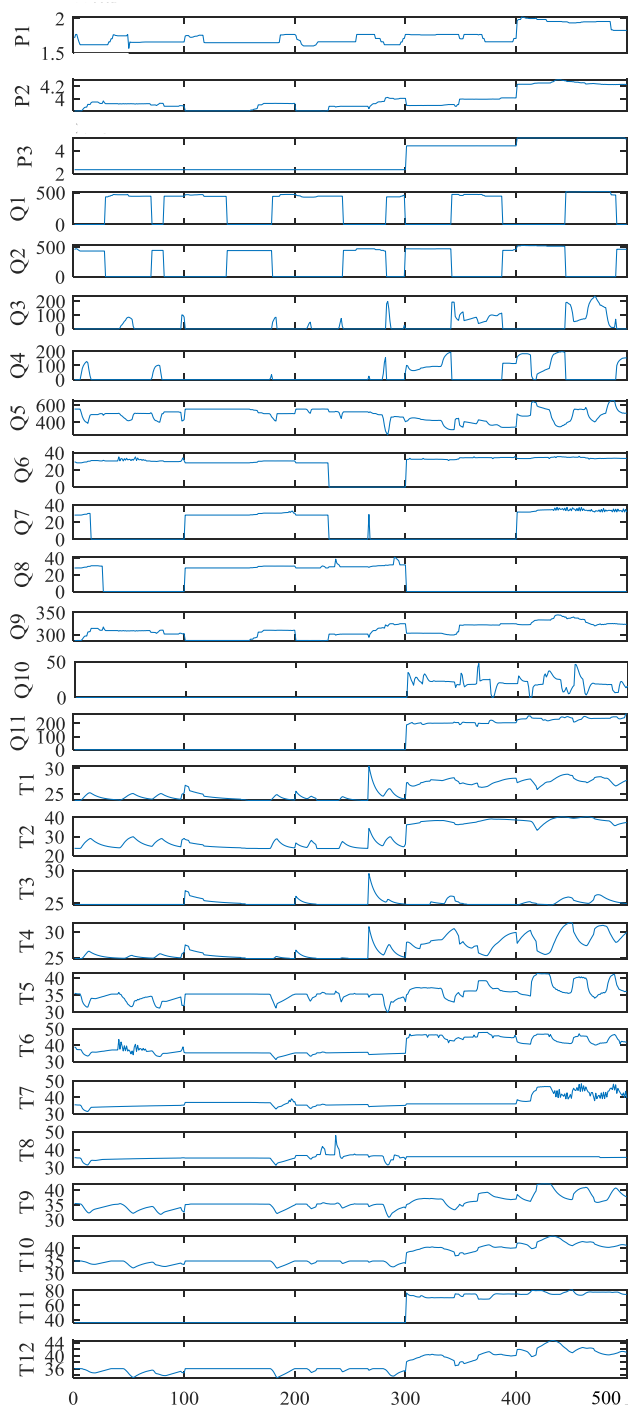model testing are shown in Fig. 8, and more details are shown in Table 4.

FIGURE 10. Sample data trends with 5 states.

TABLE 5. Parameter description and unit.

ID	Parameter description	Unit
P_1	Main Sea water pumps outlet pressure	bar
P_2	L.T.F.W. pumps outlet pressure	bar
P_3	H.T.F.W. pumps outlet pressure	bar
Q_1	Sea water flow of No.1 central cooler	m³/h
Q_2	Sea water flow of No.2 central cooler	m³/h
Q_3	Fresh water flow of No.1 central cooler	m³/h
Q_4	Fresh water flow of No.2 central cooler	m³/h
Q_5	Fresh water bypass flow of central coolers	m³/h
Q_6	No.1 A/E cooling water flow	m³/h
Q_7	No.2 A/E cooling water flow	m³/h
Q_8	No.3 A/E cooling water flow	m³/h
Q_9	Low temperature water flow of H.T.F.W. cooler	m³/h
Q_{10}	High temperature water flow of H.T.F.W. cooler	m³/h
Q_{11}	Fresh water bypass flow of H.T.F.W. cooler	m³/h
T_1	Sea water outlet temperature of No.1 central cooler	°C
T_2	Sea water outlet temperature of No.2 central cooler	°C
T_3	Fresh water outlet temperature of No.1 central cooler	°C
T_4	Fresh water outlet temperature of No.2 central cooler	°C
T_5	Outlet temperature of L.T.F.W temperature regulator	°C
T_6	Fresh water outlet temperature of A/E 1 plant	°C
T_7	Fresh water outlet temperature of A/E 2 plant	°C
T_8	Fresh water outlet temperature of A/E 3 plant	°C
T_9	Fresh water outlet temperature of main L.O. cooler	°C
T_{10}	L.T.F.W outlet temperature of H.T.F.W cooler	°C
T_{11}	M/E jacket water outlet temperature	°C
T_{12}	H.T.F.W outlet temperature of H.T.F.W cooler	°C

relationship with the following system parameters: the fluid pressure monitoring points P_1, P_2, P_3; the flow monitoring points Q_1, Q_2, Q_3, Q_4, Q_5, Q_6, Q_7, Q_8, Q_9, Q_{10}, Q_{11}; the temperature monitoring points T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8, T_9, T_{10}, T_{11}, T_{12}. The acquisition points of all parameters are shown in Fig. 9, and the description of each parameter is shown in Table 5.

2) DATA ACQUISITION AND PREPROCESSING
According to the 5 system states, after abnormal data filtering via 3σ rule, there are 500 samples acquired through DAS left, where the data trends according to 5 states is shown in Fig. 10. The static feature statistics is shown in Table 6. By data splitting, the number of samples for each state is 100, including 70 (70%) training samples and 30 (30%) testing samples, and due to the different dimensions of the state data, used (1) to further achieve data normalization.

3) IABC INITIALIZATION
Set $N_I = 100$, $N_R = 50$, $N = 40$, and the optimization range of C, σ is 0–100. Generate employed bees and onlookers, and adopt the unlearning improvement strategy to generate 40 initial nectar sources as shown in Fig. 11.

4) SVM INITIALIZATION
Set the K-folding parameter $k = 6$, adopt OAO classifier, and select the Gaussian function as the kernel function.

5) MODEL TRAINING AND PARAMETERS OPTIMIZING
The initial nectar sources selected are shown in Fig.11 D. The trends of the classification accuracy between IABC-SVM and ABC-SVM are shown in Fig. 12, and more details are shown in Table 7.

B. WORKING STATE IDENTIFICATION OF CW
1) STATE DEFINITION AND PARAMETER SELECTION
The working states of the CW were be summarized as S1 to S5: mooring state with A/E 1 running (S1), mooring state with A/E 2 running (S2), mooring state with A/E 3 running (S3), low load at sea with a single A/E running (S4), and high load at sea with two A/Es running (S5). Based on the system working principle and the abstract system layout diagram, it can be derived that the system states have a strong

A. Initial nectar sources produced.

B. Initial inverse nectar sources.

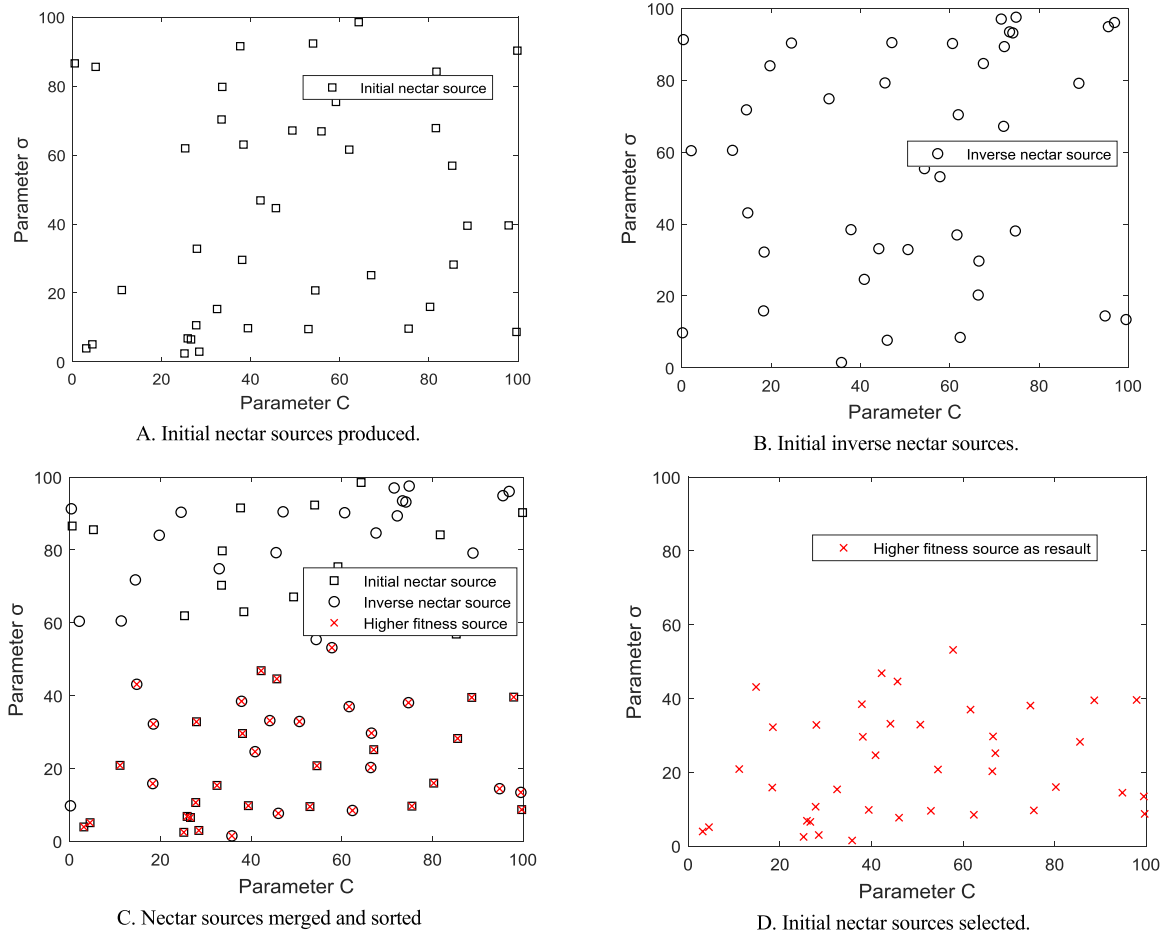C. Nectar sources merged and sorted

D. Initial nectar sources selected.

FIGURE 11. Initial nectar sources producing and selecting.

TABLE 6. Descriptive statistics of the data samples.

	P_1	P_2	P_3	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9	Q_{10}
						Part A							
count	500	500	500	500	500	500	500	500	500	500	500	500	500
mean	1.739	3.975	3.316	249.43	211.57	23.57	29.2	483.49	27.1	15.18	13.43	309.19	8.18
std	0.117	0.15	1.201	229.37	232.18	50.05	55.56	72.05	11.17	15.56	14.86	15.31	11.31
min	1.566	3.814	2.353	0	0	0	0	254.96	0.0005	0.0005	0	286.5	0
25%	1.652	3.889	2.353	0	0	0	0	428.6	28.24	0.0013	0.0013	301.25	0
50%	1.704	3.929	2.353	431.83	0	0.0005	0	501.26	30.49	0.002	0.002	309.07	0
75%	1.768	4.013	4.42	453.84	449.13	5.43	26.26	549.1	33.43	30.49	28.95	322.39	18.39
max	2.02	4.305	5.106	513.15	527.18	236.95	197.64	653.16	35.91	37.24	41.14	344.44	47.36

	Q_{11}	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}
							Part B						
count	500	500	500	500	500	500	500	500	500	500	500	500	500
mean	88.29	25.63	30.66	25.1	26.61	35.43	38.95	36.96	35.61	35.71	36.9	51.39	37.56
std	108.96	1.69	6.34	0.52	1.95	2.06	4.66	3.29	1.31	2.2	3.37	18.98	2.87
min	0	23.71	23.71	24.76	24.82	29.9	31.18	31.37	31.18	30.78	32.05	36	33.15
25%	0	24.06	24.76	24.82	24.99	34.36	35.26	34.99	35.21	34.35	34.55	36	35.66
50%	0	24.96	27.69	24.82	25.66	35.25	36.21	36	36	35.26	34.89	36	36
75%	205.53	27.35	37.95	25.17	27.98	36.07	44.1	36.88	36	36.91	40.27	74.49	40.26
max	272.95	30.46	40.33	29.54	31.79	41.32	47.77	48.28	47.88	42.05	44.74	80.77	44.74

6) CLASSIFICATION PERFORMANCE TESTING

The obtained optimal parameters C, σ were used to construct the SVM model. The testing (30%) sub-dataset was used to evaluate the generalization ability of the model. The results of model testing are shown in Fig. 13, and more details are shown in Table 8.
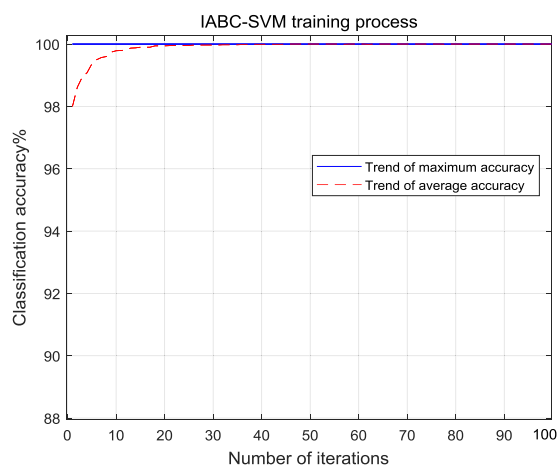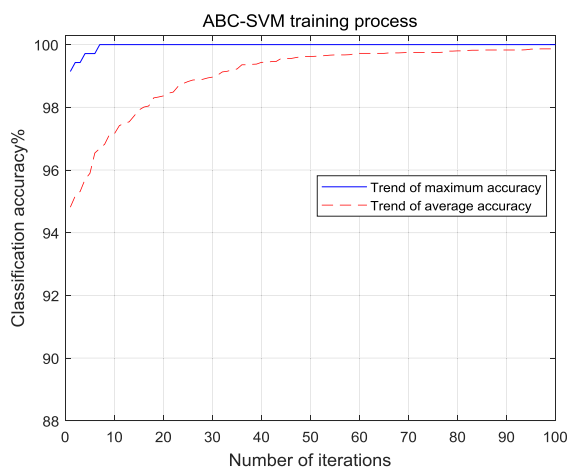
FIGURE 12. Parameter optimization process.

TABLE 7. Comparison of training performance.

	ABC-SVM	IABC-SVM
Training time	100.37s	75.28s
Initial mean accuracy	94.81%	98.01%
The best mean accuracy	99.86%	100%
Cross validation accuracy	99.14%	100%
Best C	7.11	35.76
Best σ	2.22	1.49

TABLE 8. Comparison of testing performance.

	ABC-SVM	IABC-SVM
Optimal parameters	(7.11, 2.22)	(35.76, 1.49)
S1 estimation/label	30/30	30/30
S2 estimation/label	21/30	26/30
S3 estimation/label	28/30	30/30
S4 estimation/label	30/30	30/30
S5 estimation/label	30/30	30/30
Accuracy	92.67% (139/150)	97.33% (146/150)

C. RESULTS

From the two examples, it can be seen from Fig. 7 (example 1) and Fig. 12 (example 2) that the IABC-SVM can reach the global optimization at a faster speed than the ABC-SVM
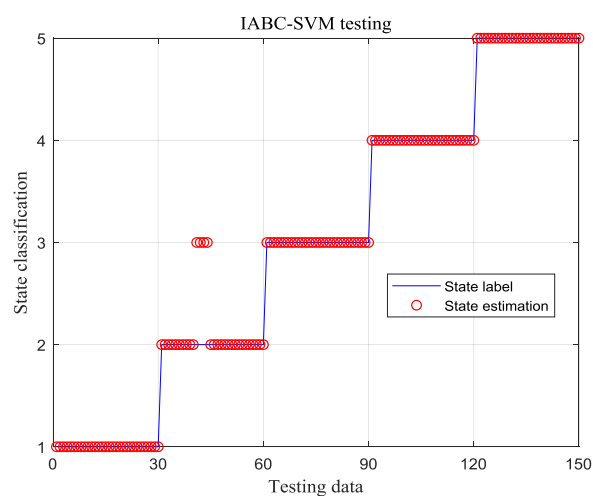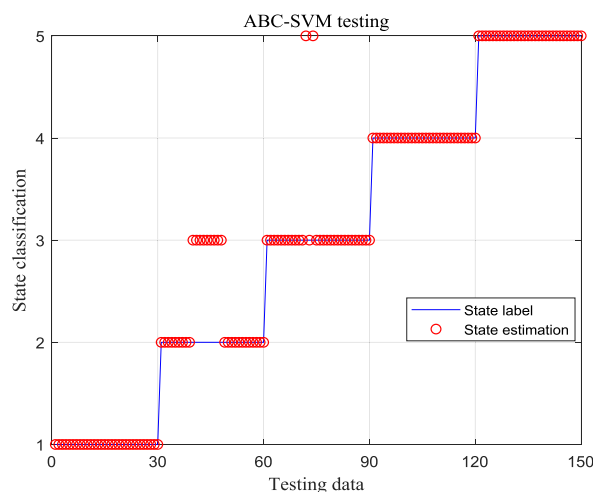

FIGURE 13. Multi-class SVM state estimation results in testing.

in the model training process. In addition, it can also be seen from the trend of their average accuracy that the IABC method can help the training start with better initial nectar sources with 96.52% (example 1) and 98.01% (example 2) average accuracy than the traditional ABC with 91.02% (example 1) and 94.81% (example 2). More comparative parameters such as the training time, the best mean accuracy, etc. are shown in Table 3 (example 1) and Table 7 (example 2). Form the comparative information, we can conclude that the performance of the IABC-SVM is better than the ABC-SVM.

From the results of model testing in Fig. 8 (example 1) and Fig. 13 (example 2), The IABC-SVM can get a slightly higher classification accuracy rate than the ABC-SVM, although the performance of the two methods tends to be the same. More evidence about the testing performance are shown in Table 4 (example 1) and Table 8 (example 2). Therefore, it can be derived that the IABC-SVM can show a slight advantage in the testing process compared to the ABC-SVM.

According to the analysis above, compared with traditional ABC optimized SVM, the IABC optimized SVM can significantly improve the model training efficiency, quickly obtain

the global optimal parameters, and obtain a slightly better state estimation performance in the process of model testing.

IV. CONCLUSION

The elaborated method can achieve the parameters optimization for the multi-class SVM model and improve the model training efficiency. Because this method can make full use of the training dataset to find the global best parameters in a SVM model, and avoid the uncertainty and local optimization, it can perform a multi-states estimation for some marine engine room systems, and get a good performance, which can be verified in the given examples.

There is still some improvement space for the proposed method: more different ship systems, especially some special systems, will be tested to verify the classification accuracy and the generalization performance; for some complex ship systems or combined large-scale ship systems, if the number of system parameters, which have a strong relationship with each system state, is too large and exceeds the capacity of the model, an effective data dimension reduction method can be further applied to the data preprocessing stage. In addition, an effective state prediction model can be further introduced in the method to achieve the performance of state prediction.

With the maintenance mode upgrading of large commercial ships, the method can further perform a large-scale system state evaluation by combining multiple marine engine room systems, and even form a unified function unit to conduct intelligent state evaluation and prediction for the entire mechanical and electrical system of large commercial ships.

REFERENCES

[1] I. Lazakis, O. Turan, and S. Aksu, "Increasing ship operational reliability through the implementation of a holistic maintenance management strategy," *Ships Offshore Struct.*, vol. 5, no. 4, pp. 337–357, Oct. 2010.

[2] I. Lazakis, C. Gkerekos, and G. Theotokatos, "Investigating an SVM-driven, one-class approach to estimating ship systems condition," *Ships Offshore Struct.*, vol. 14, no. 5, pp. 432–441, Jul. 2018.

[3] D. T. Hountalas, "Prediction of marine diesel engine performance under fault conditions," *Appl. Thermal Eng.*, vol. 20, no. 18, pp. 1753–1783, Dec. 2000.

[4] G. Chandroth, "Condition monitoring: The case for integrating data from independent sources," *J. Mar. Eng. Technol.*, vol. 3, no. 1, pp. 9–16, Jan. 2004.

[5] D. Watzenig, M. S. Sommer, and G. Steiner, "Engine state monitoring and fault diagnosis of large marine diesel engines," *e i Elektrotechnik Und Informationstechnik*, vol. 126, no. 5, pp. 173–179, May 2009.

[6] V. T. Lamaris and D. T. Hountalas, "A general purpose diagnostic technique for marine diesel engines—Application on the main propulsion and auxiliary diesel units of a marine vessel," *Energy Convers. Manage.*, vol. 51, no. 4, pp. 740–753, Apr. 2010.

[7] P. Li, L. Liu, and H. Gong, "The research of the intelligent fault diagnosis optimized by ACA for marine diesel engine," in *Advancing Computing, Communication, Control and Management*, vol. 56, 2010, pp. 174–181.

[8] K. K. Dikis and I. Lazakis, "Dynamic risk and reliability assessment of ship machinery and equipment," in *Proc. 26th Int. Ocean Polar Eng. Conf., Rhodes, Greece*, vol. 2016, pp. 969–976.

[9] C. Gkerekos, I. Lazakis, and G. Theotokatos, "Ship machinery condition monitoring using vibration data through supervised learning," in *Proc. Int. Conf. Maritime Saf. Oper.*, 2016, pp. 103–110.

[10] C. Gkerekos, I. Lazakis, and G. Theotokatos, "Ship machinery condition monitoring using performance data through supervised learning," in *Proc. Smart Ships Tech. Conf.*, London, U.K., 2017, pp. 105–111.

[11] J. Kowalski, B. Krawczyk, and M. Wozniak, "Fault diagnosis of marine 4-stroke diesel engines using a one-vs-one extreme learning ensemble," *Eng. Appl. Artif. Intell.*, vol. 57, pp. 134–141, Jan. 2017.

[12] S. Wang, B. Ji, J. Zhao, W. Liu, and T. Xu, "Predicting ship fuel consumption based on LASSO regression," *Transp. Res. D, Transp. Environ.*, vol. 65, pp. 817–824, Dec. 2018.

[13] A. Coraddu, L. Oneto, A. Ghio, S. Savio, D. Anguita, and M. Figari, "Machine learning approaches for improving condition-based maintenance of naval propulsion plants," *Proc. Inst. Mech. Eng., M, J. Eng. Maritime Environ.*, vol. 230, no. 1, pp. 136–153, Feb. 2016.

[14] F. Cipollini, L. Oneto, A. Coraddu, A. J. Murphy, and D. Anguita, "Condition-based maintenance of naval propulsion systems with supervised data analysis," *Ocean Eng.*, vol. 149, pp. 268–278, Feb. 2018.

[15] F. Cipollini, L. Oneto, A. Coraddu, A. J. Murphy, and D. Anguita, "Condition-based maintenance of naval propulsion systems: Data analysis with minimal feedback," *Rel. Eng. Syst. Saf.*, vol. 177, pp. 12–23, Sep. 2018.

[16] Y. Raptodimos and I. Lazakis, "Using artificial neural network-self-organising map for data clustering of marine engine condition monitoring applications," *Ships Offshore Struct.*, vol. 13, no. 6, pp. 649–656, Mar. 2018.

[17] Y. W. Zhang, X. L. Sun, Y. W. Ding, and P. T. Sun, "Design of intelligent diagnosis system for ship power equipment," *Chin. J. Ship Res.*, vol. 13, no. 6, pp. 140–146, Dec. 2018.

[18] M. Cheliotis, C. Gkerekos, I. Lazakis, and G. Theotokatos, "A novel data condition and performance hybrid imputation method for energy efficient operations of marine systems," *Ocean Eng.*, vol. 188, Sep. 2019, Art. no. 106220.

[19] W. B. Zhang, W. F. Shi, Y. Lan, S. Y. Gu, and J. B. Zhuo, "Realtime power quality evaluation system of the electric propulsion ship based on AHP-fuzzy comprehensive evaluation method," *Chin. J. Ship Res.*, vol. 14, pp. 48–87, Dec. 2019.

[20] G. Q. Liu, Y. J. Lin, Z. Z. Zhang, and S. Pang, "Main marine engine fault diagnosis method based on rough set theory and optimized DAG-SVM," *Chin. J. Ship Res.*, vol. 15, pp. 69–73, Feb. 2020.

[21] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2560–2574, Aug. 2007.

[22] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. 14th Int. Joint Conf. Artif. Intell.*, Montreal, QC, Canada, 1995, pp. 1137–1143.

[23] L. Zhang, B. Chao, and X. Zhang, "Modeling and optimization of microbial lipid fermentation from cellulosic ethanol wastewater by rhodotorula glutinis based on the support vector machine," *Bioresource Technol.*, vol. 301, Apr. 2020, Art. no. 122781.

[24] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Dept. Comput. Eng., Erciyes Univ., Eng. Faculty, Kayseri, Turkey, Tech. Rep. TR-06, Oct. 2005.

[25] S. Jiang, C. Lu, S. Zhang, X. Lu, S.-B. Tsai, C.-K. Wang, Y. Gao, Y. Shi, and C.-H. Lee, "Prediction of ecological pressure on resource-based cities based on an RBF neural network optimized by an improved ABC algorithm," *IEEE Access*, vol. 7, pp. 47423–47436, 2019.

[26] X. Pan, Q. Zhang, and H. Pan, "Improved artificial bee colony algorithm and its application to fundus retinal blood vessel image binarization," *IEEE Access*, vol. 8, pp. 123726–123734, 2020.

[27] Z. Wang, H. Ding, B. Li, L. Bao, and Z. Yang, "An energy efficient routing protocol based on improved artificial bee colony algorithm for wireless sensor networks," *IEEE Access*, vol. 8, pp. 133577–133596, 2020.

[28] J. Pian, G. Wang, and B. Li, "An improved ABC algorithm based on initial population and neighborhood search," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 251–256, 2018.

[29] C. Zhao, H. Zhao, G. Wang, and H. Chen, "Improvement SVM classification performance of hyperspectral image using chaotic sequences in artificial bee colony," *IEEE Access*, vol. 8, pp. 73947–73956, 2020.

[30] H. Shah, N. Tairan, H. Grag, and R. Ghazali, "Global gbest guided-artificial bee colony algorithm for numerical function optimization," *Computers*, vol. 7, no. 4, pp. 1–17, Dec. 2018.

[31] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[32] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annu. Workshop Comput. Learn. Theory COLT*, 1992, pp. 144–152.

[33] Z. Qiu, Y. L. Qian, Y. Zhang, and X. Zhang, "Gas turbine fault diagnosis based on improved support vector machine," *J. Eng. Thermal Energy Power*, vol. 33, no. 9, pp. 39–43, Sep. 2018.

[34] F. F. Chamasemani and Y. P. Singh, "Multi-class support vector machine (SVM) classifiers—An application in hypothyroid detection and classification," in *Proc. 6th Int. Conf. Bio-Inspired Comput., Theories Appl.*, Sep. 2011, pp. 351–356.

[35] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 415–425, Mar. 2002.

[36] N. J. Xue, "Comparison of multi-class support vector machines," *Comput. Eng. Des.*, vol. 32, no. 5, pp. 1792–1795, 2011.

[37] Z. Yang, Q. Zhou, X. Wu, and Z. Zhao, "A novel measuring method of interfacial tension of transformer oil combined PSO optimized SVM and multi frequency ultrasonic technology," *IEEE Access*, vol. 7, pp. 182624–182631, 2019.

XU CAO received the B.Sc. degree in naval architecture and ocean engineering from Dalian Ocean University, Dalian, China, in 2010, and the M.Sc. degree in naval architecture and ocean engineering from Dalian Maritime University, Dalian, in 2019. He is currently an Assistant Lecturer with the Dalian Shipping Vocational and Technical College. His research interests include marine engineering automation and control, ship hydrodynamic, ship building, and ship structure.

HUI CAO received the B.Sc. and D.Sc. degrees in marine engineering from Dalian Maritime University, Dalian, China, in 2003 and 2008, respectively. He is currently an Associate Professor with Dalian Maritime University. His current research interests include marine engineering automation and control, intelligence evaluation algorithm, marine engine room simulation, application of computer and networks, artificial intelligence, and smart ship.

RAN LI received the B.Sc. degree in industrial design from the Jilin Normal University of Engineering and Technology, Changchun, China, in 2012. She is currently a Lecturer with the Dalian Shipping Vocational and Technical College. Her research interests include marine engineering automation and control, intelligent ship fault diagnosis, and ship structure and marine engineering.

JUNDONG ZHANG received the B.Sc., M.Sc., and D.Sc. degrees in marine engineering from Dalian Maritime University, Dalian, China, in 1989, 1992, and 1998, respectively. He is currently a Full Professor with Dalian Maritime University. His research interests include marine engineering automation and control, integrated supervision, application of computer and networks, marine engineering education, and electrical system design.

YIRU WANG received the B.Sc. degree in naval architecture and ship engineering from Chongqing Jiaotong University, Chongqing, China, in 2011. She is currently a Lecturer with the Dalian Shipping Vocational and Technical College. Her research interests include marine engineering automation and control, ship hydrodynamic, new energy vessels, and marine engineering education.

• • •