

AIS-based operational phase identification using Progressive Ablation Feature Selection with machine learning for improving ship emission estimates

Kuiquan Duan, Qingbo Li, Shangheng Liu, Yanxin Liu, Shuang Wang, Shuang Li, Xiaochuan Wang, Nan Ma & Ye Ma

To cite this article: Kuiquan Duan, Qingbo Li, Shangheng Liu, Yanxin Liu, Shuang Wang, Shuang Li, Xiaochuan Wang, Nan Ma & Ye Ma (2024) AIS-based operational phase identification using Progressive Ablation Feature Selection with machine learning for improving ship emission estimates, *Journal of the Air & Waste Management Association*, 74:2, 100-115, DOI: [10.1080/10962247.2023.2274348](https://doi.org/10.1080/10962247.2023.2274348)

To link to this article: <https://doi.org/10.1080/10962247.2023.2274348>



[View supplementary material](#)



Published online: 22 Jan 2024.



[Submit your article to this journal](#)



Article views: 809



[View related articles](#)



[View Crossmark data](#)



[Citing articles: 6](#) [View citing articles](#)

TECHNICAL PAPER



AIS-based operational phase identification using Progressive Ablation Feature Selection with machine learning for improving ship emission estimates

Kuiquan Duan^a, Qingbo Li^{ID,a}, Shangheng Liu^a, Yanxin Liu^a, Shuang Wang^a, Shuang Li^a, Xiaochuan Wang^b, Nan Ma^c, and Ye Ma^a

^aGreen Shipping and Carbon Neutrality Laboratory, College of Environmental Science and Engineering, Dalian Maritime University, Dalian, People's Republic of China; ^bDanger Prevention Division, Qinhuangdao Maritime Safety Administration of People's Republic of China, Qinhuangdao, People's Republic of China; ^cTransportation Safety Research Center, China Academy of Transportation Sciences, Beijing, People's Republic of China

ABSTRACT

The work status of ships' engines and boilers has a significant impact on emission estimates, which are closely related to ships' operational phases. To improve the accuracy of emission estimates, this study proposed a machine learning-based classification model for identifying operational phases. We proposed 12 operational phase relevance features by analyzing motion behavior-related and geospatial characteristics-related features from the Automatic Identification System (AIS) data from the two bulk carriers. The random forest (RF) model showed the best performance in identifying one of the bulk carrier's operational phases among the five machine models, with the accuracy, F_1 score and Area Under Curve (AUC) of 96.66%, 93.34% and 99.93%, respectively. By adopting the Progressive Ablation Feature Selection (PAFS) method with RF, the number of features was reduced from 12 to 8, and the accuracy (96.38%), F_1 score (92.70%), and AUC (98.81%) were almost same with that obtained from all 12 features. Additionally, the effectiveness of the RF model was validated on the other bulk carriers. Compared with the traditional algorithms, the RF model showed better performance in ship operational phase identification and improved the average accuracy of NOx emission estimation for the main engine and auxiliary engine by 57.83% and 93.89%, respectively, under different operational phases. These results provide the basis for port traffic management and ship emission control.

Implications: A new ship operational phase identification approach was proposed in this study. If the proposed approach is adopted by International Maritime Organization, it will improve the accuracy of ship emission estimates and bring new insights into global shipping greenhouse gas (GHG) emissions and their impact on global change. The port authorities could benefit from the proposed approach, which can be extended to ship types with similar behavior to bulk carriers, such as containers and general cargoes. This can reveal patterns of ship behavior in specific areas, which helps to identify potential collision risks, channel blockages, and other safety issues and take appropriate management measures to ensure the safe operation of the port. The proposed approach can help shipping companies to accurately estimate the GHG emissions of their fleets and to accurately predict carbon tax costs. Base on that, carbon emissions and carbon tax burden can be reduced by adopting corresponding management control measures.

PAPER HISTORY

Received June 17, 2023
Revised September 28, 2023
Accepted October 5, 2023

Introduction

The exhaust emissions from ships can cause pollution and contribute to global warming, which can affect air quality in port areas (Ramacher et al. 2019; Wang et al. 2021). According to recent studies, global shipping emissions were responsible for 3% of human-caused carbon dioxide emissions in 2017 (IMO 2021), while it accounted for a higher proportion of the reactive gas NOx of about 20% (McDuffie et al. 2020). The main engines, auxiliary engines, and boilers equipped in ships are the main sources of exhaust emissions, accounting for 67%, 22%, and 10% of global shipping CO₂

emissions, respectively (IMO 2021). The work status of these devices determines the ships' exhaust emissions, but collecting the work status of these devices across the fleet is costly, inefficient, and challenging to achieve. It is worth noting that the work status of these devices varies with the ships' operational phases, such as cruise, maneuvering, anchor, and berth, which significantly affect emission estimates. Therefore, it is necessary to establish accurate operational phase identification methods for improving ship emission estimates.

There are different behavior patterns when ships load and unload cargo at port, wait to enter port at anchorage

or cruise at sea, which results in the Automatic Identification System (AIS) data showing different characteristics (Guo et al. 2020). The current research used ship movement information provided by AIS data to identify operational phases and estimate emissions for each phase. Studies have adopted ship speed to assign AIS trajectories to cruise, slow cruise, maneuvering, and hoteling phases (Chen et al. 2016; Ng et al. 2013; Wan et al. 2019, 2020). In addition to ship speed, some studies also considered the main engine load when identifying maneuvering and cruise phase (Chen et al. 2017; Fu et al. 2017; IMO 2015; Li et al. 2023). The International Maritime Organization (IMO) and the United States Environmental Protection Agency (USEPA) identify ships' operational phases by comprehensively considering the ship speed, main engine load, and location (IMO 2021; USEPA 2020). However, these methods use thresholds and may not be able to identify the deceleration and acceleration processes of ships before and after stopping as maneuvering phases (Zhang et al. 2022). As a result, estimated emissions during maneuvering phases may be inaccurate. Additionally, how these traditional methods affect ship emission estimates is unknown.

Machine learning involves using labeled data for supervised classification and regression and unlabeled data for unsupervised clustering (Kumar, Amgoth, and Annavarapu 2019). Since AIS data contains very rich maritime traffic information, using machine learning approaches to process and analyze massive ship trajectory data is the appropriate option. By adopting ship motion-related features from AIS data, many studies have used the machine learning approach for ship collision risk assessment (Gao and Shi 2020), ship types classification (Sheng et al. 2018) and behavior pattern mining (Dogancay, Tu, and Ibal 2021). Similarly, the machine learning model can effectively classify operational phases by learning ships' motion characteristics from AIS trajectories. For example, the Convolutional Neural Network-Ship Movement Modes Classification (CNN-SMMC) model was proposed to identify AIS trajectories as static, normal navigation, and maneuvering phases by considering the change rate of ships' heading and speed (Chen et al. 2020). Besides, the Back Propagation Neural Networks (BPNN) model was applied to distinguish the berth and anchor phases in AIS data using the variations of distance and course (Huang et al. 2018). The operational phases can be affected by various factors. However, the above researches only considered the limited factors when identifying the operational phases. To overcome these limitations, it is necessary to involve the analysis of domain knowledge in the machine learning process to understand the ships' motion behavior and geospatial

characteristics under different operational phases. This allows for the identification of a comprehensive set of features from AIS data that can effectively capture operational phase characteristics. More importantly, the operational phase data used for the machine learning in the above researches is defined by the ship motion characteristics, which may be different from the actual operational phases. Therefore, to accurately identify ships' operational phases, it is necessary to further mine the features from the AIS data labeled with the actual recorded operational phases.

Feature selection is an important data preprocessing technique that helps identify significant predictors and eliminate unnecessary features, playing a vital role in the machine learning process (Kalousis, Prados, and Hilario 2007). There are four categories of feature selection algorithms: filter, wrapper, embedded, and hybrid (Hoque, Bhattacharyya, and Kalita 2014). Filters evaluate the relationships between features and the target without involving any machine learning methods (Homayouni, Liu, and Thieu 2022). For example, to improve the accuracy and efficiency of model classification, researchers have explored teaching-learning-based optimization (TLBO) and Heat Transfer Search (HTS) algorithms to select the optimal feature subsets. Subsequently, the chosen feature subsets were input into machine learning models (Suthar et al. 2023; Vakharia et al. 2023). Wrapper and embedded methods require learning methods to conduct feature selection. The main drawback of wrapper methods is the number of computations required to obtain the feature subset and the possibility of overfitting (Kohavi and John 1997). The embedded method incorporates feature selection as part of the training process to reduce computation time and improve generalization performance (Chandrashekhar and Sahin 2014). In terms of the embedded method, the study employed the LASSO regression algorithm to implement feature selection in predicting ship fuel consumption, which helps to address the problem of multiple collinearities caused by highly correlated features (Wang et al. 2018). Besides, a previous study applied Recursive Feature Elimination (RFE) using a Decision Tree-based model to determine the optimum number of features for analyzing the effect of fouling on shaft power (Laurie et al. 2021). However, embedded methods include feature selection as part of the training process, which means that once the model is trained, the selected features cannot be changed. An ablation study is a type of hybrid method that sheds light on model interpretation by removing portions of input to determine their significance. A method called "Progressive Ablation Feature Selection (PAFS)" with XGBoost was proposed to select

features, which is a comprehensive approach to select the most representative features for modeling. The PAFS outperformed traditional feature selection methods, which is attributed to its superiority in correlation analysis and feature ablation based on feature importance (Homayouni, Liu, and Thieu 2022). The method provides broader applications in other fields where feature selection is necessary.

This study developed a machine learning-based classification approach to automatically identify operational phases from the ships' AIS trajectories. The main objectives of this study were (i) to propose new multi-constraint features by analyzing ships' motion behavior and geospatial characteristics under recorded operational phases, (ii) to minimize the number of features by adopting the PAFS method in the machine learning model and (iii) to compare the developed the machine learning model with the traditional algorithms in terms of improvements in ship operational phases identification and emission estimation.

Materials and methods

Figure 1 illustrates the approach proposed in this study. The AIS data from one of the research ships (BC1) was split into a training set to initially develop the operational phase identification models, and a testing set to optimize the models by using techniques like hyperparameter tuning, performance evaluation, and feature selection. All of the above processes were evaluated using ten-fold cross-validation. Then, all data from BC1 were adopted as the training set, while the data from the other research ship (BC2) was adopted as a validation set to validate the model's generalization capability on new unseen ship routes and its improvement compared to traditional algorithms. Additionally,

a portion of the validation set with NOx emission measurement data was used to evaluate the improvement of the proposed approach on ship emission estimates.

Ship and data description

Bulk carriers

The study was performed on two bulk carriers operating offshore China. One of them is a Panamax-class bulk carrier (BC1), built in 1999. The other is a Handymax-class bulk carrier (BC2), built in 2014. The detailed specifications of the ships and their engines are provided in Table S1.

AIS dataset

In this study, AIS datasets from the two ships were collected to develop operational phase classification models and estimate NOx emissions. The collected AIS datasets for the BC1 covered the period from December 31, 2020, to November 9, 2021, including 15 voyages offshore China and 64,860 data points and that for the BC2 covered the period from November 25, 2020, to December 31, 2021, including 29 voyages offshore China and 110,274 data points (Figure 2).

Operational phase record

All the collected AIS trajectories have been manually labeled with specified operational phases in advance, which serve as the basis for machine learning models for classification. According to the navigation report of the ship, the complete voyage can be generally divided into six segments: (i) maneuvering to accelerate out of the starting port, (ii) normal cruising, (iii) maneuvering to decelerate into the anchorage, (iv) anchoring, (v) maneuvering into the destination

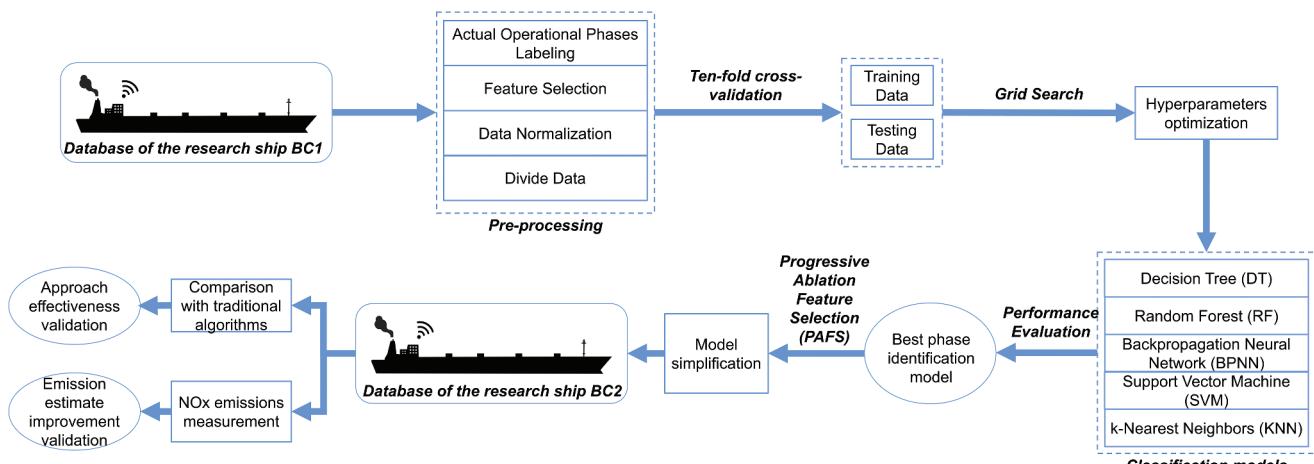


Figure 1. Flowchart of the machine learning procedure.

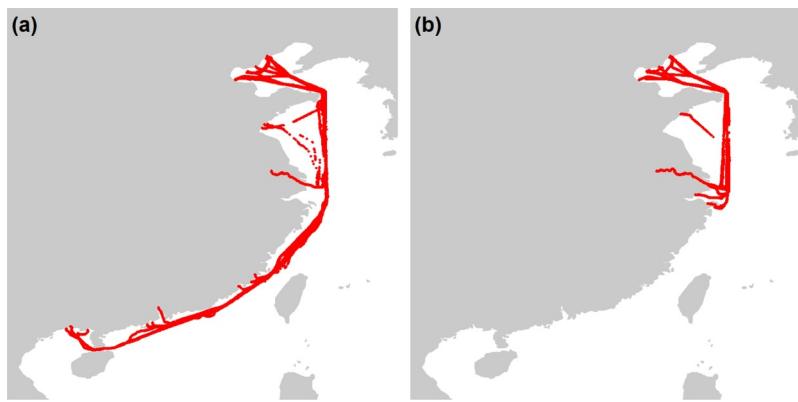


Figure 2. Trajectories of the (a) BC1 and (b) BC2, respectively.

port, and (vi) berthing. Therefore, Segments (i), (iii), and (v) are labeled as the maneuvering phase, Segment (ii) is labeled as cruise phases, Segment (iv) is labeled as anchor phases and Segment (vi) is labeled as berth phase. The spatial distribution of the AIS trajectory of the ship after labeling the four operational phases is obtained (Figure 3). Finally, the operational phases of the ship's AIS trajectory points at the corresponding time were labeled according to the operational phases' information recorded in the navigation report.

NOx emission measurement

This study utilized a continuous emission monitoring system to measure the NOx concentration and

exhaust flow from the BC2. The measurements have been carried out on the main engine and auxiliary engine of the BC2 during different operational phases such as cruise, maneuvering, anchor, and berth phases. Measurements for the main engine were conducted between November 25, 2020, and March 5, 2021, covering four voyages. Measurements for the auxiliary engine were conducted between June 12, 2021, and August 7, 2021, covering 10 voyages. The detailed voyage information during exhaust measurement is represented in Table S2. The measured NOx emissions from the ships' engines at a specific time interval are calculated by Eq. (1). More detailed information on the measurement process is provided in Duan et al. (2022).

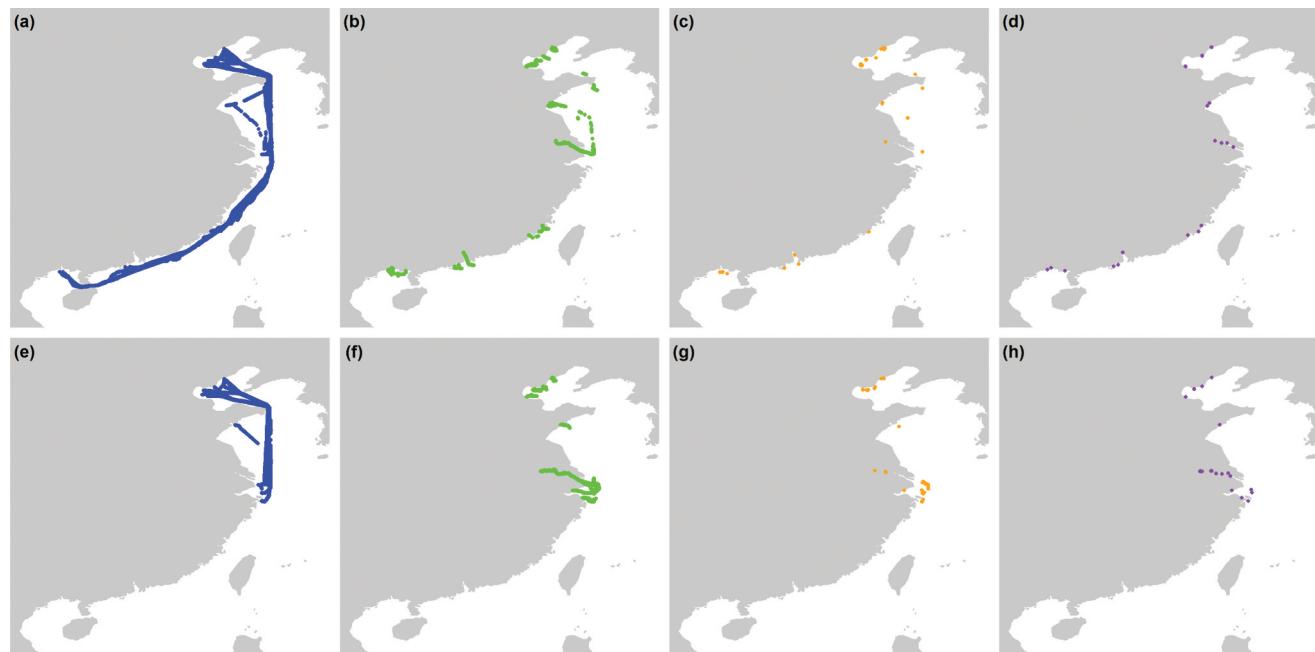


Figure 3. The spatial distribution of the AIS trajectory points for BC1 ((a)–(d)) and BC2 ((e)–(h)) during cruise ((a) and (e)), maneuvering ((b) and (f)), anchor ((c) and (g)), and berth ((d) and (h)) phases.

$$E_{NO_x} = \sum_{i=1}^n C_{NO_x,i} \times Q_i \times \Delta T \times 10^{-6} \quad (1)$$

where E_{NO_x} is the NOx emissions from an engine during a specific period (kg), $C_{NO_x,i}$ is the instantaneous NOx concentration (mg/m^3), Q_i is the instantaneous exhaust flow (m^3/h), ΔT is the time interval of the measurement data (h).

Feature engineering

In this study, a comprehensive feature set is utilized to capture the motion and geospatial characteristics of different operational phases based on domain knowledge analysis. Among them, the first four features were obtained directly from the ship's AIS data, and the last eight features were further mined from motion behavior and geospatial characteristics from the AIS data:

Location

The latitude and longitude in AIS data provide location information of ships, which can be selected as a classification feature to identify ships' operational phases (Chen et al. 2021). The longitude and latitude in the AIS trajectories were considered as identifiers (format: longitude – latitude), which were then converted to the corresponding hash values using a hash function and input to the machine learning model as location features.

Speed over ground (SOG)

During cruise phases, ships typically operate at a speed of 94% of their maximum speed, which is known as the service speed. However, when ships are maneuvering in ports, they tend to maintain a slower speed (USEPA 2009). Besides, it is common for ships to experience drifting during anchor phases (Guo et al. 2020), while ships would be fixed with the cable and thus the SOG is almost zero during berth phases.

Heading

When a ship is cruising, it typically navigates in a straight line with a constant heading. However, during the maneuvering phase, the ship's heading may need to be adjusted. Therefore, heading can be used to analyze differences in ship behavior (Xiao et al. 2015).

Course

The course information provided by AIS data depicts the motion state of ships, it provides an effective way to detect ship behaviors (Zhang et al. 2022).

SOG variance (SV)

The cruise phase refers to the ship sailing along a predetermined route without changing the course or speed, while the maneuvering phase means that the ship needs to adopt variable speed or change its course to avoid collision (Chen et al. 2020). Therefore, fluctuations in the ship's SOG are relatively stable during cruise phases but are more frequent during maneuvering phases. To quantify this, the SV was selected as one classification feature.

SOG difference (SD)

Drifting is a common phenomenon for ships during anchor phases, but during berthing, they are kept stationary in a fixed position (Guo et al. 2020). To reflect the equilibrium degree of the ship's SOG distribution in each phase, the SD between adjacent AIS trajectory points was calculated.

Distance difference (DD)

Due to the higher SOG during cruise phases, the distance between adjacent two AIS trajectory points is greater than during the maneuvering phase. On the other hand, considering that the distribution of AIS trajectory points during berth phases is more concentrated than during anchor phases (Guo et al. 2020), DD is selected as a classification feature.

Heading difference variance (HDV)

When a ship is berthing, its heading typically remains constant or fluctuates within a small range. However, due to frequent variations in the ship's heading during anchor phases, the heading difference between two adjacent AIS trajectory points is relatively large (Guo et al. 2020). During cruise phases, ships follow a predetermined route without changing their heading. To measure this, the HDV was calculated to show how much the ship's heading varies.

Distance to coast (DC)

As berths are located near the coast, ships will be closer to the coast during berth phases. In contrast, ships are typically farther away from the coast during cruise phases. Therefore, the proximity of a ship to the coast or port can provide valuable information about its current operational phase (IMO 2021). Geographic Information System (GIS) technology can be used to calculate the distance between ship AIS trajectory points and the coastline.

Time interval (TI)

According to the International Telecommunication Union (ITU), the transmission frequency of AIS data

varies from 2 s to 3 min, depending on the navigation status of the ship (ITU 2010). Thus, the time interval between two adjacent AIS trajectories of a ship can be chosen as a classification feature.

The change rate of course (CRC)

The change rate of course (ΔCOG) between two adjacent AIS trajectories (p_i, p_{i+1}) can be used to determine the operational phase (Chen et al. 2020). The Eq. (2) is used to calculate the change rate of course for each AIS point:

$$\Delta COG = \frac{Cog_{i+1} - Cog_i}{\Delta t_i} \quad (2)$$

where Cog_i and Cog_{i+1} are the ship's course at p_i and p_{i+1} , respectively, °; Δt_i is the time interval between p_i and p_{i+1} , s.

The change rate of SOG (CRS)

The operational phase can be determined by calculating the change rate of SOG (ΔSOG) between two adjacent AIS trajectories (p_i, p_{i+1}) (Chen et al. 2020). The change rate of SOG for each AIS point can be calculated by Eq. (3).

$$\Delta SOG = \frac{Sog_{i+1} - Sog_i}{\Delta t_i} \quad (3)$$

where Sog_i and Sog_{i+1} are the ship's SOG at p_i and p_{i+1} , respectively, kn.

Feature selection and processing

Data normalization

The operational phase classification features were normalized between 0 and 1 using min-max scaling in Eq. (4).

$$x_m = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (4)$$

where x_m is the dimensionless normalized values of the actual variable x , x_{min} and x_{max} are the minimum and maximum values of the variable x , respectively.

Machine learning models

The identification ability of ship operational phases was evaluated using the Decision Tree (DT), Random Forest (RF), Backpropagation Neural Network (BPNN), Support Vector Classifier (SVC), and k-Nearest Neighbors (KNN) models. The implementation of the machine learning models was done using the scikit-learn python library. GridSearchCV of scikit-learn python library was used to optimize the hyperparameters for the models, in which the best set of hyperparameters were searched by ten-fold cross-validation accuracy over a high-dimensional parameter grid (Table 1). All the experiments were conducted on a 12th Gen Intel(R) Core(TM) i5-1240 1.70 GHz processor with 16 GB RAM under Windows 11 (64 bits).

DT. DT is a commonly used classification and regression algorithm that utilizes a tree-like structure to describe the classification decisions of real-world examples. The algorithm has advantages such as fast learning speed and high interpretability (Kashifi et al. 2022).

$$\hat{f} = \sum_{b=1}^B \eta f_b(x) \quad (5)$$

where \hat{f} is the ensemble DT model; b is the number of the DT model; B is the quantity of the DT models; η is the learning rate.

RF. RF is an ensemble learning method that utilizes the bagging technique to train multiple decision trees in parallel. By aggregating the prediction results of all trees, it obtains an average output prediction, thereby achieving

Table 1. Hyperparameter optimization for various machine learning models in ten-fold cross-validation accuracy.

Model	Hyperparameters	GridSearchCV parameter grid	Optimal values
DT	max_depth min_samples_split	{None, 10, 20} {2, 5, 10}	10 2
RF	max_depth min_samples_split n_estimators	{None, 10, 20} {2, 5, 10} {50, 100, 200}	None 2 200
BPNN	alpha hidden_layer_sizes	{0.0001, 0.001, 0.01} {10, 20, 30}	0.0001 30
SVM	C kernel	{0.1, 1, 10} {"linear", "rbf"}	10 rbf
KNN	algorithm n_neighbors weights	{"auto", "ball_tree", "kd_tree"} {3, 5, 7} {"uniform", "distance"}	auto 5 distance

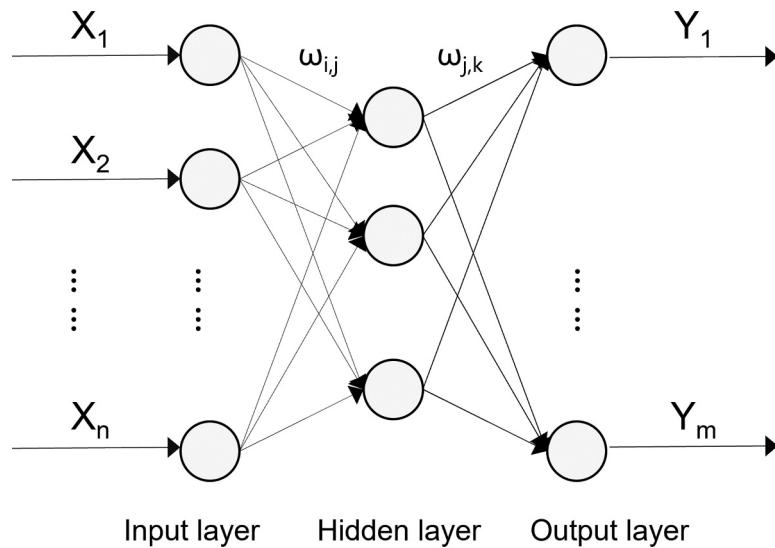


Figure 4. Topology of BPNN.

prediction of class labels or responses with high accuracy and generalization capability (Witten and James 2013).

$$\hat{f} = \frac{1}{B} \sum_{b=1}^B f_b(x) \quad (6)$$

BPNN. BPNN is a multilayer feed-forward neural network that uses forward signal transmission and backward error propagation to approximate continuous derivable functions accurately. The BPNN topology is shown in Figure 4, where the inputs of the BPNN are represented by \$X_1, X_2, \dots, X_n\$; the forecast values of the BPNN are represented by \$Y_1, Y_2, \dots, Y_m\$; and the

weights of the BPNN are represented by \$\omega_{ij}\$ and \$\omega_{jk}\$ (Wang et al. 2015).

SVC. SVC is an effective classification algorithm based on statistical learning theory (Cortes and Vapnik 1995). SVC defines an optimal hyperplane represented by \$\omega \in \mathbf{R}^n\$ as the weight vector and \$b \in \mathbf{R}\$ as the offset, with support vectors being crucial sample points on the boundary lines \$\omega x + b = 1\$ and \$\omega x + b = -1\$ (Figure 5) (Liu, Shi, and Zhu 2022).

KNN. KNN defines the connection between past and future values through similarity relationships

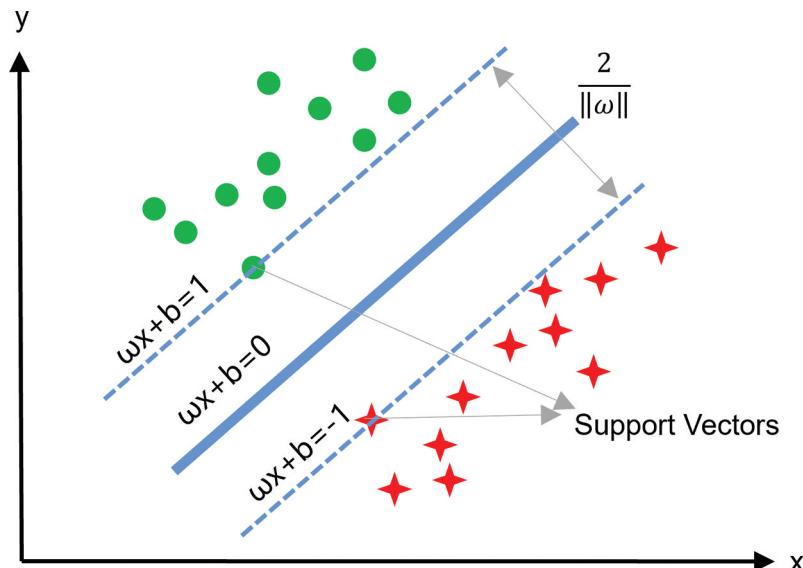


Figure 5. Schematic of SVC.

(Altunkaynak and Kartal 2021). It is a non-parametric methodology that is widely used as a classification algorithm (Suthar et al. 2023). KNN determines the class of a test sample by applying the majority voting principle, i.e., selecting the class most prevalent among the K nearest neighbors.

Feature importance and elimination

Feature importance. The gain represents the relative contribution of each feature to the machine learning models, which is determined by computing the contribution of each feature to each tree in the model (Eq. (7)) (Chen, Guestrin, and Assoc Comp 2016).

$$\text{Gain} = \frac{1}{2} \left[\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} + \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma \quad (7)$$

Feature elimination. The PAFS with machine learning model was used to conduct features in this study. Algorithm PAFS with machine learning model indicates how this algorithm works (Table 2). The detailed steps are described as follows:

- (i) Removing highly correlated pairs of features and training the machine learning model to obtain gain values for each feature base on training set;
- (ii) Iteratively eliminating the most important feature and storing the eliminated feature;
- (iii) Keeping track of the removal iterations until the desired number of features is selected according to the machine learning model performance;
- (iv) Using the machine learning model of these finally selected features for comparison with the results of the full features.

Evaluation metrics

To verify the approach, the developed machine learning model was conducted to compare with the traditional

algorithms using test dataset. The performance of the machine learning model and traditional algorithms can be evaluated by comparing the predicted operational phase labels to the actual operational phase labels. Let i be a class from the dataset of classes C , T be the test dataset, and c be a classification model, such that $c(t) = l$, where t is an element of the test dataset T and $l \in C$ is the label corresponding to the class in C assigned to t by c . Let $g(t)$ be the ground truth class label of t . Regarding the classification model c is defined as follows:

- *True Positives of class i* , represented by TP_i , is the number of elements in T that are correctly labeled as class i by c , i.e., $TP_i = |\{t \in T | c(t) = g(t) = i\}|$.
- *False positives of class i* , represented by FP_i , is the number of elements in T that are incorrectly classified by c as belonging to class i . Formally, $FP_i = |\{t \in T | c(t) = i \wedge g(t) \neq i\}|$.
- *False Negatives of class i* , represented by FN_i , is the number of elements in T that belong to class i but are classified by c with labels different from i , that is, $FN_i = |\{t \in T | c(t) \neq i \wedge g(t) = i\}|$.

The above numbers are used to define the performance of the classification models. These evaluation metrics are Precision, Recall, F₁score, and Area Under Curve (AUC). To avoid the overfitting of the machine learning models, ten-fold cross-validation was adopted in model performance evaluation process (Mohan and Saranya 2019; Suthar et al. 2023).

F₁score

Precision $p(c,i)$ of classification model c with respect to class i is defined in Eq. (8):

$$p(c,i) = \frac{TP_i}{TP_i + FP_i} \quad (8)$$

Informally, precision is the ratio of the number of test elements correctly labeled with class label i by c to the number of all elements labeled (correctly or incorrectly) by c .

Table 2. Algorithm PAFS with machine learning model.

Sr No.	Descriptions
1	Input: FS = [], $i = 1$, k , $A = (a_1, \dots, a_m)$ denotes set of features, $D = (X, y)$ where $X \in \mathbb{R}^{n \times m}$ and $y \in \mathbb{R}^n$.
2	Conduct a correlation analysis (in each highly correlated pair ($p \geq 0.8$), remove the feature that appears more frequently in all highly correlated pairs), and then update A .
3	Split D into training set (X_{train}, y_{train}) and test set (X_{test}, y_{test}).
4	$s = A $
5	Train machine learning model on X_{train} to obtain the Gain values of features, $G = \{g_1, \dots, g_s\}$.
6	Remove the argmax _A (G) feature from A and append that feature to FS.
7	if $i \leq k - 1$:
8	Go to step 4.
9	return FS

The recall of classification model c with respect to class i , represented by $r(c,i)$, is defined by Eq. (9):

$$r(c,i) = \frac{TP_i}{TP_i + FN_i} \quad (9)$$

Therefore, the recall is the ratio of the number of test elements belonging to class i that are correctly labeled by c to the total number of test elements of class i .

The F_1 score is the harmonic mean of precision and recall with equal weights, and represents the relations between data's positive labels and those given by a model, as defined by Eq. (10). The F_1 score is between "0" and "1", and the higher the value, the better the performance of the classification model.

$$F_1(c,i) = \frac{2p(c,i)r(c,i)}{p(c,i) + r(c,i)} \quad (10)$$

AUC

AUC represents the classification models' ability to avoid false classification. Usually, the AUC should be between "0.5" and "1", and the higher AUC of a classification model indicates that the model's predictions can be matched perfectly with the actual labels (Zhou et al. 2019).

Accuracy

When comparing the effectiveness of classification models generated from different learning methods, a global measure of effectiveness is usually used (Ribeiro et al. 2021). The global measure aims to resume the effectiveness of the classification model for all classes in the test dataset. This study used the following global measure to compare the performance of the classification model: Accuracy, Mac F_1 score, and MacAUC. The accuracy of classification model c is the fraction of test elements that are correctly labeled as class i by c , as defined by Eq. (11):

$$\text{Accuracy} = \frac{\sum_{i=1}^{|C|} TP_i}{\sum_{i=1}^{|C|} (TP_i + FP_i)} \quad (11)$$

Mac F_1 score

The Macro measures (MacPrec, MacRec, and Mac F_1 score) are the average of the corresponding measures (Precision, Recall, and F_1 score) for all classes, as defined by Eqs. (12)–(14).

$$\text{MacPrec}(c) = \frac{\sum_{i=1}^{|C|} p(c,i)}{|C|} \quad (12)$$

$$\text{MacRec}(c) = \frac{\sum_{i=1}^{|C|} r(c,i)}{|C|} \quad (13)$$

$$\text{Mac}F_1\text{score}(c) = \frac{\sum_{i=1}^{|C|} F_1(c,i)}{|C|} \quad (14)$$

MacAUC

Since AUC can only be used for binary classification, AUC was calculated for all four binary problems that distinguished between one operational phase class and other operational phase classes and then took the macro average (MacAUC) (Wu et al. 2018).

Ship emission estimates

To verify the improvement of the approach for ship emission estimates, this study used the machine learning model and traditional algorithms to identify the operational phases of the test dataset and conduct emission estimates.

Main engine

Main engines normally work at cruise and maneuvering phases, and the NOx emissions from the ship's main engine can be estimated as follows:

$$E_{ME, NOx} = E_{ME, Cruise} + E_{ME, Man.} \quad (15)$$

where $E_{ME, NOx}$ is the NOx emissions from the main engine (t); $E_{ME, Cruise}$ and $E_{ME, Man.}$ are the NOx emissions from main engine during cruise and maneuvering phases (t).

$E_{ME, Cruise}$ and $E_{ME, Man.}$ can be estimated as follows:

$$E_{ME,j} = P_{ME} \times LF \times EF_{ME, NOx} \times LLAF_{NOx} \times T_{ME,j} \times 10^{-6} \quad (16)$$

where $E_{ME,j}$ is the NOx emissions from the main engine during operational phase j (t); P_{ME} is the installed power of main engines (kW); LF is the load factor of the main engine; $EF_{ME, NOx}$ is the NOx emission factors of main engine (g/kW·h); $LLAF_{NOx}$ is the low-load adjustment factor for $EF_{ME, NOx}$; $T_{ME,j}$ is the working hours of the main engine during operational phase j (h).

LF can be estimated as follows:

$$LF = \left(\frac{V_{sog}}{V_{max}} \right)^3 \times DAF \times HFF \quad (17)$$

where V_{sog} is the speed over ground (knots); V_{max} is the maximum ship speed (knots); DAF is the draught adjustment factor; HFF is the hull fouling factor.

The detailed equations of DAF and HFF can be retrieved from the International Council on Clean Transportation study (ICCT 2017).

Auxiliary engine

Auxiliary engines work in all phases, and the NOx emissions from the ship's auxiliary engine can be estimated as follows:

$$E_{AE, NOx} = E_{AE, Cruise} + E_{AE, Man.} + E_{AE, Anchor} + E_{AE, Berth} \quad (18)$$

where $E_{AE, NOx}$ is the NOx emissions from the auxiliary engine (t); $E_{AE, Cruise}$, $E_{AE, Man.}$, $E_{AE, Anchor}$, and $E_{AE, Berth}$ are the NOx emissions from the auxiliary engine during cruise, maneuvering, anchor, and berth phases, respectively (t).

$E_{AE, Cruise}$, $E_{AE, Man.}$, $E_{AE, Anchor}$, and $E_{AE, Berth}$ can be estimated as follows:

$$E_{AE,j} = P_{AE,j} \times EF_{AE, NOx} \times T_{AE,j} \times 10^{-6} \quad (19)$$

where $E_{AE,j}$ is the NOx emissions from the auxiliary engine during operational phase j (t); $P_{AE,j}$ is the load power of the auxiliary engine during operational phase j (kW); $EF_{AE, NOx}$ is the NOx emission factor of the auxiliary engine (g/kW·h); $T_{AE,j}$ is the working hours of the auxiliary engine during operational phase j (h).

Results and discussion

Model comparison

The linear relationship between each feature in the dataset is indicated by the Pearson correlation coefficient (Figure S1). There were no highly correlated pairs for the BC1 ($\rho < 0.8$), and all the features could be used for modeling without causing multicollinearity problems. The best set of hyperparameters were searched by ten-fold cross-validation accuracy over a high-dimensional parameter grid (Table 1). Based on that, the model performance in operational phase identification for the BC1 in ten-fold cross-validation accuracy is presented in (Table 3). The results indicated that the two tree-based models showed a better performance on identifying operational phases when compared to BPNN, SVC, and KNN. Among the tree-based models, the RF model showed the best performance, with the macro values of accuracy, F₁score, and AUC of 96.66%, 93.34%, and 99.93%, respectively. This result is consistent with a previous study that involved a comparative analysis

Table 3. Macro values of evaluation metrics for the BC1 in ten-fold cross-validation.

Model	Accuracy (%)	F ₁ score (%)	AUC (%)
DT	95.46 ± 1.32	89.84 ± 3.59	98.39 ± 0.16
RF	96.66 ± 1.35	93.34 ± 2.86	99.93 ± 0.02
BPNN	94.17 ± 1.53	84.90 ± 5.22	98.28 ± 0.22
SVC	93.33 ± 1.61	81.87 ± 5.77	96.45 ± 0.30
KNN	92.07 ± 1.59	76.11 ± 5.35	96.12 ± 0.37

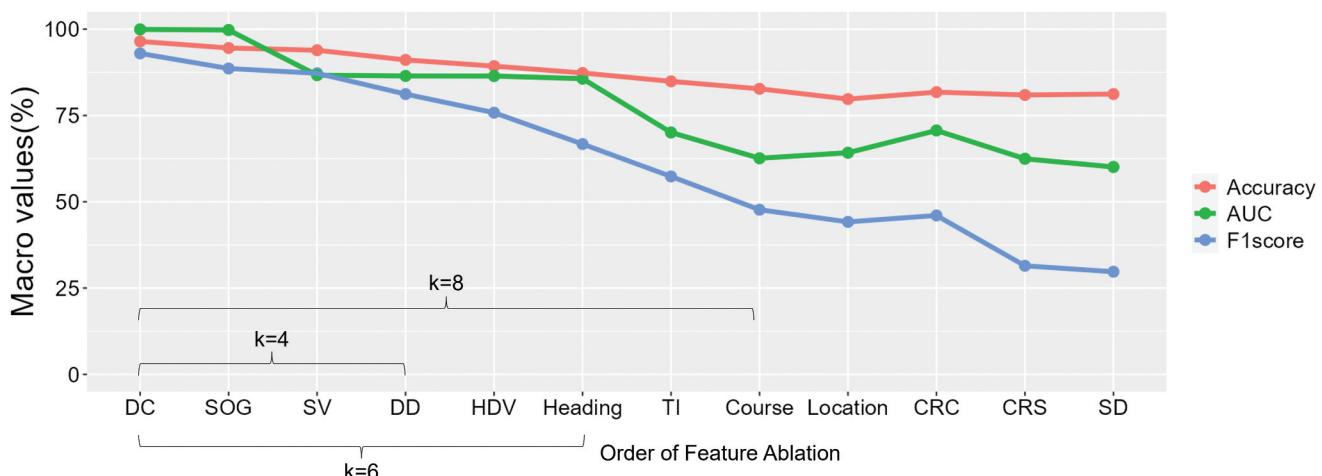


Figure 6. The evaluation metrics of each RF model using three threshold setups ($k = 4$, $k = 6$, $k = 8$) with respect to selected features at each step for the BC1.

of 179 different classifiers across 121 datasets (Fernandez-Delgado et al. 2014).

Considering that RF performed the best in this application, feature simplification will be conducted to enhance interpretability, generalization, reduce computation, and suit operational phase identification applications.

Progressive Ablation Feature Selection (PAFS) method with RF

Using the entire set of features (12 features) from the BC1, the RF model achieved an Accuracy, F₁score, and AUC of 96.45%, 92.97%, and 99.93%, respectively, and the DC was the feature with the highest gain value of 0.22 (Figure 6). After removing the DC, the RF model was trained with the remaining feature set. SOG showed to be the most important among the current feature set, with a gain value of 0.36. The removal process was repeated until the desired number of features is obtained. In this study, three different values for k (4, 6, 8) were used to study the impacts of k .

The RF model with eight features selected using the PAFS method was also close to the entire feature dataset for the BC1. The macro values of accuracy, F₁score, and AUC decreased by 0.07%, 0.29%, and 1.12%, respectively (Table 4). Therefore, when only eight significant features were included, the performance of the RF model was almost the same as that obtained from all 12 features. Additionally, the running time of the RF model was decreased by 25%. This allowed to create a more user-friendly operational phase identification model without sacrificing too much accuracy and provided an insight into each selected feature and its associations with ships' operational phases. Thus, a small set of available features can be used to identify ships' operational phases.

Comparison with traditional algorithms

In this section, a validation set from the BC2 was adopted to validate the model's generalization capability on new unseen ship routes and its improvement compared to traditional algorithms. In detail, this study evaluated the performance of traditional ship

Table 5. The evaluation results of the BC2's operational phase classification.

		Operational phase	Accuracy (%)	F ₁ score (%)	AUC (%)
RF	Cruise	91.65	91.72	92.12	
	Maneuvering	90.17	86.12	88.02	
	Anchor	95.09	72.88	82.99	
	Berth	96.48	65.59	84.55	
	Macro-avg.	93.35	79.08	86.92	
IMO3	Cruise	70.11	75.70	71.97	
	Maneuvering	68.86	33.79	60.09	
	Anchor	89.97	4.01	50.78	
	Berth	89.82	47.89	94.53	
	Macro-avg.	79.69	40.35	69.34	
IMO4	Cruise	63.75	72.02	66.03	
	Maneuvering	62.55	7.40	51.92	
	Anchor	95.37	80.71	97.35	
	Berth	96.61	43.55	63.92	
	Macro-avg.	79.57	50.92	69.80	
USEPA	Cruise	70.12	75.70	71.98	
	Maneuvering	69.23	35.41	60.62	
	Anchor	95.19	67.19	75.35	
	Berth	99.35	93.14	96.52	
	Unidentified	None	None	None	
	Macro-avg.	None	None	None	

operational phase classification algorithms presented in the Third IMO Greenhouse Gas Study (IMO3) (IMO 2015), the Fourth IMO Greenhouse Gas Study (IMO4) (IMO 2021), and the USEPA ship emission inventory study (USEPA 2020). Both IMO4 and USEPA algorithms consider geospatial data, SOG, and main engine load to identify ships' operational phases, while IMO3 only takes into account the latter two factors (Table S3). Various metrics were used to evaluate the performance of the developed RF model and traditional algorithms (Table 5). The MacAUC values of the RF model were found to be greater than 80%, indicating its good predictive ability and better performance compared to random class assignments (Zhou et al. 2019). The macro-average values of Accuracy F₁score and AUC for the RF model were 93.35%, 79.08%, and 86.92%, respectively, which showed better performance than traditional algorithms. It should be noted that the F₁score and AUC values for the anchor phase (72.88% and 82.99%) and the berth phase (65.59% and 84.55%) were lower on the validation ship BC2. This may be due to the imbalance of operational phases in the training data of BC1. In detail, the reporting frequency of ship AIS

Table 4. Model performance of the PAFS method for the BC1.

Number of features	Model	List of selected features	Accuracy (%)	F ₁ score (%)	AUC (%)	Running time (s)
12	RF	All features	96.45	92.97	99.93	49.62
8	PAFS with RF (Selection of 8 features)	DC, SOG, SV, DD, HDV, TI, Heading, Course	96.38	92.70	98.81	37.25
6	PAFS with RF (Selection of 6 features)	DC, SOG, SV, DD, HDV, TI	95.28	91.32	97.20	44.90
4	PAFS with RF (Selection of 4 features)	DC, SOG, SV, DD	94.87	89.40	96.78	42.01

(a)		Predicted classification					(b)		Predicted classification				
Actual classification	Classes	Cruise	Man.	Anchor	Berth	Total	Actual classification	Classes	Cruise	Man.	Anchor	Berth	Total
	Cruise	7294	64	0	0	7358		Cruise	7340	18	0	0	7358
	Man.	1253	4810	76	3	6142		Man.	4694	1253	85	110	6142
	Anchor	0	150	1040	340	1530		Anchor	0	3	33	1494	1530
	Berth	0	4	208	529	741		Berth	1	1	0	739	741
	Total	8547	5028	1324	872	15771		Total	12035	1275	118	2343	15771

(c)		Predicted classification					(d)		Predicted classification					
Actual classification	Classes	Cruise	Man.	Anchor	Berth	Total	Actual classification	Classes	Cruise	Man.	Anchor	Berth	/	Total
	Cruise	7358	0	0	0	7258		Cruise	7340	18	0	0	0	7358
	Man.	5711	236	195	0	6142		Man.	4693	1330	4	53	62	6142
	Anchor	3	0	1527	0	1530		Anchor	0	21	776	0	733	1530
	Berth	2	0	532	207	741		Berth	1	1	0	692	47	741
	Total	13074	236	2254	207	15771		/	0	0	0	0	0	0

Figure 7. Confusion matrix with (a) RF, (b) IMO3, (c) IMO4, and (d) USEPA for the BC2's operational phase identification.

data under anchor and berth phases is low (ITU 2010), resulting in a lack of sufficient training data for these phases, making it challenging for the model to capture the features and patterns associated with the two phases on BC2. This implies more cross-ship training data would improve generalization.

The F_1 score of the IMO3 algorithm was lower in identifying anchor phases (4.01%). This was because the IMO3 algorithm defined the SOG during the anchor phases as 1–3 knots, while most AIS trajectories during the anchor phases had SOG below 1 knot. As a result, it was found that a majority of anchor phases were identified as berth phases (Figure 7(b)). It is worth noting that the IMO4 algorithm had a relatively low F_1 score (7.40%) in identifying maneuvering phases, as most of these phases were identified as cruise phases (Figure 7(c)). Similarly, the USEPA algorithm had a lower F_1 score in identifying maneuvering and anchor phases. This was due to the algorithm's limitation of main engine load during maneuvering to below 20%, which caused some AIS trajectories with maneuvering phases being identified as cruise phases. On the other hand, the USEPA algorithm used GIS technology to define relevant areas within the port as polygons and identified corresponding operational phases by limiting SOG and main engine load in different areas. However, this method led to some AIS trajectories during anchor and berth phases not being located within the corresponding areas. For

Table 6. Comparison of operational phase identification and NOx emission estimation during main engine measurements for BC2.

	Phase	Actual value				
		RF	IMO3	IMO4	USEPA	
The number of AIS trajectory points	Cruise	5134	4979	5607	5864	5607
	Maneuvering	749	904	272	15	276
	Anchor	- ^a	-	-	-	-
	Berth	-	-	4	4	-
Estimated NOx emissions (t)	Total	5883	5883	5883	5883	5883
	Cruise	3.36	3.20	3.56	3.70	3.56
	Maneuvering	0.39	0.50	0.14	0.01	0.14
	Anchor	-	-	-	-	-
	Berth	-	-	-	-	-
	Total	3.74	3.70	3.70	3.70	3.70

Notes. ^a Not involved.

instance, a ship may anchor outside the area marked as an anchorage in the electronic chart.

Emission estimates

Main engine NOx emissions

Estimated NOx emissions from the main engine by the RF model were compared to actual NOx emissions (Table 6), and the relative errors of the cruise and maneuvering phase were 4.67% and 29.98%, respectively. By adopting the operational phase identification algorithms recommended by the IMO3, IMO4, and USEPA, the relative errors of cruise phases were 6.05%, 10.11%, and 6.05%, respectively, and that for maneuvering phases were 63.14%, 98.25%, and 62.92%, respectively. Thus, the RF model showed better

Table 7. Comparison of operational phase identification and NOx emission estimation during auxiliary engine measurements for BC2.

	Phase	Actual value	RF	IMO3	IMO4	USEPA
NOx emissions (t)	Cruise	164	289	4763	6483	4747
	Maneuvering	7183	7201	2251	531	2384
	Anchor	1728	3773	137	12804	1352
	Berth	10823	8635	12747	80	10992
	Total	19898	19898	19898	19898	19475
	Cruise	0.02	0.04	0.60	0.82	0.60
	Maneuvering	0.49	0.49	0.15	0.04	0.16
	Anchor	0.19	0.42	0.02	1.41	0.15
	Berth	1.12	0.90	1.32	0.01	1.14
	Total	1.82	1.84	2.09	2.28	2.05

performance than the traditional algorithms for NOx emission estimates in different operational phases. The total estimated NOx emissions for the RF model, IMO3 algorithm, IMO4 algorithm, and USEPA algorithm were similar, with the relative errors of 1.09%, 1.11%, 1.11%, and 1.09%, respectively. This indicates that the total estimated NOx emissions from the main engine were not determined by the results of the operational phase identification.

Auxiliary engine NOx emissions

Estimated NOx emissions from the auxiliary engine by the RF model were compared to actual NOx emissions (Table 7). Compared to traditional algorithms, the RF model improved the average accuracy of NOx emission estimation for the auxiliary engine by 93.89%, under different operational phases. This is due to the fact that there is superior performance of the RF model (Figure 7(a)). As mentioned before, since the relatively poor performance of the RF model in identifying anchor and berth phases in BC2, the NOx emissions for these two phases were overestimated by 0.23 tons and underestimated by 0.23 tons, respectively. By adopting the operational phase identification algorithms recommended by the IMO3, IMO4 and USEPA, the estimated NOx emissions during the cruise phases were 29, 40 and 29 times higher than the actual phase, respectively. However, NOx emissions during maneuvering phases were underestimated with relative errors of 68.66%, 92.61%, and 66.81%, respectively. This is because the three traditional algorithms identified most of the maneuvering phases as cruise phases (Figure 7(b)–(d)). The relative errors of the IMO3, IMO4, and USEPA algorithms for NOx emission estimation during the anchor phase were 92.07%, 640.96%, and 21.76%, respectively, and that for berth phases were 17.78%, 99.26%, and 1.56%, respectively. This is because the IMO3 algorithm identified most of the anchor phase as the berth phase (Figure 7(b)), the NOx emissions during the anchor phases were significantly underestimated. In contrast, the IMO4 algorithm identified most of the berth phases

as the anchor phases (Figure 7(c)), and the NOx emissions during the anchor phases were significantly overestimated. Besides, since the USEPA algorithm failed to identify a part of anchor phases (Figure 7(d)), the NOx emissions during anchor phases were underestimated. The total estimated NOx emissions from auxiliary engines of the RF model were close to the measured NOx emissions, with a relative error of 0.88%. Since the IMO3, IMO4, and USEPA algorithm significantly overestimated the NOx emissions in cruise phases, which led to an overestimation of the total NOx emissions with a relative error of 14.83%, 25.02%, and 12.59%, respectively. This is because the demand power of the measured auxiliary engine during cruise phases (300 kW) is higher than that of other phases (200 kW–250 kW).

Conclusion

The key novelty of this study is the development of a supervised machine learning approach using a comprehensive set of AIS-derived motion and geospatial features to accurately identify ships' operational phases. In contrast to prior threshold-based methods and unsupervised machine learning applications, this study trained models on manually verified operational phase data and validated the best performance model on multiple ship routes, demonstrating the generalizability of the approach. This achieves significantly higher accuracy than traditional algorithms, as evidenced by the comparison of the IMO and EPA algorithms. The study goes beyond modeling to link operational phase classification to emission estimation, demonstrating the practical utility of the approach for maritime management and air quality protection. Overall, this study makes a significant contribution to the research and practice of AIS-based maritime emission modeling through a novel machine learning framework, extensive feature engineering, and linking operational phase prediction to ship emission estimates. Future studies need to focus on the following areas: (i) testing the RF model with more AIS datasets to validate the scalability of the

model and its potential to handle complex maritime traffic scenarios; (ii) extending the model to identify different maneuvering situations, such as crossing, head-on, overtaking and low-visibility navigation situations; (iii) collecting recorded anchor and berth phase data from multiple ships to better train the model to generalize across ships.

Acknowledgment

The authors gratefully acknowledge the support received from all employees and crew of Taihang Shipping Co., Ltd.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China [No. 42277410] and the Fundamental Research Funds for the Central Universities [No. 3132023520].

About the authors

Kuiquan Duan is a PhD candidate in the College of Environmental Science and Engineering, Dalian Maritime University, Dalian, People's Republic of China.

Qingbo Li is a professor in the College of Environmental Science and Engineering, Dalian Maritime University, Dalian, People's Republic of China.

Shangheng Liu is a master candidate in the College of Environmental Science and Engineering, Dalian Maritime University, Dalian, People's Republic of China.

Yanxin Liu is a PhD candidate in the College of Environmental Science and Engineering, Dalian Maritime University, Dalian, People's Republic of China.

Shuang Wang is a PhD candidate in the College of Environmental Science and Engineering, Dalian Maritime University, Dalian, People's Republic of China.

Shuang Li is a master candidate in the College of Environmental Science and Engineering, Dalian Maritime University, Dalian, People's Republic of China.

Xiaochuan Wang is a leader of the Danger Prevention Division, Qinhuangdao Maritime Safety Administration, Qinhuangdao, People's Republic of China.

Nan Ma is a staff of Transportation Safety Research Center, China Academy of Transportation Sciences, Beijing, People's Republic of China.

Ye Ma is an associate professor in the College of Environmental Science and Engineering, Dalian Maritime University, Dalian, People's Republic of China.

ORCID

Qingbo Li  <http://orcid.org/0000-0002-1735-1013>

Data availability statement

The data that support the findings of this study are available from the corresponding author Qingbo Li (drlqb@dlnu.edu.cn) upon reasonable request.

References

- Altunkaynak, A., and E. Kartal. 2021. Transfer sea level learning in the Bosphorus Strait by wavelet based machine learning methods. *Ocean Eng.* 233:109116. doi:[10.1016/j.oceaneng.2021.109116](https://doi.org/10.1016/j.oceaneng.2021.109116).
- Chandrashekhar, G., and F. Sahin. 2014. A survey on feature selection methods. *Comput. Electr. Eng.* 40 (1):16–28. doi:[10.1016/j.compeleceng.2013.11.024](https://doi.org/10.1016/j.compeleceng.2013.11.024).
- Chen, D., X. Wang, Y. Li, J. Lang, Y. Zhou, X. Guo, and Y. Zhao. 2017. High-spatiotemporal-resolution ship emission inventory of China based on AIS data in 2014. *Sci. Total Environ.* 609:776–87. doi:[10.1016/j.scitotenv.2017.07.051](https://doi.org/10.1016/j.scitotenv.2017.07.051).
- Chen, D., Y. Zhao, P. Nelson, Y. Li, X. Wang, Y. Zhou, J. Lang, and X. Guo. 2016. Estimating ship emissions based on AIS data for port of Tianjin, China. *Atmos. Environ.* 145:10–18. doi:[10.1016/j.atmosenv.2016.08.086](https://doi.org/10.1016/j.atmosenv.2016.08.086).
- Chen, S., Q. Meng, P. Jia, and H. Kuang. 2021. An operational-mode-based method for estimating ship emissions in port waters. *Transp. Res. Part D Transp. Environ.* 101:103080. doi:[10.1016/j.trd.2021.103080](https://doi.org/10.1016/j.trd.2021.103080).
- Chen, T., C. Guestrin, and Assoc Comp, M. 2016. Xgboost: A scalable tree boosting system. In *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 785–94. San Francisco, CA.
- Chen, X., Y. Liu, K. Achuthan, and X. Zhang. 2020. A ship movement classification based on Automatic Identification System (AIS) data using Convolutional Neural Network. *Ocean Eng.* 218:108182. doi:[10.1016/j.oceaneng.2020.108182](https://doi.org/10.1016/j.oceaneng.2020.108182).
- Cortes, C., and V. Vapnik. 1995. Support-vector networks. *Mach Learn.* 20 (3):273–97. doi:[10.1007/BF00994018](https://doi.org/10.1007/BF00994018).
- Dogancay, K., Z.M. Tu, and G. Ibal. 2021. Research into vessel behaviour pattern recognition in the maritime domain: Past, present and future. *Digit. Signal Process.* 119:103191. doi:[10.1016/j.dsp.2021.103191](https://doi.org/10.1016/j.dsp.2021.103191).
- Duan, K., Q. Li, H. Liu, F. Zhou, J. Lu, X. Zhang, Z. Zhang, S. Wang, Y. Ma, and X. Wang. 2022. Dynamic NOx emission factors for main engines of bulk carriers. *Transp. Res. Part D Transp. Environ.* 107:103290. doi:[10.1016/j.trd.2022.103290](https://doi.org/10.1016/j.trd.2022.103290).
- Fernandez-Delgado, M., E. Cernadas, S. Barro, and D. Amorim. 2014. Do we need hundreds of classifiers to solve real world classification problems? *J. Mach. Learn. Res.* 15 (1):3133–81.
- Fu, M., H. Liu, X. Jin, and K. He. 2017. National- to port-level inventories of shipping emissions in China. *Environ. Res. Lett.* 12 (11):114024. doi:[10.1088/1748-9326/aa897a](https://doi.org/10.1088/1748-9326/aa897a).
- Gao, M., and G.-Y. Shi. 2020. Ship collision avoidance anthropomorphic decision-making for structured learning based

- on AIS with Seq-CGAN. *Ocean Eng.* 217:110479. doi:[10.1016/j.oceaneng.2020.107922](https://doi.org/10.1016/j.oceaneng.2020.107922).
- Guo, W., Z. Zhao, Z. Zheng, and Y. Xu. **2020**. A cloud-based approach for ship stay behavior classification using massive trajectory data. In *2020 International Conference on Service Science (ICSS)*, Xining, China, 82–89. IEEE.
- Homayouni, A., T. Liu, and T. Thieu. **2022**. Diabetic retinopathy prediction using Progressive Ablation Feature Selection: A comprehensive classifier evaluation. *Smart Health* 26:100343. doi:[10.1016/j.smhl.2022.100343](https://doi.org/10.1016/j.smhl.2022.100343).
- Hoque, N., D.K. Bhattacharyya, and J.K. Kalita. **2014**. MIFS-ND: A mutual information-based feature selection method. *Expert Syst. Appl.* 41 (14):6371–85. doi:[10.1016/j.eswa.2014.04.019](https://doi.org/10.1016/j.eswa.2014.04.019).
- Huang, L., Y. Wen, X. Geng, C. Zhou, and C. Xiao. **2018**. Integrating multi-source maritime information to estimate ship exhaust emissions under wind, wave and current conditions. *Transp. Res. D. Transp. Environ.* 59:148–59. doi:[10.1016/j.trd.2017.12.012](https://doi.org/10.1016/j.trd.2017.12.012).
- ICCT. **2017**. *Greenhouse gas emissions from global shipping, 2013–2015 detailed methodology*. Washington, DC: International Council on Clean Transportation (ICCT).
- IMO. **2015**. *Third IMO Greenhouse Gas Study 2014: Safe, secure and efficient shipping on clean ocean*. London: International Maritime Organization (IMO).
- IMO. **2021**. *Fourth IMO Greenhouse Gas Study 2020: Safe, secure and efficient shipping on clean ocean*. London: International Maritime Organization (IMO).
- ITU. **2010**. *Technical Classifications of Recommendations ITU-R. 1371-5. Technical characteristics for an automatic identification system using time division multiple access in the VHF maritime mobile frequency band*. Geneva, Switzerland: International Telecommunication Union (ITU).
- Kalousis, A., J. Prados, and M. Hilario. **2007**. Stability of feature selection algorithms: A study on high-dimensional spaces. *Knowl. Inf. Syst.* 12 (1):95–116. doi:[10.1007/s10115-006-0040-8](https://doi.org/10.1007/s10115-006-0040-8).
- Kashifi, M. T., A. Jamal, M.S. Kashefi, M. Almoshaqeh, and S.M. Rahman. **2022**. Predicting the travel mode choice with interpretable machine learning techniques: A comparative study. *Travel Behav. Soc.* 29:279–96. doi:[10.1016/j.tbs.2022.07.003](https://doi.org/10.1016/j.tbs.2022.07.003).
- Kohavi, R., and G.H. John. **1997**. Wrappers for feature subset selection. *Artif. Intell.* 97 (1–2):273–324. doi:[10.1016/s0004-3702\(97\)00043-x](https://doi.org/10.1016/s0004-3702(97)00043-x).
- Kumar, D. P., T. Amgoth, and C.S.R. Annavarapu. **2019**. Machine learning algorithms for wireless sensor networks: A survey. *Inform. Fusion* 49:1–25. doi:[10.1016/j.inffus.2018.09.013](https://doi.org/10.1016/j.inffus.2018.09.013).
- Laurie, A., E. Anderlini, J. Dietz, and G. Thomas. **2021**. Machine learning for shaft power prediction and analysis of fouling related performance deterioration. *Ocean Eng.* 234:108886. doi:[10.1016/j.oceaneng.2021.108886](https://doi.org/10.1016/j.oceaneng.2021.108886).
- Li, H., P. Jia, X. Wang, Z. Yang, J. Wang, and H. Kuang. **2023**. Ship carbon dioxide emission estimation in coastal domestic emission control areas using high spatial-temporal resolution data: A China case. *Ocean Coast Manag.* 232:106419. doi:[10.1016/j.ocecoaman.2022.106419](https://doi.org/10.1016/j.ocecoaman.2022.106419).
- Liu, J., G.-Y. Shi, and K.-G. Zhu. **2022**. A novel ship collision risk evaluation algorithm based on the maximum interval of two ship domains and the violation degree of two ship domains. *Ocean Eng.* 255:111431. doi:[10.1016/j.oceaneng.2022.111431](https://doi.org/10.1016/j.oceaneng.2022.111431).
- McDuffie, E. E., S.J. Smith, P. O'Rourke, K. Tibrewal, C. Venkataraman, E.A. Marais, B. Zheng, M. Crippa, M. Brauer, and R.V. Martin. **2020**. A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017): An application of the Community Emissions Data System (CEDS). *Earth Syst. Sci. Data* 12 (4):3413–42. doi:[10.5194/essd-12-3413-2020](https://doi.org/10.5194/essd-12-3413-2020).
- Mohan, S., and P. Saranya. **2019**. A novel bagging ensemble approach for predicting summertime ground-level ozone concentration. *J. Air Waste Manag. Assoc.* 69 (2):220–33. doi:[10.1080/10962247.2018.1534701](https://doi.org/10.1080/10962247.2018.1534701).
- Ng, S. K. W., C. Loh, C. Lin, V. Booth, J.W.M. Chan, A.C. K. Yip, Y. Li, and A.K.H. Lau. **2013**. Policy change driven by an AIS-assisted marine emission inventory in Hong Kong and the Pearl River Delta. *Atmos. Environ.* 76:102–12. doi:[10.1016/j.atmosenv.2012.07.070](https://doi.org/10.1016/j.atmosenv.2012.07.070).
- Ramacher, M. O. P., M. Karl, J. Bieser, J.-P. Jalkanen, and L. Johansson. **2019**. Urban population exposure to NO_x emissions from local shipping in three Baltic Sea harbour cities - a generic approach. *Atmos. Chem. Phys.* 19 (14):9153–79. doi:[10.5194/acp-19-9153-2019](https://doi.org/10.5194/acp-19-9153-2019).
- Ribeiro, A. P., N.F.F. da Silva, F.N. Mesquita, P.D.C.S. Araujo, T.C. Rosa, J.N. Mesquita-Neto, N.L. Komarova, D. Stowell, N.L. Komarova, D. Stowell, et al. **2021**. Machine learning approach for automatic recognition of tomato-pollinating bees based on their buzzing-sounds. *PLoS Comput. Biol.* 17 (9):e1009426. doi:[10.1371/journal.pcbi.1009426](https://doi.org/10.1371/journal.pcbi.1009426).
- Sheng, K., Z. Liu, D. Zhou, A. He, and C. Feng. **2018**. Research on ship classification based on trajectory features. *J. Navig.* 71 (1):100–16. doi:[10.1017/s0373463317000546](https://doi.org/10.1017/s0373463317000546).
- Suthar, V., V. Vakharia, V.K.K. Patel, and M. Shah. **2023**. Detection of compound faults in ball bearings using multiscale-SinGAN, heat transfer search optimization, and extreme learning machine. *Machines* 11 (1):29. doi:[10.3390/machines11010029](https://doi.org/10.3390/machines11010029).
- USEPA. **2009**. Current methodologies and best practices in preparing port emission inventories: Final report.
- USEPA. **2020**. *Methodologies for estimating port-related and goods movement mobile source emission inventories*. Office of transportation air quality. Washington, DC: US Environmental Protection Agency.
- Vakharia, V., M. Shah, V. Suthar, V.K. Patel, and A. Solanki. **2023**. Hybrid perovskites thin films morphology identification by adapting multiscale-SinGAN architecture, heat transfer search optimized feature selection and machine learning algorithms. *Phys. Scr.* 98 (2):025203. doi:[10.1088/1402-4896/aca49](https://doi.org/10.1088/1402-4896/aca49).
- Wan, Z., S. Ji, Y. Liu, Q. Zhang, J. Chen, and Q. Wang. **2020**. Shipping emission inventories in China's Bohai Bay, Yangtze River Delta, and Pearl River Delta in 2018. *Mar. Pollut. Bull.* 151:110882. doi:[10.1016/j.marpolbul.2019.110882](https://doi.org/10.1016/j.marpolbul.2019.110882).
- Wan, Z., Q. Zhang, Z. Xu, J. Chen, and Q. Wang. **2019**. Impact of emission control areas on atmospheric pollutant emissions from major ocean-going ships entering the Shanghai port, China. *Mar. Pollut. Bull.* 142:525–32. doi:[10.1016/j.marpolbul.2019.03.053](https://doi.org/10.1016/j.marpolbul.2019.03.053).
- Wang, Q., Y. Li, M. Diao, W. Gao, and Z. Qi. **2015**. Performance enhancement of INS/CNS integration navigation system based on particle swarm optimization back

- propagation neural network. *Ocean Eng.* 108:33–45. doi:[10.1016/j.oceaneng.2015.07.062](https://doi.org/10.1016/j.oceaneng.2015.07.062).
- Wang, S., B. Ji, J. Zhao, W. Liu, and T. Xu. 2018. Predicting ship fuel consumption based on LASSO regression. *Transp. Res. D. Transp. Environ.* 65:817–24. doi:[10.1016/j.trd.2017.09.014](https://doi.org/10.1016/j.trd.2017.09.014).
- Wang, X., W. Yi, Z. Lv, F. Deng, S. Zheng, H. Xu, J. Zhao, H. Liu, and K. He. 2021. Ship emissions around China under gradually promoted control policies from 2016 to 2019. *Atmos. Chem. Phys.* 21 (18):13835–53. doi:[10.5194/acp-21-13835-2021](https://doi.org/10.5194/acp-21-13835-2021).
- Witten, D., and G. James. 2013. *An introduction to statistical learning with applications in R*. New York: Springer Publication.
- Wu, N., K.J. Geras, Y. Shen, J. Su, S.G. Kim, E. Kim, S. Wolfson, L. Moy, and K. Cho. 2018. Breast density classification with deep convolutional neural networks. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6682–86. Calgary, Canada: IEEE.
- Xiao, F., H. Ligteringen, C. van Gulijk, and B. Ale. 2015. Comparison study on AIS data of ship traffic behavior. *Ocean Eng.* 95:84–93. doi:[10.1016/j.oceaneng.2014.11.020](https://doi.org/10.1016/j.oceaneng.2014.11.020).
- Zhang, Z., L. Huang, X. Peng, Y. Wen, and L. Song. 2022. Loitering behavior detection and classification of vessel movements based on trajectory shape and convolutional neural networks. *Ocean Eng.* 258:111852. doi:[10.1016/j.oceaneng.2022.111852](https://doi.org/10.1016/j.oceaneng.2022.111852).
- Zhou, Y., W. Daamen, T. Vellinga, and S.P. Hoogendoorn. 2019. Ship classification based on ship behavior clustering from AIS data. *Ocean Eng.* 175:176–87. doi:[10.1016/j.oceaneng.2019.02.005](https://doi.org/10.1016/j.oceaneng.2019.02.005).