TRANSNAV
http://www.transnav.eu
the International Journal on Marine Navigation and Safety of Sea Transportation
Volume 17
Number 1
March 2023
DOI: 10.12716/1001.17.01.13

Prediction of Ship's Speed Through Ground Using the Previous Voyage's Drift Speed

D. Yamane & T. Kano
NPO Marine Technologist, Tokyo, Japan

ABSTRACT: In recent years, 'weather routing' has been attracting increasing attention as a means of reducing costs and environmental impact. In order to achieve high-quality weather routing, it is important to accurately predict the ship's speed through ground during a voyage from ship control variables and predicted data on weather and sea conditions. Because sea condition forecasts are difficult to produce in-house, external data is often used, but there is a problem that the accuracy of sea condition forecasts is not sufficient and it is impossible to improve the accuracy of the forecasts because the data is external. In this study, we propose a machine learning method for predicting speed through ground by considering the actual values of the previous voyage's drift speed for ships that regularly operate on the same route, such as ferries. Experimental results showed that this method improves the prediction performance of ship's speed through ground.

1 INTRODUCTION

In this world of advanced shipbuilding technology, large volumes of cargo are transported worldwide by ships. Recently, in addition to cost reduction, efficient operation is required to reduce the global environmental burden. Consequently, there is a growing interest in weather routing, which can provide energy-saving voyage plans by considering the ship's performance and condition and forecast information on weather and sea conditions encountered during the voyage. In weather routing, an optimal route is determined before the voyage using optimization methods such as dynamic programming, considering the forecast information and the ship's propulsive performance in actual sea conditions [1,2]. Therefore, high-quality weather routing requires accurate prediction of a ship's fuel consumption and ship's speed. However, these predictions are not straightforward and have been the subject of extensive research. For instance, a method of modelling using statistical regression equations based on multiple tests of a ship's performance under specific environmental conditions is available [3,4]. This makes it possible to estimate the fuel consumption and speed of the vessel by inputting the control variables such as main engine revolution, propeller blade angel used to operate the vessel, in addition to the forecast information on the weather and sea conditions of the voyage. The advantage of this method is its high interpretability because the modelling is based on physical knowledge. However, this method has limitations because it simplifies the weather, sea conditions, and other conditions of actual voyages. Consequently, the obtained models do not represent all aspects of ship performance, and there are issues with the accuracy of predicting fuel consumption and ship's speed. Therefore, machine learning has been adopted for these predictions in recent years [5]. This method models a ship's navigation performance implicitly by training actual navigation data on a machine learning model such as a neural network (NN) [6]. Although this method lacks interpretability, unlike modelling in simplified situations, it can evaluate ships more accurately because it considers actual voyage data. Therefore, this method may also be effective in the current situation where the data collection infrastructure during voyages is becoming easier with the development of IT and other technologies. However, caution is required when applying machine learning models as they may not be effective depending on the target of prediction and the number of data available for training. Also, one of the explanatory variables used as input to the model is the forecast information on weather and sea conditions provided by meteorological agencies, but it should be noted that there are forecast errors. Thus, the authors pay attention to the fact that captains use the state of the previous day's current for navigation. Since currents generally change slowly over time, the previous day's currents are expected to contribute to predicting ship's speed on the current day's voyage. Although it is not possible to directly measure the actual values of the ocean currents, the drift speed v_d, which indicates the magnitude of the ship's current caused mainly by the ocean currents, can be calculated using the ship's speed through ground v_g and speed through water v_w, as shown in equation (1).

$$v_d = v_g - v_w \tag{1}$$

Based on the mentioned above, this study proposes a method for estimating the fuel consumption and the ship's speed using LightGBM [7], which can incorporate information on the drift speed of the previous voyage. This method improves the prediction accuracy of ship's speed through ground, which is significantly affected by sea conditions. The proposal method was also found to be effective in predicting fuel consumption and ship's speed through water.

Finally, this paper is organized as follows: Section 2 introduces the research on the prediction of fuel consumption and ship's speed and machine learning-based forecasting, while Section 3 identifies the dataset's characteristics and the task to be solved. Section 4 describes the proposal method, and Section 5 presents numerical experiments on the drift speed and the proposal method. Finally, Section 6 summarizes the paper and discusses future prospects.

2 RELATED WORKS

When predicting a value in various fields, it is a common practice to express the relationship between explanatory variables (inputs) and objective variables (outputs) using mathematical formulas. Similar methods have also been applied in the shipping field. For instance, in [8], the relationship between hourly fuel consumption l/h and ship's speed v_s per hour is expressed using a component-separated physical model, as shown in equation (2).

$$\frac{l/h}{v_s} = a_1 W^a v_s^2 + a_2 f_{wind} + \sum_{i=1}^{n} a_{3,i} f_{wave,i} \tag{2}$$

The first term on the right side represents the resistance in calm seas, the second term represents the resistance due to wind, and the third term represents the resistance due to waves. Here, W is displacement, f_{wind} is the resistance component due to wind, and $f_{wave,i}$ is the resistance component due to waves. These constants, a,a_1,a_2, and $a_{3,i}$, can be estimated using regression analysis or other methods based on the ship's operating variables such as main engine revolution and forecast information on weather and sea conditions. Also, if the fuel consumption per hour in equation (2) is assumed to be proportional to the cube of the main engine revolution, as shown in equation (3), the ship's speed through water can be predicted.

$$l/h = \gamma N_e^3 \tag{3}$$

On the other hand, in recent years, machine learning models, particularly neural networks, have increasingly been utilized to make predictions across various fields [9,10]. One of the advantages of using machine learning models is their high expressive power. Unlike a modeling using mathematical formulas, where the regression function is set based on assumptions of the relationship between explanatory variables and objective variables, machine learning models automatically learn the relationship from the data and thus can capture more detailed and complex phenomena. Although the modeling using mathematical formulas, as shown in equations (2) and (3), is highly interpretable, its expressive power is limited. Moreover, the function used for regression must be manually selected, making it prone to errors. In contrast, machine learning models are more complex and difficult for humans to interpret, but they have the potential to achieve high accuracy due to their high expressive power. Therefore, in the shipping field, machine learning models are being used to predict fuel consumption and ship's speed. For instance, in [11], a neural network is used to predict fuel consumption based on seven explanatory variables, including ship's speed, main engine revolution, average draft, trim, cargo volume, and wind and sea effects. Similarly, [5] employs a neural network to predict ship's speed through ground using explanatory variables such as main engine revolution, wing angle, wind direction, wind strength, sea current direction, sea current strength, and elapsed time since entering the dock. These studies indicate that the most common variables used as explanatory variables in predicting fuel consumption and ship's speed are operating variables of the ship, variables representing weather and sea conditions, and elapsed time since docking. It should be noted that the elapsed time since docking is used as an explanatory variable since a ship's performance tends to decrease over time due to the attachment of marine organisms after it enters the dock and is cleaned. Prediction using Transformer or LSTM, types of neural network models that can consider time series characteristics, has also been investigated [12]. With this model, data obtained during the voyage can be used as explanatory variables, enabling more accurate predictions. However, while this model is effective in situations where real-time ship's speed prediction is required, it is not suitable when the model is intended to be used before sailing, as in this study. Also, there is a study that machine learning models, such as XGBoost[13] and LightGBM[7], to predict and compare fuel consumption of ships[14]. Both models have the advantage of achieving high accuracy and speed even without a large amount of data. However, as shown in [14], LightGBM is generally preferred over XGBoost because it is faster and more accurate. Although neural networks are commonly used in machine learning studies due to their versatility and name recognition, decision tree-based machine learning models are more effective in certain cases. As demonstrated in [15], decision tree models perform better than neural network models when the data is limited and in tabular form, as in this study. Therefore, we use LightGBM, one of the decision tree models, in this study.

3 DESCRIPTION OF DATA CHARACTERISTICS AND CHALLENGES IN DEVELOPING THE METHOD

The target ship used in this study is sailing Japanese southern part of Pacific Ocean, and from January 3, 2022 to May 11, 2022, monitoring data such as datetime, fuel consumption, ship's speed (water/ground), position (latitude, longitude), wind direction and speed (relative), main engine revolution, propeller blade angle, direction of course, and direction of moving are automatically collected every 10 minutes. The data is then transmitted to and stored on a server on land via ship-to-shore communications. In this study, we performed spatiotemporal corrections to the grid point value (GPV) data provided by the Japan Meteorological Agency (JMA), and calculated and appended the wind (wind direction and wind speed), waves (wave height, wave direction and wave period), and Sea and tidal currents (current speed) corresponding to the time and ship position of the collected data. The number of data recorded during ship operations is approximately 80-100 per day. Table 1 summarizes the information on the data, and Table 2 shows an image of the data obtained.

Note that if the same data is used for both training and testing the model, the model already knows the answers, resulting in a high prediction accuracy. In the field of machine learning, this problem is known as leakage, and measures need to be taken to prevent it during model evaluation. On the other hand, in the data analysed in this study, measurements taken during the same voyage tend to be similar. Thus, indirect leakage may occur even if the measurements were taken at different times, and the data from the same voyage needs to be handled with care.

Table 1. Information on data

Period	Number of data	Measurement data used in this study	Forecast data used in this study
2022/1/3-2022/5/11	10329 (Data while the ship is moving during the measurement period)	datetime, latitude/ longitude(deg), fuel consumption(ℓ), ship's speed through water (knot), ship's speed through ground (knot), main engine revolution(rpm), propeller blade angle(deg), direction of course(deg), direction of moving(deg), inlet and outlet	wind direction(deg), wind speed(m/s), wave height(m), wave direction (deg), wave period (s), wave height (m), wave period(s), current speed (m/s)

Table 2. Image of data

datetime	latitude	longtitude	…	current speed
2022/1/3 0:00	33.4334	131.7932	…	0.22
2022/1/3 0:10	33.4003	131.7121	…	0.17
:	:	:	:	:
2022/5/11 23:50	32.7342	132.4278	…	0.21

Although the data used in this paper is comparable to data typically measured on a ship, there are three issues that need to be addressed. The first issue is that the ship's speed through water includes errors. While the ship's speed through ground can be measured almost accurately using GPS, the ship's speed through water is measured using a sensor installed on the bottom of the ship. These errors include bias and change over time and can affect the value of the prevailing current, which is calculated using equation (1). Therefore, it is necessary to develop a method that is somewhat robust to errors.

The second issue involves devising a method to incorporate information on the previous voyage's drift speed. As previously mentioned, data is measured every 10 minutes, and a certain amount of data from the previous voyage has been accumulated. While it is possible to use all of the data for forecasting, there is a lot of unnecessary information, and the computation time increases significantly. Thus, it is necessary to appropriately extract the necessary information and use it for forecasting.

The third issue pertains to the limited information available from the previous day's data. While data from the previous day can be obtained, only data from on the actual navigated route can be collected (refer to Figure 1). Therefore, it is necessary to supplement the drift speed for coordinates other than the route taken. However, the completion of coordinates for which no data is available may introduce noise. Thus, a method needs to be developed to incorporate the drift speed information as noise-free as possible.

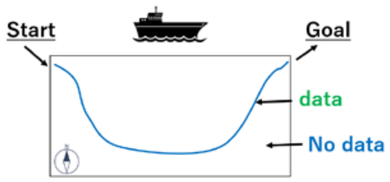

Figure 1. Relationship between the route and the data

4 PROPOSAL METHOD

In this section, the proposal method is explained in detail. Section 4.1 provides an overview of the proposal method, while Section 4.2 describes the module designed to extract the necessary information on the drift speed from the previous day's voyage data. Section 4.3 explains LightGBM machine learning method and the advantages of using it in the proposed model.

4.1 *Overview of the Proposal method*

Figure 2 illustrates the overall diagram of the proposal method: LightGBM receives inputs that include the features extracted by the module that appropriately extracts drift speeds from the previous voyage's data, ship coordinates, ship operating variables, and forecast data on weather and sea conditions. Based on these feature values, the system estimates and outputs the target variables, namely fuel consumption, ship's speed through water, and ship's speed through ground.

As discussed in Section 3, there were three main issues in this research: (1) extracting necessary information from a large amount of previous voyage data, (2) accounting for errors in ship's speed through water, and (3) appropriately supplementing the drift speed at coordinates not covered by the previous voyage. The proposal method addresses the first issue by using a module to extract information on drift speed from the previous voyage's data, and addresses the second and third issues by using LightGBM.

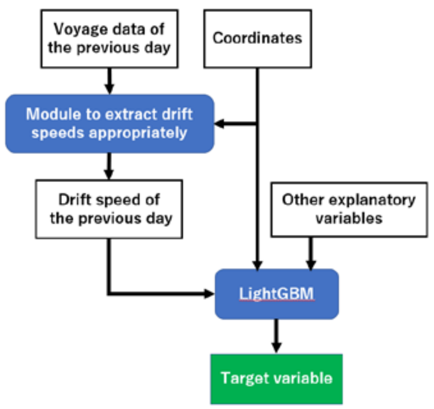

Figure 2. Overall diagram of the proposal method

4.2 *Module to properly extract drift speeds from the data of the previous voyage*

As mentioned earlier, it is necessary to extract only the necessary information from the data from previous voyages since there is a lot of unnecessary information. Given that the drift speed typically depends on the coordinates, we believe that the drift speeds around the coordinates to be predicted are useful for the forecast, and conversely, there is no need to consider the drift speeds at coordinates further away from the coordinates to be predicted. Therefore, we extract data as follows. First, we perform meshing (N × M) on the square region containing the coordinates of the possible paths during the voyage. Next, if the data corresponding to each cell of the meshed coordinates is available in the previous voyage data, we calculate the drift speed using equation (1) and embed the information in the cell. Then, the system extracts drift speeds contained in the h cells on the left, right, top, and bottom of the cell whose coordinates correspond to the coordinates to be predicted. Cells with no information are also extracted as None, indicating that they have no information. The above flow is summarized in Algorithm 1, and Figure 3 illustrates the algorithm when N=6, M=12, and h=1. When inputting features to LightGBM, they are named with the cell to be predicted as the center (0, 0), and the positive directions are right and down. An example of this is shown in Figure 4.

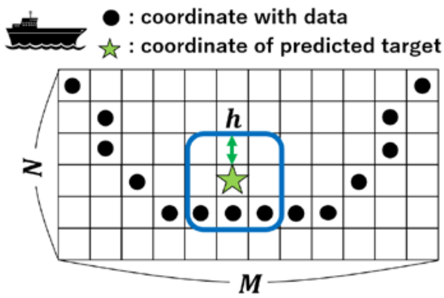

Figure 3. Image of the algorithm when N=6, M=12, and h=1

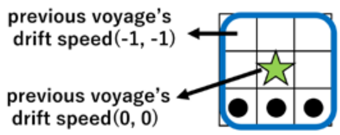

Figure 4. Image of feature naming

Algorithm 1 Extraction of drift speed from data of previous voyage

Require: Coordinate of target $\boldsymbol{c}$, Data of previous voyage $\boldsymbol{D}$

$\boldsymbol{map}$ ≔ Initialize an $\boldsymbol{N} \times \boldsymbol{M}$ array with None
for $\boldsymbol{d}$ in $\boldsymbol{D}$ **do**
  $\boldsymbol{i,j}$ ≔ Calculate $\boldsymbol{i,j}$ corresponding to the squares of map from the upper table of coordinates contained in $\boldsymbol{d}$
  $\boldsymbol{map[i,j]}$ ≔ the value of the drift speed using equation (1) from the ship's speed through ground and water included in v_d
end for
$\boldsymbol{drifts}$ ≔ Initialize an (2$\boldsymbol{h}$+1) × (2$\boldsymbol{h}$+1) array with None
$\boldsymbol{s,t}$ ≔ Calculate $\boldsymbol{s,t}$ corresponding to the squares in $\boldsymbol{map}$ from $\boldsymbol{c}$
for $\boldsymbol{i}$ in -(2$\boldsymbol{h}$+1)...(2$\boldsymbol{h}$+1) **do**
  for $\boldsymbol{j}$ in -(2$\boldsymbol{h}$+1)...(2$\boldsymbol{h}$+1) **do**

```
        drifts[i,j] ≔ map[s+i,t+j]
    end for
end for
drifts ≔ Flatten drifts to one dimension
return drifts
```

4.3 *LightGBM*

LightGBM is a machine learning method that employs an ensemble of multiple decision trees and is widely recognized for its speed and accuracy. To provide an overview of LightGBM, we briefly describe the decision tree on which LightGBM is based (for a detailed description of LightGBM, refer to [7]). The trained decision tree is a tree structure, in which non-terminal nodes have rules expressed in terms of specific features and threshold values, and terminal nodes have predicted values of the objective variable. When data is inputted into the model, a rule-based decision is made to determine if a particular explanatory variable in the data exceeds the threshold value set at the node or not. The process is repeated until the terminal node is reached, and the objective variable value of the terminal node is output as a predicted value. During training, the feature values and threshold values of each node are determined to ensure accurate predictions for the training data. Figure 5 illustrates an example of a decision tree and data for predicting ship's speed through ground. This figure is used to explain the forecasting process. At the root node, the branching rule is whether the main engine revolution is greater than or less than 620. The example data has a main engine revolution of 630, so it branches to the right child node. At the branched child node, the rule for branching is whether the wave height is greater than or less than 0.5, and since the wave height is 0.199 in the data, the node branches to the left child node. Since the branched child node is the terminal node, the predicted ship's speed through ground is output as 28.31, which is set as the predicted objective variable value. Although a detailed explanation is omitted here, LightGBM generates multiple decision trees and outputs the predicted value of each tree. The model outputs a final prediction that takes the predictions of each tree into account.

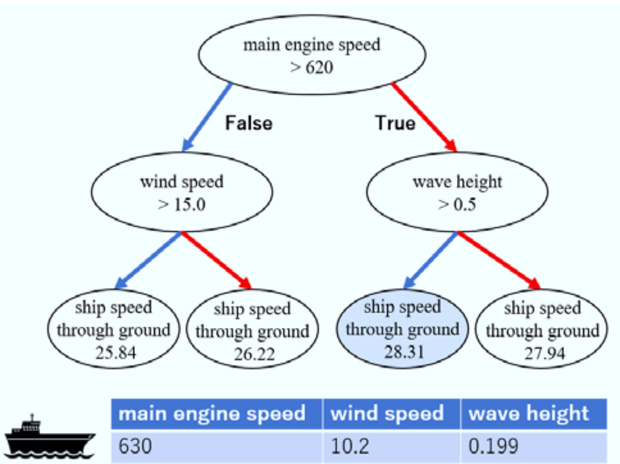

Figure 5. Image of forecasting method using decision trees

In this study, LightGBM was utilized to predict the objective variable, as shown in Figure 2. Drift speeds from previous voyages and other variables, such as ship coordinates, ship operating variables, and forecast data on weather and sea conditions, were included as explanatory variables for the model.

As discussed in Section 2, machine learning methods have recently been applied to the shipping field. However, many of the machine learning methods employed are neural networks. Among them, LightGBM was chosen for this research for three reasons.

Firstly, LightGBM is more effective in learning for the present data. Neural networks are better suited for tasks with unstructured data, such as images and sound, rather than structured data like table data used as input in this study. Additionally, a large amount of data is typically required for training neural networks, as explained in Section 3. However, the present data comprises only 10,000 cases, which is not large enough to train a neural network effectively. On the other hand, LightGBM can learn sufficiently with this amount of data.

Secondly, LightGBM is robust to a certain degree of error. This model is independent of the numerical scale of explanatory variables due to its rule-based module for determining outputs, as explained earlier. Therefore, it is unaffected by implicit errors in drift speeds computed in equation (1). For instance, even if the data contains an error of ε, a rule with a threshold shifted by ε is automatically learned when a rule is created for a certain node.

Thirdly, there is no need for supplementation of the drift speeds. As explained in Section 4.2, the drift speeds around the coordinates to be predicted are input to LightGBM. These features include None, which indicates that there is no data. In general machine learning methods, including neural networks, when None is given as input, the user must appropriately supplement the numerical values, which can easily lead to a decrease in prediction accuracy. However, LightGBM can treat None as input as it is, and it learns by taking the None information into account. This eliminates the need for the user to be involved in the completion of numerical values and has the advantage of achieving higher prediction accuracy compared to other methods.

5 NUMERICAL EXPERIMENTS

This section aims to validate the accuracy of the drift speed calculated by equation (1) and the proposal method. Section 5.1 details the learning and evaluation procedures utilized in numerical experiments. In Section 5.2, experiments were conducted to verify the accuracy of the drift speed, and in Section 5.3, experiments were conducted to validate the proposal method.

5.1 *Model training and evaluation method*

To conduct the training and testing, a walk-forward method is utilized. Specifically, training is initially performed using data collected from January 3, 2022, to April 15, 2022, while the ship is in motion. Subsequently, data from April 16, 2022, is used as the test set to predict the fuel consumption and ship's speed of the ship during its operation. The training and test sets are then shifted by one day, as depicted in Figure 6. The accuracy of the predictions for all data in the test period is evaluated using this method. Note that the number of the data between January 3 and April 15 is 8602.

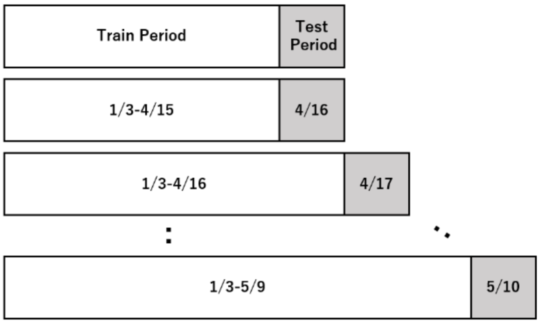

Figure 6. Walk-forward method

The accuracy of the error of the prediction for data *i*, $Accuracy_i$ is calculated as in equation (4). Note that the average value of the objective variable for all data while the ship is moving is $\overline{y}$, the true value of the objective variable in the data is y^i, and the value of the objective variable predicted by entering the explanatory variables for data *i* in the model is y^i.

$$Accuracy_i = \frac{\left|\hat{y}^i - y^i\right|}{\overline{y}} \qquad (2)$$

Since this index is the accuracy of the error, the unit is %, and the closer to 0, the better the value. We also compute $Accuracy_i$ for all data used in the test and evaluate its mean value and standard deviation.

Note that the machine learning model has several hyperparameters that must be set by the analyst during training. These hyperparameters significantly affect the performance of the model, and hyperparameter tuning is a crucial step in model learning. There are three main methods for hyperparameter tuning: random search, grid search, and Bayesian optimization. In this study, we use Bayesian optimization, which efficiently searches for the optimal parameters using Gaussian process regression, and allows for efficient hyperparameter tuning in a limited amount of time. Specifically, we use the best parameters among 30 iterations for the training of our model.

5.2 *Experiment 1: Validity of the drift speed*

In this study, we examined the effectiveness of this method for predicting drifting speed based on the captain's use of the previous day's sea conditions for voyage planning. However, it is uncertain whether the value is accurate enough for forecasting as it is not directly measured. Therefore, in this section, we conduct an experiment to verify the validity of the predictions by comparing the prediction accuracy of LightGBM with and without using the drift speed as an explanatory variable, assuming that the prediction of the ship's speed through ground, which is most affected by the sea conditions, will be performed. Specifically, we compared the accuracy of two models: LightGBM that predicts the ship's speed through ground by entering the main engine revolution, propeller blade angle, latitude, longitude, direction of course, direction of proceeding, wind direction, wind speed, wave height, wave direction, wave period, ocean current, port arrival and departure, and time since the last dock entry(calculated from the date and time of the modeling data) as explanatory variables, and a LightGBM model that inputs drift speeds calculated using equation (1) as additional explanatory variables, the variables described above. If the prediction error accuracy of the latter model is sufficiently smaller than that of the former model, then the drift speed is shown to be an effective feature. Note that ship's speed through water and ground are not used as inputs, so no leakage occurs. The drift speeds are calculated using the ship's speed through water and ground measured at the same time as the data to be predicted, so they cannot be used in actual operation. Therefore, this experiment was conducted to evaluate the degree to which the characteristic drift speeds reflect the state of the oceanographic phenomena.

The results of the experiment are shown in Table 2. The results indicate that the use of drift speeds improves the error accuracy by more than 1% on average, and the standard deviation is also improved by more than 0.5%. This suggests that the drift speed is a good feature that captures the sea condition.

Table 2. Results of Experiment 1.

Method	Mean of $Accuracy_i$	Standard deviation of $Accuracy_i$
LightGBM without drift speeds	2.311	1.782
LightGBM with drift speeds	1.234	1.200

5.3 *Experiment 2: Effectiveness of the proposal method*

This section presents experimental results conducted to confirm the effectiveness of the proposal method. Four methods, namely neural networks, LightGBM, LightGBM with the drift speeds from the previous voyage (proposal method), and a component-separated physical model (for reference), are used to predict fuel consumption, ship's speed through water and ground. Their evaluated values are compared. The component-separated physical model is a model that predicts using mathematical equations with parameters estimated based on physical findings. However, as introduced in [8], this experiment uses a model with parameters identified by NPO Marine Technologist using collected data from July 2, 2021, to October 6, 2021, when the ship was put into service. Therefore, this model is different from the machine learning model evaluated using data up to the previous day, and cannot be compared simply. However, this comparison is made for reference purposes in this experiment. Note that the model outputs predicted values for the objective variable by inputting the main engine revolution, propeller blade angle, latitude, longitude, direction of course, direction of proceeding, wind direction, wind speed, wave height, wave direction, wave period, and ocean current. In addition to the explanatory variables used in the component-separated physical model, the neural network and LightGBM make predictions by inputting the elapsed time since the last docking and the port arrival/departure as explanatory variables. Note that the neural network is implemented using scikit-learn's MLPRegressor. The hidden layer of the neural network uses relu as the activation function and Adam is used as the optimization algorithm for the neural network. In addition to these variables, the proposal method also inputs as explanatory variables the drift speed extracted from the data of the previous voyage using the module described in Section 4.2. In the experiments described in this section, test data for which no previous voyage's data existed were excluded from the evaluation. Moreover, while machine learning models can be updated easily, the component-separated physical model of [8] is difficult to update frequently due to practical works. For this reason, the parameter of the model already settled and used in actual weather routing operations was used in this experiment.

Table 3 presents the mean and standard deviation of the errors in predicting fuel consumption l^i, ship's speed through water v_w^i, and ship's speed through ground v_g^i for each method. The frequency distribution of the errors in the predictions is shown in Figures 7-9. Note that the neural network predictions are excluded from each figure due to their large errors.

Table 3. Results of Experiment 2

Method	Fuel consumption	Ship's speed through water	Ship's speed through ground
Neural Network	7.492+6.273	57.36±42.17	32.66±25.60
LightGBM	0.666±0.594	1.316±1.354	2.225±1.761
LightGBM with drift speeds (proposal method)	0.651±0.617	1.300±1.408	2.033±1.666
component-separated physical model (for reference)	1.102±1.002	1.778±1.550	3.426±2.860

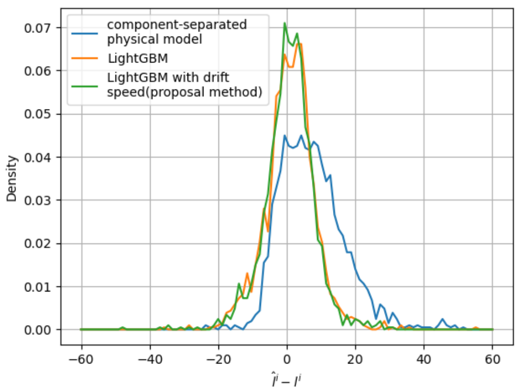

Figure 7. Frequency distribution of errors related to forecasting fuel consumption

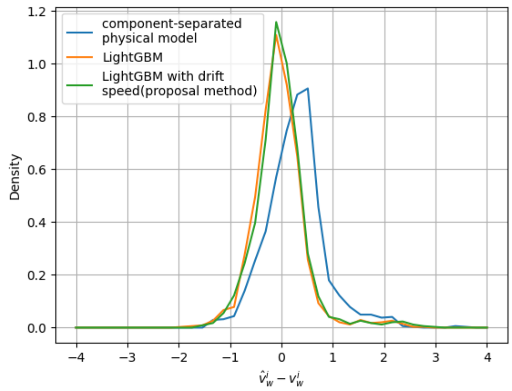

Figure 8. Frequency distribution of errors related to the prediction of ship's speed through water

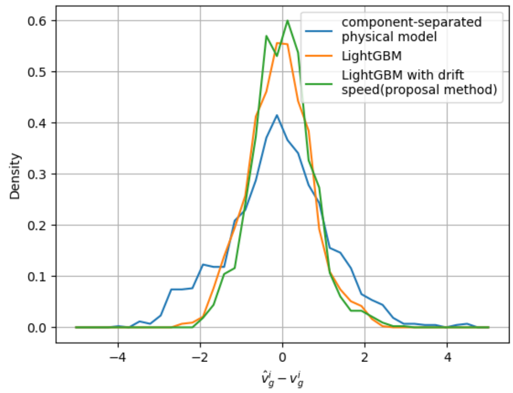

Figure 9. Frequency distribution of errors related to the prediction of ship's speed through ground

Following three findings were introduced from the experimental results. Firstly, the neural network had the worst performance for all prediction targets and was unable to make any predictions. There are two possible reasons for this outcome. The first is the small amount of data available. The second is the limited information relative to the number of data. We did not prepared enough to work with in this experiment for the neural network. In addition, as explained in Section 2, the feature values of data from same-day voyages tend to be similar, resulting in insufficient learning due to a small amount of substantially different data.

Secondly, the prediction accuracy of all methods except neural networks is generally acceptable. Table 3 demonstrates that even in the worst case, the error accuracy of the component-separated physical model for predicting ship's speed through ground is approximately 3.4%, suggesting that all methods have an accuracy that is generally acceptable in practical terms.

Thirdly, the proposal method's predictions are the most accurate for all prediction targets, and it performed well even when the neural network could not learn. This result indicates that the proposal method can handle complex events in actual voyages with not so much data and demonstrates its superiority. However, as previously mentioned, it is not possible to make a fair comparison with the component-separated physical model since it has been some time since its set parameters of the model. Figure 10 shows the results obtained when the component-separated physical model was applied to data at the beginning of its creation to predict the ship's speed through water (referred to as the component-separated physical model (original)), added to the results in Figure 8. It can be seen that the prediction for the ship's speed at the beginning of the model has a smaller error. This suggests that the parameter of the component-separated model tends to change over time due to hull fouling and other factors, and that the estimated ship's speed tend to be larger than the actual measured values. Therefore, it is essential to update the parameters as required to make accurate estimates. Note that the proposal method is completely black-boxed, and it is not possible to explain the reasons for the outputs generated from a given input. However, the proposal method, which can be updated easily, is believed to offer practical advantages.

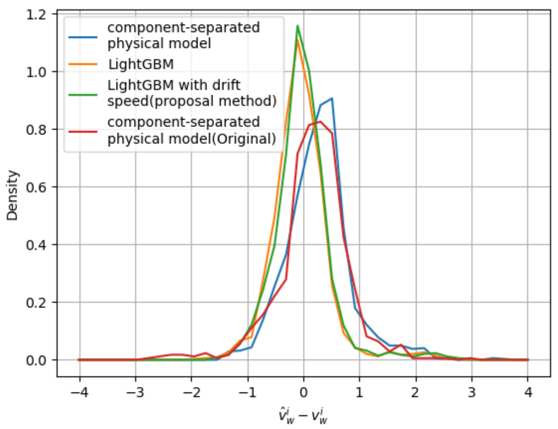

Figure 10. Graph of component-separated physical model (Original) added to the results in Figure 8

6 CONCLUSION

In this study, we aimed to enhance the prediction accuracy of fuel consumption, ship's speed through water, and ground in order to achieve highly accurate weather routing. In previous studies, many features used to predict them were commonly measured and readily available, such as ship operating control variables and predicted values of weather and sea conditions, while few other features were used. Additionally, neural networks have often been employed in machine learning models. In this study, we focused on the fact that captains and crews use the sea conditions from the previous voyage and proposed a method that combines LightGBM with a module for integrating the drift speed from the previous voyage as feature. In experiments, after confirming that the drift speed calculated using equation (1) is an effective feature for predicting the ship's speed over ground, we compared the prediction accuracy of the neural network, LightGBM, the proposal method, and the component-separated physical model introduced in [8] as a reference for comparison. The results showed that the proposal method was more accurate than the other methods, especially in predicting the ship's speed through ground. In addition, considering changes in hull performance over time, it is desirable to update the model frequently, but he proposal method has the advantage that the model can be easily updated, and is found to be useful in practice. However, the proposal method lacks the ability to explain the prediction results, and in practice, it is considered effective when used in combination with a component-separating physical model.

Although the proposal was made with pre-voyage use in mind, as shown in experiment 1, if the drift speed is an effective characteristic that represents the state of the sea conditions, data measured during the voyage several tens of minutes or hours in advance can be used for forecasting as in [12]. Thus, it is possible to optimize the route sequentially based on the data measured during the voyage by extending this study. Additionally, as described in Section 2, a ship's operational performance temporarily improves when it enters a dock due to cleaning, after which its performance gradually declines due to the attachment of marine organisms. Therefore, the time elapsed after a ship enters the dock plays a critical role in predicting the ship's speed through ground. Hence, creating a machine learning model using data that includes the entire period from the day the dock ends to the day the ship enters the next dock would be desirable. However, the data used in this study were not so much, and the period of data used for training was only about four months, making such training impossible. Thus, the accuracy of the proposal method could be further improved by using several years' worth of data for training.

Based on the above, future work will include sequential route optimization and the creation of more accurate models with more data for practical use.

REFERENCES

[1] W.Laura, R.Anisa, W.Mareike and J.Carlos, "Modeling and optimization algorithms in ship weather routing," International journal of e-navigation and maritime economy, vol.4, pp.31-45, 2016.
[2] J.Szlapczynska, "Multi-objective weather routing with customised criteria and constraints," The journal of navigation, vol.68, pp.338-354, 2015
[3] U.Hollenbach , "Estimating resistance and propulsion for single-screw and twin screw ships.," International conference on computer applications in shipbuilding, vol.2, pp 237–250, 1999
[4] J.Holtrop and G.G.J.Mennen, "An approximate power prediction method," International shipbuilding progress, vol.29, pp.166-170, 1982
[5] T.Wieslaw and R.Krzysztof, "Applying artificial neural networks for modelling ship speed and fuel consumption," Neural computing and applications, vol.32, pp.17379-17395, 2020
[6] R.Frank, "The perceptron: a probabilistic model for information storage and organization in the brain.," Psychological review, vol.65, pp.386-408, 1958
[7] K.Guolin, M.Qi, F.Thomas, W.Taifeng, C.Wei, M.Weidong, Y.Qiwei, L.Tie-Yan, "LightGBM: a highly efficient gradient boosting decision tree," Advances in neural information processing systems, vol.30, pp. 3149–3157, 2017
[8] K.Sato and T.Kano, "Eco-shipping project with speed planning system for japanese coastal ships," Scientific journals of the maritime university of szczecin, vol.46, pp.147-154, 2016
[9] K.Konstantina, P.E.Themis, P.E.Konstantinos, V.K.Michails and I.F.Dimitrios, "Machine learning

applications in cancer prognosis and prediction," Computational and structural biotechnology Journal, vol.13, pp.8-17, 2015
[10] S.Jonathan, R.G.M. Mário, B.Silvana and A.L.M.Miguel, "Recent advances and applications of machine learning in solid-state materials science," npj computational materials, vol.5, 2019
[11] E. Bal Beşikçi, O. Arslan, O. Turan and A.I. Ölçer, "An artificial neural network based decision support system for energy efficient ship operations," Computers & operations research, vol.66, pp. 393-401, 2016
[12] E.M.Sara, B.Loubna, C.Stéphane and B.Abdelaziz, "Deep learning-based ship speed prediction for intelligent maritime traffic management," Journal of marine science and engineering, vol.11, 2023
[13] C.Tianqi and G.Carlos, "XGBoost: A scalable tree boosting system," Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp.785-794, 2016
[14] K.J.Yoon, L.H.Seok and O.J.Seok, "Study on prediction of ship's power using light GBM and XGBoost," Journal of advanced marine engineering and technology, vol.44, pp.174-180, 2020
[15] G.Léo, O.Edouard and V.Gaël, "Why do tree-based models still outperform deep learning on tabular data?," arXiv, 2022
[16] B.James, B. Remi, B.Yoshua and K.Balazs, "Algorithms for hyper-parameter optimization," Proceedings of the 24th international conference on neural information processing systems, pp.2546-2554, 2011