International Journal of Naval Architecture and Ocean Engineering 13 (2021) 641–649

Contents lists available at ScienceDirect

International Journal of Naval Architecture and Ocean Engineering

journal homepage: http://www.journals.elsevier.com/international-journal-of-naval-architecture-and-ocean-engineering/

Prediction of ship power based on variation in deep feed-forward neural network

June-Beom Lee [a], Myung-Il Roh [b, *], Ki-Su Kim [c]

[a] Department of Naval Architecture and Ocean Engineering, Seoul National University, Republic of Korea
[b] Department of Naval Architecture and Ocean Engineering, and Research Institute of Marine Systems Engineering, Seoul National University, Seoul, Republic of Korea
[c] Research Institute of Marine Systems Engineering, Seoul National University, Republic of Korea

ARTICLE INFO

Article history:
Received 24 February 2021
Received in revised form
10 June 2021
Accepted 7 August 2021
Available online 12 August 2021

Keywords:
Ship power prediction
Preprocessing
Deep learning
Deep feed-forward neural network (DFN)
Hyperparameter optimization
K-means clustering

ABSTRACT

Fuel oil consumption (FOC) must be minimized to determine the economic route of a ship; hence, the ship power must be predicted prior to route planning. For this purpose, a numerical method using test results of a model has been widely used. However, predicting ship power using this method is challenging owing to the uncertainty of the model test. An onboard test should be conducted to solve this problem; however, it requires considerable resources and time. Therefore, in this study, a deep feed-forward neural network (DFN) is used to predict ship power using deep learning methods that involve data pattern recognition. To use data in the DFN, the input data and a label (output of prediction) should be configured. In this study, the input data are configured using ocean environmental data (wave height, wave period, wave direction, wind speed, wind direction, and sea surface temperature) and the ship's operational data (draft, speed, and heading). The ship power is selected as the label. In addition, various treatments have been used to improve the prediction accuracy. First, ocean environmental data related to wind and waves are preprocessed using values relative to the ship's velocity. Second, the structure of the DFN is changed based on the characteristics of the input data. Third, the prediction accuracy is analyzed using a combination comprising five hyperparameters (number of hidden layers, number of hidden nodes, learning rate, dropout, and gradient optimizer). Finally, k-means clustering is performed to analyze the effect of the sea state and ship operational status by categorizing it into several models. The performances of various prediction models are compared and analyzed using the DFN in this study.

1. Introduction

1.1. Research motivation

With stricter environmental regulations and increasing fuel prices, economic routes that minimize the fuel oil consumption (FOC) of ships have garnered increasing interest. To plan a ship route, it is necessary to predict the ship power, which is essential for determining the ship's FOC (Roh, 2013; Lee et al., 2018). In general, the method suggested in the ISO (International Organization for Standardization) 15,016 (ISO 15016, 2015) is widely used to predict ship power. However, it is known that this method is not suitable for practical use because it fits well when the weather condition is good (Kim and Roh, 2020; Kim et al., 2020c). Therefore, a method that can predict ship power more accurately is necessitated.

Traditionally, a numerical method using the results of a model test has been widely used to predict ship power (Kristensen and Lützen, 2012; Rakke, 2016). However, it is difficult to predict ship power owing to the uncertainty of the model test. An onboard test should be conducted to solve this problem. Nonetheless, it requires considerable resources and time; hence, it is challenging to use numerical methods to predict the power of operating ships. In several studies, ship power was predicted using a data-driven model (Ahlgren and Thern, 2018; Liang et al., 2019; Panapakidis et al., 2020; Uyanik et al., 2019; Yoo and Kim, 2017). As data-driven methods used for predicting ship power, two representative methods exist: regression analysis and deep learning methods. Regression analysis is a method for identifying correlations

* Corresponding author.
E-mail address: miroh@snu.ac.kr (M.-I. Roh).
Peer review under responsibility of The Society of Naval Architects of Korea.

https://doi.org/10.1016/j.ijnaoe.2021.08.001
2092-6782/

between multiple variables. This method can rapidly generate a prediction model for simple problems. Deep learning, a type of machine learning based on artificial neural networks, is an effective method for predicting data patterns. If sufficient training data are available, deep learning can solve complex problems more effectively than regression analysis. In addition, a deep learning model can be tuned based on the complexity of a problem. Therefore, we selected a deep learning model to predict ship power in this study.

1.2. Related studies

Numerical methods are typically used to predict ship power. Rakke (2016) calculated the resistance of a ship using the Holtrop–Mennen method (Holtrop and Mennen, 1982). Subsequently, they predicted the ship power and ship emissions. Kristensen and Lützen (2012) predicted ship resistance and power using the Harvald method (Harvald, 1992). The methods mentioned above are not applicable to actual ships because they deduce ship power from ship resistance. Therefore, data-driven models based on ocean environmental data and ship operational data have garnered significant attention recently. One of the traditional representative data-driven methods is regression analysis. Yoo and Kim (2017) predicted ship speed using linear regression analysis. Using the predicted speed, the authors predicted the ship power. Szelangiewicz and Żelazny (2017) proposed an approximation method using basic geometric parameters to predict ship power at the initial design stage. Uyanik et al. (2019) predicted a ship's FOC using multilinear regression (MLR) analysis. However, it is noteworthy that when the prediction of ship power exhibits significant nonlinearity, the use of the regression analysis method may be restricted. Lang and Mao (2020) proposed a semi-empirical model for predicting speed loss of the existing ship. For an accurate model, wave reflection-induced resistance was additionally considered.

Therefore, ship power has been predicted using deep learning models to overcome these limitations. Ahlgren and Thern (2018) selected a prediction model based on data characteristics and predicted a ship's FOC using AutoMachineLearning (AutoML). Panapakidis et al. (2020) predicted a ship's FOC using long short-term memory (LSTM), a type of deep-learning model primarily used for time-series data. They developed several prediction models by combining input data (ocean environmental data and engine data). In addition, they optimized the model by analyzing the accuracy of each model. Liang et al. (2019) used AIS (Automatic Identification System) data, ship performance measurement data, and weather data to predict ship power. They proposed a predictive model using the MLP (Multi-Layer Perceptron) method. They fixed the number of hidden layers to 3 and compared models with various combinations of neurons when using MLP. In addition, they compared the results with the prediction results of DNV GL. However, they used only the fully connected layer, which is the general structure of the MLP. Also, they only performed optimization for a few hyperparameters, such as the learning rate and loss. Abebe et al. (2020) used a decision tree regressor and four ensemble methods to predict ship speed. And by comparing each prediction result, it was confirmed that the extra tree regressor had the best performance. Kim et al. (2020a) used the Support Vector Regression (SVR) to predict the propulsion power of a ship. They optimized the hyperparameters of the SVR model and also improved the accuracy by preprocessing the data using Chauvenet's criterion.

Therefore, in this study, five hyperparameters were optimized. In addition, a prediction model considering various sea states and ship operational statuses was developed by changing the configuration of the DFN and k-means clustering. Table 1 shows a comparison of the characteristics of this study and other studies.

2. Method for predicting ship power

Herein, we propose a model for predicting ship power. The model was developed using a DFN, a type of deep-learning network that effectively predicts numerical data. In this study, three types of ship's operational data (draft, speed, and heading) and six types of ocean environmental data (wave height, wave period, wave direction, wind speed, wind direction, and sea surface temperature) were used to predict ship power. A total of 240,000 datasets for 13,000 TEU class container ships was used to train the model. In addition, suitable hyperparameters were selected for each model through hyperparameter optimization. In order to use the trained model, nine types of input parameters at the desired time and location, which are mentioned above, should be input to get the ship power. The configuration of the DFN was changed and analyzed based on the data characteristics. Finally, k-means clustering was performed. In this section, the DFN configuration for ship power prediction and a method for applying the various treatments introduced above to the model are described.

2.1. Preprocessing of ocean environmental data

Among the ocean environmental data, those related to wind and waves (wave height, wave period, wave direction, wind speed, and wind direction) must be applied differently in the ship's stationary state and operating state because they are vectors. Therefore, ocean environmental data related to wind and waves were converted into relative values for the velocity of the ship.

The wave height and wave period used in this study were calculated from the wave spectra. The wave spectrum must be converted in consideration of the ship's velocity to calculate the relative ocean environmental data for the ship (SNAME, 1989). According to ship's velocity, the formulas for converting the wave spectrum are shown in Equations (1) and (2) (SNAME, 1989).

$$w_e = \left| w - \left(\frac{w^2 U_0}{g}\right) \cos \mu_0 \right| \tag{1}$$

$$S_\varsigma(w_e) = S_\varsigma(w) \Big/ \left| 1 - \left(\frac{2wU_0}{g}\right) \cos \mu_0 \right| \tag{2}$$

In Eqs. (1) and (2), w_e is the encounter frequency, U_0 the velocity of the ship, μ_0 the angle between the ship and wave, $S_\varsigma(w)$ the existing wave spectrum, and $S_\varsigma(w_e)$ the wave spectrum considering the ship's velocity. The formulas for calculating wave height and wave period in the converted wave spectrum are shown in Eqs. (3)–(5) (SNAME, 1989).

$$m_{nj} = \int_0^\infty w^n S_j(w) dw \tag{3}$$

$$H_{1/3} = 4\sqrt{m_{oj}} \tag{4}$$

$$T = 2\pi \sqrt{\frac{m_{0j}}{m_{2j}}} \tag{5}$$

In Eqs. (3)–(5), $H_{1/3}$ is the significant wave height, T the mean wave period, and m the wave spectrum moment. We define $H_{1/3}$ and T as the preprocessed wave height and wave period, respectively. Meanwhile, the wind speed and wind direction were converted directly using the ship's velocity. In this study, the converted ocean environmental data were used as input data for training the DFN model.

2.2. DFN for predicting ship power

Deep learning models can be classified into various models based on the data characteristics and problem types. Among them, the DFN is primarily used for predicting numerical data. Because ship power is calculated based on various inputs, the DFN is regarded as the most suitable network. A DFN is composed of an input layer, a hidden layer, and an output layer. Each layer is composed of many nodes and nodes connected by a weight. In general, ocean environmental and ship operational data are related to the ship power (Liang et al., 2019). Therefore, the input layer of the DFN was composed of ocean environmental data and the ship's operational data. The ship power was selected as the label (output data). Fig. 1 shows the two configurations of the DFN model used in this study. In general, the DFN model was configured as DFN1 to consider the association of all input data. In this study, DFN2 was configured and analyzed simultaneously to evaluate the learning case by considering the two types of input separately.

The training of the DFN involves updating the hidden layer weights (θ_j) to minimize the loss function (L). During this process, the model performance of the model differs significantly owing to the hyperparameters. Therefore, hyperparameter optimization should be performed to fine-tune the models. This is because even when the same DFN model is used, the accuracy varies significantly depending on the combination of hyperparameters. Fig. 2 shows the five hyperparameters used for the optimization in this study.

Among the five hyperparameters optimized in this study, the first and second hyperparameters were the numbers of hidden layers and hidden nodes. As the problem becomes more complex, the feature to be expressed increases. Therefore, it is generally advantageous to increase the numbers of hidden layers and hidden nodes. However, if the complexity of the model increases excessively, overfitting may occur; hence, appropriate adjustments are necessary. The third hyperparameter is the learning rate (α), which determines the amount of weight to be updated during training. If the learning rate is extremely high, then the final solution will be difficult to obtain. On the contrary, if the learning rate is extremely low, it may fall into the local minimum and require a long training time. The fourth hyperparameter is dropout (Hinton et al., 2012), which is a method of limiting nodes participating in training to prevent overfitting. Dropout significantly affects the prevention of overfitting (regularization). However, because the training nodes are limited, the training time may be longer. The last hyperparameter was the gradient optimizer. The gradient optimizer updates the hidden layer's weight during training, and the model performance differs significantly depending on the gradient optimizer. In this study, three gradient optimizers were used: stochastic gradient descent (SGD) (Bottou and Bousquet, 2009), root mean square propagation (RMSProp) (Ruder, 2016), and adaptive moment estimation (ADAM) (Kingma and Ba, 2015). The SGD is the typically used gradient optimize. RMSProp and ADAM are gradient optimizers that reflect the inertia of the previous weights to weight updates.

To optimize these hyperparameters, grid search and random search methods are primarily used (James and Yoshua, 2012). The grid search method evaluates the model performance based on combinations of all hyperparameters within a specified range and selects hyperparameters that reflect the best performance among them. The random search method differs from the grid search method in that it evaluates the model performance by randomly selecting combinations rather than using all combinations. It is known that the random search method performs searches more efficiently in a limited time than the grid search method. The grid search method, which is more accurate than the random search method, was used in this study as it does not require a significant amount of time to train the prediction model proposed herein.

2.3. k-means clustering

Ship power is significantly affected by the sea state and operational status of the ship. Perera and Mo (2018) confirmed that the ship power exhibited different patterns based on the sea state. Previous studies indicated the necessity to appropriately partition data based on the sea state and ship operational status. Hence, we

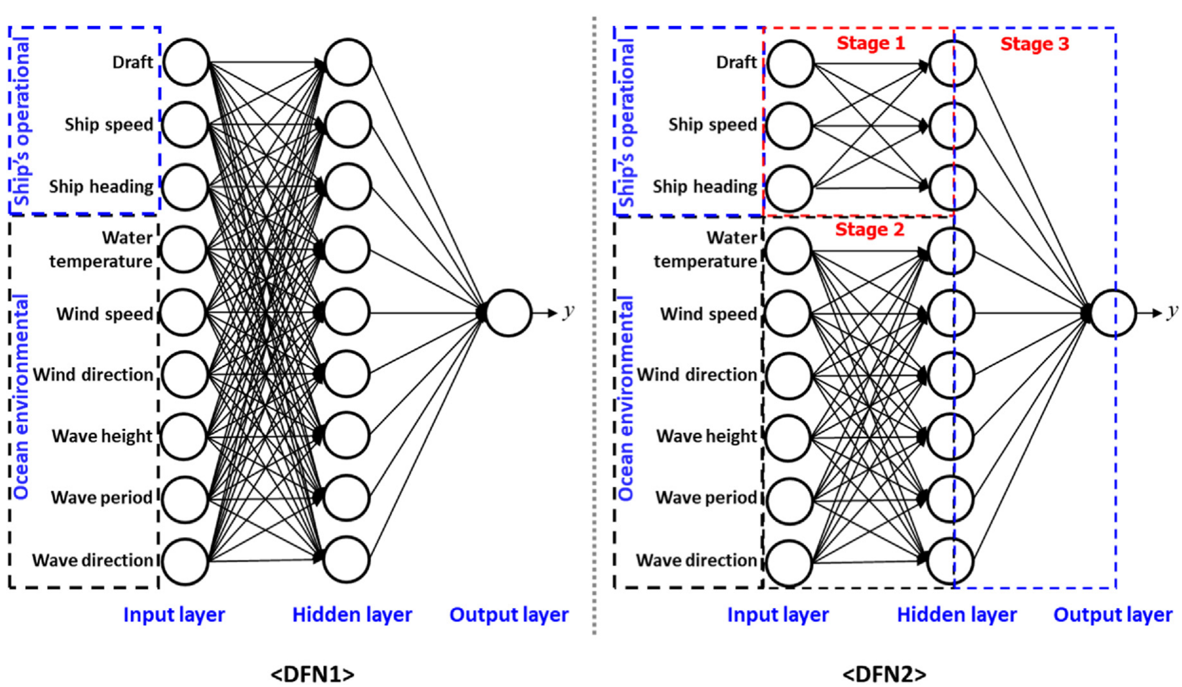

Fig. 1. DFN structure.

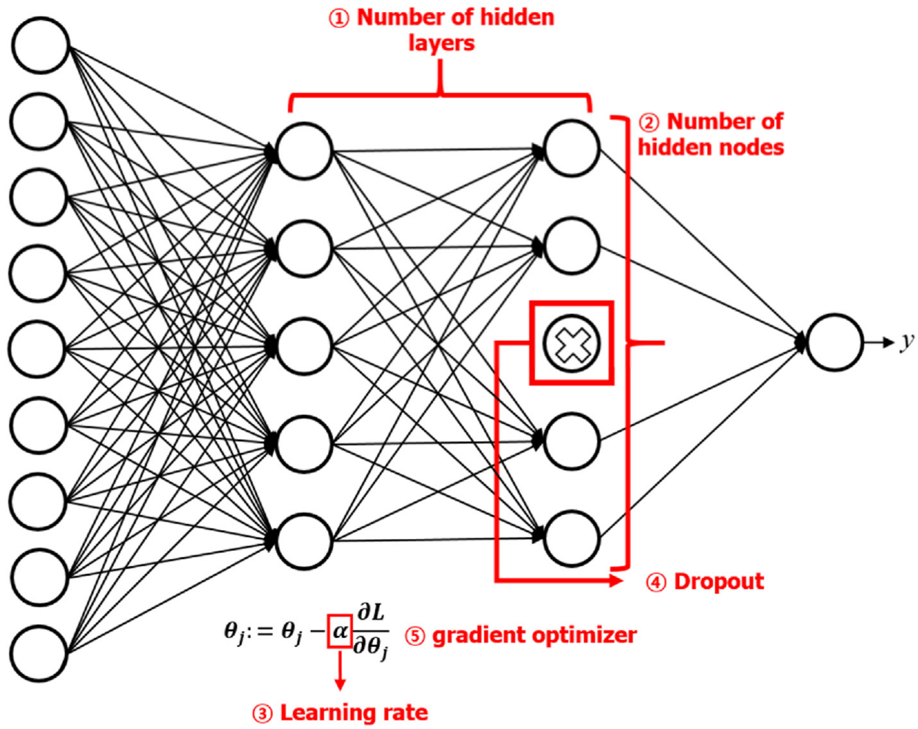

Fig. 2. Hyperparameters of DFN

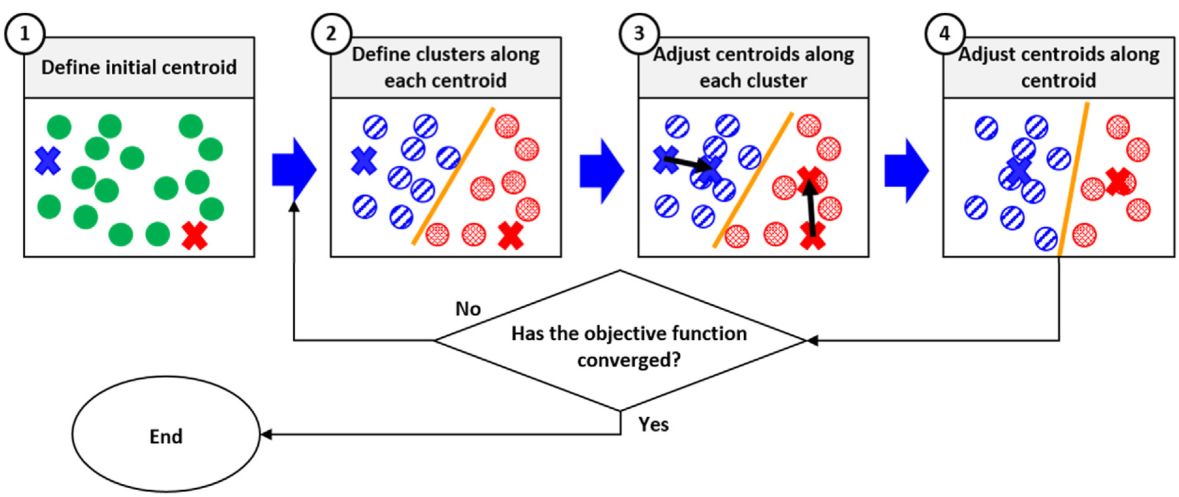

Fig. 3. k-means clustering process.

clustered the training data using k-means clustering (Lloyd, 1982), which is a type of unsupervised learning for autonomous clustering, in which the user determines the data and the number of clusters. In k-means clustering, Eq. (6) (Lloyd, 1982) was used as an objective function to determine whether clustering is successful.

$$J = \sum_{n=1}^{k} \sum_{x_i \in c_n} \|x_i - u_n\|^2 \tag{6}$$

In Eq. (6), u_n is the centroid of the nth cluster, and x_i indicates the data belonging to each cluster. As shown in Eq. (6), k-means clustering involves identifying the centroid that minimizes the objective function. Fig. 3 shows the k-means clustering process. The first step is to determine the number of clusters and then arbitrarily determine each cluster's centroid (Figs. 3–1). After determining the centroid, each datum is assigned to the appropriate cluster (Figs. 3–2). Subsequently, the centroid of each cluster is adjusted by recalculating the objective function (Fig. 3–). Finally, clustering is performed again based on the newly defined centroid (Figs. 3–4). The process above is repeated until the objective function converges.

3. Verification

3.1. Comparison model

Because it was difficult to obtain the actual ship data of a 13,000 TEU class container ship used in the application, we trained the

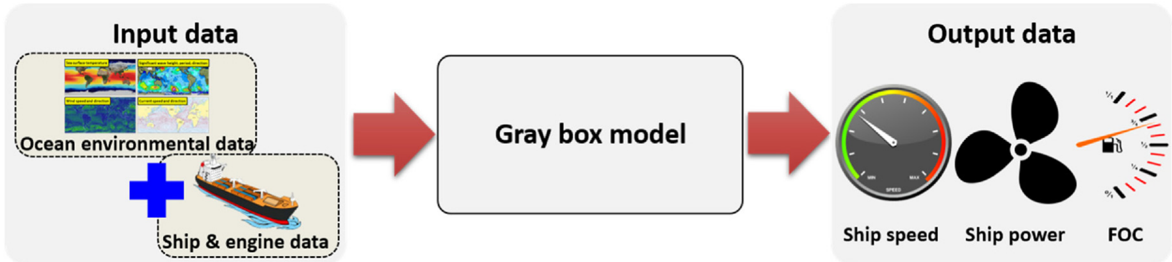

Fig. 4. Process of ship power prediction using gray box model (Kim and Roh, 2020).

Table 1
Summary of related studies.

Study	Prediction target	Input data	Prediction method			
			Prediction algorithm	Preprocessing	Hyperparameter optimization	Clustering
Rakke (2016)	Ship emission	AIS data, world fleet data, ship & engine data	Holtrop–Mennen method			
Kristensen and Lützen (2012)	Ship power	Ship data	Harvald method			
Yoo and Kim (2017)	Ship power	Ship & engine data, ocean environmental data	Linear regression analysis	No	N/A	No
Szelangiewicz and Żelazny (2017)	Ship power	Geometric parameters	Approximation method			
Uyanik et al. (2019)	FOC	Ship & engine data, ocean environmental data	Multilinear regression analysis	No	N/A	No
Lang and Mao (2020)	Ship speed loss	Geometric parameters, ocean environmental data	Semi-empirical model			
Ahlgren and Thern (2018)	FOC	Engine data	Linear regression analysis, AutoML	No	No	No
Liang et al. (2019)	Ship power	Ship operational data, ocean environmental data	MLP, Physics-based model	Yes	Loss, learning rate, number of estimators, minimum sample split	No
Panapakidis et al. (2020)	FOC	Ship operational data, ocean environmental data	LSTM	No	No	No
Abebe et al. (2020)	Ship speed	Ship operational data, ocean environmental data	Decision tree regressor, ensemble methods	Yes	Yes	No
Kim et al., 2020a, 2020b, 2020c	Ship power	Ship and engine data, ocean environmental data	SVR	Yes	No	No
This study	Ship power	Ship operational data, ocean environmental data	DFN	Yes	Number of hidden layers, number of hidden nodes, learning rate, dropout, gradient optimizer	Yes

Table 2
Hyperparameters of DFN model for verification.

Number of hidden layers	Number of hidden nodes	Learning rate	Dropout	Gradient optimizer
5	40	0.001	0	ADAM

DFN using the ship power generated using the ISO 15016 method. To verify this method, we compared the prediction model of a 4600 TEU class container ship and the DFN model proposed herein. The prediction model of a 4600 TEU class container ship was developed using the gray box model (GBM) (Samsung Heavy Industries, 2017). The ocean environmental and ship operational data were used as input data for the GBM, and the output of the GBM was the ship speed, ship power, and ship FOC. Fig. 4 shows the process for predicting ship power using the GBM.

ISO 15016 is originally a method for calibrating the performance of a ship using sea trial data. Since the sea trial is usually done in calm water, it is relatively accurate when weather conditions are good. However, it has the disadvantage of being inaccurate when weather conditions are relatively bad (Kim and Roh, 2020). Nevertheless, according to the study of Kim and Roh (2020), it can be seen that the fuel consumption error is relatively low at an average error of about 3 %, even under realistic weather conditions. Therefore, we tried to verify the efficiency of our method by comparing both the GBM (gray box model) and ISO 15016 methods.

3.2. Verification results

To train the DFN model of 4600 TEU class container ships, ocean environmental and ship operational data were used in the input layer. The ship power generated using the ISO 15016 method was used as the label. Table 2 shows the hyperparameters of the DFN used to predict the power of a 4600 TEU class container ship.

Table 3 shows a comparison of the mean absolute errors (MAEs) of the ship powers predicted using the ISO 15016 method (ship power from the ISO) and DFN (ship power from the DFN), as well as that of the actual ship power (ship power from the GBM).

The average error between the ship powers obtained from the ISO and DFN was 426 kW (5.91 %), and the average error between the ship power obtained from the GBM and DFN was 462 kW

Table 3
Result of ship power prediction.

Comparison target	Output data	MAE (%)
Ship power by the ISO	Ship power of the DFN	426 KW (5.91 %)
Ship power by the GBM	Ship power of the DFN	462 KW (6.41 %)

(6.41 %). After comparing the MAE obtained from the DFN with those from the ISO and the GBM, a difference of 0.5 % was confirmed. Owing to this slight difference of 0.5 %, the DFN model based on the ISO data can be used to replace the model that was used to create the actual data.

4. Applications

We analyzed the accuracy of the prediction model for a 13,000 TEU class container ship based on a DFN. For training, a total of 240,000 datasets (period: January 1, 2017 to December 31, 2017) of 13,000 TEU class container ships were used. The datasets included ocean environmental data and ship's operational data (draft, speed, and heading) used for the input layer of the DFN, and the ship power used for the label. The ocean environmental data were obtained from the National Oceanic and Atmospheric Administration, and the ship operational data were obtained from an automatic identification system (AIS) (IMO, 2015). The ship power data were generated via the ISO 15016 method using ocean environmental and ship operational data. To evaluate the accuracy of the prediction model, the dataset was segmented into training, validation, and test sets. A training set is a dataset used to train a model, for which 80 % of the total dataset was used. A validation set is used to evaluate the model performance based on the hyperparameters, for which 10 % of the total dataset was used. A test set is used to evaluate the model's final performance, for which 10 % of the total dataset was used. In addition, because the input data ranges were different, they were converted to a value between 0 and 1 through min-max scaling.

4.1. Hyperparameter optimization

To select the optimal prediction model, hyperparameter optimization was performed, as described in Section 2.2. For the five hyperparameters, each tuning set was defined using through trial-and-error. Table 4 shows the tuning set for the hyperparameter optimization.

The model performance was evaluated based on the average validation loss. The validation loss was set as the mean square error, which a widely used indicator for evaluating model performance. The validation loss means the loss function value for the validation dataset. First of all, in this study, in order to prevent the overfitting of the prediction model, 10 % of the total data was used as the validation dataset as mentioned above. Unlike the training dataset, the validation dataset is not directly used in the training process but is used for monitoring and serves to evaluate how accurately the current model actually predicts. Therefore, by identifying the

Table 4
Hyperparameter set for optimization.

Hyperparameters	Tuning set
Number of hidden layers	3, 4, 5, 6, 7
Number of hidden nodes	10, 20, 30, 40
Learning rate	0.0001, 0.001, 0.01
Dropout	0, 0.25, 0.5
Gradient optimization	ADAM, SGD, RMSProp

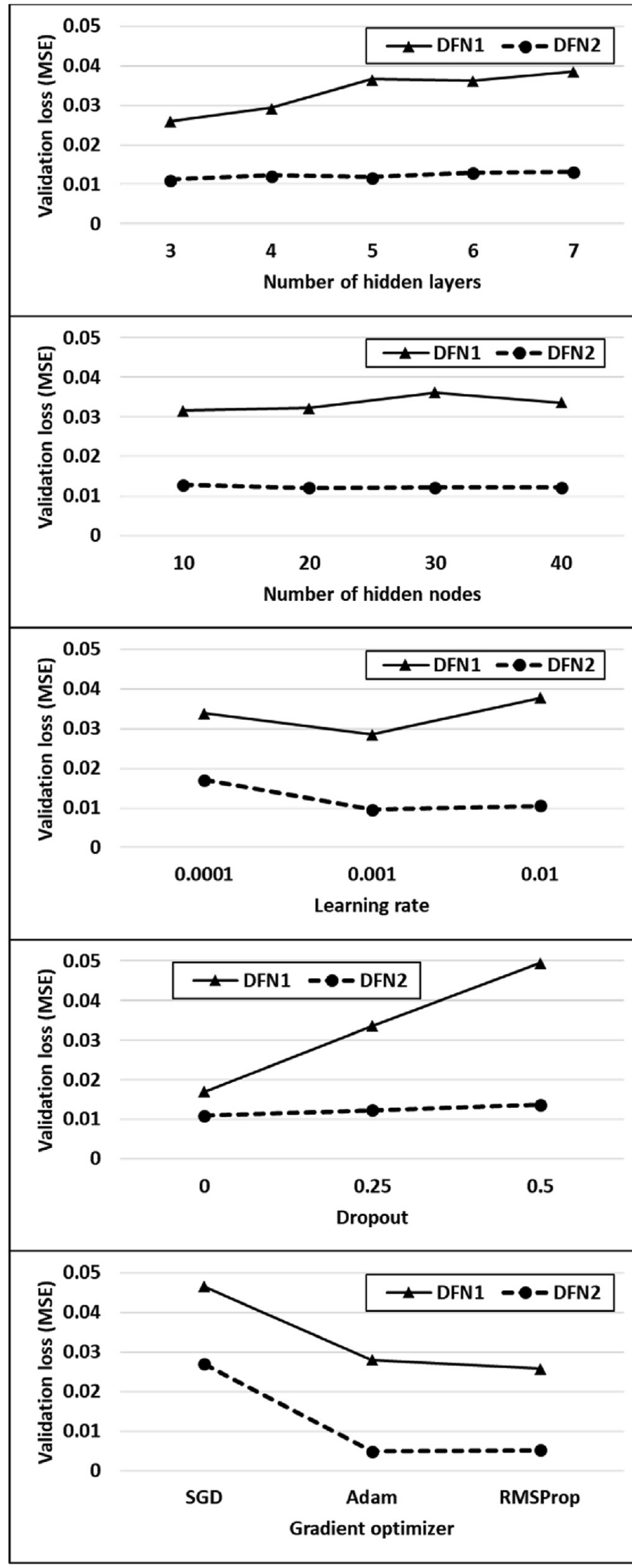

Fig. 5. Validation loss based on hyperparameters.

Table 5
Result of hyperparameter optimization.

Model	Number of hidden layers	Number of hidden nodes	Learning rate	Dropout	Gradient optimizer
DFN1	6	40	0.001	0	RMSProp
DFN2	4	40	0.001	0	RMSProp

validation loss, we can indirectly assess whether the current model is overfitted. Fig. 5 shows the average validation loss of DFN1 and DFN2 based on each hyperparameter.

After analyzing the accuracy by changing the number of hidden layers, it was confirmed that the average validation loss increased with the number of hidden layers. If the number of hidden layers is extremely high, then the number of unnecessary features will increase. Therefore, an accurate prediction will not be possible. Furthermore, we changed the hidden node and analyzed its accuracy. The average validation loss was the lowest when the number of hidden nodes was the smallest. However, no significant pattern was observed. The learning rate confirmed that the average validation loss was the lowest when the learning rate was 0.001. This result was obtained because an optimal solution could not be identified when a high value was selected; furthermore, it would fall into the local minimum when an extremely small value is selected. Subsequently, it was confirmed that the average validation loss increased with the dropout. In general, dropout is known to affect model regularization. However, the model proposed herein this study did not require regularization. Finally, by comparing the gradient optimizers, the SGD indicated the highest average validation loss, whereas ADAM and RMSProp indicated the lowest validation loss. It appeared that ADAM and RMSProp reduced the risk of falling into a local minimum.

In addition to the results shown in Fig. 5, the validation loss was calculated for all combinations of hyperparameters, as shown in Table 4, and the results were compared. A total of 540 cases ($5 \times 4 \times 3 \times 3 \times 3$) were compared. Table 5 shows the optimal hyperparameter results for DFN1 and DFN2. Comparing the overall results for hyperparameter optimization, it was discovered that the validation loss of DFN2 was generally lower than that of DFN1. This may be because the configuration of DFN2 is more suitable for predicting ship power than using a combination of hyperparameters.

4.2. Comparison with other prediction methods

We compared DFN models with other data-driven methods described in Section 1.2. The DFN models proposed herein were compared with prediction models using MLR and support vector regression (SVR). MLR (Draper and Smith, 1998), a representative regression analysis used for numerical prediction, predicts the results by the linear operation of independent variables. SVR (Vapnik, 2000), a regression analysis based on a support vector machine, is primarily used to predict nonlinear data. Table 6 shows a comparison of the accuracy of the prediction model using the DFN and prediction models using MLR and SVR. MLR-P, SVR-P, DFN1-P, and DFN2-P denote models trained using preprocessed input data. As described in Section 2.1, the preprocessing was carried out to convert ocean environmental data given as absolute values into relative values for the ship. Since this is to prepare the dataset before training, it does not affect the learning time.

As shown in Table 6, the DFN model (DFN2-P) is more accurate by 1.51 % (MLR) and 0.83 % (SVR). Comparing DFN1 and DFN2, it can be confirmed that it is more efficient to classify data using domain knowledge and then learn than to learn all the data as a fully connected layer at once. The complexity of predicting ship power cannot be expressed using a general regression model; hence, the DFN model is more suitable for predicting ship power. In addition, comparing the predicted results using DFN1-P and DFN2-P, the error of DFN2-P decreased by 0.13 % compared with that of DFN1-P. Hence, the configuration of DFN2 is more suitable for predicting ship power.

Table 6
Comparison of prediction results.

Prediction method	Pre-processing	MAE (%)
MLR	X	**2754 KW (5.12 %)**
MLR-P	O	2698 KW (5.01 %)
SVR	X	2326 KW (4.32 %)
SVR-P	O	2338 KW (4.34 %)
DFN1	X	1951 KW (3.62 %)
DFN1-P	O	1949 KW (3.62 %)
DFN2	X	1898 KW (3.52 %)
DFN2-P	O	**1876 KW (3.49 %)**

4.3. Effects of k-means clustering

Ship power is affected significantly by the sea state and ship operational status. They are typically classified based on the ship speed, ship draft, and Beaufort scale (wind speed) (Bialystocki and Konovessis, 2016). Therefore, k-means clustering was performed based on the three parameters above. Table 7 shows the results of k-means clustering for three clusters based on each parameter.

As shown in Table 7, k-means clustering was performed with k values set to 3, 4, 5, and 6. Next, the accuracy of the prediction model for each cluster was analyzed. Table 8 shows the prediction results for the ship power using k-means clustering. In Table 8, the MAE shown is the average MAE of each cluster model. For example, if the DFN model is segmented into three clusters, then the

Table 7
Result of k-means clustering ($k = 3$).

Clustering parameter	Cluster 1	Cluster 2	Cluster 3
Beaufort scale	0–4	5–7	8–12
Ship speed (knot)	0.1–12.3	12.4–18.0	18.1–29.7
Ship draft (m)	5.3–9.3	9.4–13.3	13.4–21.1

Table 8
Result of prediction along with clusters.

Number of clusters	Clustering criteria	MAE (%)
3	Beaufort scale	1925 KW (3.58 %)
4		2193 KW (4.07 %)
5		2024 KW (3.76 %)
6		2301 KW (4.27 %)
3	Ship speed	1950 KW (3.62 %)
4		1978 KW (3.67 %)
5		1996 KW (3.71 %)
6		2245 KW (4.17 %)
3	Ship draft	**1884** KW (3.50 %)
4		**2511** KW (4.67 %)
5		2034 KW (3.78 %)
6		2158 KW (4.01 %)

accuracy of the model is calculated as the average prediction accuracy of each cluster.

As the number of clusters increased, each cluster's training data decreased; hence, the accuracy decreased. Among the prediction results using k-means clustering, the best result (1884 KW (3.50 %)) was obtained when the data were segmented into three clusters based on the ship draft. However, the accuracy did not improve significantly compared with the results of DFN2-P in Section 4.2. As described above, the low accuracy might be attributed to the insufficient training data of each cluster as the cluster increased. Furthermore, it might have occurred because the effect to be obtained through k-means clustering had already been reflected in the DFN.

5. Conclusions and future works

Herein, we proposed a DFN model, which is a data-driven method, to predict ship power. In addition, four methods were used to improve the prediction accuracy, and a comparative analysis was performed. For training the DFN model, data from 13,000 TEU class container ships were used. By performing four treatments, the accuracy of the DFN model was improved. The first treatment was input data preprocessing. It was discovered that the prediction accuracies of the preprocessed and general prediction models were similar. The model with the most significant impact of preprocessing was MLR (0.11 % difference), and the model with the most negligible impact was DFN1 (no difference in accuracy). Therefore, it is concluded that using preprocessed data is efficient in terms of accuracy, but it is unnecessary to use it in terms of practicality. The second treatment was changing the DFN model configuration. When the DFN was changed based on the data characteristics (DFN2), the prediction accuracy increased slightly by 0.13 %. Hence, it was confirmed that the model configuration based on the data characteristics was more appropriate. In addition, the prediction results were compared with the regression analysis methods. The prediction error of the DFN model decreased by 1.51 % and 0.83 % compared with the results predicted using MLR and SVR, respectively. Hence, the DFN model exhibited better performance than the existing numerical prediction method. The third treatment was hyperparameter optimization. Among the various elements constituting the DFN, five hyperparameters that significantly affected the prediction result were defined, and optimization was performed. It was discovered that the hyperparameters that exhibited significant patterns were dropout and the number of hidden layers. The prediction accuracy was the best when dropout was not used, and the number of hidden layers was small. The last treatment was k-means clustering. As a result of k-means clustering, the prediction accuracy was the best when the data were segmented into three clusters for the ship draft, and the prediction error was 3.50 %. Therefore, k-means clustering did not significantly affect the power prediction of a 13,000 TEU class container ship.

To develop a more realistic prediction model in the future, the data of ship power during actual operation, including weather data (Kim et al., 2020a, Kim et al., 2020b, Kim et al., 2020c), must be obtained to be used as a label. Furthermore, a model suitable for predicting ship power should be developed by further improving the structure of a simple DFN model.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was partially supported by (a) the Korean Register, (b) the Research Institute of Marine Systems Engineering at Seoul National University, Republic of Korea, and (c) the Institute of Engineering Research at Seoul National University, Republic of Korea.

References

Abebe, M., Shin, Y., Noh, Y., Lee, S., Lee, I., 2020. Machine learning approaches for ship speed prediction towards energy efficient shipping. Appl. Sci. 10 (7). https://doi.org/10.3390/app10072325.
Ahlgren, F., Thern, M., 2018. Auto machine learning for predicting ship fuel consumption. Proceedings of the ECOS 2018. Guimaraes, Portugal.
Bialystocki, N., Konovessis, D., 2016. On the estimation of ship's fuel consumption and speed curve: a statistical approach. J. Ocean Eng. Sci. 1, 157–166. https://doi.org/10.1016/j.joes.2016.02.001.
Bottou, L., Bousquet, O., 2009. The tradeoffs of large scale learning. Proceedings of the 20th International Conference on Neural Information Processing, pp. 161–168. New York, United States.
Draper, N., Smith, H., 1998. Applied Regression Analysis. John Wiley & Sons.
Harvald, S.A., 1992. Resistance and Propulsion of Ships. Krieger Publishing Company.
Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R., 2012. Improving Neural Networks by Preventing Co-adaptation of Feature Detectors, 0580v1 arXiv:1207.
Holtrop, J., Mennen, G.G.J., 1982. An approximate power prediction method. Int. Shipbuild. Prog. 29, 166–170.
IMO, 2015. Revised guidelines for the onboard operational use of shipborne automatic identification systems (AIS). Resolution A 1106 (29).
International Maritime Organization, 2015. Ships and Marine Technology Guidelines for the Assessment of Speed and Power Performance by Analysis of Speed Trial Data.
James, B., Yoshua, B., 2012. Random search for hyper-parameter optimization. J. Mach. Learn. Res. 13, 281–305.
Kim, K.S., Roh, M.I., 2020. ISO 15016:2015-based method for estimating the fuel oil consumption of a ship. J. Mar. Sci. Eng. 8. https://doi.org/10.3390/jmse8100791.
Kim, D., Lee, S., Lee, J., 2020a. Data-driven prediction of vessel propulsion power using support vector regression with onboard measurement and ocean data. Sensors 20 (6). https://doi.org/10.3390/s20061588.
Kim, S.H., Roh, M.I., Oh, M.J., Park, S.W., 2020b. Estimation of ship operational efficiency from ais data using big data technology. Int. J. Nav. Archit. Ocean Eng. 12, 440–453. https://doi.org/10.1016/j.ijnaoe.2020.03.007.
Kim, K.S., Lee, J.B., Roh, M.I., Han, K.M., Lee, G.H., 2020c. Prediction of ocean weather based on denoising autoencoder and convolutional LSTM. J. Mar. Sci. Eng. 8. https://doi.org/10.3390/jmse8100805.
Kingma, D.P., Ba, J.L., 2015. Adam: a method for stochastic optimization. Proceedings of the International Conference on Learning Representations (San Diego, United States).
Kristensen, H.O., Lützen, M., 2012. Prediction of Resistance and Propulsion Power of Ships. Project No. 2010-56. Emissionsbeslutningsstøttesystem.
Lang, X., Mao, W., 2020. A semi-empirical model for ship speed loss prediction at head sea and its validation by full-scale measurements. Ocean. Eng. 209, 107494. https://doi.org/10.1016/j.oceaneng.2020.107494.
Lee, S.M., Roh, M.I., Kim, K.S., Jung, H.R., Park, J.J., 2018. Method for a simultaneous determination of the path and the speed for ship route planning problems. Ocean Eng. 157, 301–312.
Liang, Q., Tvete, H.A., Brinks, H.W., 2019. Prediction of vessel propulsion power using machine learning on AIS data, ship performance measurements and weather data. J. Phys.: Conf. Ser. Proceedings of the Journal of Physics: Conference. Trondheim, Norway 1357. https://doi.org/10.1088/1742-6596/1357/1/012038.
Lloyd, S.P., 1982. Least squares quantization in PCM. IEEE Trans. Inf. Theor. 28, 129–137. https://doi.org/10.1109/TIT.1982.1056489.
Panapakidis, I.P., Sourtzi, V.M., Dagoumas, A.S., 2020. Forecasting the fuel consumption of passenger ships with a combination of shallow and deep learning. Electronics 9. https://doi.org/10.3390/electronics9050776.
Perera, L.P., Mo, B., 2018. Ship speed power performance under relative wind profiles in relation to sensor fault detection. J. Ocean Eng. Sci. 3, 355–366. https://doi.org/10.1016/j.joes.2018.11.001.
Rakke, S.G., 2016. Ship Emissions Calculation from AIS-Annotated. M.Sc. Thesis. Norwegian University of Science and Technology, Norway.
Roh, M.I., 2013. Determination of an economical shipping route considering the effects of sea state for lower fuel consumption. Int. J. Nav. Archit. Ocean Eng. 5 (2), 246–262. https://doi.org/10.2478/IJNAOE-2013-0130.
Ruder, S., 2016. An Overview of Gradient Descent Optimization Algorithms, 04747 arXiv:1609.
Samsung Heavy Industries, 2017. Energy Efficiency Management System (Seoul, Korea).
SNAME, 1989. Principles of naval architecture. The Society of Naval Architects and Marine Engineers.
Szelangiewicz, T., Żelazny, A., 2017. Prediction power propulsion of the ship at the

stage of preliminary design part II: mathematical model ship power propulsion for service speed useful in the preliminary design. Manag. Syst. Prod. Eng. 25 (4), 231–236. https://doi.org/10.1515/mspe-2017-0032.

Uyanik, T., Arslanoglu, Y., Kalenderli, O., 2019. Ship fuel consumption prediction with machine learning. Proceedings of the 4th International Mediterranean Science and Engineering Congress, pp. 757–759.

Vapnik, V.N., 2000. The Nature of Statistical Learning Theory, the Nature of Statistical Learning Theory. https://doi.org/10.1007/978-1-4757-3264-1.

Yoo, B., Kim, J., 2017. Powering performance analysis of full-scale ships under environmental disturbances. IFAC-PapersOnLine 50, 2323–2328. https://doi.org/10.1016/j.ifacol.2017.08.474.