Journal of Korea Trade Vol. 25, No. 4, June 2021, 17-36
https://doi.org/10.35611/jkt.2021.25.4.17
ISSN 1229-828X

A Baltic Dry Index Prediction using Deep Learning Models*

Sung-Hoon Bae
Department of Trade and Logistics, Chung-Ang University, South Korea
Gunwoo Lee
Department of Transportation and Logistics Engineering, Hanyang University, South Korea
Keun-Sik Park†
Department of International Logistics, Chung-Ang University, South Korea

JKT 25(4)

Received 1 April 2021
Revised 7 May 2021
Accepted 14 May 2021

Abstract

Purpose – This study provides useful information to stakeholders by forecasting the tramp shipping market, which is a completely competitive market and has a huge fluctuation in freight rates due to low barriers to entry. Moreover, this study provides the most effective parameters for Baltic Dry Index (BDI) prediction and an optimal model by analyzing and comparing deep learning models such as the artificial neural network (ANN), recurrent neural network (RNN), and long short-term memory (LSTM).

Design/methodology – This study uses various data models based on big data. The deep learning models considered are specialized for time series models. This study includes three perspectives to verify useful models in time series data by comparing prediction accuracy according to the selection of external variables and comparison between models.

Findings – The BDI research reflecting the latest trends since 2015, using weekly data from 1995 to 2019 (25 years), is employed in this study. Additionally, we tried finding the best combination of BDI forecasts through the input of external factors such as supply, demand, raw materials, and economic aspects. Moreover, the combination of various unpredictable external variables and the fundamentals of supply and demand have sought to increase BDI prediction accuracy.

Originality/value – Unlike previous studies, BDI forecasts reflect the latest stabilizing trends since 2015. Additionally, we look at the variation of the model's predictive accuracy according to the input of statistically validated variables. Moreover, we want to find the optimal model that minimizes the error value according to the parameter adjustment in the ANN model. Thus, this study helps future shipping stakeholders make decisions through BDI forecasts.

Keywords: Artificial Neural Network, Baltic Dry Index, Big Data, Long Short-Term memory, Recurrent Neural network
JEL Classifications: C45, F17, L91

1. Introduction

The tramp shipping market is a perfectly competitive market with a relatively low entry barrier. the market flourished in the mid-2000s until right before the financial crisis due to competitive new building development, China's massive absorption of freight volume for raw materials, and intensified demurrage from lack of port facilities. The shipping market then rapidly declined due to the global financial crisis in the United States in 2008. The decline was

* This research was supported by the 4th Educational Training Program for the Shipping, Port and Logistics from the Ministry of Oceans and Fisheries.
† **Corresponding author:** pksik0371@cau.ac.kr

caused by oversupply from the accumulated ship supply because the market has a low entry barrier. Meanwhile, the Baltic Dry Index (BDI), which had made slow progress, constantly decreased due to the effect of the oversupply, hitting an all-time low in 2016.

The tramp shipping market, which had recorded a low market in the 2010s, is hitting the highest level in 11 years after resolving the typical oversupply due to COVID-19. The BDI then has been rising to its maximum value since 2010.

Moreover, the tramp shipping market is one in which ships do not regularly sail fixed routes. They must respond to irregular freight requests of various shippers and match the right vessels by meeting the shipping demand. In addition, they are affected by other exogenous variables such as macroeconomic variables (i.e., LIBOR interest rate, exchange rate, and international oil price). In other words, the shipping market shows high volatility (Stopford, 2008), and the decision-making of shipping companies and other relevant firms participating in the shipping market is based on such movements of market conditions, making them highly sensitive to the prediction and forecasts of market conditions. Various factors instead of one specific factor have complex effects on tramp shipping market conditions; therefore, not only predicting market conditions but also making decisions in work-site operations are difficult. Even if demand exceeds supply, ship orders cannot increase, and rising raw material prices inevitably do not have a positive effect on the shipping market. All of these are mutually influenced, and the factor that can best represent these complex factors is BDI.

Therefore, forecasting market conditions with high accuracy is crucial for shipping companies. Moreover, market forecasting models perform a key role in corporate management and investment strategies (Yu and Bulut, 2019; Shin Sung-Ho, Lee Paul Tae-Woo and Lee Sung-Woo, 2019). In particular, industrial and academic circles showed significant interest because the volatility of market conditions directly affects the profitability of market participants, and various efforts have been made in studies predicting market conditions (Celik et al., 2009).

The BDI is a typical index used to predict ocean freight market conditions. An increase in the ocean freight index is accompanied by an increase in profits of shipping companies and ship demand and orders. It also exerts a considerable impact on the domestic real economy by improving the performance of the shipbuilding and steel industries. Therefore, both institutions doing business in the shipping finance and relevant institutions that must capture global real economy trends should enhance the competency to analyze and forecast the ocean freight market.

Econometric analysis, which is used as a predictive analysis method, has been widely used with price fluctuations and many variables. For example, in the trend of price fluctuations, models such as autoregressive integrated moving average (ARIMA) and generalized autoregressive conditional heteroskedasticity (GARCH) were used. Moreover, various variables such as international oil prices, exchange rates, world industrial output, and raw material prices were used to improve the accuracy of prediction.

The time series analysis using BDI, the typical index for tramp shipping market conditions, has been a major concern of several studies for a long time. The ARIMA model was mostly used in time series analysis, along with regression analysis for variable selection and vector autoregressive (VAR) and vector error correction model (VECM) models to increase the predictive power.

Several studies have considered univariate and multivariate models provided that the latter show better performance in terms of prediction accuracy (Tsioumas et al., 2017; Kagkarakis, Merikas and Merika, 2016). Multivariate models such as VAR and VECM that predict BDI with several variables were examined (Franses and Veenstra, 1997; Pelagidis and Tsahali, 2019; Yin, Luo and Fan, 2017; Xu, Yip and Marlow, 2011; Lin, Chang and Hsiao, 2019; Zhang and Zeng, 2015). Moreover, various studies on the relationship between freight and charter rates in tramps and between charter rate and FFA verified the effect of each variable on the freight rate using the VECM (Franses and Veenstra, 1997; Zhang and Zeng, 2015; Yin, Luo and Fan, 2017).

However, classifying various factors affecting shipping market conditions and collecting data are difficult; thus, the current study uses deep learning models such as artificial neural networks (ANN) (Jain, 2011) with high predictive power that can be analyzed and predicted. This study aims to predict a BDI using deep learning models based on the selected factors affecting BDI.

2. Literature review

2.1. Prediction Using Deep Learning Models in Other Industries

Studies on predicting shipping industries using deep learning models are still limited. Therefore, this study reviews and summarizes previous research in other fields such as stock price prediction, tourism demand forecast, and energy consumption prediction (Selvin et al., 2017; Cheng and Yang, 2021; Vidya and Prabheesh, 2020; Yucesan et al., 2021; Khadhir, Kumar and Vanajakshi, 2021).

Selvin et al. (2017) examined the stocks of companies listed on the New York Stock Exchange at 100-minute intervals to predict stock prices after 10 minutes. They used Infosys, TCS, and CIPLA data from July to November 2014, and selected ARIMA, RNN, LSTM, and CNN. The results showed the superiority of the CNN models. Meanwhile, Cheng and Yang (2021) used LSTM and GRU models that can handle both time series and nonlinear problems instead of the Arps decline curve, which is commonly used for oil well prediction. Results of their study revealed that LSTM shows better performance when the variables were many and data were large, whereas the GRU model shows excellent performance when the data were small.

Vidya and Prabheesh (2020) measured the trade interconnection between countries before and after COVID-19 and predicted future trade directions using ANN. Their study results showed that China continues to maintain its trade center despite COVID-19. Moreover, ANN predicted that both exports and imports would decline in all countries by December 2020. Meanwhile, Yucesan et al. (2021) studied an optimal deep learning model to predict the demand for natural gas to increase and minimize economic losses, such as storage costs and contracts in the future. Among the various models, the hybrid model of SARIMAX–ANN exhibits best performance, followed by the hybrid model of ARIMA–ANN. However, the worst result was the hybrid model of the genetic algorithm–ANN. In the future, the temperature, wind speed, and industrial production index were added as additional explanatory variables to provide implications for improving the model's performance.

Khadhir, Kumar and Vanajakshi (2021) conducted a study to determine the location of the vehicle using the GPS system installed on the bus. Results of the study of the spatio-temporal LSTM model verified that the prevalence of traffic delays is highest at six intersections, bus terminals (spatio), and morning/afternoon peaks (temporal). Therefore, traffic congestion and delay are expected to be eliminated by introducing excess buses to other routes, according to the visualization tool presented in the study.

Moreover, LSTM has shown the best performances among deep learning models with time-series data such as forecasting tourists in Jiuzhaigou, China, predicting the trend of nuclear power plant parameters, and predicting ship tracks (Zhang et al., 2020; Bae Jun-Yong, Ahn Jee-Yea and Lee Seung-Jun, 2019; Tang, Yin, and Shen, 2019).

2.2. Prediction Using Deep Learning Models in The Logistics-Shipping Industry

The following studies compared and verified the forecast accuracy of econometric models and deep learning models (Mostafa, 2004; Yun Hee-Sung, Lim Sang-Seop and Lee Ki-Hwan, 2018; Zhang, Xue, and Stanley, 2018).

Mostafa (2004) used monthly net tonnage data from June 1975 to June 1998 to predict the throughput of the Suez Canal. Comparing the result of using the traditional ARIMA and ANN models, he determined that the RMSE of the ANN models is lower than that of the ARIMA models, thereby proving superiority. Furthermore, the predictive power of the ANN models can vary by determining the number of input layers. Meanwhile, Yun Hee-Sung, Lim Sang-Seop and Lee Ki-Hwan (2018) investigated the valuation of options for time charter party extension. They used the Black–Scholes model (BSM) and ANN. Their result revealed that ANN shows higher correlation and lower root mean square error than the BSM, thereby confirming ANN's superiority.

In addition, Zhang, Xue, and Stanley (2018) compared and verified econometric and ANN models using the daily, weekly, and monthly BDI data from 1999 to 2018. The econometric models had greater predictive power than the ANN models in the daily (t+1) prediction, but the ANN models showed relatively better results in the daily (t+7) and weekly/monthly (t+1, t+7) prediction. Moreover, the Back Propagation Neural Network (BPNN) model showed superiority in both short and long-term predictions. Furthermore, Şahin et al. (2018) used three methods for BDI prediction. The first method was to verify the BDI by applying the past observation value of BDI(BDI_{t-1}), and the second was to apply the last two observation values of the BDI(BDI_{t-1} and BDI_{t-2}) and compare with BDI. The third method was to compare the index of the previous observation of the BDI and Brent oil price (BDI_{t-1} and $Brent\ oil\ price_{t-1}$). The results proved the superiority of the second model with the lowest MAPE. Meanwhile, Kamal et al. (2019) extracted weekly data from BCI, BPI, BSI, bunker price, and charter rate per route for BDI prediction. They used the Pearson correlation analysis to classify variables with a correlation of more than 0.7, and they examined DNN and LSTM using verified data. The results demonstrated the superiority of DNN with lower RMSE than the LSTM, implying the need to develop various models using ensemble methods and support vector machines (SVM).

Finally, the following studies used ANN models to predict other dependent variables besides BDI (Gurgen, Altin, and Ozkok, 2018; Tsai and Huang, 2017).

Gurgen, Altin, and Ozkok (2018) used ANN to compare the actual and predictive values of five output variables (LOA, LBP, Breadth, Draught, and Freeboard) of chemical tankers using deadweight tonnage and vessel speed, which are valued by ship-owners, as the input variables. The comparison result revealed that the actual and the predicted values are similar, suggesting the usability of two input variables (e.g., deadweight tonnage and vessel speed) for the preliminary design of vessels in the future. Meanwhile, Tsai and Huang (2017) selected 10 major ports of Asia for the prediction of container volume. The analysis used a multilayer perceptron (MLP), with variables such as GDP, exchange rate, economic growth rate, industrial production index, GDP per capita, import volume, and export volume. The results revealed the fewest errors in predictive values of import/export container volume in the Port of Hong Kong, proving the high potential of use.

2.3. Implications

The following are the limitations of previous studies. First, many studies have used BCI, BPI, BSI, and BHSI, which are sub-elements of BDI, as variables used to predict BDI. Second, a period exceeding at least one week must be forecasted in consideration of freight rate liquidity in the shipping market. However, in previous studies, many predictions were made one day later. Third, various variables were used for BDI prediction without a detailed explanation of the variable selection process. Fourth, all independent variables were collectively inputted to the BDI prediction; therefore, the effect of the variable combination cannot be presented. Lastly, the use of the model in the future is limited because overfitting of the selected model is not verified.

To overcome the limitations of previous studies, the present study is conducted in five aspects. First, this study excluded BCI, BPI, BSI, and BHSI that are announced simultaneously as subcomponents of BDI. Moreover, this study used external variables that affect market conditions. Second, considering that direct transactions are difficult in the shipping market, this study used the method for predicting a week after, thereby providing more useful and available models for stakeholders in the shipping market. Third, this study used items that have significance with BDI among independent variables used in previous studies and from which data can be collected. The variables were statistically analyzed using correlation analysis, multiple regression analysis, and Granger causality test. Fourth, by conducting additional research selecting the combinations of optimal variables, this study first selected the hyperparameter value with the lowest prediction error in each model, based on which, it established the final model through combinations of variables. Finally, to validate the overfitting of the selected model, this study compared the recent actual and predictive values of BDI to increase usability.

3. Methodology

Time series forecasting using machine learning methods such as ANN based on big data has begun to be implemented to compensate for the deficiencies of the traditional time series models (Zhang, Xue, and Stanley, 2018; Kamal et al., 2019; Tsai and Huang, 2017; Gurgen, Altin and Ozkok, 2018; Zhang et al., 2020).

Unlike the traditional time series models, forecasting time series data using deep learning models enables analysis and prediction, without considering the constraints on the distributions of error terms, assumptions of linearity among variables, and identification issues, thereby having a wide application of the model and high predictive power (Jain, 2011). The typical ANN model shows superior prediction performance compared to the VAR model in predicting BDI freight rate (Batchelor, Alizadeh and Visvikis, 2007). Additionally, BDI prediction using the RNN model shows better prediction performance to the classic econometric models, and time-series data by LSTM rather than other models (e.g., random decision trees and random walk model) show better prediction performance (Nelson, Pereira, and de Oliveira, 2017).

Based on these advantages, deep learning models can be a powerful modeling tool to examine the complicated nonlinearity issues of time series by learning the weights through the learning process of data provided and forecasting the future. Therefore, this study considers ANN, RNN, and LSTM model for the BDI forecast.

3.1. ANN Model

ANN refers to a structure that analyzes data using a computer in a similar way as the neural network inside human brains. In other words, ANN's role is to determine a certain pattern by analyzing what is hidden in the data. The greatest advantage of ANN is that it has a superior capability of learning hidden patterns in the data compared to traditional methods.

Neurons create output values using a function f that is referred to as the activation function. In the end, the activation function f becomes the function of the value obtained by multiplying input data (x) by weight (w) and adding deviation (b). In other words, the output value $f(K)$ is as follows.

$$f(K) = f(w_i x_i + b) \tag{1}$$

3.2. RNN Model

Recurrent neural network (RNN) is one of the various methods of ANN developed to handle time series data. Traditional ANN has the assumption that the input data are independent from one another, whereas RNN is a neural network that can learn the corresponding relationship between output and input data according to time while handling time series data. The hidden state value of time t in RNN is the function of the hidden state value of time t-1 and the input value of time 1.

ANN receives only one unit of data as input value, whereas RNN receives information from the present and the past, using both data to create output value. Therefore, RNN can be regarded as a neural network that analyzes the present data with past memories. However, RNN is in a structure that has one tanh or ReLU activation function, and thus has long-term dependency issues in which a longer chain leads to vanished results learned from the past.

RNN operates according to the time flow, comprising input, hidden, and output layers. In other words, it calculates time t using the calculated result of time t-1 and calculates time t+1 using the calculated result. In other words, the input value of the hidden layer in RNN can be obtained as follows.

$$h_t = g_n(W_{xh}X_t + W_{hh}h_{t-1} + b_n) \tag{2}$$

Here, h_t is the hidden layer of time t, g_n is the activation function, W_{xh} is the weight matrix in which the input value is sent to the hidden layer, W_{hh} is the weight matrix in which h_{t-1} is sent to h_t in the hidden layer, $X_t$is the input value of time t, h_{t-1} is the hidden layer of time t-1, and b_n is the deviation or limit.

The following is the final output value:

$$Z_t = g_n(W_{hz}h_t + b_z) \tag{3}$$

Here, $Z_t$is the output vector, W_{hz} is the weight matrix when sent from the hidden layer to the output layer, and b_z is the deviation or limit.

RNN learns through the process known as propagation through time (BPTT). However, BPTT has the vanishing or exploding gradient problem. In other words, the vanishing gradient problem exists when the individual gradient of the hidden layer for previous values is smaller than 1, and the exploding gradient problem exists when the individual gradient is greater than 1.

In other words, the hidden layer of RNN remembers data from the past, but cannot remember data selectively. Memories fade with time because only the inputs of all moments with the same weight are remembered. This problem is known as the vanishing gradient problem, whereas the opposite case is called the exploding gradient problem (Gulli and Pal, 2017).

3.3. LSTM model

The LSTM (Hochreiter and Schmidhuber, 1997) network is an algorithm that solves the vanishing or exploding gradient problem of RNN. LSTM can overcome the vanishing gradient problem that may occur in learning long-term patterns and thus can handle larger circulation networks and difficult sequence problems.

The LSTM network has three types of gates to control cell state information, such as input, forget, and output gates. The input gate determines which new data to store in the cell state, and the forgotten gate determines which data to discard at the previous time. Finally, the output gate determines the output data.

$$\begin{aligned} f_t &= k_z\left(W_f x_t + U_f h_{t-1} + b_f\right) \\ i_t &= k_z(W_i x_t + U_i h_{t-1} + b_i) \\ o_t &= k_z(W_o x_t + U_o h_{t-1} + b_o) \\ c_t &= f_t \times c_{t-1} + i_t \times k_c(W_c x_t + U_c h_{t-1} + b_c) \\ h_t &= o_t \times k_h(C_t) \end{aligned} \tag{4}$$

Here, x_t is the input value of time t. f_t, i_t, o_t each indicates forgotten, input, and output gate of time t. C_t indicates cell state. The LSTM network solved the vanishing gradient problem and thus is suitable for predicting long-range dependent time series.

4. Dataset and Empirical results

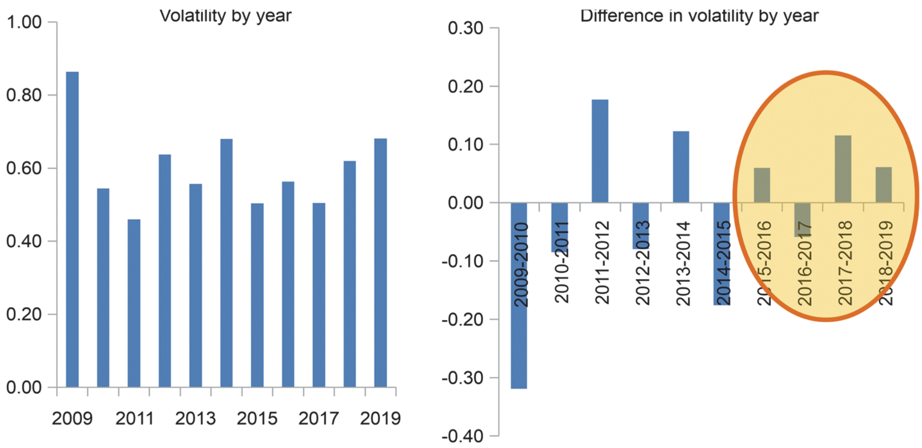
4.1. Dataset

The tramp shipping market conditions (BDI) have been stable at the BDI below 2,500, except for a rapid increase due to China's effect since the mid-2000s. In other words, the market showed a relatively stable balance from the 1980s until 2003. September 2019 shows a temporary increase due to the installation of scrubbers to comply with IMO 2020 taking effect starting 2020. Therefore, data from the period with high volatility (2009–2014) do not have to be used. This period is when the aftermath of the unprecedented situation (2003–2009) in the history of the shipping market conditions remained and when a bias existed. In contrast, to increase the predictive power in a situation where the market is stabilized, it may be more suitable to use the data from 2015 in which the bias has been eliminated as the market is becoming stable (see Fig. 1).

Moreover, in predicting shipping market conditions, the prediction for at least one week of time has significance in decision making. This is because, considering the fluidity of the freight market, direct transaction is highly unlikely to occur unlike stock trading, and thus, at least one week of precedence must be secured (Cooke et al., 2014).

Moreover, previous studies showed that the results of deep learning models showed the highest superiority in 8:2 data partitioning (Kamal et al., 2019; Zhang, Xue and Stanley, 2018). Therefore, from a total period of 25 years, the years 1995–2014 are classified as training data and 2015–2019 as validation data to predict a week after.

Fig. 1. BDI Volatility by Year Analysis



Source: Authors' calculation using Clarkson BDI data.

4.2. Variable Selection Process

The models used in studies predicting the shipping index can be classified into three types. First, a statistical approach constructs statistical function with supply and demand factors that affect the shipping index, while using the shipping index as a dependent variable. Second, the time series analysis uses autoregressive variables. Finally, the machine learning model exists, and most studies use ANN. Most studies support prediction performance superiority of the machine learning model (Eslami et al., 2017). Therefore, this study extracts and validates variables through a statistical approach and uses deep learning models for BDI forecast.

The variable selection process design for the empirical analysis of BDI is as follows. First, this study examines the level and direction of change among variables through correlation analysis using variables used in previous studies. Second, this study primarily selects variables when the significance level and VIF meet certain values through multiple regression analysis and multicollinearity test. Third, this study reviews stability for time series analysis with the selected variables and selects the final variables by checking whether they have causality with BDI through Granger causality analysis.

Among various factors proved to affect BDI, the current study used iron ore freight volume (Yin, Luo, and Fan, 2017), China's steel production (Tsioumas et al., 2017), new building development (Lee Sung-Yhun and Ahn Ki-Myung, 2018; Tsioumas et al., 2017), coal price (Bae Sung-Hoon, Ha Young-Mok and Park Keun-Sik, 2018), Brent oil price (Choi Ki-Hong and Kim Dong-Yoon, 2018), charter rate (Pelagidis and Tsahali, 2019), scrap price (Kagkarakis, Merikas, and Merika, 2016). Dow Jones Index and dollar/yen exchange (Kim Chang-Beom, 2011), LIBOR (Lee Sung-Yhun and Ahn Ki-Myung, 2018), China's GDP growth rate (Kim Do-Hee et al., 2019), China's industrial production index (Kim Chang-Beom, 2008), and the Clarkson Index (Han Min-Soo and Yu Song-Jin, 2019) as explanatory variables for multivariate analysis. Table 1 summarizes the data sources for the 16 independent variables and the dependent variable BDI.

Table 1. Data Sources

Variable	Definition	Unit	Source
BDI	Baltic Dry Index	Index	Baltic Exchange
ND	New Building Development	DWT/Mil	Clarkson
II	Iron Ore Export Volume (Korea, China, Japan Sum)	Ton/Mil	
CP	Coal Price	US$/Mil	Reuters
Bre	Brent Oil Price	US$/bl	Korea National Oil Corporation
Dow	Dow Jones Index	Index	Yahoo Finance
Dyen	Dollar/Yen Exchange	US$/¥	Reuters
Libor	Libor Index	%	
IPC	China Industry Production Index	%	Clarkson
CI	Clarkson Index	US$/daily	
CGDP	China GDP Growth Rate	%	
CSP	China Steel Production	Ton/Mil	
CT	Cape Time Charter	US$/daily	
PT	Panamax Time Charter	US$/daily	
ST	Supramax Time Charter	US$/daily	
CS	Cape Scrap Price	US$/Mil	
PS	Panamax Scrap Price	US$/Mil	

Notes: Dependent Variable= Baltic Dry Index.

For the relationship among factors affecting freight rate volatility of the tramp shipping market, this study first examined the correlation and direction among variables using correlation analysis and validated the effectiveness through multiple regression analysis. To solve the multicollinearity problem, VIF is validated to eliminate variables with VIF over 10. Finally, only the variables affecting BDI as antecedent variables in the Granger causality test are selected for analysis. The empirical analysis is conducted at the significance level of $p<0.01$ and $p<0.05$.

All selected variables had different units such as index, %, price, ton, DWT, and so on. As shown in Table 2, the natural log is used to for descriptive statistics. The analysis shows that LIBOR had the highest standard deviation and thus had a high dispersion, whereas dollar/yen exchange had the lowest standard deviation among the variables, thereby showing low volatility.

Table 2. Descriptive Statistics

Classification	Mean	Standard deviation	Skewness	Kurtosis	Jarque-Bera	Significance level
BDI	7.400	0.676	0.703	0.169	107.54	0.00***
II	3.782	0.694	-0.141	-1.514	127.33	0.00***
CSP	3.382	0.826	-0.365	-1.468	144.27	0.00***
ND	0.820	0.753	0.136	-0.466	15.793	0.0015***
CT	9.961	0.713	1.057	0.629	260.63	0.00***
PT	9.517	0.606	1.043	0.695	258.86	0.00***
ST	9.430	0.559	1.158	0.861	326.84	0.00***
CS	1.792	0.518	-0.417	-1.220	117.48	0.00***

Table 2. (Continued)

Classification	Mean	Standard deviation	Skewness	Kurtosis	Jarque-Bera	Significance level
PS	1.274	0.472	-0.391	-1.175	106.97	0.00***
CP	4.001	0.523	-0.011	-1.171	73.788	0.00***
Bre	3.910	0.622	-0.583	-0.528	88.126	0.00***
Dow	9.363	0.415	-0.014	0.075	0.306	0.86
Dyen	4.674	0.132	-0.792	0.204	136.78	0.00***
Libor	0.664	0.995	-0.380	-1.241	113.81	0.00***
CGDP	2.173	0.210	0.287	-0.501	31.315	0.00***
IPC	2.355	0.411	-0.419	-0.469	49.589	0.00***
CI	9.591	0.419	0.908	0.020	177.16	0.00***

Notes: 1. Dependent Variable: Baltic Dry Index.
2. $^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.001$.

The results of the Jarque-Bera test of normality revealed that most variables except the Dow Jones index do not follow a normal distribution at the 5% significance level. This indicates that the variables have high volatility in the market. Moreover, BDI and China's GDP growth rate, Clarkson Index, Cape, Panamax, and Supramax charter rates have positive skewness, indicating that the samples are leaning toward the left from the mean. In other words, the right tail is longer. Additionally, the kurtosis of all variables is smaller than 3, and thus, a fat-tailed distribution exists rather than the normal distribution.

4.3. Variable Selection of Shipping Market

To test the models, correlation analysis was first conducted among independent variables. According to analysis results, most variables had low correlation with BDI below 0.3. However, BDI had a correlation of 0.85 with the Clarkson Index, 0.71 with China's GDP, 0.68 with China's industrial production index, and 0.9 with a charter rate by ship type, all showing high correlation.

Second, according to the results of multiple regression analysis, iron ore freight volume, Cape charter rate, Panamax charter rate, Cape scrap price, Panamax scrap price, dollar/yen exchange, China's GDP growth rate, China's industrial production index, and Clarkson Index showed significance at the p-value of 5% with BDI.

However, multicollinearity among variables must be reviewed before the final variable selection. In other words, even though a high correlation exists between the dependent variable and other variables, certain variables must be excluded if multicollinearity exists through the validation of VIF. Therefore, iron ore freight volume, dollar/yen exchange, Cape charter rate, Panamax charter rate, Cape scrap price, and Panamax scrap price must be excluded from the analysis because they have VIF higher than 10 (Myers, 1990; Chatterjee and Price, 1991). In conclusion, China's industrial production index, Clarkson Index, and China's GDP were selected as variables with VIF lower than 10 and p-value below 5%.

Third, for statistical estimation with multiple observed time-series data, the samples must be assumed to be stationary. Stationary means that the stochastic properties of the time series model do not change according to time, indicating that the mean and variance of data are consistent, and the difference in data must depend only on time lag regardless of the point of time.

This study verified the stationarity of time series variables using an augmented Dickey–

Fuller (ADF) unit root test. If the stationarity was not confirmed, the data were verified again through the first derivative.

According to the unit root test results, the variable test results are greater than the threshold in level variables, and are non-stationary time series data. Therefore, we conducted a unit root test again after the difference-stationary process securing stationarity through the first derivative of these variables; the test results of all variables were smaller than the threshold with 1% significance level, thereby satisfying stationarity.

Finally, the Granger causality analysis, which is used to study the lead–lag relationship among variables using the distributed lag model, can be used when the relationship between explanatory and dependent variables is difficult to analyze using traditional regression analysis due to uncertain cause and effect.

This study conducted the Granger causality test to analyze causality by examining the relationship between BDI and the selected independent variables.

$$X_t = C_1 + \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=1}^{q} \beta_j Y_{t-j} + W_{1t} \tag{5}$$

$$Y_t = C_2 + \sum_{i=1}^{\gamma} \gamma_i Y_{t-i} + \sum_{j=1}^{q} \delta_j X_{t-j} + W_{2t} \tag{6}$$

To determine causality of X and Y, regression analysis is first conducted on X_t and Y_t with the past values and constant terms observed. Coefficients are identified using F statistics with the null hypothesis that no Granger causality exists among variables. Thus, rejecting the null hypothesis indicates no causality among variables and that they are mutually independent.

The Granger causality test results between explanatory variables and BDI are as described. The Clarkson Index impacted an antecedent variable on BDI after 1 week, and China's industrial production index and China's GDP impacted an antecedent variable on BDI after 2 weeks. Therefore, as antecedent variables that affect BDI, China's industrial production index, Clarkson Index, and China's GDP are used to conduct a multivariate analysis.

4.4. Univariate Analysis

For BDI analysis using deep learning models, out of a total 1,291 time-series data from January 1995 to December 2019, this study used 1,033 data samples from January 1995 to December 2014 as the training data. The data from January 2015 to December 2019 were then validated. To predict BDI, the ANN model used Deepnet library, and RNN and LSTM used Keras deep learning library and were developed through the R program (Arnold, 2017).

To fix the models when building RNN and LSTM, we must transform input data in three stages (Brownlee, 2017). The first stage is to align the price index of the past time stage (t-1) as input and the price index of this time stage (t) as output, thereby transforming time series input data into supervised learning data. The second stage is to transform input data to a measure of 0 to 1 using the Min-Max Scaler to improve the performance of predictive accuracy. The final stage is determining the variables of neural networks to find the optimal model using the transformed input data. Here, hyperparameter values of the models are estimated, which is an important process that determines the prediction performance. However, no theoretical method is provided, and it depends on the researcher's experience in repeated experiments and given data (Li and Parsons, 1997; Fan et al., 2013).

The deep learning models undergo the process in which the input signal of the hidden layer nodes delivered from the input layer node is transformed into the input variable of the output layer through the transfer function, and the function used is referred to as the activation function. This study used the sigmoid function, which is most widely used in ANN, as the activation function. In RNN and LSTM model design, research was conducted using the hyperbolic tangent (tanh) function verified as the activation function in previous studies (Gensler et al., 2016; LeCun, Bengio, and Hinton, 2015; Kim Do-Hee et al., 2019). The Adam algorithm was used as the optimization program.

Table 3 presents the design scope of hyperparameter values used in the analysis of the deep learning models. Moreover, the prediction model evaluation index is compared among values obtained for each case.

Table 3. Hyper–parameter Estimates

	ANN	RNN	LSTM
Normalization	MinMaxScaler	MinMaxScaler	MinMaxScaler
Learning Rate	0.01	0.01	0.01
Hidden layers	1-4	1-4	1-4
Epochs	1,000	1,000	1,000
Neurons	10-30	10-30	10-30
Batches	1-200	1-200	1-200
Activation Function	Sigmoid	Tanh	Tanh
Optimizer	Adam algorithm	Adam algorithm	Adam algorithm
Dropout Rate	0.2	0.2	0.2

Different results may occur due to different initial conditions in ANN; thus, this study compared how prediction errors (RMSE, MAE, MSE, and MAPE) are changed by changing the number of nodes (10, 20, and 30) and hidden layers (1, 2, 3 and 4) with the learning rate, number of epochs, and repeat count fixed. Moreover, the size of batches (16, 32, 64 and 128) and the ratio of dropouts (0.2) were applied to prevent overfitting of learned data and exclude interdependency within neurons. The hyperparameter values finally selected in each model are as described. In the ANN, prediction error was lowest when the batch size was 32, the number of nodes was 30, and the hidden layer was 2. In the RNN, prediction error was lowest when the batch size was 64, the number of nodes was 30, and the hidden layer was 3. In the LSTM, prediction error was lowest when the batch size was 32, the number of nodes was 20, and the hidden layer was 3.

This study compared the predictive power of the deep learning models applying selected hyperparameter values. Table 4 reports the results of univariate analysis on BDI. A comparison of MAPE (%) of RNN and LSTM reveals that LSTM has a structural benefit of approximately 16.21% improvement compared to RNN, and 17.52% improvement compared to ANN.

Table 4. Comparison of Univariate Prediction Errors

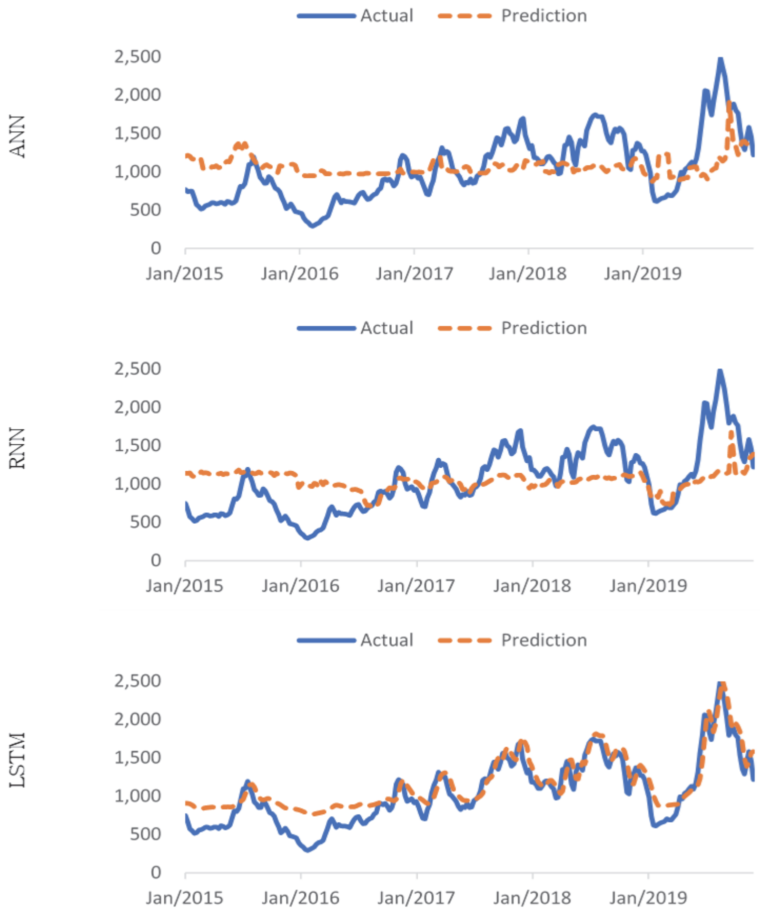
Classification	ANN	RNN	LSTM	ANN/LSTM(%)	RNN/LSTM(%)
RMSE	418.34	410.38	203.91	51.26	50.31
MAE	338.10	328.15	168.95	50.03	48.51
MSE	175,004.81	168,411.12	41.577.88	76.24	75.31
MAPE(%)	41.25	39.94	23.73	17.52	16.21

As a result, for univariate time series analysis using BDI, the predictive power of LSTM is greater than ANN and RNN.

In predicting freight rate of the tramp shipping market where most predictions are long-term forecasting, this study proved that LSTM is more suitable than the RNN model with long-term memory loss and the ANN model suitable for short-term forecasting. In other words, LSTM showed the highest predictive power in univariate BDI prediction among the three models.

Fig. 2 illustrates the predictive power of the univariate case for each model according to scenarios. Predictive values followed the pattern of actual values toward the RNN and LSTM specialized in time series forecasting, demonstrating the increasing predictive power.

Fig. 2. Comparison of Univariate LSTM Predictive Power (Unit: $/Daily)



4.5. Multivariate Analysis

In the multivariate ANN model, 1,033 samples of data were learned from January 1995 to December 2014, and the remaining 258 samples (January 2015 to December 2019) were used to test the predictive power. Normalization scope and use of hyperparameter values in the multivariate model design were applied in the same way as the univariate method. China's industrial production index, Clarkson Index, China's GDP growth rate, and BDI were combined to conduct an additional test for each model applying the finally selected hyperparameter values. The optimal model is determined through optimized hyper-parameter values for each model and combinations of variables.

The hyperparameter values of ANN, RNN, and LSTM for the finally adopted BDI prediction are as described. The following shows that the optimal hyperparameter values vary among models. In the ANN, prediction error was lowest when the batch size was 16, the number of nodes was 20, and the hidden layer was 3. In the RNN, prediction error was lowest when the batch size was 64, the number of nodes was 20, and the hidden layer was 2. In the LSTM, prediction error was lowest when the batch size was 32, the number of nodes was 30, and the hidden layer was 2.

The intention is to find the optimal combination of variables with selected hyperparameter values. The combinations of China's industrial production index, Clarkson Index, China's GDP growth rate, and BDI can be divided into seven types. Table 5 presents the predicted values of each combination. Finally, in the ANN model, optimal results were found in the combination of China's industrial production index, Clarkson Index, and BDI.

Table 5. Optimal Variable Selection Results for Each Model

Classification	Fixed value	Combination	RMSE	MAE	MSE	MAPE (%)
ANN	Number of nodes : 20 Batch : 16 Hidden lyaer : 3	IPC, CI, CGDP, BDI	243.63	183.74	59,353.79	26.61
		IPC, CI, BDI	213.06	184.60	45,395.75	25.62
		IPC, BDI	320.75	283.34	102,880.07	37.14
		IPC, CGDP, BDI	267.02	231.38	71,299.25	32.35
		CI, CGDP, BDI	372.75	306.47	138,943.13	44.44
		CI, BDI	263.95	206.20	69,670.95	31.40
		CGDP, BDI	283.64	216.70	80,449.05	33.46
RNN	Number of nodes : 20 Batch : 64 Hidden lyaer : 2	IPC, CI, CGDP, BDI	177.53	154.91	31,516.23	21.28
		IPC, CI, BDI	237.20	206.45	56,263.96	28.94
		IPC, BDI	313.01	277.74	97,978.14	38.64
		IPC, CGDP, BDI	257.07	221.12	66,086.78	31.01
		CI, CGDP, BDI	306.37	267.88	93,860.47	37.81
		CI, BDI	184.84	144.44	34,166.21	19.89
		CGDP, BDI	188.97	155.08	35,711.46	22.08
LSTM	Number of nodes : 30 Batch : 32 Hidden lyaer : 2	IPC, CI, CGDP, BDI	154.25	133.52	23,792.39	17.89
		IPC, CI, BDI	157.95	129.86	24,948.00	18.05
		IPC, BDI	225.11	181.08	50,676.26	26.88
		IPC, CGDP, BDI	188.71	159.27	35,609.91	21.29
		CI, CGDP, BDI	181.94	156.96	33,102.58	21.20
		CI, BDI	137.22	114.38	18,830.09	15.00
		CGDP, BDI	194.34	164.36	37,767.88	22.90

RNN and LSTM were also carried out in the same way as the ANN models above. The results in Table 5 reveal that RNN model shows the best results using the combination of Clarkson Index and BDI. LSTM model also showed the best results in the combination of Clarkson Index and BDI. As shown, 15.00% of LSTM's MAPE represents a predicted value of about ±15% of the actual BDI index from January 2015 to December 2019 during the test period.

A comprehensive summary based on the results of multivariate analysis is presented. Table 6 shows hyperparameter values of ANN, RNN, and LSTM for the finally selected BDI prediction and combinations of variables. The table shows that not only do the hyperparameter values vary among models, but the optimal combinations also vary among variables.

In other words, the smallest prediction error was found in the following combinations: China's industrial production index, Clarkson Index, and BDI for the ANN model; Clarkson Index and BDI for the RNN model; and Clarkson Index and BDI for LSTM.

Table 6. Finally Selected Hyperparameter Values and Combination of Variables

ANN			RNN			LSTM		
Neurons	Batch size	Hidden layer	Neurons	Batch size	Hidden layer	Neurons	Batch size	Hidden layer
20	16	3	20	64	2	30	32	2
IPC, CI, BDI			CI, BDI			CI, BDI		

Predictive power of each the deep learning models applying the selected hyperparameter values was compared. Table 7 presents the results of multivariate analysis on BDI.

Table 7. Comparison of Multivariate Prediction Errors

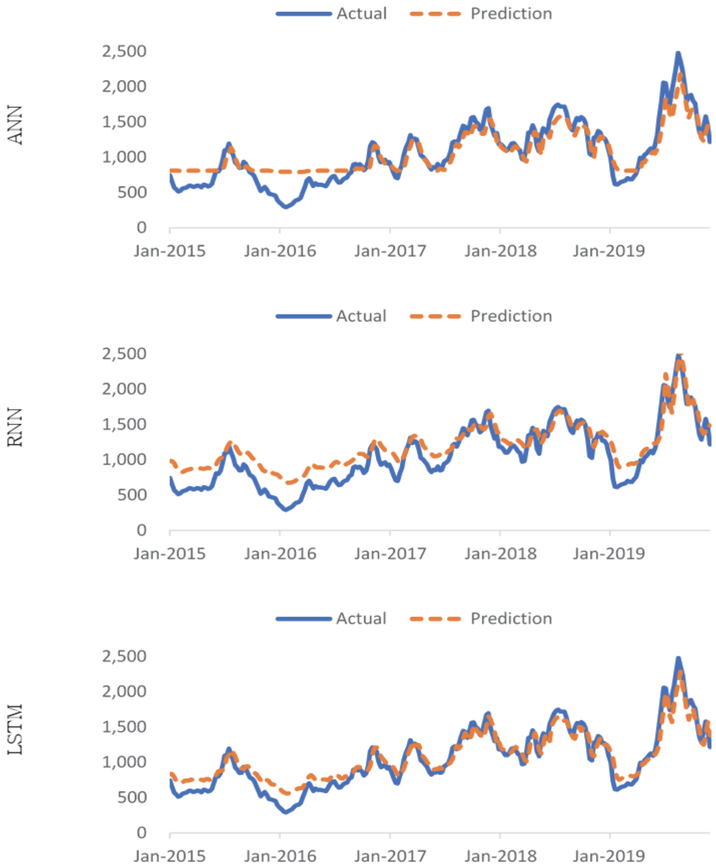
Classification	ANN	RNN	LSTM	ANN/LSTM(%)	RNN/LSTM(%)
RMSE	213.06	184.84	137.22	35.60	25.76
MAE	184.60	144.44	114.38	38.04	20.81
MSE	45,395.75	34,166.21	18,830.09	58.52	44.89
MAPE(%)	25.62	19.89	15.00	10.62	4.89

A comparison of MAPE (%) of RNN and LSTM provides that LSTM has a structural benefit of approximately 4.89% improvement, and 10.62% improvement between ANN and LSTM.

A comparison of predictive power among ANN models showed the superiority of LSTM that has an advantage in long-term forecasting in the multivariate analysis, as in the univariate analysis. Moreover, the prediction rate was highest in the combination of Clarkson Index and BDI through the comparison of the prediction error among variable combinations. In other words, predictive values vary among combinations and thus showed the importance among the combinations.

Fig. 3 illustrates the predictive power of multivariate analysis among the models. A difference in actual and predictive values exist in the ANN model, but it reduced toward the RNN and LSTM, with an increasing predictive power.

Fig. 3. Comparison of Multivariate LSTM Predictive Power (Unit: $/Daily)



4.6. Overfitting validation

To validate the finally selected model, the latest data are used to determine overfitting and compare prediction errors.

Overfitting validation is analyzed using the method of Kamal et al. (2019) in which the latest data are applied to the selected model. Therefore, the overfitting of the models is validated by comparing and analyzing the latest values of 2020 with high volatility.

Whether the selected model is overfitted is validated by predicting the latest values. As indicated in Table 8, for both the univariate and multivariate models, below the 6% difference in MAPE (%) exists when comparing the results from January 2015 to December 2019 and from January to June 2020. In other words, the slight difference of 6% implies the absence of overfitting.

January–June 2020 used in the overfitting validation section includes the period with poor market conditions due to COVID-19 and that in which they rapidly improved due to the expectation of economic recovery. This section showed a slight increase in MAPE compared to the result during the downward stabilization validation period (January 2015–December

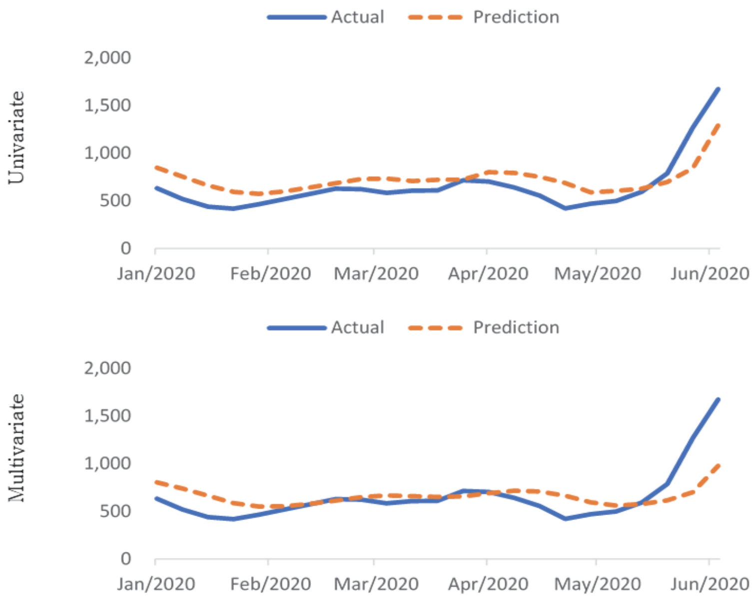
2019) with the univariate MAPE of 24.66 and multivariate MAPE of 20.92. However, although the slight decrease in the prediction rate, this model can be used as a forecasting model as it predicts the trend of rapid movement of the latest data.

Fig. 4 illustrates the results of overfitting validation.

Table 8. Comparison of Overfitting Validation Values

Model	Classification	Validation period	RMSE	MAE	MSE	MAPE (%)
LSTM	univariate	January 2015 - December 2019	203.91	168.95	41,577.88	23.73
		January 2020 -June 2020	182.97	152.64	33,479.44	24.66
	multivariate	January 2015 - December 2019	137.22	114.38	18,830.09	15.00
		January 2020 -June 2020	221.63	144.19	49,118.68	20.92

Fig. 4. Comparison of Overfitting Validation Values (Unit: $/Daily)



5. Conclusion and further research

The current study predicts BDI using deep learning models based on time series data from January 1995 to December 2019. For the analysis, 1,033 weekly data from January 1995 to December 2014 were used as learning data, and 258 weekly data from January 2015 to December 2019 were used as test data. In other words, the BDI time series data for 25 years were analyzed by dividing them into 8:2 of learning and test.

The results of this study can be summarized as follows. First, LSTM showed the highest predictive power among the three models, and the multivariate model showed better predictive power than the univariate model. Second, this study verified a difference in predictive power depending on the combinations of selected variables. Predictive power varied in the analysis through combinations of variables instead of overall variables in each model. Third, analyzing the latest values of January–June 2020 with the validated model using validation data yield a 6% difference in MAPE (%) when comparing the results from January 2015 to December 2019 and from January to June 2020, thereby passing the overfitting problem of the model.

However, from the results of this study, we would like to suggest the following limitations. First, deep learning models require many trial and errors we need to consider an alternative way to find the hyperparameter value that determines the optimal model. Therefore, the predictive accuracy is limited in that it changes according to the selection of the hyperparameter value. Second, predictive accuracy also varied among combinations selected variables. Therefore, combinations with high effectiveness and accuracy of BDI prediction must be constantly determined through simulation validation among variables in the deep learning models. Third, in addition to ANN, RNN, and LSTM, research on hybrid models such as convolutional LSTM should be actively conducted to improve the predictive accuracy. This study is anticipated to provide a reference for decision making and future investments in the tramp shipping market and the shipping industry in general.

References

Arnold, T. B. (2017), "kerasR: R Interface to the Keras Deep Learning Library", *Journal of Open Source Software*, 2(14), 296.

Bae, Jun-Yong, Jee-Yea Ahn and Seung-Jun Lee (2019), "Comparison of Multilayer Perceptron and Long Short-Term Memory for Plant Parameter Trend Prediction", *Nuclear Technology*, 206(7), 951-961.

Bae, Sung-Hoon, Young-Mok Ha and Keun-Sik Park (2018), "An Empirical Study on the Effect of the Factors Influencing on the Dry Bulk Freight Rate", *KOREA LOGISTICS REVIEW*, 28(5), 117-132.

Batchelor, R., A. Alizadeh and I. Visvikis (2007), "Forecasting spot and forward prices in the international freight market", *International Journal of Forecasting*, 23(1), 101-114.

Brownlee, J. (2017), *Long Short-term Memory Networks with Python: Develop Sequence Prediction Models with Deep Learning*, Machine Learning Mastery.

Celik, M., S. Cebi., C. Kahraman and I. D. Er (2009), "An integrated fuzzy QFD model proposal on routing of shipping investment decisions in crude oil tanker market", *Expert Systems with Applications*, 36(3), 6227-6235.

Chatterjee, S. and B. Price (1991), *Regression Diagnostics*, New York: John Wiley.

Cheng, Y. and Y. Yang (2021), "Prediction of Oil Well Production Based on the Time Series Model of Optimized Recursive Neural Network", *Petroleum Science and Technology*, 1-10.

Choi, Ki-Hong and Dong-Yoon Kim (2018), "Relationship between Baltic Dry Index and Crude Oil Market", *Journal of Korea Port Economic Association*, 34(4), 125-139.

Cooke, J., T. Young., M. Ashcroft., A. Taylor., J. Kimball., D. Martowski and M. Sturley (2014), *Voyage charters*, LLP.

Eslami, P., Ki-Hyo Jung., Dae-Won Lee and A. Tjolleng (2017), "Predicting tanker freight rates using parsimonious variables and a hybrid artificial neural network with an adaptive genetic algorithm", *Maritime Economics & Logistics*, 19(3), 538-550.

Fan, S., T. Ji., W. Gordon and B. Rickard (2013), "Forecasting Baltic Dirty Tanker Index by Applying Wavelet Neural Networks", *Journal of Transportation Technologies*, 3(1), 68-87.

Franses, P. H. and A. Veenstra (1997), “A cointegration approach to forecasting freight rates in the dry bulk shipping sector”, *Transportation Research. Part A: Policy and Practice*, 31(6), 447-458.

Gensler, A., J. Henze, B. Sick and N. Raabe (2016, Oct 9-12), “Deep Learning for solar power forecasting—An approach using Auto Encoder and LSTM Neural Networks”, In 2016 IEEE international conference on systems, man, and cybernetics (SMC), Budapest, Hungary

Gulli, A., and S. Pal (2017), *Deep Learning with Keras*, Packt Publishing Ltd.

Gurgen, S., I. Altin and M. Ozkok (2018), “Prediction of main particulars of a chemical tanker at preliminary ship design using artificial neural network”, *Ships and Offshore Structures*, 13(5), 459-465.

Han, Min-Soo and Song-Jin Yu (2019), “Prediction of Baltic Dry Index by Applications of Long Short-Term Memory”, *J Korean Soc Qual Manag*, 47(3), 497-508.

Hochreiter, S. and J. Schmidhuber (1997), “Long short-term memory”, *Neural computation*, 9(8), 1735-1780.

Jain, K. (2011), “A review study on urban planning & artificial intelligence”, *International Journal of Soft Computing and Engineering (IJSCE)*, 1(5), 101-104.

Kagkarakis, N. D., A. G. Merikas and A. Merika (2016), “Modelling and forecasting the demolition market in shipping”, *Maritime Policy & Management*, 43(8), 1021-1035.

Kamal, I. M., Hye-Rim Bae, Sung-Hyun Sim, Hye-Mee Kim, Do-Hee Kim, Yu-Lim Choi and Hee-Sung Yun (2019), “Forecasting high-dimensional multivariate regression of Baltic Dry Index (BDI) using Deep Neural Networks (DNN)”, *ICIC Express Letters*, 13(5), 427-434.

Khadhir, A., B.A. Kumar and L. D. Vanajakshi (2021), “Analysis of Global Positioning System Based Bus Travel Time Data and its Use for Advanced Public Transportation System Applications”, *Journal of Intelligent Transportation Systems*, 25(1), 58-76.

Kim, Chang-Beom (2008), “Economic Growth of China and Tramp Shipping Market”, *The Journal of Shipping and Logistics*, 56, 1-12.

Kim, Chang-Beom (2011), “The Effects of International Finance Market Shocks and Chinese Import Volatility on the Dry Bulk Shipping Market”, *Journal of Korea Port Economic Association*, 27(1), 263-280.

Kim, Do-Hee, Hye-Mee Kim, Sung-Hyun Sim, Yu-Lim Choi, Hye-Rim Bae and Hee-Sung Yun (2019), “Prediction of Dry Bulk Freight Index Using Deep Learning”, *Journal of the Korean Institute of Industrial Engineers*, 45(2), 111-116.

Lee, Sung-Yhun and Ki-Myung Ahn (2018), “Study on the Forecasting and Effecting Factor of BDI by VECM”, *Journal of Korean Navigation and Port Research*, 42(6), 546-554.

LeCun, Y., Y. Bengio and G. Hinton (2015), “Deep learning”, *Nature,* 521(7553), 436-444.

Li, J. and M. G. Parsons (1997), “Forecasting tanker freight rate using neural networks”, *Maritime Policy & Management*, 24(1), 9-30.

Lin, A. J., H. Y. Chang and J. L. Hsiao (2019), “Does the Baltic Dry Index drive volatility spillovers in the commodities, currency, or stock markets?”, *Transportation Research Part E: Logistics and Transportation Review*, 127, 265-283.

Mostafa, M. M. (2004), “Forecasting the Suez Canal traffic: a neural network analysis”, *Maritime Policy & Management*, 31(2), 139-156.

Myers, R. H (1990), *Classical and Modern Regression with Applications* (2nd ed.), Pacific Grove, CA: Duxbury Thompson Learning.

Nelson, D. M., A. C. Pereira and R. A. de Oliveira (2017, May 14-19), “Stock market's price movement prediction with LSTM neural networks”, In 2017 International joint conference on neural networks (IJCNN), Alaska, USA.

Pelagidis, T. and E. Tsahali (2019), “BDI's CORRELATION WITH LEADING ECONOMIC INDICATORS”, *Regional Science Inquiry*, 11(1), 167-189.

Şahin , B., S. Gürgen , B. Ünver and I. Altin (2018), "Forecasting the Baltic Dry Index by using an artificial neural network approach", *Turkish Journal of Electrical Engineering & Computer Sciences*, 26(3), 1673- 1684.

Selvin, S., R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon and K. P. Soman (2017, Sep 13-16), "Stock price prediction using LSTM, RNN and CNN-sliding window model", In 2017 international conference on advances in computing, communications and informatics (ICACCI), Manipal, India.

Shin, Sung-Ho, Paul Tae-Woo Lee and Sung-Woo Lee (2019), "Lessons from bankruptcy of Hanjin Shipping Company in chartering", Maritime Policy & Management, 46(2), 136-155.

Stopford, M. (2008). *Maritime economics* (3rd ed.), Routledge.

Tang, H., Y. Yin and H. Shen (2019), "A model for vessel trajectory prediction based on long short-term memory neural network", *Journal of Marine Engineering & Technology*, 1-10.

Tsai, F. M. and L. J. Huang (2017), "Using artificial neural networks to predict container flows between the major ports of Asia", *International Journal of Production Research*, 55(17), 5001-5010.

Tsioumas, V., S. Papadimitriou, Y. Smirlis and S. Z. Zahran (2017), "A novel approach to forecasting the bulk freight market", *The Asian Journal of Shipping and Logistics*, 33(1), 33-41.

Vidya, C. T. and K. P. Prabheesh (2020), "Implications of COVID-19 Pandemic on the Global Trade Networks", *Emerging Markets Finance and Trade*, 56(10), 2408-2421.

Xu, J. J., T. L. Yip and P. B. Marlow (2011), "The dynamics between freight volatility and fleet size growth in dry bulk shipping markets", *Transportation research part E: logistics and transportation review*, 47(6), 983-991.

Yin, J., M. Luo and L. Fan (2017), "Dynamics and interactions between spot and forward freights in the dry bulk shipping market", *Maritime Policy & Management*, 44(2), 271-288.

Yu, M. M. and E. Bulut (2019), "Performance evaluation of BDI forecasting models cross efficiency, the directional distance function and the AVs utility function", *International Journal of Transport Economics*, 46(1-2), 177-199.

Yucesan, M., Pekel, E., Celik, E., Gul, M. and F. Serin (2021), "Forecasting Daily Natural Gas Consumption with Regression, Time Series and Machine Learning Based Methods", *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 1-16.

Yun, Hee-Sung, Sang-Seop Lim and Ki-Hwan Lee (2018), "The value of options for time charterparty extension: an artificial neural networks (ANN) approach", *Maritime Policy & Management*, 45(2), 197-210.

Zhang, B., N. Li, F. Shi and R. Law (2020), "A deep learning approach for daily tourist flow forecasting with consumer search data", *Asia Pacific Journal of Tourism Research*, 25(3), 323-339.

Zhang, H. and Q. Zeng (2015), "A study of the relationships between the time charter and spot freight rates", *Applied Economics*, 47(9), 955-965.

Zhang, X., T. Xue and H. E. Stanley (2018), "Comparison of Econometric Models and Artificial Neural Networks Algorithms for the Prediction of Baltic Dry Index", *IEEE Access*, 7, 1647-1657.