K. Hiekata et al. (Eds.)

doi:10.3233/ATDE190113

# Predicting a Dry Bulk Freight Index by Deep Learning with Global Vessel Movement Data

Kei KANAMOTO[a,1], Yujiro WADA[b] and Ryuichi SHIBASAKI[a]
[a] *Department of Systems Innovation, the University of Tokyo, Japan*
[b,2] *Graduate School of Engineering, Hiroshima University, Japan*

**Abstract.** The Baltic Dry Index (BDI) is an indicator of freight rates of dry bulk cargo. Since freight rate fluctuates widely with reflecting a rapid change in the shipping market, it is important for a shipping company to predict the BDI. In the shipping industry, detailed data on the movement of each vessel is currently available in the automated identification system (AIS) on board, and a wide use of the AIS data is increasingly expected. This study proposes a new prediction method for freight index using deep learning and AIS data. The prediction target is the Baltic Capesize Index (BCI) representing the freight rate index of a large dry-bulk carrier over 180,000 DWT. The AIS data and various statistics are incorporated into the model to predict the rise or fall of the BCI value after 30 days. A multiple set of AIS data of the entire world and several specific regions are used. Furthermore, the number of related statistics to be incorporated is increased and a method of selecting them by introducing maximal information coefficient is shown. From the simulation results, the prediction of BCI can be performed with a certain degree of accuracy. In addition, the effect of introducing AIS data in the BCI prediction is confirmed.

**Keywords.** Shipping Freight Market, Baltic Dry Index, Satellite AIS Data, Deep Learning

## Introduction

There is a significant variation in vessel freight rates in the maritime shipping industry. Vessel orders have significantly increased between 2005 and 2008 because of increased demand, resulting in the highest profit accrued by the shipping industry throughout its history. On the other hand, the growth of the shipping industry has been slow since 2011 due to oversupply in the shipbuilding market and financial crisis. The impact of these are ongoing as shipping companies are currently faced with severe global competition. Therefore, it is important for shipping companies to develop a method of predicting freight rates and to examine decisions based on the predicted results for the sustainable development of the shipping industry.

Although many factors affect shipping market conditions, shipping demand generally depends on world economy. Additionally, weather not only directly

[1] Corresponding Author, Email: kei-kanamoto862@g.ecc.u-tokyo.ac.jp.
[2] At the time of research: Knowledge and Data System Department, National Maritime Research Institute

influences vessel operations but affects crop production as well as electricity demand. Hence, shipping demand greatly fluctuates seasonally.

Meanwhile, data usage becomes more efficient in the shipping industry as the Automatic Identification System (AIS) data is improved. Furthermore, a practical model that predicts a sharp change in transport fares due to various factors can be built based on the rapid developments of machine learning technology.

Therefore, this study aims to develop a method that predicts freight rates using deep learning and satellite AIS data. It incorporates a multiple set of AIS data for all the world as well as several specific regions, and proposes a new method for feature selection to be incorporated.

## 1. Literature Review

The prediction of freight charges is very important for the sustainable development of the shipping industry. Therefore, several researches on maritime freight prediction have been conducted.

Source [1] developed a Baltic Dry Index (BDI) predictive model using a vector autoregressive model with exogenous variables (VARX). They suggested that the VARX model performs better than ARIMA model and that the selected independent variables can significantly improve the accuracy of BDI prediction. Source [2] also developed a BDI predictive model using an improved support vector machine (SVM). They suggested that the proposed method has higher accuracy and could be used to predict the short-term trend of BDI in comparison with VAR model, ARMA model, and neural networks. Furthermore, source [3] developed a BDI predictive model based on numerical and textual data using deep learning. They suggested that the predictive accuracy is improved by combining numerical and textual data. These studies suggest that more practical predictive models can be developed using machine learning method.

Meanwhile, source [4] focused on the dry bulk market of Capesize carriers and developed a model that predicts the freight rate from three to 30 days later. The simulations showed that his proposed method is effective. In addition, he compared the results of models that incorporated satellite AIS data and models that did not incorporate it, and found that the predictive accuracy of the shipping market is significantly improved if satellite AIS data is incorporated.

In summary, several studies have predicted freight rates using deep learning. However, there is room for improvement in the amount of data to be input as well as feature selections. In this study, multiple sets of AIS data for all the world and several specific regions are used. Moreover, the number of related statistics to be used as input variables is increased and a maximal information coefficient (MIC) is introduced as a method of selecting them.

## 2. Development of BCI predictive model

### *2.1. Prediction target*

This study follows source [4] to construct a model to predict the Baltic Capesize Index (BCI) change for bulk carriers with over 180,000 deadweight tonnnage (DWT). In the tramp shipping market, a preparatory survey of shipping market conditions is often conducted one month before a regular contract. Therefore, prediction is limited to 30 days later for the BCI in this study, to develop a model with the highest accuracy based on [4] and preparatory examinations.

Furthermore, in the tramp shipping market, the fluctuations of time series BCI is more important than the value of the BCI itself. Hence, the authors focus on predicting a change in BCI from its previous status rather than its absolute value.

### *2.2. Model development*

Figure 1 shows a flow chart of the development of BCI predictive model.

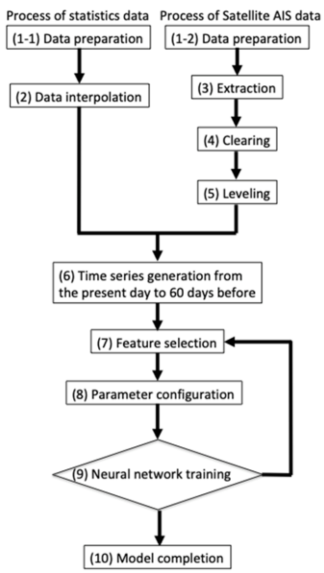



**Figure 1.** Overview of BCI prediction model development.

#### *2.2.1. Statistical data processing*

Table 1 shows the statistical data utilized in this study. Six other related statistics are examined as input values in this study in addition to the statistics used in a previous research [4] such as iron ore price, steel stock price, and West Texas Intermediate (WTI) oil price. For example, stock price is considered for China and the US since bulk carrier shipping is most active in these countries.

In case that some statistical data are missing when the Baltic or some stock exchange markets did not operate during the holidays. These missing values have been interpolated using the values of the previous day.

**Table 1.** Statistical data utilized in this study.

| **Data Name** | **Description** |
|---|---|
| Freight indicator | Baltic Capesize Index (BCI) |
| Steel stock price | Tokyo SE TOPIX17 Steel Stock Price |
| Iron ore price | Iron ore fines 62% Fe Cost and Freight (CFR) Futures |
| Crude oil price | WTI oil price |
| Crude oil price | Brent oil price |
| Shanghai stock price | Shanghai Sharp Electronics Co |
| Chinese steel index | China steel production cost index |
| Transport US equity investment result index | iShares Transportation Average Exchange Traded Funds (ETF) |
| Exchange rate | US dollar index future |
| Financial industry companies stock prices | iShares MSCI Global Metals & Mining Producers |

### *2.2.2. Satellite AIS data extraction and cleaning*

The AIS data is obtained from satellite and coastal base stations. However, the authors focused on the satellite AIS data which represents the navigation of a vessel on the sea, since it is considered related to freight rates, both of which indicate the degree of market activity. Since the number of dry bulk carriers in operation is related to the dry bulk market, the satellite AIS data for each dry bulk carrier with over 180,000 DWT are extracted from the database, where the vessel speed is more than three knot. Table 2 shows the conditions of the data extraction. Moreover, the target date of the model development is between January 1, 2016 and August 17, 2018, because of the availability of data.

The model considers the satellite AIS data not only for specific regions including Indian Ocean, Australian and Brazilian offshores, which are shown in Figure 2, but also throughout the world. The vessels moving in the target regions are filtered based on course of ground (COG), to focus on the vessels moving towards import countries such as China, Japan, and Australia.

Additionally, the satellite AIS data contains some errors such that the data needed to be processed. To increase the data reliability, the data whose vessel speed was more than 30 knot is removed. Moreover, since a large deficiency was observed for June 5, 2016, the aggregate value of the previous day is adopted.

Furthermore, the average speed, sum of DWT, and sum of product of DWT and ship speed for vessels in the target regions are calculated as input variables (items to be input) as with the previous research [4]. Additionally, the number of vessels and the sum of DWT of vessels operated all over the world are also calculated as input variables.

**Table 2.** Extraction condition of satellite AIS data.

| | |
|---|---|
| World | |
| Ship speed (knot) | $3.0 < S_{au} < 30.0$ |
| Indian Ocean (A1) | |
| Ship speed (knot) | $3.0 < S_{io} < 30.0$ |
| Course over ground (degree) | $10.0 \leq COG_{io} \leq 170.0$ |
| Latitude (degree) | $-28.58 \leq Lat_{io} \leq 2.41$ |
| Longitude (degree) | $63.37 \leq Lon_{io} \leq 83.67$ |
| Around Australia (A2) | |
| Ship speed (knot) | $3.0 < S_{au} < 30.0$ |
| Course over ground (degree) | $280.0 \leq COG_{au} < 360.0$ and $0.0 \leq COG_{au} \leq 80.0$ |
| Latitude (degree) | $-39.16 \leq Lat_{au} \leq -10.49$ |
| Longitude (degree) | $109.60 \leq Lat_{au} \leq 154.07$ |
| Around Brazil (A3) | |
| Ship speed (knot) | $3.0 < S_{au} < 30.0$ |
| Course over ground (degree) | $280.0 \leq COG_{bra} < 360.0$ and $0.0 \leq COG_{bra} \leq 80.0$ |
| Latitude (degree) | $-14.56 \leq Lat_{bra} \leq 2.79$ |
| Longitude (degree) | $-39.04 \leq Lat_{bra} \leq -24.30$ |

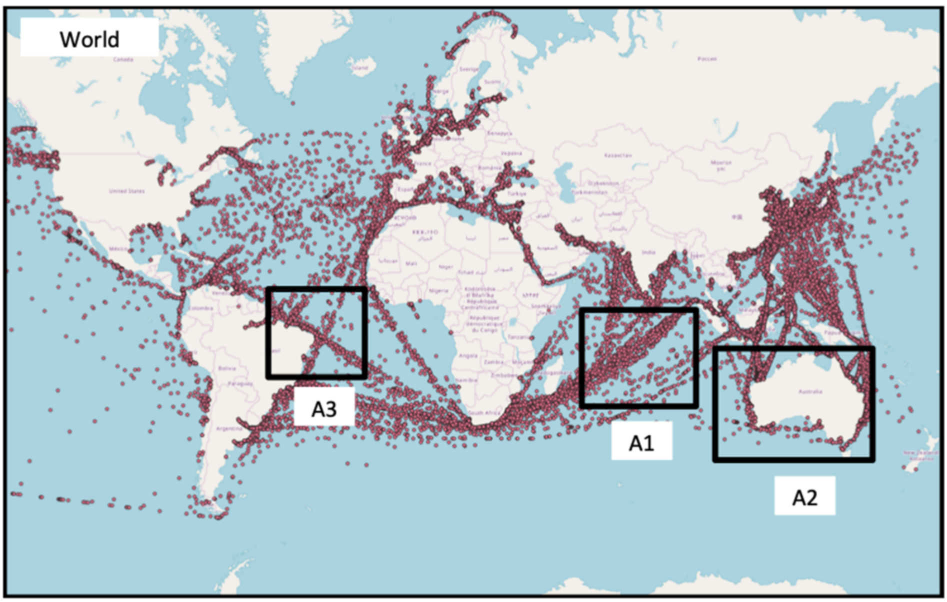



**Figure 2.** Target regions of satellite AIS data extraction.

## *2.2.3. Leveling*

It is difficult to comprehend the long-term trend of freight rate since raw satellite AIS data fluctuates daily. Therefore, the authors have incorporated the moving average of 14 days to reduce noise and ensure a longer trend.

## 3. Deep learning model for BCI prediction

### 3.1. Deep learning model

A model that predicts BCI is formulated as shown in Equations (1)-(3). The input data ($D$) comprises of statistical and vessel movement data of the present and past 60 days to reflect time series change, because one average voyage of long-distance dry bulk shipping takes about two months.

$$\Delta BCI_{t+pt} = BCI_{t+pt} - BCI_{t} \tag{1}$$

$$BCI_{t+pt} = f\left(D_{t}, D_{t-1}, \cdots, D_{t-60}\right) \tag{2}$$

$$D_{t} = \left\{S_{t}, Io_{t}, Au_{t}, Br_{t}, Wo_{t}\right\} \tag{3}$$

where $\Delta BCI$: differential value of BCI, $t$: time (day), $pt$: time difference (day), $f$: predictive model of $\Delta BCI$, $D$: set of input data, S: set of statistical data, $Io$: set of vessel movement data extracted for Indian ocean, $Au$: set of vessel movement data extracted for Australian offshore, $Br$: set of vessel movement data extracted for Brazilian offshore, and $Wo$: set of vessel movement data to cover the entire world.

Deep neural network (DNN) algorithm is applied to solve the problem outlined above. Although there are several deep learning models such as recurrent neural network (RNN) and long short-term memory (LSTM), these are not used due to insufficient amount of dataset for the time series model. Figure 3 shows the structure of DNN model. The selected statistical and vessel movement data are selected as input to predict the BCI 30 days later through hidden layers, based on the conditions shown in Section 3.3. Linear functions are used as the activation functions in the input and output layers, while a rectified linear unit (ReLU) function is used as the activation function in the hidden layer. All the hidden layers include a dropout layer, whose rate is set to 0.2, while all layers include batch normalization. Other conditions of the model calculation are shown in Table 3.

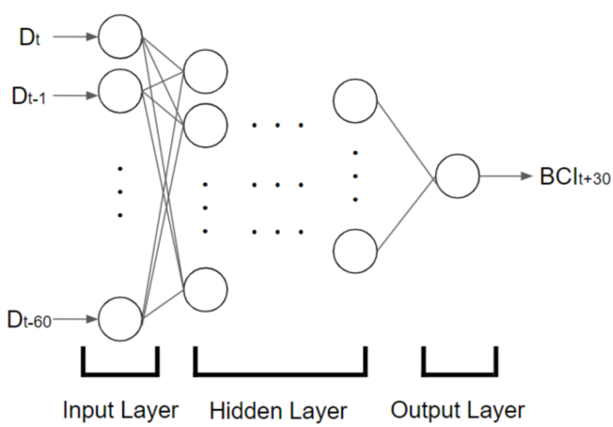



**Figure 3.** Structure of deep neural network (DNN) model.

**Table 3.** Specification of DNN model.

| | |
|---|---|
| Number of dataset for learning | 639 (selected from 2016/03/01 to 2018/05/07) |
| Number of dataset for validating | 159 (selected from 2016/03/01 to 2018/05/07) |
| Number of dataset for testing | 102 (from 2018/05/08 to 2018/08/17) |
| Method for normalization | Min-max normalization |
| Number of middle layers | 9 |
| Loss function | Root mean squared error (RMSE) |
| Optimizer | Adam |
| Batch size | 100 |
| Epoch number | 500 |

### *3.2. Model evaluation*

In the tramp market, the change in BCI is often more important than its absolute value since BCI represents a relative value based on its status as at January 4, 1985. Therefore, the authors focus on predicting the change in BCI. Correlation coefficient (CC), movement direction matched rate (MDMR), and root mean squared error (RMSE) are used as indices to evaluate the model. The MDMR is defined as the agreement rate between the predicted and observed movement direction (i.e. increasing or decreasing). For example, if the predicted $\Delta BCI$ is +100 and the observed $\Delta BCI$ is +30, then the movement directions correspond. Hence, the learning and validation datasets are separated to ensure an equal number of positive and negative directions, for avoiding bias.

### *3.3. Feature selection*

Twenty-one items consisting of 10 items from statistical data (Table 1) and 11 items from satellite AIS data are prepared. Nine items related to satellite AIS data are related to the specific regions, while two are related to the world sea region. Since each item includes 61 days data from that at the present day to 60 days before as described in Section 3.1, the maximum number of features in the time series data considered in this study is $21 \times 61 = 1281$. However, effective features for the BCI prediction should be selected to avoid overlearning the model with excessive input variables. Hence, three indices are introduced for feature selection such as CC, RMSE, and MIC. When the features are compared based on the top scores in each index, CC has a linear relationship while RMSE measures the difference between the estimated and observed values to describe their closeness. Meanwhile, MIC calculates mutual information between two features by dividing coordinate systems by grids and determine the strength of the linear or non-linear relationship. Here, mutual information $I(X,Y)$ is defined as

$$I(X,Y) = \sum_{y \in Y} \sum_{x \in X} \mathrm{p(x,y) \log p(x,y)/p(x)p(y)} \tag{4}$$

where $X$, $Y$: random variables, $p(x,y)$: joint probability distribution function, and $p(x)$, $p(y)$: marginal probability distribution functions. In the calculations shown in Section 4.2, MIC is adopted as the criteria for selecting features.

## 4. Simulation Results

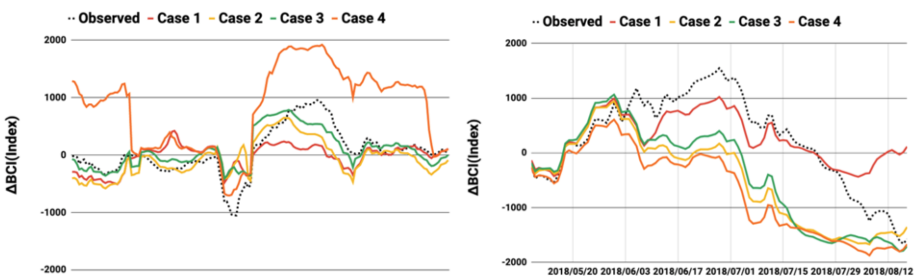
### *4.1. Impact of incorporating satellite AIS data*

The following simulations of BCI prediction are executed and their results compared, to determine the effect of satellite AIS data in BCI prediction.

Case 1: Incorporating all statistical and satellite AIS data (data related to the entire world and the specific regions).
Case 2: Incorporating all statistical and satellite AIS data, except those related to the entire world.
Case 3: Incorporating selected statistical and satellite AIS data compliant with [4].
Case 4: Incorporating all statistical data but no satellite AIS data.

Figure 4 shows the predicted results of the change in BCI 30 days later in each case for validation and test datasets. Table 4 shows the estimated evaluation indices in each case (i.e. CC, MDMR, and RMSE). The observed changes in BCI for the validation dataset are partially reproduced in Case 1, Case 2, and Case 3; however, it is not well reproduced in Case 4, judging from the worst figures on MDMR and RMSE.

Meanwhile, the three evaluation criteria indicate that Case 1 has the best reproducibility for the test dataset. It suggests that the introduction of satellite AIS data in BCI prediction is effective. Specifically, the predictive precision in Case 1 is better than in Case 2, highlighting the significance to incorporate satellite AIS data to cover the entire world. Furthermore, the predictive precision in Case 1 is better than in Case 3 (the previous study), highlighting a positive effect with an increase in the quantity of statistical data.




Validation dataset　　　　Test dataset

**Figure 4.** Comparison of the predicted results of BCI change.

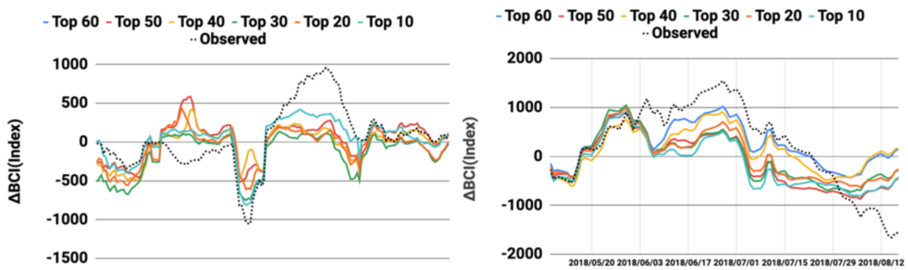
**Table 4.** Estimated evaluation indices of each case.

| | Validation dataset | | | Test dataset | | |
|---|---|---|---|---|---|---|
| **Case** | **CC** | **MDMR (%)** | **RMSE** | **CC** | **MDMR (%)** | **RMSE** |
| Case 1 | 0.436 | 67.9 | 371 | 0.807 | 95.0 | 554 |
| Case 2 | 0.805 | 65.4 | 327 | 0.688 | 67.6 | 971 |
| Case 3 | 0.820 | 84.3 | 231 | 0.755 | 80.3 | 883 |
| Case 4 | 0.821 | 54.1 | 883 | 0.691 | 49.0 | 1124 |

### *4.2. Impact of feature selection based on MIC*

The impact of feature selection based on MIC is examined after the preparatory calculations, as follows. All features with MIC score ≥ 0.5 are selected. Additionally, some features (i.e. 10, 20, 30, 40, 50, and 60) having the largest MIC score are also selected for each input item (even if their scores < 0.5), to ensure a variety of input data.

Figure 5 shows the results of predicted change in BCI 30 days later in each number of features adopted for both validation and test datasets. Table 5 shows the estimated evaluation indices in each number of features adopted. The model with top 10 features selected additionally in each input item for the validation dataset shows the best result based on CC and RMSE and the second best result based on MDMR. On the other hand, the model with top 60 features selected additionally in each input item of the test dataset has the best result based on the three evaluation indices. These results indicate that a selection based on MIC is effective for the validation process, while selection based on many features is effective for the test dataset. Meanwhile, Table 5 shows that the model with top 20 features selected additionally in each input item achieves some degree of accuracy for the validation and test datasets. These results indicate that a model with a small number of features and a certain accuracy can be developed. Moreover, a selection based on MIC can prevent model overlearning when the number of features is increased in the future.




Validation dataset  Test dataset

**Figure 5.** Comparison of predicted results in BCI change by number of features adopted.

**Table 5.** Estimated evaluation indices by number of features adopted.

| | | Validation dataset | | | Test dataset | | |
|---|---|---|---|---|---|---|---|
| **Number of additional selected features for each index** | **Total number of selected features** | **CC** | **MDMR (%)** | **RMSE** | **CC** | **MDMR (%)** | **RMSE** |
| 60 | 1187 | 0.436 | 67.9 | 371 | 0.807 | 95.0 | 554 |
| 50 | 992 | 0.453 | 64.7 | 362 | 0.758 | 80.4 | 651 |
| 40 | 798 | 0.534 | 68.5 | 340 | 0.764 | 72.8 | 612 |
| 30 | 626 | 0.570 | 59.7 | 362 | 0.685 | 80.4 | 651 |
| 20 | 464 | 0.501 | 66.0 | 367 | 0.751 | 81.4 | 622 |
| 10 | 315 | 0.826 | 66.7 | 237 | 0.703 | 67.6 | 712 |

## 5. Conclusion

In this study, a method to predict freight index using deep learning and satellite AIS data based on [4] s further improved. Freight index is important to make managerial judgement for shipping companies. In this study, the prediction target is the Baltic Capesize Index (BCI) representing the freight rate of large dry-bulk carriers over 180,000 DWT. Satellite AIS data and various statistics are incorporated into the model to predict the rise or fall of the BCI value after 30 days. The effectiveness of the proposed method is examined by comparing it with the estimated results. The key findings in this study are as follows:

(1) Simulation accuracy is improved by using satellite AIS data. Moreover, multiple set of AIS data of the entire world and added related statistics have positive effect in the prediction.
(2) Although the effect of the improvement of accuracy by incorporating a method of feature selection based on MIC is still not unclear, a model with a small amount of features to avoid overlearning with a certain accuracy can be developed.

This study also demonstrates that the expansion of statistical data will improve the predictive accuracy. Thus, the predictive accuracy may be improved by adding related data. Furthermore, although it is difficult to determine whether selection by MIC in each type of data item is effective or not, the priority of input data item indicators can be measured using MIC or other criteria. Meanwhile, not only economic indicators as introduced in this study but also other kinds of indicators such as weather information and ship accident can improve the model accuracy. Also, although the prediction is limited to 30 days later for the BCI in this research, sensitivity analysis on the duration how many days later should be examined. Another point to be further improved is that the effectiveness of deep learning method is not examined by comparing it with other methods in this study. If a time-series predictive model such as LSTM is developed for improved accuracy, higher data volume is required.

## References

[1] V. Tsioumas, S. Papadimitriou, Y. Smirlis and S.Z. Zahran, A Novel Approach to Forecasting the Bulk Freight Market, *The Asian Journal of Shipping and Logistics*, 2017, 33(1), pp. 33–41.
[2] Q. Han, B. Yan, G. Ning and B. Yu, Forecasting dry bulk freight index with improved SVM, *Mathematical Problems in Engineering*, 2014, 460684.
[3] T. Ling and T. Nagao, Using numerical data and textual data for dry bulk index forecasting, *The proceeding of the National Conventions of Information Processing Society of Japan*, 2017, 2, pp. 341–342.
[4] Y. Wada, Dry bulk freight index forecasting based on satellite AIS data using deep learning, *International Association of Maritime Economists 2019 Annual Conference*, Athene, Greece, 2019.
[5] D.N. Reshef, Y.A. Reshef, H.K. Finucane, S.R. Grossman, G. McVean, P.J., Turnbaugh and P.C. Sabeti, Detecting novel associations in large data sets. *Science*, 2011, 334(6062), pp. 1518-1524.
[6] H. Tanimoto, and R. Miyawaki, Yosenkeiyakunozhitumutekikaisetu [Practical commentary on the chartered contract] Seizando-Shoten. (In Japanese), 2016.