CrossMark

Seung Jae Lee
Kwang Hyo Jung
Namkug Ku
Jaeyong Lee

A comparison of regression models for the ice loads measured during the ice tank test

Summary

To evaluate the time-domain positioning performance of arctic marine structures, it is necessary to generate an ice load appropriate for the current position and heading of the structure. The position and orientation angle of a floating body continuously change with time. Therefore, an ice load is required for any attitude in the time-domain simulation. In this study, we present a fundamental technique for analyzing ice loads in the frequency domain based on data measured at various angles in the ice-water tank experiment. We perform spectral analysis instead of general FFT to analyze the ice load, which has the characteristics of a random signal. To generate the necessary ice load in the time domain, we must first interpolate the measured data in the frequency domain. Using the Blackman-Tukey method, we estimate the spectrum for the measured data, then process the data to generate the training set required for machine learning. Based on the results, we perform regression analysis by applying four representative techniques, including linear regression, random forest, or neural network, and compare the results with MSE. The deep neural network method performed best, but we provide further discussion for each model.

Key words: machine learning; regression; ice load, power spectral density; mean squared error

1. Introduction

Factors that determine ice load are the degree of how well the ice field is managed, ice thickness, type of ice, concentration, etc. During the design stage, one must determine the ice conditions of the sea area where the offshore structure is to be installed. This survey is then reflected in the ice tank test to observe the effects of ice on the structure. For floating structures, it is necessary to evaluate the performance related to the station-keeping in the time domain, and this requires consideration of the ice load. Mathematical models are mainly used, but there is a limit to their use in simulations that consider complex environmental loads that include wind, currents, and waves.

Calculating ice load can be broadly classified into methods based on a numerical model and methods based on an empirical formula taken from experiments or data measured in ice infested waters. Researchers have developed numerical models such as a finite element method, a discrete element method, a particle-in-cell method, a method using computational fluid dynamics, and a cohesive element method [1]. There is a well-organized overall survey on this topic in [2]. The empirical method derives a correlation between ice load and the parameters of ice and hull from experimental or measurement data from ice-covered seas. The advantage of this method is its numerical efficiency. For use in real-time simulation, a computation cannot take a long time. In addition, the development period and verification process should not be complicated, and the processing time should not be long. Numerical model–based methods generally do not satisfy these conditions.

The best way to get the ice load on a floater is to use the test results of a real sea experiment. The most widely recognized test was conducted in 2004, as part of the 302 Arctic coring expedition of the international ocean drilling program (IODP). One drillship performed drilling operations, assisted by two icebreakers. Ice load was measured with sensors and showed the difficulty of maintaining a position in the Arctic ocean [3]. Due to heavy costs and the impossibility of controlling ice conditions, it is very difficult to obtain data for certain situations. Considering these difficulties, Mikulec and Piehl [4] established a full-scale CFD model using the test results and verified the results. Tests performed in ice tanks can also generate good data to evaluate the ice load on floaters [5]. Representative examples of such tests are the drillship-related experiments carried out as part of the international joint research project DYPIC (DYnamic Positioning in Ice Condition) [6, 7] and the model tests performed by the National Research Council (NRC) of Canada [8, 9]. However, such ice tank tests are still costly, and they have limitations in reflecting other environmental loads, such as wind, currents, and waves.

Studies are ongoing to simulate the interactions between ice and structures, considering the limitations mentioned above [10]. A method of using machine learning techniques to predict ice resistance for ship performance analysis on sea ice is also presented [11]. One such trial combines existing commercial software with experimental data from the ice tank test [12]. Although it is still difficult to deduce the ice load due to the lack of data sets, the method can be used as a framework. In addition, given more data, the accuracy of the model can greatly improve. This can make use of methods of storing the characteristics of an ice load as data and then randomly generating a signal suitable for the condition.

Fig. 1 Example of dynamic motion analysis with ice and DP modules in the time domain with a motion solver [10]

Figure 1 is a framework for evaluating the station-keeping performance of a floating offshore structure for the Arctic region. The movement of the floater is analyzed by connecting the module that generates the control input for the dynamic positioning (DP) system and the module that creates the ice load to the dynamic analysis software. Connecting the DP module is not a difficult task, but estimating the ice load used for time domain analysis is not easy. This is because the ice load must be calculated while the position and direction of the floating structure are constantly changing. The reason for using commercial software is to consider the effects of wind, waves, and ocean currents in addition to the ice load. There are specialized

SWs that estimate ice resistance, but they are lacking in terms of environmental loads. Therefore, a specialized ice load generation module for precise time-zone motion analysis is required.
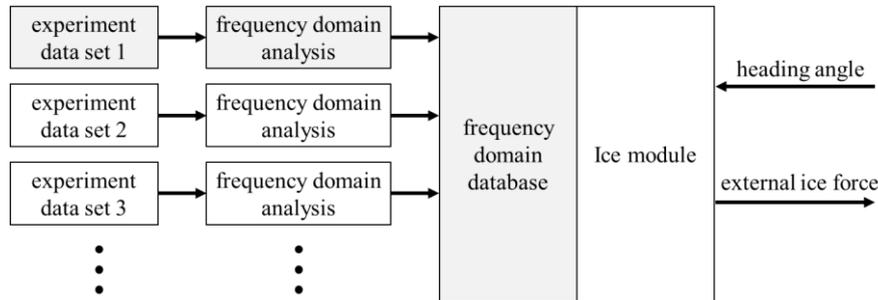

**Fig.** 2 Detailed process of the future ice module

Figure 2 is the plan for the ice module generating the ice load mentioned in Figure 1. The gray boxes are the processes that will be covered in this paper. The Ice module will generate signals based on experimental data. Data obtained from experiments performed in ice tanks are analyzed in the frequency domain. The results are collected to form a primary database, and the database provide basic data for generating the ice weight required by the ice module. More data will be acquired and made available as we continue to conduct experiments. With the results for more ice conditions, it will be possible to generate ice load for different environmental conditions.

The main objective of this research is to present a basis from which the ice load can be derived. The basic idea is to generates a random signal based on the power spectrum at an arbitrary angle. However, in the ice tank test, data can only be obtained from a limited number of angles. This paper proposes a regression model based on the acquired data to obtain a value corresponding to an arbitrary angle. The power spectrum of the ice load is analyzed by using 9 data sets performed with different orientation angles. From among the machine learning methods, we select a supervised linear regression for this application. The four most commonly used methods are applied, including linear interpolation, random forest, and artificial neural networks. Four models are applied to the measured data to construct a power spectrum that can generate a random signal at the specific angle required for a time-domain simulation. Measured data are used to train the models and compare the results between each model. Future usage of the output is discussed in the conclusion.

2. Power spectrum of ice load

2.1 Measuring ice load

The basic data we use in this paper are the experimental results from the previous research [10]. Experiments were performed in the ice tank at KRISO (Korean research institute of ships and ocean engineering). Data were collected using a 1:40 model of an FPSO designed for the Artic region. As shown in Figure 3, a captive test was performed by installing a 3-axis load cell on top of the model ship and maintaining the ice area at a constant angle. The load cell can measure F_x, F_y and M_z separately, and the maximum capacities are 4kN, 4kN, and 2kNm, respectively. After towing begins from an ice-free area, the ship enters the ice-covered area and recording begins when data is reliably acquired. The bow angle was changed from -40 degrees to 40 degrees at 10-degree intervals, and we produced data for a total of 9 cases.
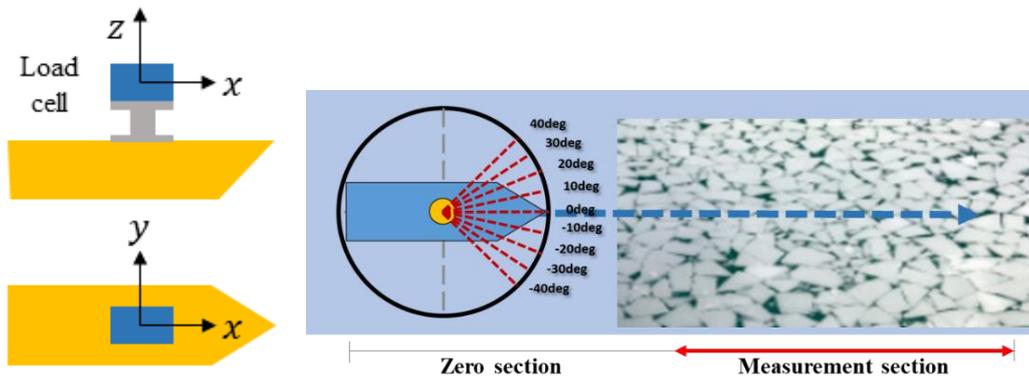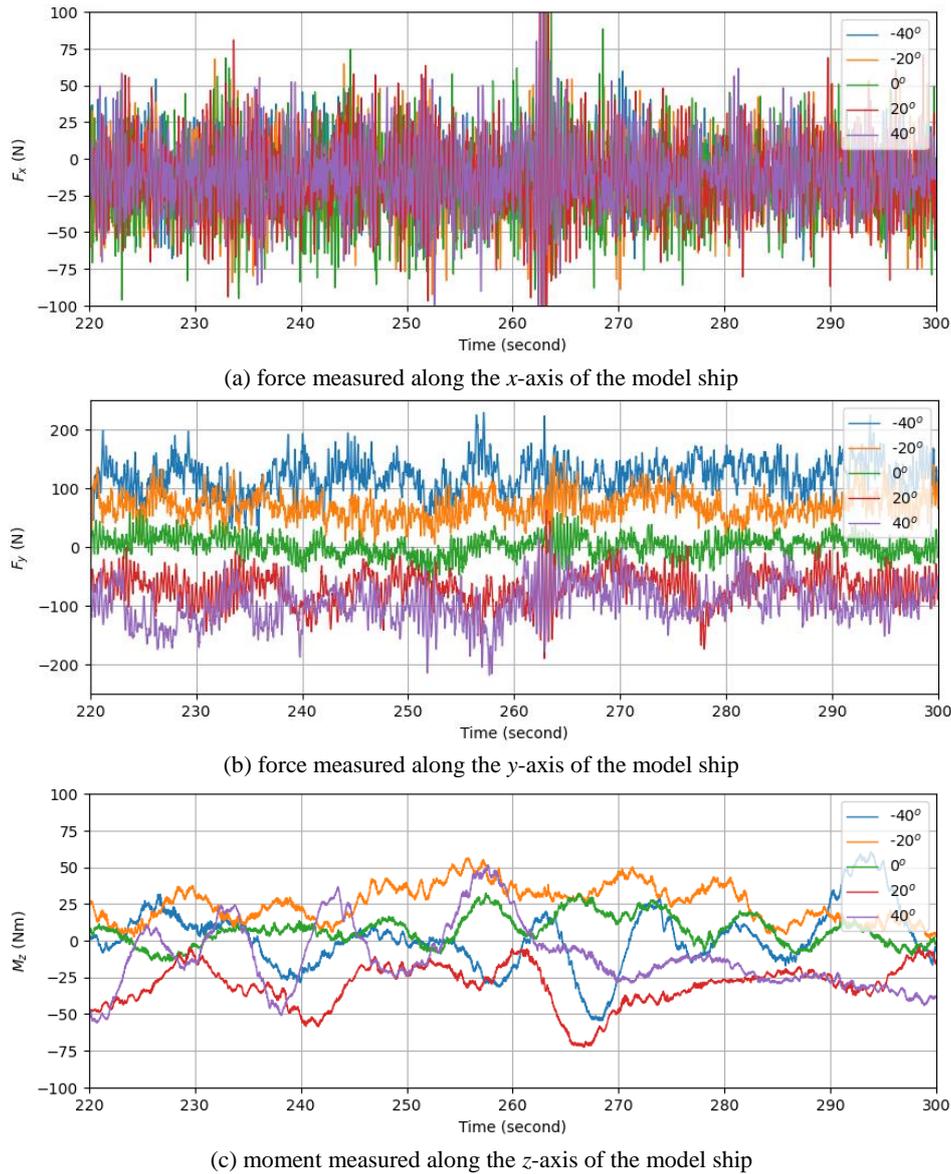
Fig. 3 Schematics of the captive test

The ice condition for the experiment is summarized in table 1. Because we assume ice breakers manage the target ice field, a level ice sheet is prepared and cut into irregularly distributed pieces.

Table 1 Ice test conditions with a 1/40 model

	Full scale	Model scale (1/40)
Ice thickness (m)	1.4	0.035
length of ice floes (m)	12	0.3
Ice drift speed	1.0 knot	0.081 m/s
Ice concentration	8/10	
Shape of ice floes	irregular	

For each case, we acquired 8,000 valid data points at a sampling rate of 100 Hz for a total of 80 seconds. Figure 4 shows the measured force and moment values for F_x, F_y and M_z acting in the x, y and z direction of the hull, respectively. The model ship passed the open water area and entered the ice area. Data was recorded after the measurement stabilized, which was at 220 seconds in this experiment. Excluding the drag component caused by the current in the actual measured value, only the pure ice load is displayed.

Since the body-fixed coordinate is used, F_x can be interpreted as the load caused by ice approaching in the bow direction. The outlier that occurred at about 263 seconds in Figure 4 (a) was caused by a problem with the rail and was excluded from the analysis. In the case of F_x and F_y, significant fluctuation is generated by ice fragments coming in the direction of the bow. The pattern depends on the ice conditions. Since FPSO for the Arctic region was considered, the heading angle was limited from -40 degrees to 40 degrees. If the heading angle gets larger than that, it is generally assumed that the structure cannot stand against the ice load. This can be confirmed in Figure 4 (b). It can be seen that the larger the heading angle, the greater the load. The increased heading angle means that more ice fragments approach the long side of the hull. Therefore, it is because the ice fragments push the hull together. Since it is impossible to maintain the position in this situation in the Arctic sea, the experiment was designed on the premise that the -40 to 40 degree of heading angle is maintained with the heading control. The moment shows a pattern in which the overall magnitude increases as the angle increases and then decreases again. Compared with the force components, low-frequency fluctuations appeared more clearly.

(a) force measured along the *x*-axis of the model ship

(b) force measured along the *y*-axis of the model ship

(c) moment measured along the *z*-axis of the model ship

Fig. 4 Selected time series data measure for F_x, F_y and M_z. Some data sets (heading angles at -30°, -10°, 10°, and 30°) are omitted from the figures to show tendencies clearly as the attack angles change.

2.2 Self-correlation function of random signal and power spectrum

The Fourier transform in a random process is difficult to define precisely because the random signal is different for each measurement and thus cannot itself be precisely defined. However, for a stationary process, it is possible to define power as a statistical expected value. That is, one can use general concepts such as autocorrelation, cross-correlation, and autocovariance. If the autocorrelation function in the random process is subjected to a Fourier transform, it becomes power spectral density (PSD), offering useful information related to the frequency spectrum. In spectrum analysis, the FFT result is multiplied by a complex component to make the amplitude into a real value, then divided by the frequency resolution. The output is independent of the frequency resolution, and it is easy to compare the vibration levels of signals having different numbers of data. For this reason, we use the power spectrum to characterize many random signals.

Let $x(\tau)$ be a stationary random process, then the self-correlation function $\gamma_{xx}(\tau)$ can be expressed as

$$\gamma_{xx}(\tau) = E[x^*(t)x(t+\tau)], \tag{1}$$

where $E[\cdot]$ denotes a statistical average. By applying the Wiener-Khintchine theorem, the Fourier transform of the autocorrelation function can provide the power spectrum. This is expressed as

$$\Gamma_{xx}(F) = \int_{-\infty}^{\infty} \gamma_{xx}(\tau)e^{-j2\pi Ft}dt. \tag{2}$$

A time-averaged autocorrelation function generates a power spectrum estimation, as expressed by the following equation:

$$R_{xx}(\tau) = \frac{1}{2T_0}\int_{-T_0}^{T_0} x^*(t)x(t+\tau)dt, \tag{3}$$

where $2T_0$ is the observation interval. If the first and second moments (mean and autocorrelation functions) of a stationary random process are ergodic random processes, we have the following relation:

$$\gamma_{xx}(\tau) = \lim_{T\to\infty} R_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{2T_0}\int_{-T_0}^{T_0} x^*(t)x(t+\tau)dt. \tag{4}$$

This means that the time-averaged autocorrelation function $R_{xx}(\tau)$ can serve as an estimate of the statistical autocorrelation function $\gamma_{xx}(\tau)$. Furthermore, the Fourier transform of $R_{xx}(\tau)$ becomes an estimate of the power spectrum density $P_{xx}(F)$, for which the equation is

$$P_{xx}(F) = \int_{-T_0}^{T_0} R_{xx}(\tau)e^{-j2\pi F\tau}d\tau$$

$$= \frac{1}{2T_0}\int_{-T_0}^{T_0}\left[\int_{-T_0}^{T_0} x^*(t)x(t+\tau)dt\right]e^{-j2\pi F\tau}d\tau = \frac{1}{2T_0}\left|\int_{-T_0}^{T_0} x(t)e^{-j2\pi Ft}dt\right|^2. \tag{5}$$

The actual power spectral density is the expected value of the limit of $P_{xx}(F)$ as time $T_0 \to \infty$:

$$\Gamma_{xx}(F) = \lim_{T_0\to\infty} E[P_{xx}(F)] = \lim_{T_0\to\infty} E\left[\frac{1}{2T_0}\left|\int_{-T_0}^{T_0} x(t)e^{-j2\pi Ft}dt\right|^2\right]. \tag{6}$$

2.3 Estimation of power spectrum using Blackman & Tukey method

To simplify the analysis, we hereafter replace observation time T_0 with $T/2$. The overall procedure is from [13]. The auto-correlation function of the time series data $x(t)$ measured during time T can be expressed as

$$R(\tau) = E[x(t)x(t+\tau)] = \frac{1}{T}\int_{T/2}^{T/2} x(t)x(t+\tau)dt, \quad \tau \geq 0. \tag{7}$$

The auto-energy density spectrum can be defined as

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} R(\tau)e^{-j\omega\tau}d\tau, \quad -\infty \leq \omega \leq \infty. \tag{8}$$

The inverse Fourier transform for Equation (8) above is

$$R(\tau) = \int_{-\infty}^{\infty} S(\omega)e^{i\omega\tau}d\omega. \tag{9}$$

Since the above auto-correlation function and auto-energy density spectrum are even functions with respect to time and frequency, respectively, the one-sided auto energy spectrum and auto-correlation function are respectively expressed as follows:

$$U(\omega) = \frac{1}{\pi}\int_0^\infty R(\tau)cos\omega\tau \, d\tau, \tag{10}$$

$$R(\tau) = \int_0^\infty U(\omega)cos\omega\tau \, d\tau. \tag{11}$$

Therefore, the autocorrelation function can be arranged in the following discretized form:

$$R_p = \frac{2}{N-p}\sum_{i=1}^{N-p} x(t_i) \cdot x(t_i + p\Delta\tau). \tag{12}$$

In equation (12), $\Delta\tau = \Delta t$, $T = N\Delta\tau$, $\tau = p\Delta\tau$, $p = 0,1,2,...m$, and m indicates the maximum lag number, decided by the guideline stated below. The number 2 in the numerator reflects the symmetry of the autocorrelation function. Therefore, equation (12) gives $m + 1$ discrete and evenly spaced values. In equation (7), the integration over time t is taken over the sample length $T - \tau = (N - p)\Delta t$. Therefore, the denominator becomes $N - p$. When $p = 0$, the autocorrelation function is the variance of $x(t)$.

To estimate the one-sided spectral density, equation (10) is discretized with $m + 1$ evenly spaced autocorrelation coefficients:

$$\Delta\omega = \frac{\omega_c}{m} = \frac{\pi}{m\Delta\tau}, \tag{13}$$

where ω_c is the Nyquist frequency. By using $\omega = h\Delta\omega$ and equation (13), we have

$$\omega\tau = (h\Delta\omega)(p\Delta\tau) = \frac{\pi ph}{m}. \tag{14}$$

After applying the trapezoidal formula to integrate equation (10), the raw spectral density is expressed as follows:

$$L_h = \frac{1}{m\Delta\omega}\left[R_0 + 2\sum_{p=1}^{m-1} R_p cos\frac{\pi ph}{m} + R_m cos\pi h\right], \quad h = 0,1,2,...,m. \tag{15}$$

The above raw spectral density is smoothed by a moving weighted average, and the Hamming process is applied.

$$U(h\Delta\omega) = U_h \equiv \sum_{i=1}^m a_{h,i}L_h, \quad h = 1,2,...,m-1 \tag{16}$$

$$\begin{cases} a_{h,h-1} = 0.23 \\ a_h = 0.54 \\ a_{h,h+1} = 0.23 \\ a_{h,i} = 0.00 \;\; \text{for} \;\; i \neq h-1, h \;\; \text{or} \;\; h+1. \end{cases} \tag{17}$$

Therefore, the overall smoothed estimates are

$$\begin{cases} U_0 = 0.54L_0 + 0.46L_1 \\ U_h = 0.23L_{h-1} + 0.54L_h + 0.23L_{h+1}, h = 1,2,...,m-1 \\ U_m = 0.46L_{m-1} + 0.54L_m. \end{cases} \tag{18}$$

2.4 Power spectral density

Figure 5 shows the PSD values for F_x at intervals of 10 degrees from +40 degrees to -40 degrees of the hull. Since the Nyquist frequency is 50 Hz, we analyzed it up to about 314 rad/s. After 50 rad/s, however, the value converges to almost 0 and is omitted. The same approach is applied to next analyses.

The distribution of PSD can be analyzed in relation to Figure 4. In the case of F_x and F_y, a peak value is shown at a frequency of about 17 rad/sec in common in all attack angles. The

peak value appears in relation to the operating speed of the model ship and the degree of ice density. In the case of M_z, this difference does not appear significantly, and there is a dominant feature in the low frequency band. Considering that the hull is a fixed coordinate system, it can be seen that the effect of the ice approaching in the direction of the bow acts differently on the force elements. As shown in Figure 4 (c), the overall magnitude of the moment indicates a pattern that gradually increases as the heading angle increases and then decreases again. That is because the model ship has a long shape in the x-axis direction, so the tendency to return to the original direction of less resistance (the heading angle of 0) becomes stronger up to a certain angle. As the angle becomes larger than a certain angle, this tendency becomes smaller again. Accordingly, the power in the low-frequency region gradually increases and then decreases again. Also, the degree of fluctuation is slightly weaker than the F_x and F_y components. The period of significant vibration appears at about 1 second, all power spectrum values are concentrated within 5 rad/sec. Therefore, as shown in the figure, the magnitude of power appears very small at frequencies greater than 5 rad/sec.

(a) PSD values (N^2/s) for F_x

(b) PSD values (N^2/s) for F_y

(c) PSD values ($N^2 m^2$/s) for M_z
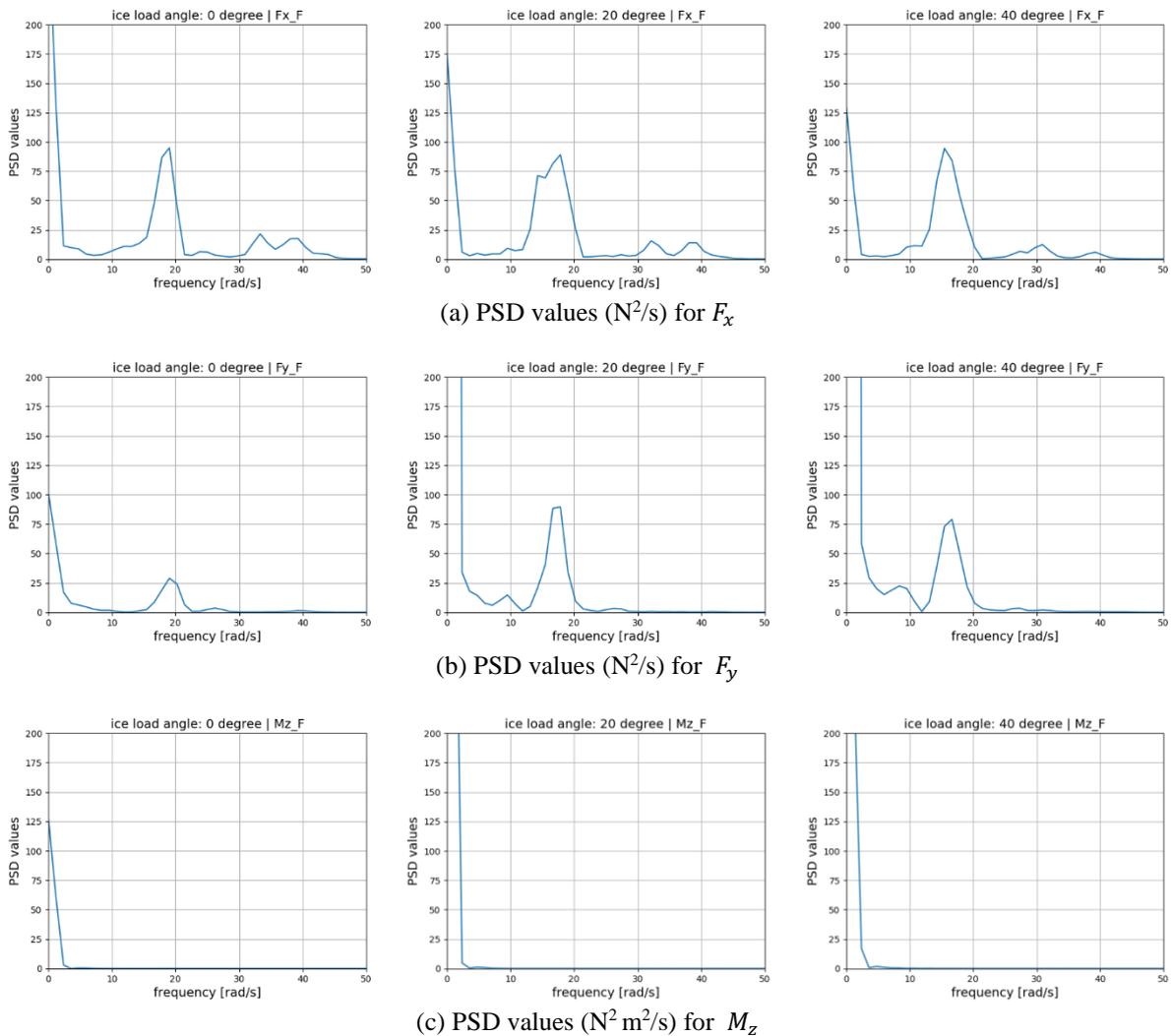
Fig. 5 Power spectral density values for F_x, F_y, and M_z at selected heading angles of 0, 20 and 40 degrees

3. Model setup and preparation of the training set

3.1 General procedure and regression models for supervised learning

This paper applies supervised learning using 4 regression methods. This is a regression analysis technique that models the correlation between the dependent variable and one or more independent variables. The use of regression aims to predict the ice load required for real-time simulation based on the data measured in the ice tank test. Regression analysis in machine learning consists of three main steps: preparing data for regression analysis, training the model, and predicting new data. We use support vector machines, the random forest method, and the neural networks method, along with basic linear interpolation. The overall sequence for applying machine learning is as follows:

```
1:   (Load) Load ice load experiment data
2:   (Build) Build machine leaning model (Linear regression, Support Vector
             Machine, Random Forest, Neural Networks)
3:   (Fit) Fit the model
4:   (Train) Start train
5:           angles = [-40, -30, -20, -10, 0, 10, 20, 30, 40]
6:           for ang in angles do
7:               predict(ang);
8:   (Test) Start test
9:           angles = [-40, -39, -38, -37, ..., 37, 38, 39, 40]
10:          for ang in angles do
11:              predict(ang);
```

Linear interpolation is the most basic regression method for finding the relationship between input and output. It assumes linearity between variables. This paper applied linear regression using the least squares method. For the Build and Fit stages, we use the experimental measurement angles and PSD values for those angles for training. This method creates a linear relationship for each frequency by comparing two PSD values for each angle. For example, with PSD values corresponding to -40 degrees and -30 degrees, it stores linear interpolation equations for each frequency. This recurs for a total of 8 repetitions, up to sets corresponding to 30 and 40 degrees. Afterwards, in the test step, for a given input angle, the method finds a corresponding section among the 8 sections divided in the training step and obtains an interpolation of the value.

The random forest (RF) method is one of the decision trees used for classification and regression, developed by Leo Breiman [14]. The RF method is known to show good results, especially for tasks related to classification and guessing [15]. In addition, RF has several advantages when it is necessary to deal with many variables with small amounts of observational data [16]. RF method shows the relationship between the dependent variable and the predictor in the form of a tree. It allows for readable and interpretable conclusions, which is why it is so popular. The estimation accuracy of a model is very important, and the RF method is accepted as estimating models with high accuracy.

The artificial neural network (ANN) was developed to analyze complex data structures or big data. It can be used for pattern recognition, parameter estimation, and classification. Burak Yildiz [17] applied the ANN in predicting residual resistance of a trimaran vessel based

on the experimental data. Several structures and algorithms for building ANNs have been developed, such as the multilayer perceptron (MLP) and radial basis function (RBF). In addition, various optimization algorithms have been applied to MLP or RBF. The ANN structure basically consists of three layers. The first is the input layer containing the independent variables. The second, called the hidden layer, has several activation functions to compute the relationship between the input (input variable) and output (dependent variable) [18]. The third is the output layer.

3.2 Generation of the training data

We create a standardized training data set for machine learning models. The shape of the data set is shown in Figure 6. Since there are models for each of the 8 sections between the 9 measurement angles for a specific frequency, we need a total of 8x500 models for 500 frequencies in the linear interpolation method. A support vector machine requires a total of 500 models because there is one model for each frequency band. This means that a regression model is created with a total of 9 data sets per model. Random forests and neural networks use a single model for each method.
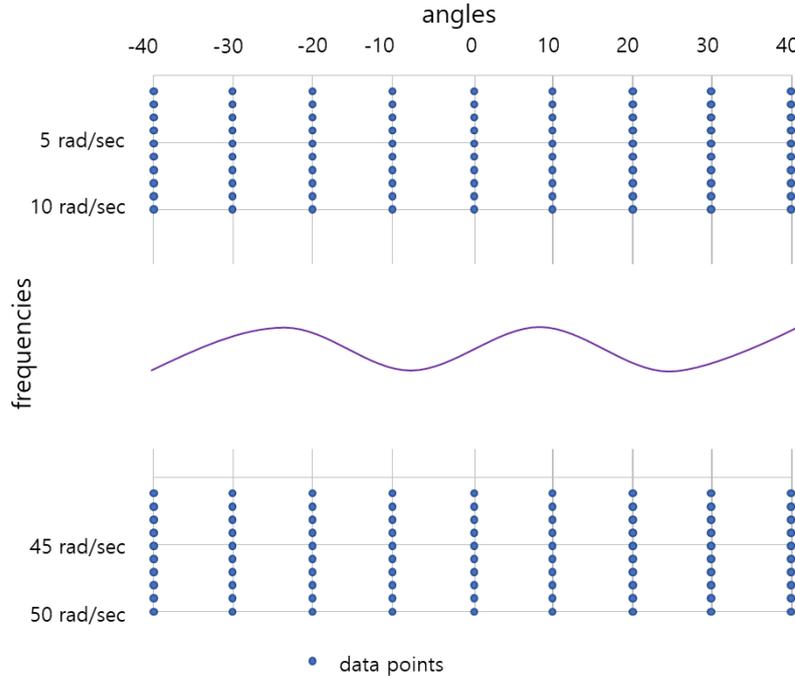

Fig. 6 The shape of the training data set
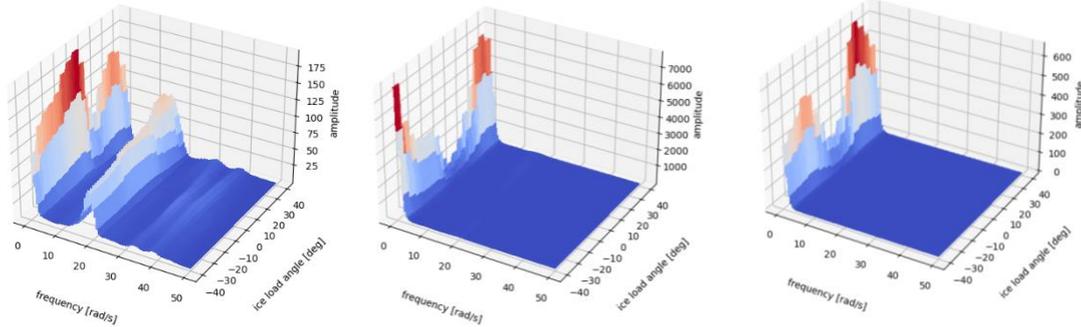
4. Regression results

Figure 7 shows the results of applying all the regression methods of linear interpolation, support vector machine, random forest, and artificial neural network. In each method, F_x, F_y, and M_z values are shown from left to right. Note that the scale of the PSD values was set automatically, and the PSD values of F_y appear relatively large. Therefore, the peak near 17 rad/sec of F_y in Figure 5 is not well visible. For support vector machines, we applied a polynomial kernel. Kernels such as linear kernels or radial basic functions are not suitable for this kind of data set due to the shape of the data. Through trial and error, we confirmed that having a higher order term gave better results. Finally, we set the degree of the polynomial kernel to 8 and the hyperparameter to 5. It can be seen that overall performance is poor because the degree of the polynomial is fixed and must be used for 500 different data sets.
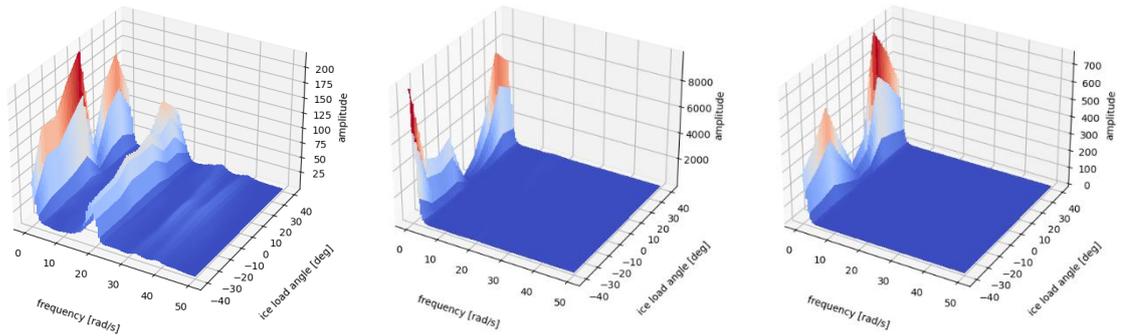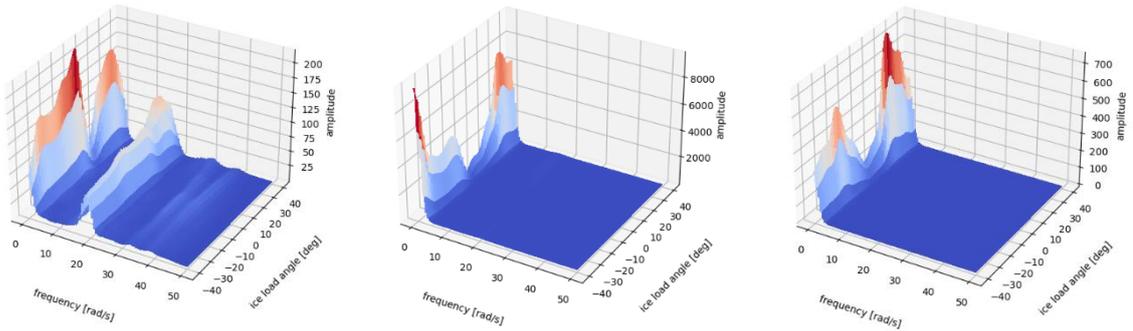
(a) linear interpolation
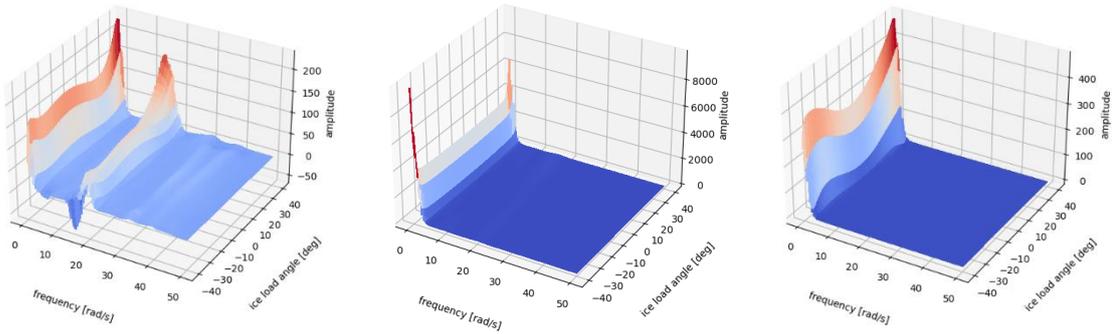
(b) support vector machine

(c) random forest

(d) neural networks

Fig. 7 3D plots with the regression results for 4 algorithms.

In the Random Forest method, we set the number of trees to 1,000, and observed no significant performance improvement beyond that. In the case of depth, since the input is only the value corresponding to the angle, we set a default value of 1; even if this increases to 2 or more, the result does not improve.

For the neural network model, we used rectified linear unit (ReLU) as the activation function and He initialization as the kernel initializer. This consists of three hidden layers, and proceeds in the order input, dense, batch normalization, and dropout.

A graph presenting the regression results for each method at a specific frequency (18 rad/sec) is shown in Figure 8. For the linear interpolation method, it can be seen that the interpolation reference values are connected to each other. For the support vector machine, as the order of the kernel increases to enhance accuracy, it can be seen that it bounces a lot within the 30 to 40 degree range. In an ideal situation, the results should be better than those seen in the figure, with bilateral symmetry around 0 degrees. For the random forest, the depth is 1, which shows the limitations of cascading results. Neural Networks showed relatively good performance for all intervals.

(a) linear interpolation

(b) support vector machine

(c) random forest

(d) neural networks
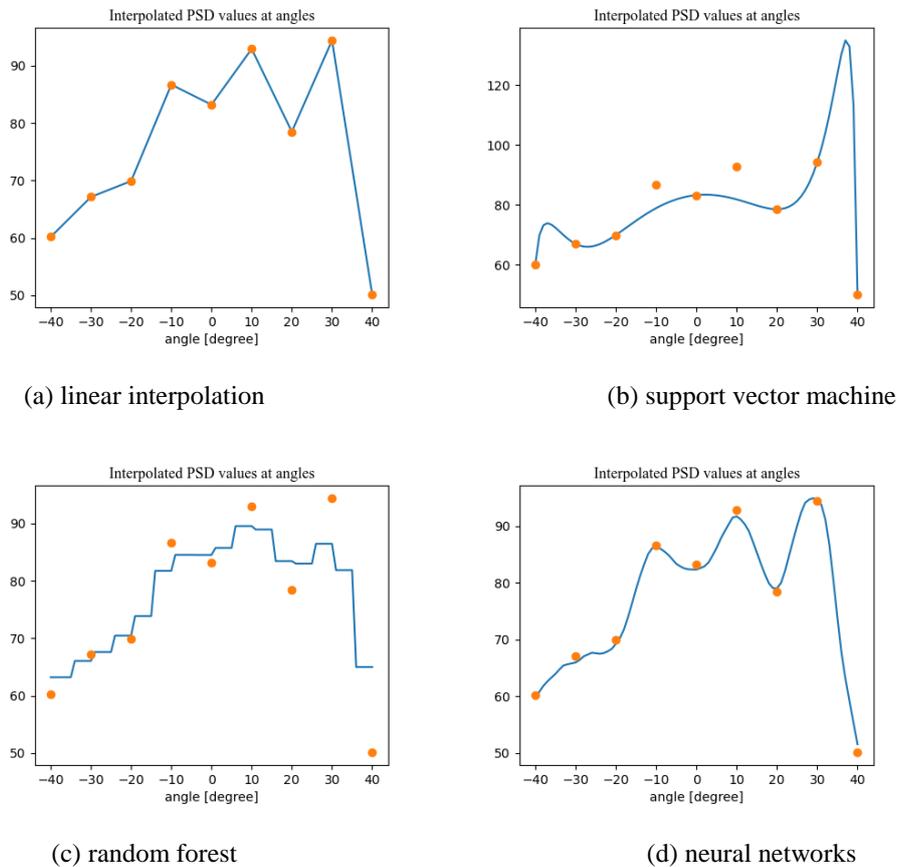
Fig. 8 Regression results for 4 algorithms at a specific frequency (18rad/sec). Orange dots are training sets, and blue solid lines are result values generated at 1-degree intervals by each interpolation method.

The overall performance results based on MSE, determined by comparing the training set and the predicted values at the same angles, appear in Table 2. For linear interpolation, the predicted value must represent the training value, so the MSE value must be zero. It is clear that this does not indicate that the prediction was perfect. For the support vector machine, we confirmed that there exists a section with a lot of deviation due to the influence of a high-order polynomial in both F_y and M_z. The reason for the large error can be confirmed in Figure 7 (b). The applied kernel for the SVM was selected to show the result value as close to the test set as possible for comparison. As a result, a regulation parameter of 5 was selected for the 8th order polynomial and applied to the data analysis. Due to the nature of MSE, the results are amplified when large errors are included. Therefore, since 500 errors appearing in the interpolation

process are accumulated for each angle, the fit of the model deteriorates, resulting in a very large error. The neural network method shows the best results for all force directions.

Table 2 Comparison of MSE values from the regression methodologies

	F_x	F_y	M_z
Linear regression	0.00	0.00	0.00
Support vector machine	689,809.65	487,344,726.71	28,777,468.15
Random forest	16.15	21,681.15	109.49
Deep neural networks	0.31	640.43	3.40

5. Conclusion

The ice tank test offers a limited amount of data due to difficulties in the time and cost required for preparation. A method of examining the results by analyzing individual pieces of experimental data has been under development for a long time. However, it is also necessary to observe patterns in the measurement data according to changes in various parameters known to affect the ice load, and machine learning techniques can be effectively applied as a tool for this purpose.

A methodology for calculating the power spectral density was presented for the data measured in the ice tank. The overall pattern according to the attack angle was observed and analyzed. The ice tank test was conducted as a captive test by changing the bow angle from -40 to 40 degrees at intervals of 10 degrees. The indirect method was used in estimating power spectral density suggested by Blackman & Tukey for each angle at which the test was conducted. In order to observe the change trend of the overall power spectrum density according to the bow angle, various regression methods were applied by subdividing the angle, and the results were compared. Four widely used methods of machine learning (linear interpolation, a support vector machine, a random forest, and a deep neural network) were applied for linear regression. 9 data sets were used as training data and tested for each angle.

Linear interpolation is the most basic method and can serve as a standard for judging the performance of other methods. In the case of the Support Vector Machine, we tried to find the optimal situation by changing the order, but the results were poor. If the measurement data has left-right symmetry, performance should be better, but our results suffered due to the asymmetry of the measurement values. The more data we have, the better the performance seems to be. However, it is difficult to tackle an increase in parameters, so there will be limitations in future usage.

The random forest method is known to be the most suitable model for multivariate regression, but it did not perform well in this data analysis. This is because the basic data itself only provides information about the angle, which limits the depth of the decision tree. However, when there are many variables to consider, such as ice thickness and degree of integration, it can be expected to show greater performance improvement than other methods. The deep neural network method showed the best performance. As it can be used for multivariate input, it will be comparable to the random forest method if given additional data. Overall, the artificial neural network method performed best. Although it takes a long time to complete, this arguably has no significant impact because this is not a problem that needs to be analyzed in real time.

If the power spectrum model is further developed as presented in this paper, it will be possible to generate realistic ice loads in a time-domain simulation. As discussed in the introduction, the ice module receives the heading angle of the floating structure. Then the module generates a corresponding random signal based on the regressed PSD values. Technique to generate adequate ice load with the PSD values will be the next step of the research. It should

be possible to obtain better results by applying the improved ice load in simulation techniques that use a complex environmental load.

ACKNOWLEDGEMENT

REFERENCES

[1]   Wang, F., Zhou, L., Zou, Z. -J., Song, M., Wang Y., Liu, Y., 2019. Study of continuous icebreaking process with cohesive element method. *Brodogradnja*, 70(3), 93-114. https://doi.org/10.21278/brod70306

[2]   Moran, K., Backman, J., Farrel, J. W., 2006. Deepwater Drilling in the Arctic Ocean's Permanent Sea Ice, Proceedings of the Integrated Ocean Drilling Program, 302. https://doi.org/10.2204/iodp.proc.302.106.2006

[3]   Islam, M., Mills, J., Gash, R., Pearson, W., 2021. A literature survey of broken ice-structure interaction modelling methods for ships and offshore platforms. *Ocean Engineering*, 221, 108527. https://doi.org/10.1016/j.oceaneng.2020.108527

[4]   Mikulec, M., Piehl, H., 2023. Verification and validation of CFD simulations with full-scale ship speed/power trial data. *Brodogradnja*, 74(1), 41-62. https://doi.org/10.21278/brod74103

[5]   Jenssen, N. A., Muddesitti, S., Phillips, D., Backstrom, K., 2009. DP in Ice. *Dynamic Positioning Conference*, Houston, USA.

[6]   Jenssen, N. A., Hals, T., Jochmann, A., Santo, X. D., Kerkeni, S., Doucy, O., Løset, S., 2012. DYPIC-A Multi-National R&D Project on DP Technology in Ice. *Dynamic Positioning Conference*, Houston, USA.

[7]   Kerkeni, S., Santo, X. D., Doucy, O., Jochmann, P., Haase, A., Metrikin, I. A., Støle-Hentschel, S., 2014. DYPIC Project: Technological and Scientific Progress Opening New Perspectives. *Arctic Technology Conference*, Houston, USA. https://doi.org/10.4043/24652-MS

[8]   Wang, J., Sayeed, T., Millan, D., Gash, R., Islam, M., Millan, J., 2016. Ice Model Tests for Dynamic Positioning Vessel in Managed Ice. *Arctic Technology Conference*, St. John's, Canada. https://doi.org/10.4043/27430-MS

[9]   Islam, M., Wang, J., Brown, J., Lau, M., Gash, R., Millan, D., Millan, J. 2018. Physical Model Testing for Supporting Ice Force Model Development of DP Vessels in Managed Ice. *Arctic Technology Conference*, St. John's, Canada. https://doi.org/10.4043/29157-MS

[10]   Xie, C., Zhou, L., Ding, S., Liu, R., Zheng, S., 2023. Experimental and numerical investigation on self-propulsion performance of polar merchant ship in brash ice channel. *Ocean Engineering*, 269, 113424. https://doi.org/10.1016/j.oceaneng.2022.113424

[11]   Sun, Q., Zhang, M., Zhou, L., Garme, K., Burman, M., 2022. A machine learning-based method for prediction of ship performance in ice: Part I. ice resistance. *Marine Structures*, 83, 103181. https://doi.org/10.1016/j.marstruc.2022.103181

[12]   Kang, H. H., Lee, D. -S., Lim, J.-S., Lee, S. J., Jang, J., Jung, K. H., Lee, J., 2021. A study on the effectiveness of the heading control on the mooring line tension and position offset for an arctic floating structure under complex environmental loads. *Journal of Marine Science and Ocean Engineering*, 9, 102. https://doi.org/10.3390/jmse9020102

[13]   Kim, C. H., 2008. Nonlinear waves and offshore structures. *Advanced Series on Ocean Engineering: Volume 27*, New Jersey, U.S.A. https://doi.org/10.1142/4906

[14]   Breiman, L., 2001. Random Forests. *Machine Learning*, 45, 5-32. https://doi.org/10.1023/A:1010933404324

[15]   Uriarte, R. D., Andrés, S. A., 2006. Gene Selection and Classification of Microarray Data Using Random Forest. *BMC Bioinformatics*, 7(3), 1-13. https://doi.org/10.1186/1471-2105-7-3

[16]   Ulrike, G., 2009. Variable Importance Assessment in Regression: Linear Regression versus Random Forest. *The American Statistician*, 63(4), 308-319. https://doi.org/10.1198/tast.2009.08199

[17]   Yildiz, B., 2022. Prediction of residual resistance of a trimaran vessel by using an artificial neural network. *Brodogradnja*, 73(1), 127-140. https://doi.org/10.21278/brod73107

[18]   Bonanno, F., Capizzi, G., Graditi, G., Napoli, C., Tina, G. M., 2012. A Radial Basis Function Neural
       Network Based Approach for the Electrical Characteristics Estimation of a Photovoltaic Module. *Applied
       Energy*, 97, 956-961. https://doi.org/10.1016/j.apenergy.2011.12.085

Submitted: 20.12.2022.   Seung Jae Lee
                                 Korea Maritime and Ocean University, Korea South
Accepted: 28.04.2023.   Kwang Hyo Jung
                                 Pusan National University, Korea South
                                 Namkug Ku
                                 Dong-eui University, Korea South
                                 Jaeyong Lee*, jlee@deu.ac.kr
                                 Dong-eui University, Korea South