APPLICATION OF AN ARTIFICIAL NEURAL NETWORK AND MULTIPLE NONLINEAR REGRESSION TO ESTIMATE CONTAINER SHIP LENGTH BETWEEN PERPENDICULARS

Tomasz Cepowski*
Paweł Chorab
Dorota Łozowicka
Akademia Morska w Szczecinie, W. Chrobrego 1-2, 70-500 Szczecin, Polska

* Corresponding author: *t.cepowski@am.szczecin.pl (T. Cepowski)*

ABSTRACT

Container ship length was estimated using artificial neural networks (ANN), as well as a random search based on Multiple Nonlinear Regression (MNLR). Two alternative equations were developed to estimate the length between perpendiculars based on container number and ship velocity using the aforementioned methods and an up-to-date container ship database. These equations could have practical applications during the preliminary design stage of a container ship. The application of heuristic techniques for the development of a MNLR model by variable and function randomisation leads to the automatic discovery of equation sets. It has been shown that an equation elaborated using this method, based on a random search, is more accurate and has a simpler mathematical form than an equation derived using ANN.

Keywords: ship design,ANN,regression,container ship,length

INTRODUCTION

The ship design process consists of three main stages: the preliminary, contract, and detailed design steps. The key characteristics of ship design are based on the main requirements of a shipowner during the preliminary design stage. A parametric and geometric design stage are the main phases of preliminary design. Watson [1], Rawson and Tupper [2], and Papanikolaou [3] have argued that the selection of a ship's main dimensions such as length, breadth, and draught are the main parametric design objectives. However, there is no detailed information about technical characteristics of a ship during the parametric design stage to accurately estimate these parameters. The ship designer should resolve this problem and select these characteristics based on the requirements of the shipowner, the character of the ship's mission, and various formal maritime rules and regulations [1-3].

Ship design is a loop iterate process which goes through a design spiral, originally introduced in 1959 by Evans [4] and modified in 1985 by Andrews [5]. The inaccurate estimation of ship dimensions during the parametric design stage increases the number of design spiral loops, and the design time and project cost. As noted by Papanikolaou [3] and Chądzyński [6] for a standard ship type, various statistical, empirical, or regression methods based on similar ship builds may be used to solve this problem.

Chądzyński [6] and Papanikolaou [3] argued that initially, the length of a cargo ship is estimated based on cargo capacity, such as its deadweight, hold, or TEU capacity. Other ship dimensions are usually estimated based on this length at later stages.

Various linear or nonlinear equations developed using a container ship database have often been used for the initial estimation of a ship's length. Piko [7], Kristensen [8], and Papanikolaou [3] prepared a set of equations for estimating a container ship's main dimensions. Piko's equations were based on the statistical data of container ships that were being built up to 1980. Papanikolaou used the data of container ships built prior to 2005 and Kristensen used a database of container ships built before 2013. Linear and nonlinear regression methods were used in these studies.

Over the past 20 years, container ship design trends have been influenced by market and trade demands. The financial crisis of 2007-2008, together with fuel price changes and strict emission requirements have all had an influence on these trends. Container ships are usually categorised as volume carriers in common design procedures. Container number and velocity are the main requirements of any container ship owner. Economic and environmental factors could have an influence on changing these requirements and later, the design process in the future. Figure 1 shows container ship age profiles which are dependent on size and mean speed values. The Sea-web Ships database [9] of all container ships built from 2000-2020 was used in this analysis. Figure 1 shows that container ship capacity and speed have fluctuated to a large degree throughout the last 6-7 years. Moreover, the latest container ships usually have a higher TEU capacity and a lower Froude number.
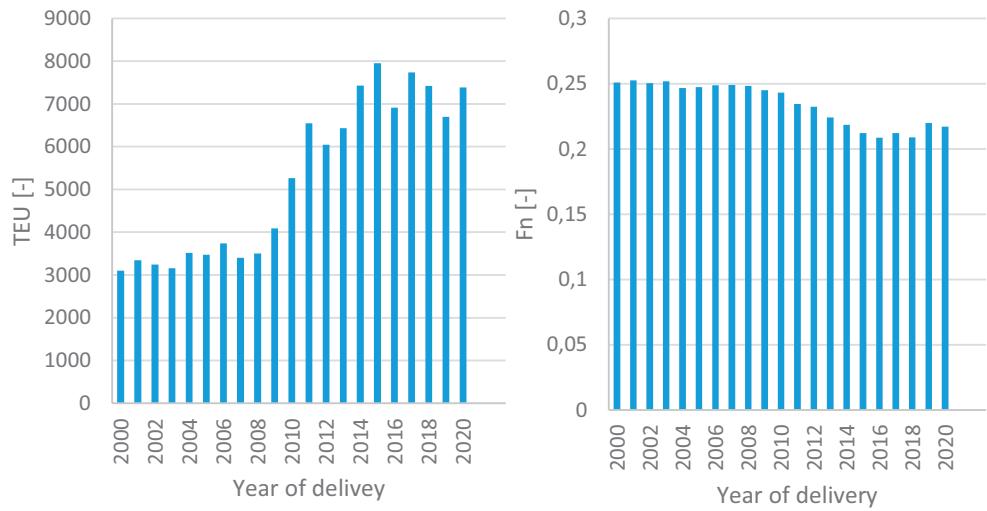
A literature review showed that key container ship characteristic equations were developed based on deadweight or container number capacity. While Piko [7] and Papanikolaou [3] used deadweight capacity, Kristensen [8] used a TEU container number for a ship's characteristic estimation. Deadweight capacity includes the mass of the cargo, ballast, and ship stores. To calculate the cargo mass, the number and mass of containers must be known. Container mass is based on the number of containers. However, a design characteristic estimation based on the number of TEU containers does not include ballast and store mass. However, a deadweight calculation using the number of containers required is unnecessary in this case.

A literature review did not show which parameter led to the smallest estimation error of a design characteristic, i.e., the deadweight or container number capacity. Piko, Papanikolaou, and Kristensen's formulas for a container ship length's estimate did not consider velocity, despite it being a main ship owner's requirement. These formulas were developed using regression methods. Initially, Piko implemented nonlinear approximation methods such as power and 2nd degree polynomial regression models. Deadweight was used as an independent variable in these equations. Piko compared the results and considered that the functions developed using the power regression model provided reliable estimates of the parameters over a wider range of deadweights.

Papanikolaou [3] presented a theory and a detailed compendium of knowledge on practical methods for the preliminary design of a ship. This book also provides an equation for estimating the length between perpendiculars of a container ship using the deadweight developed using a power regression model. In Piko and Papalikonau's studies, the exponent values of non-linear regression functions were similar and almost 0.4. The Kristensen approximations were based only on TEU capacity and were developed using linear, 2nd degree polynomial and power regression models. In this study, the exponents of the power regression models were 0.38, 0.55, and 0.34 for Small, Panamax, and Post-Panamax container ships, respectively.

There have been no publications in the scientific literature on the use of artificial neural networks (ANN) to determine the length of a container vessel. Only Gurgen et al. [10] applied an ANN to predict the main dimensions of chemical tankers using deadweight capacity and speed. An ANN based on a multi-layer perceptron structure with 13 neurons in a hidden layer was used in this research. Gurgen et al. [10] argued that

Fig. 1. Age profiles of the container ships: mean number of containers' TEU capacity and Froude number (Fn) vs delivery year

Old regression formulas developed before 2014 are inadequate for considering the design trend changes of modern container ships.

neural networks may create more accurate models for complex systems than standard statistical methods.

Therefore, the aim of this study was to develop an empirical equation to estimate the length between perpendiculars for container ships built since 2014, considering the container number and velocity.

MATERIALS AND METHODS

The container ship characteristics used in this study are defined as follows:

Length between perpendiculars (LBP) – the horizontal distance measured parallel to the baseline from the aft to forward perpendicular.

Ship velocity (V) – service speed in knots, which is less than the maximum ahead service speed.

TEU – the maximum number of 20-foot standard containers below and above the deck.

Deadweight (DWT) – the maximum deadweight of the ship immersed at the summer load line in water with a 1.025 t/m³ density.

Froude number (Fn) – a dimensional number calculated as shown in Eq.(1).

$$Fn = \frac{V}{\sqrt{g \times L}} \tag{1}$$

where

V – ship velocity in m/s,

g – standard gravity.

L – length of the ship (it was here assumed that L = LBP)

The data of the 120 latest new build container ships from 2014 to 2020 was used in this study. The source of the data was Sea-web Ships [9]. Sister ships with identical or similar characteristics were removed from the data set. The range and mean values of ship samples that remained after verification are shown in Table 1.

Tab. 1. *The mean, minimal, and maximal values of container ships, where: LBP = length between perpendiculars, B – breadth, D – side depth, T – draft, V – velocity, TEU – the number of containers, DWT – deadweight, and Fn – Froude number*

	LBP [m]	LBP/B [-]	B/D [-]	T [m]	V [kts]	TEU [-]	DWT [t]	Fn [-]
Min	69.20	3.95	1.61	3.50	9.0	100	1780	0.15
Mean	214.59	6.12	1.96	10.94	19.0	5265	60675	0.22
Max	386.23	7.26	3.67	16.50	25.0	21000	202036	0.30

In this study, ANNs and a random search method based on nonlinear regression and heuristics techniques were used to estimate the container ship's length. The second aim of the research was to compare the accuracy of these methods for estimating the container ship's length.

ARTIFICIAL NEURAL NETWORKS (ANN)

In recent years, ANNs have been used in several scientific ship design theory publications. For example, Alkan et al. [11] calculated the initial stability parameters of a fishing vessel using neural networks. Artificial neural networks were developed using sample ship data to estimate the vertical centre of gravity, the transverse metacentre height above the keel, and the vertical centre of buoyancy of the ship. Gurgen et al. [10] created an ANN to estimate chemical tanker dimensions. In this paper, the main ship parameters, such as, overall length, length between perpendiculars, breadth, draught, and freeboard were estimated based on deadweight and vessel speed. Gurgen et al. [10] argued that the initial main particulars of chemical tankers could be determined using ANNs, offering results which were much more accurate than those obtained with sample ship data. Ekinci et al. [12] used 18 computational intelligence methods (including neural network methods) to estimate the main design parameters of oil/chemical tankers. Abramowski [13] developed a model for determining the effective power of a ship using neural networks. In this publication, a mathematical model was developed using neural networks to determine the effective power of a ship. Cepowski [14] applied ANNs to estimate added resistance in regular head waves while using ship design parameters, such as length, breadth, draught, and Froude number. To create a reliable model, only experimental data determined through model test measurements was used to train the neural network. Song et al. [15] used the radial base function ANN to predict a ship's rolling motion. Based on this method, the disturbing moment and roll time series were estimated. Sahin et al. [16] used the ANN model linked to the main ship parameters to estimate the dilution factors in the preliminary design. Gross and deadweight ton, passenger number, freeboard, engine power, propeller number, and block coefficient values were used to estimate the likely dilution factors. Luan et al. [17] used ANNs to estimate the fuel consumption of container vessels. Cheng et al. [18] presented a comparative study of the sensitivity analysis and simplification of the ANN for a ship's motion prediction. Indeed, the use of ANNs has provided excellent results in several research experiments.

An ANN was created based on the functioning of the biological nervous system. The nervous system is a structure consisting of neurons and connections linking them. A numerical model of the neural network was developed based on this structure and the signal transmission method. The neural network was built from an input, output, and one or more hidden layers that consist of neurons [19]. Values from previous layers were passed through neurons which were connected with weights. These weights determined the relationships between input and output data [20]. The main

problem in developing an artificial network is selecting an optimal network structure and calculating the neuron weight values. Therefore, different types of neural networks and methods of learning can be used. Multilayer nonlinear neural networks and a learning backpropagation algorithm are often applied to predict technical parameters.

Overfitting phenomena are an additional problem in the development of neural networks. This phenomenon occurs when a statistical model has too many parameters in relation to the data sample size used in the creation of the model. A test set method is usually used to detect this phenomenon. Unfortunately, about 25-30% of randomly selected data is wasted to test the model in this method. Additionally, about 25-30% of data are used to validate the neural network during the learning process. This means that about half of the dataset is wasted throughout the neural network development process.

In this study, the search process for the best neural network included the following steps:

- creating a neural network topology,
- training a network,
- testing a network,
- making an accuracy assessment of a network's model based on the test results.

The statistical method of mean absolute error (MAE) value was used for an accuracy assessment Eq (2).

$$MAE = \frac{\sum |LBP - LBP_e|}{n} \quad \text{(2)}$$

where:

LBP – length between perpendiculars from the data set,
LBP_e – estimated length between perpendiculars using a neural network,
n – number of ships in the data set.

The following neural network types were tested for container ship length prediction:

- generalised regression neural network (GRNN),
- multilayer perceptron (MLP),
- radial basis function network (RBF),
- linear network.

To develop these neural networks the following assumptions were made:

- sum of squares as an error function,
- backpropagation as demonstrated by [19,21,22], the conjugate gradient descent [19], and Levenberga-Marquardt [23,24] as a training algorithm,
- logistic sigmoid function as activation,
- validation and test sets included 30 cases each (60 cases in total).

RANDOM SEARCH METHOD

A random search method based on a Multiple Nonlinear Regression (MNLR) and heuristic algorithm was applied to estimate the alternative length between perpendiculars. A general MNLR model is given by the relations in Eq. (3) [25,26].

$$Y = \alpha + \beta_1 X_i + \beta_2 X_j + \beta_3 X_i^2 + \beta_4 X_j^2 + \ldots + \beta_k X_i X_j \quad \text{(3)}$$

where:

Y – dependent variable,
X – independent variables,
i, j – number of independent variables,
α – intercept,
β – coefficient,
k – observation number.

The authors defined the following general model Eq. (4) to develop the relationship between the length between perpendiculars (LBP), velocity (V), and the number of containers (TEU) based on the model Eq. (3):

$$LBP = \alpha + \beta_1 f_1(V) + \ldots + \beta_n f_n(V) + \beta_{n+1} f_{n+1}(TEU) + \ldots + \beta_{n+m} f_{n+m}(TEU) + \beta_{n+m+1} f_{n+m+1}(V) f_{n+m+2}(TEU) + \ldots + \beta_{n+z} f_{n+m+k-1}(V) f_{n+m+k}(TEU) \quad \text{(4)}$$

where:

α – intercept,
β – coefficient,
f – base function, such as power, logarithmic, or exponential function,
n, m, k, z – the number of functions or β coefficient.

A set of 400 power, logarithmic, or exponential functions was used in this study. Finding the best TEU and V combinations in this model and selecting the best-fitting functions from the function set led to a large number of possible variants. For example, if we assume the simplest model Eq. (5):

$$LBP = \alpha + \beta_1 f_1(V) + \beta_2 f_2(TEU) + \beta_3 f_3(V) f_4(TEU) \quad \text{(5)}$$

which consists of four function combinations (f1, …, f4) selected from a collection of 400 base functions, we get the total combination number (n):

$$n = 400^4 = 2{,}65E + 10 \quad \text{(6)}$$

Searching through all these possible variants using an exact algorithm is computationally expensive and time-consuming, and thus, a heuristic approach was applied to solve this problem. The disadvantage of this method is that the solution is not as optimal as the exact approach. However, multiple searches allow the user to find a solution which is almost optimal.

An algorithm can be developed in which variables and base function combinations are randomised during the first step. Then, the model's fit to data is checked, and statistical errors are calculated. Finally, the best functions and variable combinations are selected through looped searching.

The authors implemented selected parts of this algorithm in the ndCurveMaster computer program [27] which was used to support equation searching.

Increasing model elements improves accuracy but may lead to overfitting with this method. Therefore, the next problem was to detect and prevent overfitting.

A test set was randomly selected from the data set to detect overfitting. The following two data sets were selected:

- data set A contained 75% of all data (90 cases) used for model development,
- test data set B contained 25% of all data (30 cases) used for overfitting detection.

For data sets A and B, the root mean squared errors RMSE(A) and RMSE(B) were calculated using the following formula:

$$RMSE(A\ or\ B) = \sqrt{\frac{\sum(LBP-LBP_e)^2}{n}} \qquad (7)$$

where:
LBP – length between perpendiculars from the data set,
LBP_e – estimated length between perpendiculars,
n – number of ships in set A or B.

In this study it was assumed that overfitting occurs when the root mean squared error related to test set B is 20% higher than the root mean squared error related to set A. To estimate the overfitting, the ratio of the error RMSE(B) to RMSE(A) was calculated.

The algorithm schemes are shown in Figures 2 and 3. As shown in Figure 2, during the first step, data for sets A and B were randomly selected, and the simplest model (5) was initially defined. Next, the best functions f1–f4 and regression coefficient values were discovered through random searching, based on the algorithm shown in Figure 3. Better functions were selected based on a higher correlation coefficient R value.

As shown in Fig 2, after the initial development of the model (5), a standard error (SE) value was checked. The research assumed that the SE value limit was 7.7 m (ship length).

If the SE was greater than this limit value, the model was randomly expanded in the next step. After this expansion, the occurrence of overfitting phenomena was checked by computing the ratio RMSE(B)/RMSE(A). If this ratio value was greater than 1.2, the least statistically significant component was removed from the model. This procedure was looped until these two conditions were met.
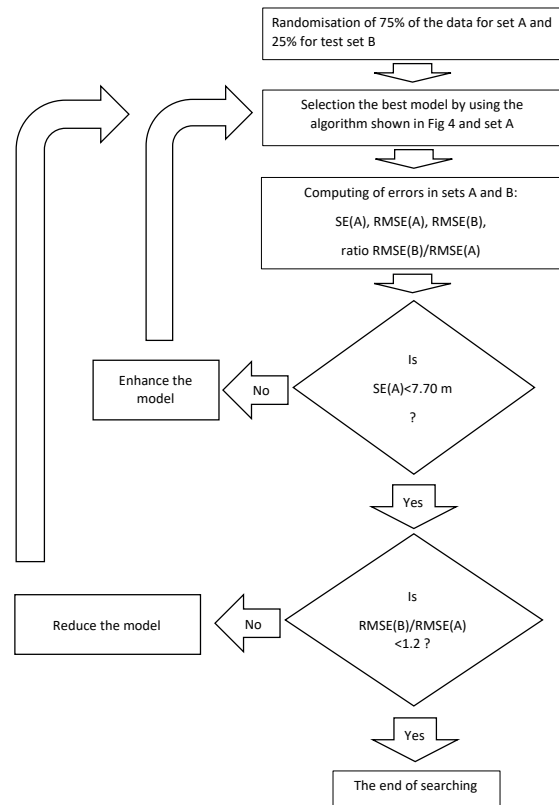

Fig. 2. The general algorithm scheme, where: A – data set, B – test data set, SE(A) - standard error related to set A, RMSE(A) – root mean squared error related to set A, RMSE(B) - root mean squared error related to test set B
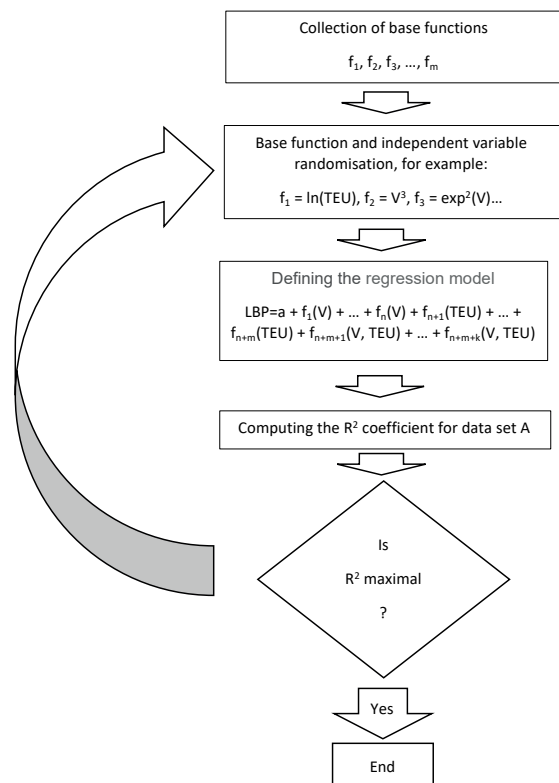

Fig. 3. The general algorithm scheme for searching the model for the maximum R-squared value, where f – base function, n, m, k – the number of functions

RESULTS

ARTIFICIAL NEURAL NETWORKS

Among all the neural network types, the multiple perceptron (MLP), which has two neurons in the input layer, 11 neurons in the hidden layer, and one neuron in the output layer, was the most accurate. Table 2 shows the statistical data of this neural network broken down by teaching, validation, and testing sets. The mathematical form of this network is given by the formulas in Eqs. (8), (9), (10), (11), (12), and (13).

$$LBP = \frac{c}{0.00318} + 0.22 \qquad (2)$$

where:
c – the variable, calculated as follows:

$$(3)$$

$$c = [0.098, 0.297, -1.485, -0.032, -0.176, -1.325, -0.529, -0.702, -0.727, -0.494, 4.206] \times A + 1.386$$

where
A – column matrix:

$$A = \begin{bmatrix} \frac{1}{1+e^{-a_1}} \\ \frac{1}{1+e^{-a_2}} \\ \frac{1}{1+e^{-a_3}} \\ \vdots \\ \frac{1}{1+e^{-a_{11}}} \end{bmatrix} \qquad (4)$$

where:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ a_{11} \end{bmatrix} = \begin{bmatrix} 0.899 & -0.219 \\ 0.416 & -0.965 \\ -0.993 & -0.426 \\ -0.036 & -0.159 \\ 0.486 & -0.604 \\ -0.739 & 0.330 \\ 0.704 & 0.001 \\ -0.575 & 0.053 \\ 0.438 & -0.448 \\ -0.672 & 0.541 \\ 7.409 & 0.280 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} - \begin{bmatrix} 0.007 \\ 0.509 \\ -0.032 \\ 0.822 \\ 0.764 \\ 0.320 \\ 0.675 \\ 0.735 \\ -0.003 \\ -1.239 \\ -2.186 \end{bmatrix} \qquad (5)$$

where:

$$b_1 = TEU \cdot 4.89 \cdot 10^{-5} + 0.00489 \qquad (6)$$

$$b_2 = V \cdot 0.0654 + 0.5882 \qquad (7)$$

Tab. 2. Statistical parameters of an ANN

	Training set	Validation set	Testing set
Mean absolute error (MAE)	8.49	8.12	7.88
Correlation	0.99	0.99	0.99

RANDOM SEARCH METHOD

Figures 4 and 5 show the process of discovering an equation for estimating the length between perpendiculars using the random search method. The SE value related to set A, the ratio RMSE(B)/RMSE(A), and the number of model elements are shown through the model evolution. In the first phase, the model was inaccurate and the SE value was high. Therefore, the model was expanded to seven elements in the next phase. This model extension reduced the SE error but increased the RMSE(B)/RMSE(A) ratio which overfitted the model. In the next phase, the model was reduced to four elements to avoid overfitting. However, the model accuracy was reduced after this procedure. Finally, increasing the model element number to five and finding the most accurate functions allowed the successful completion of this search. Eq. (14) was discovered:

$$LBP = 54.296 + 2.656 \cdot TEU^{1/2} + 1.4E\text{-}06 \cdot V^{5.6} - 2.821E\text{-}21 \cdot TEU^{5.6} \cdot V^{-1.8} - 1.116E\text{+}08 \cdot TEU^{-1.3} \cdot V^{-4} - 1.007E\text{-}04 \cdot TEU^{0.4} \cdot V^{3.1} \qquad (8)$$

Eq. (14) is characterised by high accuracy for set A, i.e.:
- R-squared = 0.993,
- Standard error SE = 7.69 m.

The error difference between set A and test set B is less than 20%, i.e.:
- Root mean squared error related to set A RMSE(A) = 7.43 m,
- Root mean squared error related to set test B RMSE(B) = 8.69 m,
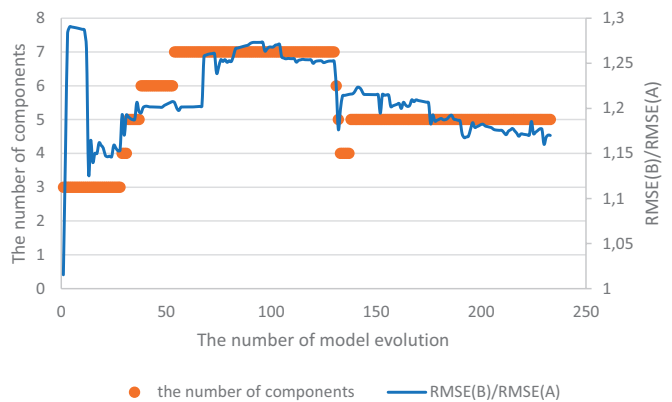- the ratio of RMSE(B) to RMSE(A) = 1.17.

Fig. 2. The ratio of the root mean squared error in test set RMSE(B) to data set RMSE(A), and the number of model elements through the model evolution
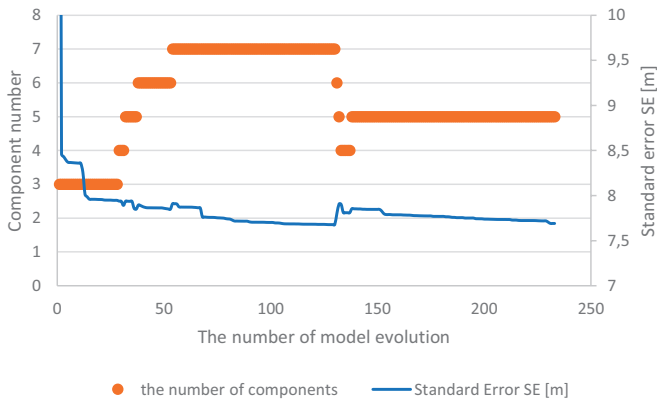
Fig. 5. Standard error (SE) value in set A, and the number of model elements through the model evolution

Fig. 3. Length (LBP) values estimated in the full range of TEU and V calculated using Eq. (8)

DISCUSSION

EVALUATION OF MODELS IN TERMS OF ESTIMATION ACCURACY

Figures 6 and 7 show the estimated length (LBP) values in the full range of TEU capacity and ship velocity calculated using Eqs. (8) and (14). The surfaces presented on these figures look similar, though for the extreme values of velocity and TEU capacity, the length calculated using an ANN was slightly larger than the one obtained using the random search method.

Figure 8 compares the length calculations using both methods for ship sample data for selected speeds. As shown in Figure 8 both methods gave similar results; Eq. (14) provided slightly more accurate results in relation to sample ship data at a speed of 11 kts.

Figure 9 illustrates the estimates obtained using both methods compared to test sample ship data. The test data was in the range of full ship length. This figure shows that the length values calculated using both methods were close to the perfect fit line.

Table 3 shows the root mean squared error (RMSE) and Pearson R-squared coefficient values relating to regression Eq. (14) and the developed neural network referenced for the entire data set (including the training, validation, and test sets). This table shows that Eq. (14) is characterised by a RMSE estimation error 2 metres smaller than Eq. (8).

Tab. 2. The values of the root mean squared RMSE error and Pearson R-squared coefficients relating to regression (14) and the developed neural network referenced for the entire data set

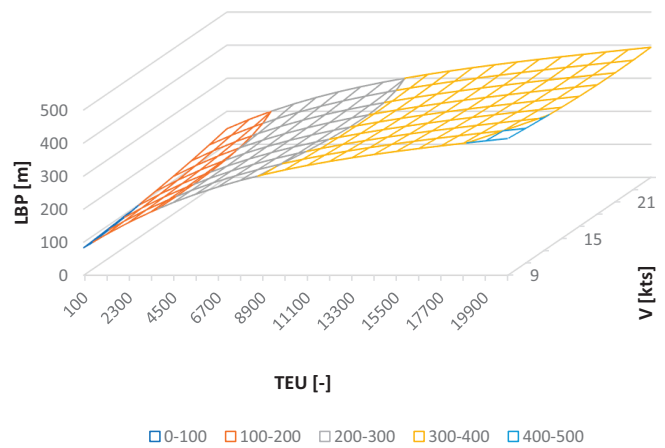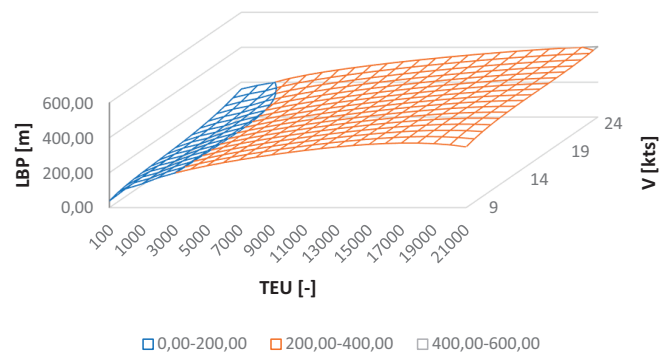| Equation	RMSE [t]	R-Pearson [-]
equation (14)	7.77	0.996
equation (8)	9.86	0.994

Fig. 4. Length (LBP) values estimated in the full range of TEU and V calculated using Eq. (14)
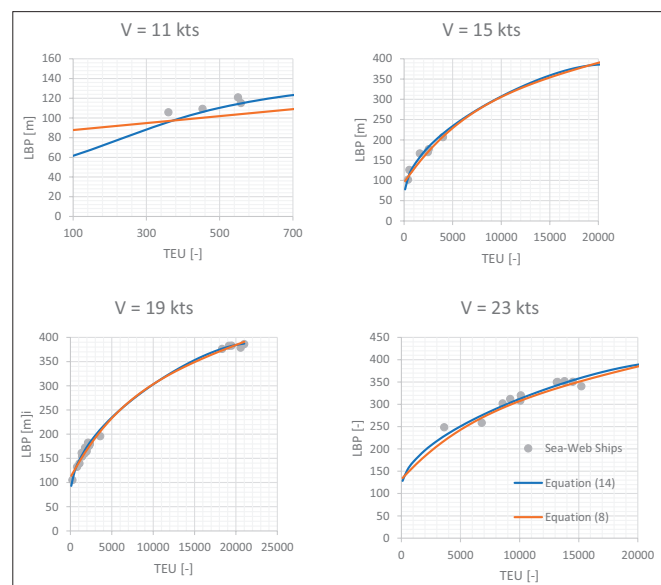

Fig. 5. Comparison of length (LBP) values estimated with sample ship data using Eqs. (8) and (14)
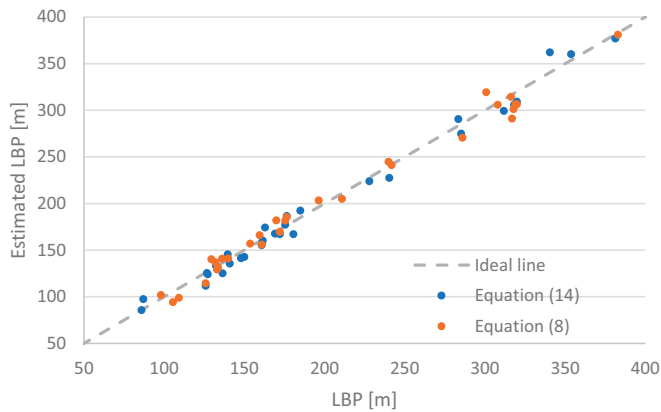
Fig. 6. Comparison of length (LBP) values calculated with sample data values using both methods

AN EVALUATION OF METHODS IN TERMS OF EASE OF COMPUTATION AND MODEL SIMPLICITY

The equations presented in this article were developed for practical use by a ship designer. In this respect, one important factor was the possibility of using these formulas for manual calculations by a person with a basic knowledge of computer techniques. For this reason, model simplicity was an important criterion.

Research showed that Eq. (14) developed using the random search method had a simpler form than Eq. (8) developed using neural networks. Eq. (14) is based on only seven basic functions and six coefficients. The analytical relationships of independent variables are also clearly shown in this equation. The practical use of this equation requires only a basic knowledge of computer techniques. A scientific calculator or a simple spreadsheet can be used for calculating the LBP length based on Eq. (14).

In contrast, Eq. (8), which was developed using an ANN, is much more complex. Several calculations should be performed using the formulas in Eqs. (8) – (13) to estimate the length between perpendiculars in this case. The relationships between TEU capacity, ship velocity, and ship length presented in the formulas of Eqs. (9) - (12) are unclear. Eq. (8) ANNis more difficult to implement by a normal user. Indeed, the user must have an advanced knowledge of computer techniques to implement the ANNANN model, or alternatively, a specialised computer software may be used. The use of a scientific calculator to estimate LBP length using Eqs. (8) - (13) is more complicated and time-consuming than simply using Eq. (14).

THE EVALUATION OF METHODS FOR THE USE OF DATA

Both methods presented here use heuristic techniques, so these methods do not offer an optimal solution. The complex equations developed using an ANN and a random search method may lead to overfitting. Therefore, overfitting was detected using a test set which included 25% of the data in both methods. An ANN and random search method are the same in this respect.

However, an additional validation data set was used to develop the ANN. This validation set also included 25% of the data. In this research, ANNs lost 50% of data in total for validation and testing.

In contrast, the random search method did not require a validation set and lost only 25% of the data during the overfitting detection. In this regard, the random search method provides a more effective use of the data set than the ANN.

Table 4 shows a summarised method comparison in terms of different properties.

Tab. 3. A summarised method comparison in terms of different properties

Method	Equation no	Accuracy	Simplicity	Ease of computation	Data loss
ANN	(8)	high	complex	difficult	50%
Author's method	(14)	high, RMSE 2 metres smaller	simple	easy	25%

CONCLUSIONS

In ship design, only DWT or TEU capacity are usually used to estimate a container ship's length. Over the last few years, economic and environmental factors have affected ship owner requirements. At present, ship velocity may be a second key design parameter in addition to TEU or DWT capacity. Therefore, alternative design equations for estimating a container ship 's length based on TEU capacity and ship velocity have been proposed in this research. This provides a new approach in naval engineering. The equations presented in this work were developed based on the data of the most recent standard container ships built since 2014.

In this article, ANNs and a random search method based on MNLR were applied to estimate a container ship's length. The conclusions drawn from these compared methods may be summarised as follows:

- Both methods are characterised by a high estimation accuracy. The random search method is slightly more accurate and offers a RMSE error value less than 2 metres in length.
- Eq. (14) developed using the random search method is simpler and easier to compute than Eq. (8) developed using an ANN.
- The random search method used the data set more effectively than the ANN.
- The random search method also only used 25% of the data for testing while the ANN needed 50% of the data for validation and testing.

Estimates acquired with an equation developed using Multiple Nonlinear Regression (MNLR) may be as accurate as ones obtained using ANNs. The application of heuristic techniques for the development of MNLR by variable and

function randomisation automatically enables the discovery of a set of equations.

The methods presented in this article may be used to estimate the parameters of a container ship, such as breadth, side depth, or draught. In general, in the case of volume carriers, these dimensions are primarily determined by the ratio of length to breadth and side depth, and the displacement of the vessel. However, it seems that the use of these estimates could be helpful to assess the accuracy of the design calculations. The results presented in this paper and in [4] indicate the possibility of developing a neural network to predict all dimensions of a container ship and, potentially, other types of ships.

The algorithms described here may have practical applications for the commercial design of container vessels. However, these formulas can be inaccurate for the design of an innovative container ship and can only be used to estimate the length of container ships with design characteristic ranges listed out in Table 1. The use of algorithms to determine a ship's length for characteristics outside these ranges may possibly be associated with less reliable calculations.

REFERENCES

1. D.G.M. Watson, "Practical Ship Design", Volume 1. Elsevier Science. 1998.

2. K.J. Rawson and E.C. Tupper, Basic Ship Theory: Ship Dynamics and Design. Butterworth-Heinemann. 2001.

3. A. Papanikolaou, Ship Design: Methodologies of Preliminary Design. Dordrecht: Springer. 2014.

4. J.H. Evans, Basic Design Concepts, Naval Engineers Journal, 1959.

5. D.J. Andrews, An Integrated Approach to Ship Synthesis, Trans. RINA. 1985.

6. W. Chądzyński, "Elements of contemporary design methods of floating objects", Scientific Reports of Szczecin University of Technology. (in Polish). 2001.

7. G.P. Piko, Regression Analysis of Ship Characteristics, Canberra: C. J. THOMPSON Commonwealth Government Printer. 1980.

8. H.O. Kristensen, "Determination of Regression Formulas for Main Dimensions of Container Ships based on IHS Fairplay data". Project no. 2010-56, Emissionsbeslutning sstøttesystem. Work Package 2, Report no. 03. Technical University of Denmark. 2013.

9. Sea-web Ships, IHS Markit, https://maritime.ihs.com, (accessed 8 Decemeber 2020)

10. S. Gurgen, I. Altin, I. and O. Murat, "Prediction of main particulars of a chemical tanker at preliminary ship design using artificial neural network", Ships and Offshore Structures, 13(5), 459-465, 2018.

11. A.D. Alkan, K. Gulez and H. Yilmaz, "Design of a robust neural network structure for determining initial stability particulars of fishing vessels", Ocean Engineering 31 761–777, 2004.

12. S. Ekinci, U.B. Celebi, M. Bal, M.F. Amasyali and U.K. Boyaci, "Predictions of oil/chemical tanker main design parameters using computational intelligence techniques", Applied Soft Computing, 11 2356–2366, 2011.

13. T. Abramowski, "Application of Artificial Intelligence Methods to Preliminary Design of Ships and Ship Performance Optimization", Naval Engineers Journal, 125 (3) 101-112, 2013.

14. T. Cepowski, "The prediction of ship added resistance at the preliminary design stage by the use of an artificial neural network", Ocean Engineering 195, 106657. 2020. doi.org/10.1016/j.oceaneng.2019.106657

15. J.G. Song, L.H. Liang, S.T. Zhang and J.M. Wang, "Design and experimental investigation of a GA-based control strategy for a low-speed fin stabilizer", Ocean Engineering, 2018. 10.1016/j.oceaneng.2020.108234

16. V. Sahin and N. Vardar, "Determination of Wastewater Behavior of Large Passenger Ships Based on Their Main Parameters in the Pre-Design Stage", Journal of Marine Science and Engineering, 8 (8), 2020. DOI: 10.3390/jmse8080546

17. Luan Thanh Le, Gunwoo Lee, Keun-Sik Park and Hwayoung Kim, "Neural network-based fuel consumption estimation for container ships in Korea", Maritime Policy and Management, 47:5, 615-632, 2020. DOI: 10.1080/03088839.2020.1729437

18. X. Cheng, S. Chen and C. Diao, "Simplifying Neural Network Based Model for Ship Motion Prediction: A Comparative Study of Sensitivity Analysis", Proceeding of the ASME 36th International Conference on Ocean, Offshore and Arctic Engineering, Vol. 1. 2017.

19. S. Haykin, "Neural Networks: A Comprehensive Foundation". New York: Macmillan Publishing. 1994.

20. J.H. Lee, T. Delbruck and M. Pfeiffer, "Training Deep Spiking Neural Networks Using Backpropagation", Frontiers in Neuroscience, 10, 508, 2016.

21. L. Fausett, Fundamentals of Neural Networks. New York: Prentice Hall. 1994.

22. D. Patterson, Artificial Neural Networks. Singapore: Prentice Hall. 1996.

23. C. Bishop, Neural Networks for Pattern Recognition, Oxford: University Press. 1995.

24. J. Shepherd, Second-Order Methods for Neural Networks. Springer: New York. 1997.

25. J. Adamowski, H.F. Chan, S.O. Prasher, B. Ozga-Zielinski and A. Sliusarieva, "Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal", Water Resources Research, 48W01528, 2012.

26. E. Tsykin, "Multiple nonlinear regressions derived with choice of free parameters, Applied Mathematical Modelling", Volume 8 (4) 288-292, 1984.

27. T. Cepowski, ndCurveMaster n-dimensional Curve and Surface Fitting Software, https://www.ndcurvemaster.com, (accessed 13 December 2020).

CONTACT WITH THE AUTHORS

Tomasz Cepowski
e-mail: t.cepowski@am.szczecin.pl

Akademia Morska w Szczecinie,
W.Chrobrego 1-2, 70-500 Szczecin,
Poland

Paweł Chorab
e-mail: p.chorab@am.szczecin.pl

Akademia Morska w Szczecinie,
W.Chrobrego 1-2, 70-500 Szczecin,
Poland

Dorota Łozowicka
e-mail: d.lozowicka@am.szczecin.pl

Akademia Morska w Szczecinie,
W.Chrobrego 1-2, 70-500 Szczecin,
Poland